

THE BOOK WAS DRENCHED

Tight Binding book.

UNIVERSAL
LIBRARY

OU_162900

UNIVERSAL
LIBRARY

TO
NEW YORK UNIVERSITY
IN HONOUR OF ITS
CENTENARY
1931

PREFACE TO THE ENGLISH EDITION

THE two following reports on the Principles of the Quantum Theory first appeared in two volumes of the well-known series "Ergebnisse der Naturwissenschaften" (Jul. Springer, Berlin). The authors have the more readily welcomed the suggestion of Messrs. Methuen & Co. Ltd. to allow them to appear in one volume in an English translation, by Professor Henry Brose, as they felt assured that the work was being placed in experienced hands. Nevertheless, the authors feel constrained to admit that they themselves entertain certain misgivings about publishing reports in book form,—misgivings which will no doubt be felt generally.

A number of text-books on the quantum theory have already appeared to date, some of which give a full and detailed account of the new theoretical ideas. The coherence and completeness achieved by such works cannot, of course, be aimed at in the present book, nor is it to be expected. Although the authors have endeavoured to bring into clear relief all the physically important ideas that have played a part in the development of the new quantum mechanics, yet the book in its present form cannot claim to do more than to record a stage in the development of the quantum theory both with respect to the view it embodies and the facts it contains.

The appearance of the book in English may perhaps be justified by the manner in which the subject is presented, but also by the interfusion of the theoretical principles with the physical principles of the theory. We believe that this has

not been carried out to the same extent in most of the other accounts that have come to our notice. May we finally express the hope that the book will also secure a place as a critical work, such as was accorded in "Nature" recently by a leading English physicist, *à propos* of a similar occasion.

OTTO HALPERN
H. THIRRING

NEW YORK
VIENNA

THE translator wishes to express his warmest thanks to the authors for their kind help in reading and checking the proofs, and also to his friend and colleague, Dr. L. G. H. Huxley, for undertaking the final reading of the proofs.

HENRY L. BROSE

UNIVERSITY COLLEGE
NOTTINGHAM

CONTENTS

PART I

THE COURSE OF DEVELOPMENT UP TO 192

CHAP.		PAGE
I.	1. Introduction	I
II.	THE FIELD OF RADIATION OF REVOLVING ELECTRONS ACCORDING TO CLASSICAL ELECTRODYNAMICS	3
	2. The Formulation of the Problem	3
	3. The Fundamental Equations of Mechanics	4
	4. Hertz's Formulæ	8
	5. Criticism of the Preceding Reflections	12
III.	BOHR'S THEORY	14
	6. The Existence of Stationary States	14
	7. The Correspondence Principle	15
	8. The Quantum Conditions	16
	9. The Achievements of Bohr's Theory	18
IV.	HEISENBERG'S QUANTUM MECHANICS	20
	10. The Fundamental Ideas of Heisenberg's Theory	21
	11. The New Kinematics	22
	12. Calculations with Matrices	25
	13. The Commutation Law	29
	14. The Mechanical Equations of Motion	30
	The Integration of the Quantum-mechanical Equations of Motion	31
	15. Systems with Several Degrees of Freedom	32
	16. Survey of the Matrix Method	32
	The Achievements of Quantum Mechanics	33
V.	THE MATERIAL WAVES OF DE BROGLIE	35
	17. The Allocation of a Vibration Frequency to the Electron	35
	18. First Step: the Phase Waves	36
	19. Second Step: the Wave Groups	38
	20. The Optical-mechanical Analogy	40
	21. The Interpretation of Bohr's Quantum Conditions	41
	22. The Occurrence of Interference Phenomena in the Case of Electrons	41
	23. The Conception of Phase-waves as a Guide to Wave-mechanics	42
VI.	THE WAVE-MECHANICS OF SCHRÖDINGER	44
	24. The Wave-equation of the One-electron Problem	44
	25. The Idea of Proper Values	46
	26. Application to the Hydrogen Atom	47

CHAP.		PAGE
27.	The Interpretation of the Process of Radiation by means of Wave-mechanics	48
28.	The Physical Meaning of the Field-scalar ψ	50
	Calculation of the Intensities of Spectral Lines	51
29.	Generalization of the Wave-equation	52
30.	Criticism of Wave-mechanics	54

PART II

THE FURTHER DEVELOPMENTS SINCE 1926

31.	Introduction	57
VII.	RELATIVITY AND ELECTRON SPIN	59
32.	Recapitulation. Application of the Wave-equation to Conservative Systems	59
33.	The Wave-equation Depending on the Time	61
34.	The Relativistic Wave-equation according to Schrödinger	65
35.	Survey of the Development of the Theory of Multiple Quantizing	69
36.	The Hypothesis of Electron Spin	76
37.	The Newer View of the Fine-structure of the Hydrogen Spectrum	81
38.	The Wave-mechanical Four-current	84
39.	Dirac's Theory of the Spinning Electron in the Absence of External Fields	90
40.	Dirac's Theory of the Spinning Electron in the Presence of External Fields	97

ACHIEVEMENTS OF THE THEORY

VIII.	ONE-BODY PROBLEMS	102
41.	Oscillator and Rotator. Methods of Calculation	102
42.	Theory of Perturbations of Conservative Systems	108
43.	The Unperturbed Hydrogen Atom	114
44.	Influence of External Fields on the Hydrogen Atom	119
45.	Problems Concerning the Diffraction of Matter	123
46.	Crossing a Potential Barrier	128
47.	Theory of Perturbations of Non-stationary Systems	131
48.	Theory of Dispersion	134
49.	Photo-electric Effect. Compton Effect	137
IX.	MANY-BODY PROBLEMS	142
50.	Exchange Degeneracy. A Special Example	142
51.	Degeneration of Particles. The General Proof for the Problem of n Bodies	146
52.	The Helium Atom	148
53.	Theory of Band-spectra and Nuclear Spin	152
54.	Calculation of the Hydrogen Molecule	155
55.	The Prospects Opened up by the Theory of Groups and Chemistry	157
56.	The Auger Effect	158
57.	Preliminary Remarks on Classical Statistics	159
58.	The New Statistics	161
59.	Statistical Methods of Calculation. The Statistics of Bose-Einstein and Fermi-Dirac	162
60.	Theory of the Ideal Gas	165
61.	Metallic Electrons as a Degenerate Fermi-gas	171
62.	Photo-electric Limit and Richardson Effect.	173

CONTENTS

xi

CHAP.	PAGE
63. The Paramagnetism of Metals	175
64. The Heat of Rotation of the Hydrogen Molecule	176
65. The Theory of Ferro-magnetism	178
 X. INTERPRETATIONS OF THE THEORY	 181
66. Enunciation of the Problem	181
67. Heisenberg's Original Formulation. Schrödinger's Wave- mechanics in its Extended Sense	182
68. The Statistical Interpretation of Quantum Mechanics. The Transformation Theory	186
69. A More Detailed Discussion of the Statistical Interpretation. An Imaginary Experiment by Einstein.	188
70. Heisenberg's Uncertainty Relation. Purely Mechanical Problems	192
71. Heisenberg's Uncertainty Relation. Mutual Action between Radiation and Matter	197
72. Heisenberg's Uncertainty Relation. Explanation of Ap- parent Paradoxes in Mechanical Problems	201
73. Concluding Remarks on Wave-mechanics	205
 BIBLIOGRAPHY AND REFERENCES	 209
 INDEX	 213

THE ELEMENTS OF THE NEW QUANTUM MECHANICS

PART I

THE COURSE OF DEVELOPMENT UP TO 1926

CHAPTER I

INTRODUCTION

1. ALTHOUGH the quantum theory of spectra founded by Niels Bohr in 1913 has impressed its characteristic form on physical research for over fifteen years, its creator has never regarded it as other than a provisional structure. By making certain postulates (the existence of stationary radiationless states, the frequency condition, quantum transitions, the correspondence principle), it became possible to interpret the genesis of the most important types of optical and Röntgen spectra and, still more, to obtain valuable information about the structure of atoms, valency, the periodic system and other questions. The fundamental postulates themselves, however, were factors alien to the structure of theoretical physics. Their crass contradiction with the ideas of classical electrodynamics and mechanics was tolerated only on account of the extraordinary successes of the theory, which had disclosed unsuspected relationships between spectral lines, chemical properties and so forth. The effect of these successes had become so powerful that the majority of physicists had become accustomed to accepting the ideas of quantum orbits and electronic transitions as obvious, and to make unhampered use of them without realizing with sufficient clearness the deep chasm which lay between these ideas and the conceptions underlying the rest of physics.

The position became changed, however, when the continued application of the original "classical" quantum theory more and more often encountered obstacles before which it showed itself simply powerless. In the problem of the double helium spectrum, in the multiple structure of alkali spectra, in the

anomalous Zeeman effect, in the problem of dispersion, difficulties of a serious nature arose which called into life again the latent need for a re-moulding of the theory from its very foundations. This process of reconstruction was envisaged in 1925 and 1926 from two totally different angles: further, the two theories proposed led, in spite of their radically different starting-points, to practically the same results. The "Quantum Mechanics" founded by Heisenberg and elaborated by Born, Jordan and Dirac may be regarded as a kind of phenomenological theory, inasmuch as it regards as its problem the establishing of relationships between *quantities which, at least in principle, are observable*. The "Wave Mechanics" founded by Schrödinger, which follows directly on the ideas of de Broglie, stands in strong contrast to Heisenberg's theory, for it operates with new conceptions and abstractions in principle inaccessible to direct observation, such as, for example, with waves of unknown physical meaning in multi-dimensional spaces, and so transgresses consciously those rules which governed Heisenberg in the formation of his theory. Nevertheless wave-mechanics leads ultimately to the same results as quantum mechanics and, indeed, in a manner which renders it much more amenable to mathematical treatment, so that the introduction of the fictions mentioned is shown to be very helpful.

The account of these theories here given is intended to meet the need felt by numerous physicists to grasp the fundamental ideas of these new and not easily understood theories. For this purpose we attach more importance to a wide and intelligible treatment rather than to an exact and complete presentation. The skeleton of the two theories is to be stripped of all encumbrances and shown in its most elementary form. In the first part of our account we restrict ourselves to the first stage of the development of the theory, which was reached about the autumn of 1926. The second stage, due to later work by Heisenberg and, in particular, Dirac, will be treated in the sequel.

CHAPTER II

THE FIELD OF RADIATION OF REVOLVING ELECTRONS ACCORDING TO CLASSICAL ELECTRODYNAMICS

2. The Formulation of the Problem

IN order better to understand the new theories we must go back a fair way and—even at the risk of repeating what is well known—begin by contrasting the rules of classical electrodynamics for calculating the radiation emitted by an atomic system, with those of the quantum theory. Suppose we have a system of point-charges, after the manner of the nuclear model of Rutherford and Bohr, consisting of an atomic nucleus and one or more electrons that execute orbits around the nucleus under the influence of their mutual forces. According to classical electrodynamics the problem of calculating the radiation is equivalent to that of determining the electromagnetic field generated by the charges. The fundamental equations of electrodynamics prescribe perfectly definite rules for calculating this field, so that for a given system of charges and under given initial conditions there always exists a definite field and therefore also a definite radiation. The actual process of calculation is, however, highly intricate and cannot be exactly performed except in the very simplest cases. For the field at a definite point of observation P depends on the position and the velocity of each individual electron at the time $t - r/c$.^{*} But the position and velocity of the individual electrons is again conditioned by the electromagnetic field yet to be calculated. Hence although solutions of the field-equations and the equations of motion always exist, and although the motion of the system in the sense of classical electrodynamics would thus be completely determined when the initial positions and velocities of the electrons are given, these solutions are not found in practice, and therefore we must resort to methods of approximation. The first step towards such an approximation consists in neglecting the electromagnetic forces

^{*} Retarded time (*Latenzzeit*): it indicates that moment of time which is previous to the moment "now" by an interval of time equal to that required by light to travel from the place at which the charge is situated to the point P under consideration.

in comparison with the purely electrostatic forces. This step is justifiable in that (as Bohr argued in his earliest papers) under the conditions that presumably obtain within the atom, the electrical forces predominate over the magnetic forces. Since Coulomb's elementary law of electrostatics is formally equivalent to Newton's law of gravitation, the problem of calculating the electronic orbits, when the magnetic forces are neglected, becomes reduced to the well-known astronomical problem of determining the orbits of planets, which, however, for its part, too, is not exactly soluble for more than two bodies.

3. The Fundamental Equations of Mechanics

For the sake of completeness we shall here recapitulate those fundamental laws of mechanics to which we shall frequently need to refer later when discussing the theories of Heisenberg, de Broglie and Schrödinger. We are essentially concerned with Hamilton's formulation of the fundamental laws of mechanics, on which these theories all rest. At the same time these are the laws which, in principle, enable us to perform the mechanical part of the problem above proposed, namely, the calculation of the motions of the electrons under the influence of the nuclear field and of the forces of the remaining electrons.

We restrict our attention to conservative systems, that is, to those whose total energy E is constant. If we call the kinetic energy of the system T and the potential energy V , we must then have

$$T + V = E. \quad (1)$$

We designate the co-ordinates of the system, which may be either Cartesian or generalized of any sort by $q_1, q_2 \dots q_f$ (where f is the number of degrees of freedom of the system). The kinetic energy $T(q, \dot{q})$ is a quadratic form of the $\dot{q}$'s. Here the dots, as usual, signify differential coefficients with respect to time, that is, $\dot{q} = \frac{dq}{dt}$. As is well known, we then introduce in place of the velocities $\dot{q}$ the generalized momenta

$$p_k = \frac{\partial T}{\partial \dot{q}_k}, \quad (2)$$

and, with the help of eqn. (2), we eliminate the velocities out of the expression for the total energy, which thus, in the case of conservative systems, passes over into the Hamiltonian function designated by $H(p, q)$:

$$T + V = H(p, q). \quad (3)$$

In other words, the Hamiltonian function in the case of a conservative system is nothing more than the expression for the total energy in terms of only the generalized co-ordinates and generalized momenta of the system.

Examples.—From the expression for the kinetic energy of a point-mass of mass μ ,

$$T = \frac{\mu}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2), \quad (4)$$

we get for the momenta in Cartesian co-ordinates $p_x = \mu\dot{x}$, etc., and the Hamiltonian function for the one-body problem is

$$H(p, q) = \frac{1}{2\mu}(p_x^2 + p_y^2 + p_z^2) + V(x, y, z). \quad (5)$$

For a point-mass that can move along a straight line and that is attracted by a force $-ax$ to its position of rest $x=0$ when no force is acting (this is the case of the linear harmonic oscillator) this equation becomes

$$H(p, q) = \frac{1}{2\mu}p^2 + \frac{a}{2}q^2. \quad (6)$$

For the corresponding three-dimensional problem (harmonic anisotropic oscillator) the corresponding formula is

$$H(p, q) = \frac{1}{2\mu}(p_x^2 + p_y^2 + p_z^2) + \frac{1}{2}(ax^2 + by^2 + cz^2). \quad (6a)$$

In polar co-ordinates the kinetic energy of the point-mass is expressed by

$$T = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2). \quad (7)$$

By using (2) we get

$$p_r = \mu\dot{r}, \quad p_\theta = \mu r^2 \dot{\theta}, \quad p_\phi = \mu r^2 \sin^2 \theta \dot{\phi}. \quad (8)$$

Hence, if e denotes the elementary electronic charge, Hamilton's function for the model of the hydrogen atom analogous to the simplest planetary system is

$$H(p, q) = \frac{1}{2\mu} \left(p_r^2 + \frac{1}{r^2} p_\theta^2 + \frac{1}{r^2 \sin^2 \theta} p_\phi^2 \right) - \frac{e^2}{r}. \quad (9)$$

It is known that the fundamental laws of mechanics may be derived from a variation principle which asserts that a certain integral taken along the actual path of the motion has an extreme value as compared with all the integrals with the same integrands

6 RADIATION OF REVOLVING ELECTRONS

taken over any other kinematically possible paths. This principle of variation may be formulated in various ways. One of the oldest formulations is the Principle of Least Action of Maupertuis, which states

$$\int_A^B T dt = \text{a minimum}, \quad (10)$$

where the paths permissible for comparison comprise those which occur between the same limits of the space co-ordinates and have the same total energy E . For a single point-mass we have

$T = \frac{\mu v^2}{2}$, so that on account of $v dt = ds$ we may also write the condition (10) in the form

$$\int_A^B \mu v ds = \text{a minimum}, \quad (11)$$

a relation which plays an important part in the theory of de Broglie.

The condition necessary for the fulfilment of this property of being an extreme value (Euler's conditions of variation) may be brought by means of the Hamiltonian function $H(p, q)$ into the form of the fundamental *canonical equations* of mechanics:

$$\dot{q}_k = \frac{\partial H(p, q)}{\partial p_k}, \quad \dot{p}_k = - \frac{\partial H(p, q)}{\partial q_k}; \quad k = 1, 2 \dots f. \quad (12)$$

To integrate these equations we often find it expedient to use the fact that the canonical equations do not change their form if in place of the co-ordinates q_k and the momenta p_k we introduce new variables $\bar{q}_k$ and $\bar{p}_k$ by means of a transformation which fulfils certain conditions.* After such a *canonical transformation*

$$(p, q) \rightarrow (\bar{p}, \bar{q}) \quad (13)$$

has been effected the transformed equations run

$$\dot{\bar{q}}_k = \frac{\partial \bar{H}(\bar{p}, \bar{q})}{\partial \bar{p}_k}, \quad \dot{\bar{p}}_k = - \frac{\partial \bar{H}(\bar{p}, \bar{q})}{\partial \bar{q}_k}. \quad (12a)$$

The canonical transformations then lead directly to an integration of the fundamental equations (12) or (12a), if we can arrange for the transformed Hamiltonian function $\bar{H}$ to depend only on the $\bar{p}_k$'s and not on the $\bar{q}_k$'s as well. Thus

$$\bar{H} = \bar{H}(\bar{p}_1 \dots \bar{p}_f). \quad (14)$$

* Concerning the characteristics of these canonical transformations, see M. Born, "Lectures on Atomic Mechanics," § 7.

In accordance with (12a) we then have

$$\dot{\bar{p}}_k = 0 \quad \text{or} \quad \bar{p}_k = \text{const.} = \alpha_k. \quad (15)$$

Further,

$$\dot{\bar{q}}_k = \frac{\partial \bar{H}}{\partial \bar{p}_k}(\bar{p}_1 \dots \bar{p}_f). \quad (16)$$

The right-hand side of this equation is constant on account of (15). We call this constant ν_k , thus

$$\frac{\partial \bar{H}}{\partial \bar{p}_k} = \nu_k \quad (17)$$

and the integrals of the equations of motion are then

$$\bar{q}_k = \nu_k t - \beta_k, \quad \bar{p}_k = \alpha_k. \quad (18)$$

The considerations that now follow are restricted to *separable conditionally periodic motions*: in these motions the original co-ordinates q_k of the system change periodically with the time, but each one changes with a different length of period so that in general the initial point is never traversed a second time. (*Example*.—the Lissajou figures of the plane anisotropic oscillator

$$x = a \cos \omega_1(t - \delta_1), \quad y = b \cos \omega_2(t - \delta_2). \quad (19)$$

If ω_1 and ω_2 stand in no rational ratio to each other (that is, if they are incommensurable) the motion does not repeat itself; if their ratio is rational the orbit will pursue its old track after a finite number of revolutions. In the latter case the system is called *degenerate*.) The transformed variables that lead to the solutions (18) do not, however, change periodically with the time, since according to (18) the $\bar{p}_k$'s are constants and the $\bar{q}_k$'s are linear functions of the time. Hence in view of the periodicity of the motion the $\bar{q}_k$'s have the character of angles or of functions of angles, for example the area swept out by the radius vector. If the initial co-ordinates are to be periodic functions of the time, then in conformity with (18) they must also be periodic functions of the $\bar{q}_k$'s, and we can choose the units of the $\bar{q}_k$'s so that the length of the periods becomes equal to one. The quantities $\bar{q}_k$ that have been normalized in this way are then called *angle variables* w_k (also called cyclic co-ordinates), and the associated "canonically conjugate" momenta $\bar{p}_k$ are called the *action variables* I_k . The quantities ν_k are then the frequencies of the motion of revolution, and, in accordance with (17), we have

$$\nu_k = \frac{\partial H(I_1 \dots I_f)}{\partial I_k}. \quad (17a)$$

This relation plays an important part in Bohr's form of the quantum theory.

In the case of the planetary problem (hydrogen atom) the function for $H(p, q)$ given in (9) must be substituted in the fundamental equations (12). The integration of the equations of motion then lead to Kepler ellipses, the expression for which in polar co-ordinates may be assumed known and which, for the rest, is of no account for the sequel. On the other hand, in view of calculation of the field of radiation in the next section, the Fourier expansion of the rectangular co-ordinates of the electron is of interest. For this we get *

$$\left. \begin{aligned} \frac{x}{a} &= -\frac{3}{2}\epsilon + \sum_{n=1}^{\infty} A_n \cos n\omega t \\ \frac{y}{a} &= \sqrt{1-\epsilon^2} \sum_{n=1}^{\infty} B_n \sin n\omega t \end{aligned} \right\} \quad (20)$$

Here a denotes the semi-major axis of the ellipse; ϵ is the numerical value of the eccentricity; ω is the angular velocity with which each revolution is performed (called *Kreisfrequenz* in German, being the number of radians described per second); A_n and B_n are coefficients that depend only on the number n and the value of the eccentricity ϵ . The rectangular co-ordinates have, of course, been so placed that the X - and the Y -axes are respectively parallel to the major and minor axis of the ellipse, whereas the Z -axis is perpendicular to the orbital plane.

4. Hertz's Formulæ

After the solution of the problem of motion the next step of the task set in § 2 is to calculate the field generated by the revolving electrons. In this process we may restrict ourselves to those approximations which describe the field at a distance from the centre of the atom which is great compared with the wave-length of the radiation. Further, we make use of the principle of superposition, which states that on account of the linearity of the fundamental electrodynamic equations we obtain the field generated by several point-charges or by several components of motion of one point-charge by calculating the field due to each individual component of motion and forming the sum of all these fields. But the resolution of the motion into single components

* See M. Born, "Lectures on Atomic Dynamics."

† See M. Born, loc. cit.

is equivalent to replacing the moving charge by point-charges that vibrate linearly (that is, along a straight line); and the radiation due to such linearly vibrating point-charges may be calculated from formulæ given by Heinrich Hertz. To illustrate these formulæ let us consider Fig. 1, which is taken from Sommerfeld's "Atomic Structure and Spectral Lines" (p. 24). Here the origin or "source-point" (*Quellpunkt*) O denotes the position, and $\mathbf{v}$ denotes the velocity vector of the point-charge, while P is the reference point (*Aufpunkt*) at which the strength of field is to be determined. The directions of the electric intensity of field $\mathbf{E}$ and of the magnetic intensity of field $\mathbf{H}$ are evident from the figure; their contributions at the time t are given by

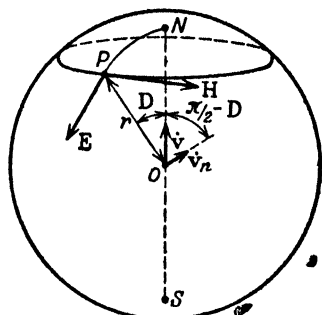


FIG. 1.—Field of a moving point-charge in the wave-zone.

$$|\mathbf{E}| = |\mathbf{H}| = \frac{e}{c^2 r} \dot{\mathbf{v}}_n \left(t - \frac{r}{c} \right). \quad (21)$$

Here e denotes the charge, c the velocity of light, r the distance between the source-point and the point P , $\dot{\mathbf{v}}_n$ the component of acceleration perpendicular to the direction of vision. The argument $\left(t - \frac{r}{c} \right)$ denotes that the value of the component of acceleration is to be taken at the retarded time $\left(t - \frac{r}{c} \right)$. The amount of radiation that goes through the point P (energy flux per unit time per unit area) is given by Poynting's vector

$$|\mathbf{S}| = \frac{c}{4\pi} |\mathbf{E}| |\mathbf{H}| = \frac{e^2}{4\pi c^3 r^2} \dot{\mathbf{v}}_n^2 \left(t - \frac{r}{c} \right) = \frac{e^2 \sin^2 \theta}{4\pi^2 c^3 r^2} \dot{\mathbf{v}}^2 \left(t - \frac{r}{c} \right). \quad (22)$$

The radiation is polarized, the plane of polarization being given in conformity with the convention usual in optics by the magnetic vector $\mathbf{H}$ and the radius vector connecting the source-point with $\mathbf{P}$. Hence for the intensity as perceived through a polarization apparatus the effective component of the acceleration of the point-charge is that which is perpendicular to the line of vision and the plane of polarization of the polarizer. The change of the field in time comes about owing to the time change of acceleration, $\dot{\mathbf{v}}_n$; the change with position, on the other hand, arises through

the time $t - \frac{r}{c}$ changing as the distance r between the source of light and P changes. Moreover, in accordance with (22), the intensity of radiation decreases proportionately with the inverse square of the distance.

Equation (22) gives the field of a single linearly vibrating point-charge. The field of a system of electrons that execute motions that do not occur in a straight line is obtained by superposing the fields of each individual component of motion, these fields being formed in accordance with (22). The dependence of such a field on space and time would in general be very complicated. Spectral analysis, however, shows that the radiation emitted by luminous gases (that is, from a system of atoms that interfere little with one another) consists of single lines of definite wave-length. Now the production of a spectrum is nothing other than a sort of harmonic analysis, a resolution of the field of radiation into individual components depending according to the sine law on space and time. The presence of spectral lines, therefore, leads to the inference—also suggested by analogy with planetary motion—that periodic terms occur in the motion of the system of electrons. Further, according to Fourier a periodic motion is capable of being developed as a series of pure sine vibrations. Let q be the co-ordinate of the vibrating charge, such that for the acceleration we have $|\ddot{\mathbf{v}}| = \ddot{q}$. Now if q is a periodic function with the fundamental frequency ω , it may be represented by

$$q = \sum_{n=1}^{\infty} A_n \sin n\omega(t - \delta_n) \quad (23)$$

where δ_n is a phase-constant that changes as we pass from one member to the next of the series. The partial vibrations given by the individual terms of the series (25) are called *harmonic components* of the motion.

Two successive differentiations of the expression (23) with respect to the time give the acceleration

$$\ddot{v} = \ddot{q} = -\omega^2 \sum_{n=1}^{\infty} n^2 A_n \sin n\omega(t - \delta_n) \quad (24)$$

which occurs in the equations (21) and (22). According to the above-mentioned principle of superposition the total wave-field of the point-charge is composed linearly of the individual summands of the series (24). Thus, according to classical electrodynamics, due account being taken of the approximations here applied, the spectrum that we may expect of point-charges in periodic motion should consist of a line of greatest wave-length ω ,

and the associated harmonic vibrations $2\omega, 3\omega, \dots$. If the motion of the charges generating the field is multiply periodic (various fundamental frequencies in various co-ordinate directions), there also occur along with the fundamental and harmonic frequencies the combination frequencies which arise from the linear superposition of different harmonic frequencies. If we pass from the angular frequencies in radians per second to the usual frequencies (vibrations per second), $\nu = \frac{\omega}{2\pi}$, the most general expression for the frequency of a combination vibration is

$$\nu = n_1\nu_1 + n_2\nu_2 + \dots \quad (25)$$

where the n_k 's are whole numbers. Taking into consideration the relations (17a) we may write this expression in the form

$$\nu = n_1 \frac{\partial H}{\partial I_1} + n_2 \frac{\partial H}{\partial I_2} + \dots \quad (26)$$

According to (22) the intensity belonging to the n th harmonic is to be set proportional to $(n^2 A_n)^2$. Observation through a polarizer would then allow the full intensity of this radiation to be perceived if the plane of polarization had the position given above (parallel to the magnetic vector). For the sequel we shall find it expedient to specify the "components of intensity" of the radiation, that is, those intensities $S_n^{(1)}$ and $S_n^{(2)}$ of the n th harmonic, which are observed in the polarization apparatus at two definite azimuths perpendicular to each other. For this purpose we must insert for $\ddot{q}$ that component of the acceleration vector of the point-charge which is normal to the plane of polarization. If we designate the Fourier coefficients of these components by $A_n^{(1)}$ and $A_n^{(2)}$ then, according to (22) and (24), we have

$$\begin{aligned} S_n^{(1)} &\sim (n^2 A_n^{(1)})^2 \\ S_n^{(2)} &\sim (n^2 A_n^{(2)})^2 \end{aligned} \quad (27)$$

This relation will be used later in calculating the intensities of spectral lines according to Heisenberg and Schrödinger. We have yet to mention that the word "component" is here to be understood simultaneously in the double sense of "vector component," that is, projection on a given direction of polarization, and secondly as a "Fourier-component," that is, as that part which belongs to a definite partial vibration. Further, it must be observed that the factor of proportionality in (27) also contains, according to (22), the square of the electric charge of the vibrating corpuscles. The quantity $e\ddot{\mathbf{v}}$ that occurs in (22) represents the second derivative with respect to the time of the

electric moment of the corpuscle with reference to the centre of vibration. The components of intensity of the field of radiation are therefore proportional to the squares of the second differential coefficients with respect to the time of the component under consideration of the electric moment of the charged corpuscles.

We next apply formula (27) to the radiation which, according to classical electrodynamics, should be emitted by an electron that executes an elliptic orbit. The corresponding Fourier expansions are given by the equations (20). The following table gives the components of intensity that should be observed according to theory, if the model is observed along the three co-ordinate axes :

Direction of Observation.	Polarized Character of the Radiation.	Plane of Polarization.	Components of Intensity.
X	linearly polarized	ZX	$S_n^x \sim (n^2 a \sqrt{1 - \epsilon^2} B_n)^2$
Y	" "	YZ	$S_n^y \sim (n^2 a A_n)^2$
Z	elliptically polarized	—	$\begin{cases} S_n^x \sim (n^2 a \sqrt{1 - \epsilon^2} B_n)^2 \\ S_n^y \sim (n^2 a A_n)^2 \end{cases}$

Since according to classical theory the eccentricity of the orbit is subject to no restrictions it can also have the value zero. In that case $A_n = B_n$ and the radiation emitted in the Z -direction would be circularly polarized.

5. Criticism of the Preceding Reflections

The approximate calculation of the classical theory, which we have just outlined, would thus lead to a spectrum that would consist of the lines with the frequencies $\nu, 2\nu, 3\nu \dots$, which would therefore be equidistant in the scale of frequencies, and whose components of intensity for different directions of observation and different azimuths of polarization would be given by the above table. The second part of the calculations sketched out above, in which with the help of Hertz's formulæ the field of radiation is worked out from the given components of motion of the electron, is sufficiently exact; on the other hand, the calculation of the orbit of the electron itself implies that something has been neglected in that the equations of motion (12) with the energy function (9) involve only the electrostatic attraction of the nucleus, whereas the electrodynamic reaction of the field of radiation on the electron itself was not taken into account. This reaction, moreover, causes secular disturbances in the motion which, although they are hardly

appreciable in intervals of time of the order of magnitude of the period of revolution of the electron, yet play an appreciable part in the times required for practical observation. For the radiation continually takes energy from the system, with the result that the electron continually approaches the nucleus in a sort of spiral orbit, the frequency of revolution in conformity with Kepler's third law becoming increasingly greater as the diameter of the orbit decreases. Corresponding to this the field of radiation above defined, having fixed fundamental and harmonic frequencies, will subsist only for very short intervals of time, of the order of magnitude of a few periods of revolution. In the length of time required to observe a spectral line the orbital diameter, and hence the frequency of revolution, would on the other hand change so rapidly that in place of each line a broad band shading off towards the violet would be emitted. The result, according to classical electrodynamics, would thus be as follows. A super-snapshot (time of exposure being about 10^{-13} sec.) of the spectrum of a Rutherford model of the hydrogen atom would have to show lines equidistant on the scale of frequencies; a "time" exposure would, on the other hand, give a continuous spectrum.

The hydrogen spectrum actually observed is, as we know, quite different. It contains series of sharp lines which are not spaced at equal distances in the scale of frequencies but accumulate at the short wave end of each series. It would be possible to regard the hydrogen spectrum as derived from a system of harmonic vibrations by a sort of stereographic projection which transforms the point at infinity to a point finitely distant.

CHAPTER III

BOHR'S THEORY

6. The Existence of Stationary States. Frequency Condition

THE arguments in the preceding sections show that the nuclear model and electrodynamics together lead to definitely false results so far as spectra are concerned. Now whereas the older attempts to find a theoretical interpretation of the hydrogen spectrum betrayed the desire to retain the electrodynamic ideas, and to adapt the atomic model to correspond with the spectrum (J. J. Thomson), Bohr decided in favour of Rutherford's nuclear model and, taking his stand on the older quantum assumptions of Planck and Einstein, disregarded the electrodynamic view-point altogether. The fundamental postulates of his theory are so well known that we need do no more than recall them in a few brief words.

Existence of stationary states of the atom.—There are a series of stationary orbits with discrete energy values, in which the electrons travel without emitting radiation (radiationless orbits); for the radiation of the electron in the hydrogen atom, in particular, the energy assumption (9) holds exactly except for the relativity correction for mass in motion.

Frequency condition.—In the transition from one state of energy E_m to another of energy E_n monochromatic radiation is emitted or absorbed, as the case may be, and its frequency is given by

$$\nu(nm) = \frac{1}{h}(E_n - E_m). \quad (28)$$

The fact, established long ago by spectroscopists, that the frequencies of the lines of the series allow themselves to be represented as the differences of two members of series of *spectral terms*, here receives an easily pictured interpretation since the spectral terms represent nothing other than the energy values of the individual stationary states divided by Planck's constant h .

7. The Correspondence Principle

The rules of the quantum theory for calculating the frequencies and intensities of the spectral lines are to be such that in the limiting case of long wave-lengths (large quantum numbers) they merge into those of classical electrodynamics. The fulfilment of this postulate with regard to frequencies leads, if we use equation (28) for periodic and conditionally periodic motion, to the quantum conditions for stationary orbits that are given in the next section. With regard to the application of the correspondence principle to the components of intensity, it must be remarked that so far as questions of intensity are concerned this principle cannot relate to the radiation from a single atom, for the radiation emitted in a single act of emission causes the removal of the energy $h\nu$ in conformity with (28). What is actually observed is, indeed, not the radiation of a single atom but that of a great number of atoms. Hence the correspondence principle restricts itself to an assertion of a statistical character. It gives us points of attack for dealing with the *probabilities of transition* between the individual stationary states. The intensities of the individual lines in the spectrum of a luminous gas are, according to the quantum theory, equal to the product of the energy $h\nu$ of a single light quantum (photon) with the number of emitted quanta of the frequency in question. But this number is equal to the number of actual transitions executed. On the other hand, according to classical electrodynamics, the intensity components of the individual lines depend in accordance with (27) on the harmonic components of the motion. Bohr had shown that with the use of the quantum conditions to be given in § 8 the following relation holds. If n and m are two integers, which are large compared with their difference $p = n - m$, the frequency of the radiation emitted in the transition between the m th and the n th state is approximately equal to that of the p th harmonic of the radiation that the electron would generate according to classical electrodynamics in revolving in one of these paths. The application of this correspondence, which holds with regard to frequencies, to the intensities yields, after what has been said, the following result. The relative frequency with which transitions occur (transition frequencies) between the m th and the n th orbit, depends on the p th harmonic component of the motion of the electron in one of these orbits. The nature of the functional dependence is given by (27), in which the index n is to be replaced by the index p . In particular the complete absence of certain harmonics in the motion of the electron allows us to infer that the corresponding transitions will not occur. In this way the correspondence

principle has been intensified so as to yield *selection rules* and has proved itself an excellent guide for understanding certain peculiarities of spectra. Special mention must be made of the occurrence of "forbidden" transitions in cases where the atom is perturbed by an external field, and which could be explained by supposing the action of the field to change the form of the orbit in such a way that harmonic components of the motion which were absent before are now able to assert themselves.

Nevertheless, there is a certain element of uncertainty in the assertions of the correspondence principle. A quantum transition or switch, as it is often called, involves two orbits, the initial and the final orbit. Of which of these two paths do the harmonic components determine the transition probabilities? In default of an unambiguous answer to this question attempts have been made to overcome the difficulty by taking a mean of the initial and the final orbits so as to get a value for these quantities.* The need for a more precise formulation, however, remained, and has now actually been met by the newer forms of the quantum theory proposed by Heisenberg and Schrödinger.

8. The Quantum Conditions

The central problem in Bohr's form of the quantum theory is the fixing of the electronic orbits (quantum orbits) corresponding to stationary states of the atom. In those systems whose motion calculated on classical mechanics is periodic or conditionally periodic definite conditions result for the stationary orbits if we apply the postulate, expressed in the correspondence principle, that for large quantum numbers the radiation, calculated according to the quantum theory by (28), is to be the same as that which is to be expected according to classical electrodynamics. To find these *quantum conditions* we follow on the dynamical reflections given in § 3. We imagine the equations of the mechanical system integrated by the introduction of angle and action variables. In the case of a conditionally periodic motion a definite orbit has a definite system of action variables I_k which remain constant during the whole motion, in accordance with (15) and (16). The values of the I_k 's themselves depend on the initial conditions of the problem: according to the classical theory they should be able to have arbitrary values. On the quantum theory, however, only a discontinuum (discrete series) of orbits is able to exist. Corresponding with this the action variables I_k will also be capable of assuming only discrete

* H. A. Kramers, "Dissertation, Kgl. Danske Videnk. Selsk. Skrifter, Naturvidensk.," Afd. 8, Raekke III, 3, 287, 1919.

values. It can easily be shown that the correspondence principle is satisfied if we assume for these discrete values that

$$I_k = n_k h \quad (29)$$

in which the n_k 's are positive integers, which we call the *quantum numbers* of the state in question. For we imagine, in agreement with the ideas of Bohr's theory, that the system performs a transition from the state given by the quantum numbers $m_1, m_2, m_3 \dots$ to another with the quantum numbers $n_1, n_2, \dots$, in which process all the differences $m_k - n_k$ are to be small compared with the values of the quantum numbers themselves. According to the frequency condition (28) radiation is emitted during this transition, and its frequency is given by

$$h\nu = \Delta E = \Delta H = \frac{\partial H}{\partial I_1} \Delta I_1 + \frac{\partial H}{\partial I_2} \Delta I_2 + \dots \quad (30)$$

On account of $\Delta I_k = (m_k - n_k)h$ this equation becomes

$$\nu = \frac{\partial H}{\partial I_1} (m_1 - n_1) + \frac{\partial H}{\partial I_2} (m_2 - n_2) + \dots$$

This expression * is now identical with the most general expression (26) for the radiation that is to be expected according to the classical theory. (Concerning the coefficients n_k that there occur, it was only assumed that they were whole numbers; that is why we can here identify them at once with the differences $m_k - n_k$ that now occur.)

Through the assumption (29) for the quantum conditions of the stationary states the postulate of the correspondence principle is satisfied so far as the radiation frequencies are concerned.

The quantum conditions (29) may also be expressed in the older form given by Sommerfeld and Wilson,

$$\oint p_k dq_k = n_k h \quad (29a)$$

in which, however, the choice of the co-ordinates q_k and p_k are subject to certain definite conditions (separation variables).

Hence, according to Bohr the calculation of a spectrum is to be effected in the following steps. The energy expression $H(p, q)$ and the equations of motion (12) for the atomic model in question must first be set up. Action variables must be introduced, and stationary orbits must be selected according to (29). The energy-

* The expression of ΔH in terms of a Taylor series, which ceases after the linear terms, is permissible only on the assumption made above, namely, that the differences $m_k - n_k$ are small compared with the quantum numbers themselves.

values of the individual orbits are then calculated and, from them, by means of (28), the frequency of the lines emitted in the transitions are obtained. The intensities are estimated by harmonic analysis of the forms of motion of the stationary states.

9. The Achievements of Bohr's Theory

The application of the quantum conditions (29) to the atomic model of hydrogen, whose motions are given by using (12) in conjunction with (9), leads, as we know, to the result that the energy-values of the individual stationary orbits are inversely proportional to the squares of the quantum numbers n . We have

$$E_n = - \frac{2\pi^2 m e^4}{h^2} \cdot \frac{1}{n^2}. \quad (31)$$

From this we get, by using the frequency condition (28), the well-known Balmer series formula

$$\nu = R \left(\frac{1}{n^2} - \frac{1}{m^2} \right), \quad (32)$$

which not only has the correct form required by experimental results but also gives the correct numerical value of the Rydberg constant R , evaluated from the mass and charge of the electron and the radiation constant h . Another consequence is that all the finer deviations of the values from those of the energy assumption (9) are correctly reproduced in the observed spectra. By taking into account the fact that the nuclear mass is not infinitely great but that rather the nucleus and the electron revolve about the common centre of gravity, we get a quantitatively correct value for the difference between the hydrogen spectrum and that of ionized helium. If further we add to the energy assumption (9) the potential energy of an electric field, we get the characteristic resolution of the lines which are observed in the Stark effect, and so forth. We make emphatic reference to these well-known facts in order to point out clearly that the application of the fundamental mechanical equations (12) to the calculation of the stationary states is justified without doubt in the cases here given. A reform of the theory of the hydrogen spectra can therefore relate only to the *pictorial* interpretation of the events discussed (stationary orbits, quantum switches and so forth). The functional relationships between the observed spectral frequencies and the fields that excite them are given absolutely accurately by applying the fundamental mechanical equations in conjunction with the frequency condition (28), so

that every new theory will with advantage take its stand on these very relationships. This, it is true, holds only for systems with one electron (hydrogen, ionized helium and approximately also for the alkalis and the Röntgenspectra), whereas the attempts at an exact treatment of the simplest of systems having more than one electron, such as neutral helium, led to crass contradictions with experiment. The results of the calculations always came out as if the motion of an electron in a static field (for example, in that of the nucleus plus that of, say, an externally superposed field such as in the Stark effect) were described perfectly accurately by the fundamental mechanical equations, whereas these equations fail even in the simple case of describing the motion of an electron under the influence of the nucleus and a second electron revolving round it.

This failure was to be anticipated from the very outset in view of the fundamental ideas of the Bohr theory. According to these ideas an electron pursuing an orbit in an atom and exposed to a field of radiation excited from without remains undisturbed in its stationary orbit until it is raised through the absorption of a quantum to a higher quantum orbit. Thus it responds to the action of the *alternating field* in a manner quite different from that of classical theory. And so it was not to be expected that it would react to the alternating field of a second electron, moving in the same atomic configuration, according to the prescripts of classical electrodynamics and mechanics. Thus the theory of stationary states had already reached the limits of its possible achievements at this point. On the other hand, the frequency condition (28) seems to be of universal importance.

CHAPTER IV

HEISENBERG'S QUANTUM MECHANICS

IN spite of the new and strange character of Heisenberg's quantum mechanics and its mathematical apparatus we may nevertheless regard it as a rational product of reflections based on Bohr's Correspondence Principle. Although, as we have shown in the preceding section, we have been able, by means of semi-exact considerations in the sense of the Correspondence Principle, to draw inferences from the mechanics of the system in question—some of which were even far-reaching—we have nevertheless had to admit to ourselves that this method of describing the processes involved contained, apart from its incompleteness and ambiguity, too many superfluous “wheels,” so to speak, to furnish a final trustworthy basis for atomic physics: its failure in practice strengthened this view. We shall illustrate these shortcomings for a special case, inasmuch as historically it led directly to Heisenberg's new quantum mechanics.

The treatment of the *dispersion phenomena* of any arbitrary atom, such as the sodium atom if we wish to take a concrete instance, offered Bohr's theory a characteristic difficulty which could not be overcome by semi-mechanical explanations. It is known that the dispersive properties of a linear harmonic oscillator, or of any other mechanical system which can execute vibrations, are characterized by the appearance of the characteristic denominator

$$\frac{C}{\nu_0^2 - \nu^2}.$$

Here C is a constant which represents the response of a mechanical system to an incident wave of frequency ν , while ν_0 represents the frequency of the *orbital motion* of the irradiated mechanical system. If we apply these formulæ to Bohr atoms we find that the dispersion equations contain the orbital frequency but not the *line-frequencies* which, emitted according to the Frequency Condition, are in no way connected with the orbital frequencies. The orbital and the line-frequencies coincide or agree very approximately only in the special case of harmonic oscillators

or for very large quantum numbers. Now, experiment shows that the dispersion of Bohr atoms can be represented excellently by means of the above classical formula so long as we replace the orbital frequency ν_0 by the transition frequency ν_{nm} .*

In some respects this behaviour is not surprising, since the application of quantum-theory reasoning to non-closed systems in the sense of the original form of Bohr's theory hardly seemed possible; for all "quantum conditions" rested on the existence of an energy integral and the possibility of obtaining action variables by means of the corresponding phase-integrals. On the other hand, this example shows clearly that it is not possible to make progress with "semi-conservative" methods; rather, we must follow Heisenberg and regard the good agreement which occurs in the case of closed systems as a sort of "accident." So, to set up a theory that would be applicable really generally we should certainly have to introduce far more definite departures from the concepts of classical mechanics than those exemplified, say, in the supplementary quantum conditions.

This conclusion was not entirely negative. The Correspondence Principle suggests to us to make the formal structure of the new theory conform as closely as possible with that of Bohr's theory, which had proved its value up to a certain point, and so to take over the above-mentioned "accidental" agreement into the foundation of the new theory. In this connexion we can do no more here than to refer to a note by Kramers³², in which, for the special case of dispersion, certain formal generalizations led to the classical theory of dispersion being satisfactorily translated into the quantum theory. But even so it was necessary to take an independent decisive step to arrive at the new quantum mechanics of Heisenberg.

10. The Fundamental Ideas of Heisenberg's Theory

In setting up his new quantum theory Heisenberg was guided essentially by the following considerations. In the application that had so far been made of mechanics to atomic theory the macroscopic quantities, position, orbit, momentum, and so forth, had been used. Yet, so far as the theory of spectra was concerned, these concepts were of no account for the final result. In spectroscopy we deal only with energy-levels, whose differences indicate the frequencies, and with the intensity and polarization of the different spectral lines. The circumstance that the material particles also have, say, mechanical co-ordinates,

* Cf., for example, the dispersion of sodium.

position, momentum, and so forth, does not arise for the final results. Hence economy of thought leads us to try and formulate the theory in such a way that at any rate no superfluous kinematical and mechanical quantities enter into the discussion of the spectroscopic phenomena that are being treated, but only those physically observable quantities which interest us at the end of the calculations. Of course, this does not relieve the theory of the responsibility of formulating the precise connexion between these quantities and the macroscopic quantities of space, time, orbit, and so forth, but this is a second problem. The fact that such an elimination of the quantities that are irrelevant in this respect would have its purpose is shown particularly well by the example of the dispersion formula above-mentioned, into which the irrelevant orbital frequencies intrude undesirably instead of the line-frequencies. On the other hand, the Correspondence Principle shows that the quantities that are characteristic for the frequency, intensity and polarization must be sought in forms which are analogous to mechanical and kinematical quantities.

This is the sense in which we must interpret Heisenberg's assertion that he wishes to take up into his new quantum mechanics only quantities which, "in principle, are observable." This is not to mean that we are not in the position to determine the position of an electron or its impulse and so forth. This question is in itself very complicated. In Ch. X we shall discuss exhaustively the problems connected with it. Heisenberg's point of view is, rather, that the concepts of mechanics and kinematics are unnecessary in the framework of Bohr's original spectral theory and must be replaced by a general formal scheme.

11. The New Kinematics *

Following directly on Heisenberg's original investigation by considering the special case of a system of one degree of freedom, we shall show how the classical mechanical quantities are to be translated into quantum-mechanical quantities. As we know, the motion of a system bounded in space may be represented as a function of the time by a Fourier series of the form

$$x = \sum_{-\infty}^{+\infty} C_r(\mathcal{F}) e^{2\pi i r \nu(\mathcal{F}) t}.$$

* We shall refrain here from giving an account which is based on independent axioms. Some of the laws which are specially postulated in the sequel follow completely or in part from assumptions that have already been made, but it seems to us to be more important here to give a clear picture of the important features than to present an axiomatic structure which is compact as possible.

Here the Fourier coefficients C_r are functions of the action variables $\mathcal{F}$. The motion is singly periodic, but in general the period depends on the action variables, that is, on the integration constants. The different frequencies

$$1 \cdot \nu(\mathcal{F}), \quad 2 \cdot \nu(\mathcal{F}) \quad \dots \quad \tau \cdot \nu(\mathcal{F}) \quad \dots$$

are harmonically related to one another. We can regard each frequency as a function of two whole numbers $\mathcal{F}$, n , so long as we quantize according to

$$\mathcal{F} = nh.$$

In the classical case the one variable τ always occurs as a multiplicative factor

$$\nu(\tau, n) = \tau \nu(n).$$

For great quantum numbers the application of the Correspondence Principle has shown that the quantum-mechanical frequency that belongs to the transition $n \rightarrow n - \tau$ agrees very approximately with the τ th harmonic. Hence according to the quantum theory for great quantum numbers vibrations occur which can be characterized symbolically by the expression

$$A(n, n - \tau) e^{2\pi i \nu(n, n - \tau) t}.$$

The significance of this expression, which will be explained in detail later, consists in the fact that it contains all the determining elements that come into question for characterizing the radiation, namely, the frequency (in the exponential) and an amplitude-factor A which when more closely defined is to furnish a measure for the intensity and polarization of the radiation.

The most important point is the difference between the classical and the quantum-mechanical frequencies. For, while the classical frequencies are always harmonic, the relation

$$\nu(n, n - \sigma) + \nu(n - \sigma, n - \sigma - \tau) = \nu(n, n - \sigma - \tau)$$

(Ritz's Principle of Combination) holds for the quantum-mechanical frequencies; this relation gives harmonic overtones only in the case of great quantum numbers.

The decisive difference between the classical and the quantum-mechanical method of description consists in the following. In the classical theory there is sense in representing the "motion" of the mechanical system by means of a sum of different terms in the Fourier series. Moreover, according to the classical theory all frequencies are emitted simultaneously with the relative intensities given by the Fourier coefficients. On the other hand, according to the quantum theory only *one* frequency is to be

allocated to every elementary process. Although the Fourier coefficients of the classical vibrations are still of importance in view of the Correspondence Principle, they now only make assertions about relative frequencies of occurrence of the many elementary processes that constitute a manifold, in other words, we are dealing with a system of statistics. Accordingly it would be meaningless to form a sum of the quantum-mechanical quantities

$$A(n, n - \tau) e^{2\pi i \nu(n, n - \tau)t}.$$

Every quantum-mechanical quantity corresponds to an *elementary process* which has a definite frequency, intensity and polarization. The problem of quantum mechanics consists in finding laws which will enable us to calculate for every given "mechanical model" the quantum-mechanical quantities characteristic of each elementary process.

Thus in the quantum mechanics of a system of one degree of freedom each quantum-mechanical quantity q has a two-dimensional discrete manifold allocated to it which may be depicted most conveniently in the form of a scheme:

$$\begin{array}{ccc} a_{11}e^{2\pi i \nu_{11}t} & a_{12}e^{2\pi i \nu_{12}t} & \dots \dots \dots \\ a_{21}e^{2\pi i \nu_{21}t} & a_{22}e^{2\pi i \nu_{22}t} & \dots \dots \dots \\ \dots \dots \dots & \dots \dots \dots & a_{nn}e^{2\pi i \nu_{nn}t} \\ a_{n1}e^{2\pi i \nu_{n1}t} & \dots \dots \dots & \dots \dots \dots \end{array}$$

Schemes of this kind have long been known in the theory of determinants. They are called *quadratic matrices*. In particular the matrices which are to be considered here are called *infinite matrices*, since even in Bohr's theory the number of quantum states are infinitely great. We denote the matrix q by the abbreviated symbol (q_{nm}) .

Since a Fourier's series of the form

$$f(t) = \sum_{-\infty}^{+\infty} C_{\tau} e^{2\pi i \tau t},$$

whose coefficients C_{τ} are in general complex, represent a *real quantity* only if

$$C_{\tau} = C_{-\tau}^*$$

(where C_{τ}^* denotes the complex quantity conjugate to C_{τ}) we likewise postulate, in order to maintain correspondence, that in the case of the matrices the relation

$$q_{nm} = q_{m\eta}^*$$

must hold, and hence that

$$\begin{aligned} \nu_{nm} &= -\nu_{mn}, \\ a_{nm} &= a_{mn}^*. \end{aligned}$$

A matrix which satisfies this condition is called a *Hermitean matrix*.

12. Calculations with Matrices

Matrices represent a type of hyper-complex numbers. Just as, for example, the realm of real numbers has been extended to that of complex numbers by introducing certain calculations with imaginary numbers which for the special case of real numbers merge into the well-known rules of arithmetic, so also for matrices we must agree upon certain rules if we are to be able to operate with them. These rules of calculation can in many cases be endowed with a direct physical meaning.

(a) Equality of Matrices

Two matrices **f** and **g** are called equal if each element f_{nm} of the one matrix is equal to the corresponding matrix-element g_{nm} of the other :

$$f_{nm} = g_{nm}. \quad (33)$$

This is analogous to the statement that two complex numbers are equal to each other only if the real and imaginary parts of the one are respectively equal to the real and imaginary part of the other. This statement has also a physical foundation, since every matrix-element f_{nm} is allocated to an elementary physical process and has no direct concern with any other elementary process.

(b) Addition and Subtraction of Matrices

The matrix **h** = **f** ± **g** stands for the matrix whose elements h_{nm} satisfy the relation

$$h_{nm} = f_{nm} \pm g_{nm}.$$

(c) Multiplication of Matrices

To illustrate this process we first refer to the rules for multiplying Fourier series. If two quantities $f(t)$ and $g(t)$ are given in the form of Fourier series thus :

$$\begin{aligned} f(t) &= \sum_{-\infty}^{+\infty} C_r e^{2\pi i r t} \\ g(t) &= \sum_{-\infty}^{+\infty} D_r e^{2\pi i r t} \end{aligned}$$

then their product, $f(t)g(t) = h(t)$ is given by the expression

$$h(t) = \sum \tau \sum \sigma C_{\tau} D_{\sigma} e^{2\pi i \tau \nu t}.$$

In this expression we always have

$$f(t)g(t) = g(t)f(t);$$

that is, the commutative law holds. Further, it must be observed that the frequencies of the quantity $h(t)$ agree with the frequencies of the functions f and g , respectively. This is necessary physically since otherwise different physical quantities would exhibit different periodicities.

We analogously define the matrix $\mathbf{h}(t) = \mathbf{f}(t)\mathbf{g}(t)$ as that matrix whose elements h_{nm} satisfy the relation

$$h_{nm} = \sum_k f_{nk} g_{km}. \quad (34)$$

This definition is familiar to us from the theory of determinants. It is here shown that the product of two determinants defined by the matrices $\mathbf{f}$ and $\mathbf{g}$ is equal to the determinant defined by the matrix $\mathbf{h}$ as determined by the law given just above for h_{nm} .

But there is an essential difference in this case compared with the classical theory. For whereas in the classical theory we always had

$$fg = gf,$$

here in general

$$\mathbf{fg} \neq \mathbf{gf}.$$

Thus, in general, multiplication is not commutative in quantum mechanics. The difference

$$\mathbf{fg} - \mathbf{gf}$$

is called the *commutator* or *multiplication difference* of $\mathbf{f}$ and $\mathbf{g}$.

In this definition we find Bohr's frequency condition entering as an essential element. For we obtain as the frequencies of the matrix $\mathbf{h}$ terms of the form

$$\nu_{nk} + \nu_{km} = \nu_{nm}.$$

But these are already contained in the old frequencies, so long as the relation

$$\nu_{nl} + \nu_{lm} = \nu_{nm}$$

holds. Bohr's frequency condition or Ritz's Combination Principle, respectively, expresses that all the physical quantities of a mechanical system are described by a single group of frequencies.

The idea of a diagonal matrix is in many ways useful for the sequel. This term denotes a matrix for which all the elements in which the index n has a different value from the index m

are zero, so that only the diagonal elements $n = m$ remain. If, in particular, all the elements of the diagonal matrix are equal to 1, the matrix is called a *unit matrix*. It is characterized by the fact that if it is multiplied by an arbitrary matrix $\mathbf{f}$, the matrix $\mathbf{f}$ again results.

(d) Division of Matrices

The division of matrices may be defined very simply with the help of the multiplication rule. For this purpose we need only introduce the concept of the reciprocal of a matrix. If the product of two matrices $\mathbf{f}$ and $\mathbf{g}$ yields the unit matrix $\mathbf{1}$, then we say that $\mathbf{f}$ is the reciprocal of $\mathbf{g}$. It is easily seen that then the relation

$$\mathbf{fg} = \mathbf{gf} = \mathbf{1}$$

also holds, that is a matrix can always be commutated with its reciprocal.* The reciprocal of $\mathbf{f}$ is symbolically denoted by $\mathbf{f}^{-1}$. We now take the quotient $\mathbf{f} : \mathbf{g}$ as simply standing for the expression

$$\mathbf{f} : \mathbf{g} = \mathbf{f} \cdot \mathbf{g}^{-1}.$$

(e) Differentiation with respect to the Time

The differentiation of a matrix with respect to a pure number is carried out fully analogously to the process used in the mathematics of commutative quantities. We take the matrix $\frac{d\mathbf{f}}{dt} = \mathbf{g}$ to stand for the matrix whose components are given by the relation

$$g_{nm} = \frac{df_{nm}}{dt}. \quad (35)$$

On account of Bohr's Frequency Condition, which always allows us to represent f_{nm} in the form

$$f_{nm} = a_{nm}e^{2\pi i\nu_{nm}t},$$

we may also write the new matrix $\mathbf{g}$ in the form

$$g_{nm} = 2\pi i\nu_{nm}f_{nm}. \quad (35a)$$

This last expression may also be represented in another way. To do this we introduce a diagonal matrix $\mathbf{W}^\dagger$ and form the matrix product

$$(\mathbf{W}\mathbf{f} - \mathbf{f}\mathbf{W}).$$

* We shall not deal here with the complicated mathematical question as to whether a reciprocal exists at all.

† In the case of a diagonal matrix it is sufficient to give one index.

Applying the multiplication rule and the definition of a diagonal matrix we get

$$(\mathbf{W}\mathbf{f} - \mathbf{f}\mathbf{W})_{nm} = (W_n - W_m)f_{nm}.$$

If we now set

$$\nu_{nm} = \frac{W_n - W_m}{h},$$

we have

$$\frac{d\mathbf{f}}{dt} = \frac{2\pi i}{h}(\mathbf{W}\mathbf{f} - \mathbf{f}\mathbf{W}). \quad (36)$$

This allows us to state the following theorem. Differentiating a matrix with respect to t is equivalent to forming the multiplication difference of this matrix and that which, according to Bohr's frequency condition, determines the frequency.

Hence the momentum of classical mechanics $m\dot{\mathbf{x}}$ is represented by a matrix $\mathbf{p}_x$ which is connected with the co-ordinate matrix $\mathbf{x}$ by the relation

$$\mathbf{p}_x = \frac{m2\pi i}{h}(\mathbf{W}\mathbf{x} - \mathbf{x}\mathbf{W}).$$

(f) Differentiation of a Matrix with respect to another Matrix

By means of the rules so far given we have defined what is to be conveyed by the function of a matrix $\mathbf{f}(\mathbf{q})$, so far as this functional sign only implies a succession of elementary arithmetical operations. Thus, expressed briefly, we may treat any function capable of being represented as a power series in the form of a matrix. We define the process of differentiating a matrix, $\frac{d\mathbf{f}(\mathbf{q})}{d\mathbf{q}}$, by means of the equation

$$\frac{d\mathbf{f}(\mathbf{q})}{d\mathbf{q}} = \lim_{\alpha \rightarrow 0} \frac{\mathbf{f}(\mathbf{q} + \alpha \mathbf{1}) - \mathbf{f}(\mathbf{q})}{\alpha}. \quad (37)$$

Here $\mathbf{1}$ stands for the unit matrix. This definition of differentiation produces in the case of all simple functions the relations which are well known from the ordinary differential calculus.

For example, we get

$$\begin{aligned} \frac{d\mathbf{x}^2}{d\mathbf{x}} &= \lim_{\alpha \rightarrow 0} \frac{(\mathbf{x} + \alpha \mathbf{1})^2 - \mathbf{x}^2}{\alpha} \\ &= \lim_{\alpha \rightarrow 0} \frac{\mathbf{x}^2 + 2\alpha\mathbf{x} + \alpha^2\mathbf{1} - \mathbf{x}^2}{\alpha} = 2\mathbf{x}. \end{aligned}$$

or, in general,

$$\frac{d\mathbf{x}^n}{d\mathbf{x}} = n\mathbf{x}^{n-1}. \quad (37a)$$

In the same way the following rule for the differentiation of a product also holds :

$$\frac{d}{d\mathbf{q}}(\mathbf{f} \cdot \mathbf{g}) = \frac{d\mathbf{f}}{d\mathbf{q}}\mathbf{g} + \mathbf{f}\frac{d\mathbf{g}}{d\mathbf{q}},$$

but we must observe that, in general,

$$\frac{d}{d\mathbf{q}}(\mathbf{f}\mathbf{g}) \neq \frac{d}{d\mathbf{q}}(\mathbf{g}\mathbf{f}).$$

13. The Commutation Law

The matrix $\mathbf{q}$ can be commutated with every matrix which is only a function of $\mathbf{q}$. But besides the co-ordinate matrix we also have the momentum matrix. As yet we know nothing of the possibility of commutating the co-ordinate matrix with the momentum matrix. We now agree to decide that the following relation must hold for them :

$$\mathbf{pq} - \mathbf{qp} = \frac{h}{2\pi i} \cdot \mathbf{1}. \quad (38)$$

Here h is for the present a constant which has been introduced for dimensional reasons, its dimensions being $\text{cm.}^2 \text{ grm. sec.}^{-1}$. We shall find later that it is identical with the constant which enters into the frequency condition, namely, with Planck's constant.

As the reason for introducing this relation we merely state here, without giving the details of calculation, that for *high quantum numbers* it merges into the quantum condition of the old Bohr theory :

$$\oint p dq = nh.$$

It is justified otherwise only by its success. With the help of the commutation law we can yet derive another form of the definition given above for the differentiation of matrices. For every function $f(\mathbf{p}, \mathbf{q})$ which can be represented by means of power series the two following relations hold :

$$\left. \begin{aligned} \frac{\partial \mathbf{f}}{\partial \mathbf{p}} &= \frac{2\pi i}{h} (\mathbf{f}\mathbf{q} - \mathbf{q}\mathbf{f}), \\ \frac{\partial \mathbf{f}}{\partial \mathbf{q}} &= \frac{2\pi i}{h} (\mathbf{p}\mathbf{f} - \mathbf{f}\mathbf{p}). \end{aligned} \right\} \quad (39)$$

We prove this relationship by means of mathematical induction. If it is assumed to hold for two functions $\mathbf{g}$ and $\mathbf{h}$, then it also holds for the sum and product of the two functions. The fact that it holds for the sum needs no further proof. To show that it holds for the product we proceed as follows :

$$\begin{aligned}\frac{\partial}{\partial \mathbf{p}}(\mathbf{g}\mathbf{h}) &= \frac{\partial \mathbf{g}}{\partial \mathbf{p}}\mathbf{h} + \mathbf{g}\frac{\partial \mathbf{h}}{\partial \mathbf{p}} \\ &= \frac{2\pi i}{h}[(\mathbf{g}\mathbf{q} - \mathbf{q}\mathbf{g})\mathbf{h} + \mathbf{g}(\mathbf{h}\mathbf{q} - \mathbf{q}\mathbf{h})] \\ &= \frac{2\pi i}{h}[(\mathbf{g}\mathbf{h})\mathbf{q} - \mathbf{q}(\mathbf{g}\mathbf{h})],\end{aligned}$$

and an analogous result holds for the differentiation with respect to $\mathbf{q}$.

Now the two relations for $\mathbf{q}$ and $\mathbf{p}$ hold in view of the validity of the commutation rule. Hence they hold generally.

14. The Mechanical Equations of Motion

A mechanical system is characterized in the classical theory by its Hamiltonian function $H(p, q)$. We take the expression for Hamilton's function over into quantum mechanics, but we now regard the quantity H as a matrix $\mathbf{H}$ which is itself a function of the matrices $\mathbf{p}$ and $\mathbf{q}$.*

The equations of motion then run :

$$\left. \begin{aligned}\frac{d\mathbf{p}}{dt} &= -\frac{\partial \mathbf{H}}{\partial \mathbf{q}}, \\ \frac{d\mathbf{q}}{dt} &= +\frac{\partial \mathbf{H}}{\partial \mathbf{p}}.\end{aligned}\right\} \quad (40)$$

These equations of motion are explained unambiguously by the rules of calculation given above. The following fundamental law is a consequence of them : *if the quantity $\mathbf{H}$ is not an explicit function of the time t , $\mathbf{H}$ is a diagonal matrix.* To prove this theorem we first transform the equations of motion. On account of (36) and (39) we now have

$$\begin{aligned}\frac{d\mathbf{p}}{dt} &= \frac{2\pi i}{h}(\mathbf{W}\mathbf{p} - \mathbf{p}\mathbf{W}) = -\frac{\partial \mathbf{H}}{\partial \mathbf{q}} = \frac{2\pi i}{h}(\mathbf{H}\mathbf{p} - \mathbf{p}\mathbf{H}), \\ (\mathbf{W} - \mathbf{H})\mathbf{p} &= \mathbf{p}(\mathbf{W} - \mathbf{H}).\end{aligned} \quad (41)$$

* Difficulties which may occur in certain Hamiltonian functions which have the form pq , say, in classical theory, are treated by Born, Heisenberg and Jordan (18a).

Similarly,

$$\begin{aligned}\frac{d\mathbf{q}}{dt} &= \frac{2\pi i}{h}(\mathbf{W}\mathbf{q} - \mathbf{q}\mathbf{W}) = \frac{\partial \mathbf{H}}{\partial \mathbf{p}} = \frac{2\pi i}{h}(\mathbf{H}\mathbf{q} - \mathbf{q}\mathbf{H}), \\ (\mathbf{W} - \mathbf{H})\mathbf{q} &= \mathbf{q}(\mathbf{W} - \mathbf{H}).\end{aligned}\quad (42)$$

Thus the quantity $\mathbf{W} - \mathbf{H}$ can be commutated with $\mathbf{p}$ and $\mathbf{q}$, and therefore with every function $f(\mathbf{p}, \mathbf{q})$. Hence we also have

$$\begin{aligned}\frac{d}{dt}(\mathbf{W} - \mathbf{H}) &= \frac{2\pi i}{h} [\mathbf{W}(\mathbf{W} - \mathbf{H}) - (\mathbf{W} - \mathbf{H})\mathbf{W}] \\ &= -\frac{2\pi i}{h}(\mathbf{W}\mathbf{H} - \mathbf{H}\mathbf{W}) = -\frac{d}{dt}\mathbf{H} = \mathbf{D}.\end{aligned}\quad (43)$$

Thus $\mathbf{H}$ is a diagonal matrix, and, moreover, on account of (41) it is identical with the diagonal matrix $\mathbf{W}$ that was introduced earlier.

The Integration of the Quantum-mechanical Equation of Motion

The matrices $\mathbf{p}$ and $\mathbf{q}$ or functions of them are determined as follows. We form the two equations of motion which hold for each matrix-element; these are in reality, then, an infinite system of equations with an infinite number of unknowns. We next make use of the commutation relation. If we are then successful in finding matrices $\mathbf{p}$ and $\mathbf{q}$ which satisfy this system of equations together with the commutation relation, the quantum-mechanical problem is solved. From these matrices $\mathbf{p}$ and $\mathbf{q}$ we form Hamilton's function $\mathbf{H}(\mathbf{p}, \mathbf{q})$. On account of (41) and (42) this becomes a diagonal matrix:

$$\begin{aligned}H_{nm} &= H_n \delta_{nm}, \\ \mathbf{H} &= \mathbf{W},\end{aligned}\quad (44)$$

and the difference of the diagonal elements $H_n - H_m$ gives us the frequency $\nu_{nm} = \frac{H_n - H_m}{h}$ associated with the transition

$n \rightarrow m$.

When we have formed the Hamiltonian function in co-ordinates which correspond to the Cartesian co-ordinates and momenta of the classical theory we can then also derive expressions for the intensity of a line. For this purpose Heisenberg, by appropriately generalizing the classical expressions (24) and (27), sets the intensity of the line allocated to the transition $n \rightarrow m$ proportional to the expression

$$(\ddot{x}_{nm})^2.$$

For high quantum numbers this relation is in agreement with the Correspondence Principle and refines it for the region of small quantum numbers.

15. Systems with Several Degrees of Freedom

A system with several degrees of freedom is characterized by the co-ordinates $q_1, q_2, \dots, q_f$ and the momenta $p_1, p_2, \dots, p_f$ together with the Hamiltonian function $H(q_i, p_i)$. When translated into the matrix form the whole treatment is fully analogous, the only difference being that the commutative properties become changed.

At the outset we are fully ignorant of the nature of the commutation of

$$\begin{array}{ll} & q_i \text{ with } q_m, \\ \text{that is,} & \mathbf{q}_i \mathbf{q}_m - \mathbf{q}_m \mathbf{q}_i, \\ \text{or of} & q_i \text{ with } p_m, \\ \text{that is,} & \mathbf{q}_i \mathbf{p}_m - \mathbf{p}_m \mathbf{q}_i, \\ \text{or of} & p_i \text{ with } p_m, \\ \text{that is,} & \mathbf{p}_i \mathbf{p}_m - \mathbf{p}_m \mathbf{p}_i. \end{array}$$

We are helped by the following hypothesis, which is in agreement with the quantum conditions of the original Bohr theory :

$$\mathbf{q}_i \mathbf{q}_m - \mathbf{q}_m \mathbf{q}_i = 0, \quad (45)$$

$$\mathbf{p}_i \mathbf{p}_m - \mathbf{p}_m \mathbf{p}_i = 0, \quad (46)$$

$$\mathbf{p}_i \mathbf{q}_m - \mathbf{q}_m \mathbf{p}_i = \frac{h}{2\pi i} \mathbf{1} \delta_{im} \quad (47)$$

$$(\delta_{im} = 1 \text{ for } i = m; \delta_{im} = 0 \text{ if } i \neq m.)$$

This hypothesis is evidently justified since at least in those cases where the Hamiltonian function *can* be separated, the motions in the individual degrees of freedom occur independently of one another, so that we cannot here expect definite commutation relations to hold. Apart from formal completeness practical success is here again ultimately the surest support for the generalized commutation rule.*

16. Survey of the Matrix Method

The mechanical quantities become replaced by matrix-elements which are characteristic of the properties of a spectral line which has been assigned to an elementary process. The form of the Hamiltonian function is taken over unchanged into matrix

* The formulation (47) is also suggested by some important correspondence considerations.

mechanics. An essentially new datum appears with the commutation relation between the co-ordinates and the momenta. The calculation of the matrices $\mathbf{q}$ and $\mathbf{p}$ is then performed by combining the infinite system of equations (40) with the commutation relation (38) and calculating from them the individual matrix-elements p_{nm} and q_{nm} . If with these matrix-elements we then form the Hamiltonian function $\mathbf{H}$, this is a diagonal matrix. The diagonal members of Hamilton's function become identical with the individual energy-levels of the atom and the frequency ν_{nm} , which is associated with two levels, is calculated with the help of the frequency condition, which has been taken over unchanged. With certain restrictions the quantities q_{nm} then give a measure of the intensity of the line emitted during the transition $n \rightarrow m$.

When applied in practice the matrix method often fails except in quite simple cases owing to the impossibility of solving the systems of equations (40) and (38) with their infinite number of equations and unknowns. This is where Schrödinger's wave-mechanics has been able to play a helping part.

The Achievements of Quantum Mechanics

We shall comment on the successful results of quantum mechanics at this early stage although the more flexible method of proper functions (cf. Ch. IV) serves much more conveniently for deriving them. In all cases where the original theory of Bohr was confirmed by experiment the results of the new theory agree with those of the old theory, but in addition it overcomes a great number of difficulties which the original theory could not conquer. In the first place, it automatically yields half-integral quantum numbers, which had previously presented considerable difficulties, particularly in the case of molecular spectra. In the same way the theory of the anomalous Zeeman effect with all its details results from it, for example, those which occur in the transition to the normal triplet, as caused by a Paschen-Back effect. The difficulties in dispersion theory mentioned at the beginning of this section are removed by means of the new quantum mechanics in agreement with the formula of Kramers which was based on the Correspondence Principle. A particularly valuable feature, which speaks for the logical consistency of the theory, is that several supplementary decrees (*Verbote*) which had to be invoked to impose restrictions on quantum states allowed by the older theory (pendulum orbits and adiabatic invariance) now disappear. The fact that with the help of the

new theory we are able to deal successfully with many-body problems by making certain refinements, will be shown in Ch. IX. Finally, in questions of intensity we are no longer dependent on the unsatisfactory qualitative statements of the Correspondence Principle. Rather, as mentioned above, the quantities $(\ddot{x}_{nm})^2$ give an exact measure of the intensity of the associated spectral lines.

In this connexion we shall only briefly mention aperiodic phenomena which may be treated much more simply with the help of Schrödinger's theory. We shall discuss a number of questions which link up with this subject in later sections.

Perhaps the most surprising factor in the new quantum mechanics consists in the following. Heisenberg gives up entirely the idea of mechanically interpreting atomic processes in space and time. The whole calculation is adapted to determining certain optical data. On the other hand, the Hamiltonian function and the whole formal structure is taken over from mechanics and the individual relationships are arranged in such a way that agreement with macroscopic mechanics is established for high quantum numbers. This apparent paradox has become eliminated in the later development of the theory in the following manner (see Ch. X). It is by no means necessary to restrict ourselves to optical data. We can equally well describe purely mechanical data, such as position, momentum of an electron and so forth, but we cannot intelligently talk *simultaneously* of the energy and the position of an electron in an atom. We call attention to the new interpretation discussed in detail in the last chapter merely in order to remove the feeling of dissatisfaction which the theory necessarily calls up in its present stage of development.

CHAPTER V

THE MATERIAL WAVES OF DE BROGLIE

17. The Allocation of a Vibration Frequency to the Electron

A YEAR before the first paper of Heisenberg's on quantum mechanics appeared an attack on the problem of matter had been made from quite another quarter by Louis de Broglie.²⁰ His arguments had their origin in a certain paradox that occurs when we seek to bring the laws of light-corpuscles and of the motion of elementary particles under one heading. The energy of a light-corpuscle is, according to the fundamental postulate of the quantum theory, $h\nu$; the energy of a material particle at rest is given by μc^2 according to the theory of relativity, where μ is the mass and c denotes the velocity of light. Many experimental facts, particularly that of the pressure of light, had taught us that inertia is a property of radiation, and, more recently, the Compton effect had shown quite clearly that the laws of mechanics hold for collisions of electrons and light-corpuscles, the mass $h\nu/c^2$ being ascribed to each light-corpuscle. This fact suggested to de Broglie to proceed conversely and to allocate to the material elementary particles, such as electrons and protons, a vibration frequency ν such that the relation

$$h\nu = \mu c^2 \quad (48)$$

holds. For example, we might, as a first step, imagine the electron to be a very minute kind of oscillator, which vibrates with the frequency

$$\nu = \mu c^2/h = 1.29 \cdot 10^{20} \text{ sec.}^{-1}.$$

It soon becomes manifest, however, that this method of connecting light-corpuscles and electrons would be in contradiction to the theory of relativity. The numerical value of the frequency ν just given has been chosen so that the relation (48) holds in a system of reference K in which the electron is at rest. Let us now consider the event from the point of view of a system of reference K' , with regard to which the electron moves with the velocity v rectilinearly and uniformly. According to the well-

known law of the theory of relativity, which asserts that the rate of clocks and of all regularly periodic events is slower when regarded from the moving system of reference, we should have in K'

$$\nu' = \nu \sqrt{1 - \frac{v^2}{c^2}}. \quad (49)$$

On the other hand, however, the relativity theory also states that the mass and energy of the particle must appear *greater* when regarded from the moving system, that is

$$\mu' = \frac{\mu}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (50)$$

Thus we see that the relation (48) would lose its validity in the moving system of reference, and this would conflict with the principle of relativity. This simple discussion would doubtless be regarded by many physicists as a proof that the allocation of a vibration event to a material particle is impossible and senseless. De Broglie did not, however, allow himself to be daunted by this negative result but applied himself to removing the apparent contradiction. His efforts brought to light very remarkable relationships. For it became evident that the equation (48) loses its validity in the moving system of reference only if, as was done above, the vibration event in the electron is regarded as *strictly localized*. It is therefore not permissible to imagine the electron as an oscillator of sub-atomic dimensions. On the other hand, we can arrive at a transformation formula for the frequencies ν , which agrees with the equation (50) for the change of mass, if we replace the picture of the oscillator restricted to electronic dimensions by a vibration event *extended in space*. We shall divide the next part of our argument into two steps.

18. First Step: the Phase-Waves

We first ascribe a vibration of unlimited extent to a material particle of mass μ at rest in a system of reference K . Regarded from the point of view of K let the phase of the vibration depend only on the time and not on the place. In other words, let all points of our vibrating "something" always be in the same phase at the same time (as, for example, the points of a rigid rod that is being moved up and down periodically). Thus if we designate our vibrating quantity, about whose physical character we make

no assertions whatsoever, by ψ , then the vibration event may be represented by the expression

$$\psi(t) = A \sin 2\pi\nu t. \quad (51)$$

We next postulate that the fundamental relation (48) holds in all systems moving uniformly and rectilinearly with respect to K , and thus also in a system K' moving with the velocity v in the direction of the negative X -axis. Then, on account of (50) the frequency observed from K' must be given by *

$$\nu' = \frac{\nu}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (52)$$

Further, if we denote by t' the time measured in the system of reference K' , then on account of the Lorentz-Einstein transformation for the time we have

$$t = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left(t' - \frac{vx'}{c^2} \right). \quad (53)$$

From (52) and (53) we get

$$\nu t = \nu' \left(t' - \frac{vx'}{c^2} \right). \quad (54)$$

Thus, regarded from K' , the vibration event is given by

$$\psi(x', t') = A \sin 2\pi\nu' \left(t' - \frac{vx'}{c^2} \right). \quad (55)$$

Hence the vibration event, when regarded from K' , no longer occurs synchronously at all points; rather, it represents a *wave*, which travels in the direction of the X -axis. This is at the same time the direction in which our material particle moves with respect to the observer, since we assumed above that K' moves with respect to K , and hence also with respect to the particle, in the direction of the negative X -axis. If we call the phase-velocity of the particle u , then such a wave can be represented by the expression

$$A \sin 2\pi\nu' \left(t' - \frac{x'}{u} \right), \quad (56)$$

* The transformation formula (52) appears to contradict (49). But we must bear in mind that the two equations refer to different events. In (49) we were concerned with the frequencies of revolution of an individual clock localized at one point; in (52), on the other hand, we are dealing with the frequencies of an event extended in space,—with a wave.

from which we get by comparison with (55)

$$u = \frac{c^2}{v}. \quad (57)$$

De Broglie has introduced the expression "phase-waves" (*ondes de phase*) for the waves representing the point-mass: this expression actually tells us nothing new since every vibration event manifests changing phases. But as the physical character of the quantity ψ has been left quite open and as it plays no part in the first part of our discussion, the lack of expression in the term phase-wave is fully appropriate at this juncture. The more striking term "material waves" which is now often used may give rise to misunderstanding, for what actually vibrates is certainly not matter.

Result of the first step.—We may assign to a material particle (such as an electron) a vibration event extended in space and having the frequency ν such that the relation (48) is valid in all systems moving uniformly and rectilinearly with respect to the particle. Corresponding to a particle moving with the velocity v there is a wave travelling in the direction of motion of the particle whose phase-velocity is given by $u = c^2/v$. The object in view, namely that of ascribing a frequency to a material particle, has been achieved in this first step, but at the price of losing almost all the other characteristics of a corpuscle; in place of the space-time localization we now have a spatially unlimited vibration event. An assertion such as "the point-mass is situated at a given time at a given place" becomes devoid of meaning in the relationships just given.

19. Second Step: the Wave-Groups

The complete absence of localization of corpuscles just mentioned may be rectified by allotting to a particle not a monochromatic vibration of fixed frequency ν but, instead, a narrow spectral interval, say the frequency band between ν and $\nu + \Delta\nu$. For if we imagine v in (52) to be replaced by the phase-velocity u with the help of (57), we obtain a law of dispersion: every frequency has a different but quite definite phase-velocity of its own. Now if instead of a monochromatic wave a frequency-band of vibrations passes through a dispersive medium, then according to well-known laws of classical wave-kinematics the following phenomenon presents itself. The partial vibrations of the frequency-band which lie close together will in general weaken one another by interference, and only at one place will

a region of maximum amplitudes arise, a so-called *wave-group*, in which the major part of the total energy of the frequency-band is concentrated (Fig. 2).^{*} This wave-group moves on with a velocity which is in general quite different from the phase-velocity v . It is called the *group-velocity* g . It is related to the phase-velocity in the following way [†]:

$$\frac{1}{g} = \frac{d\left(\frac{v}{u}\right)}{dv}. \quad (58)$$

If we make use of the relations between ν , u and v given by equations (55) and (53) we get from this that

$$g = v. \quad (59)$$

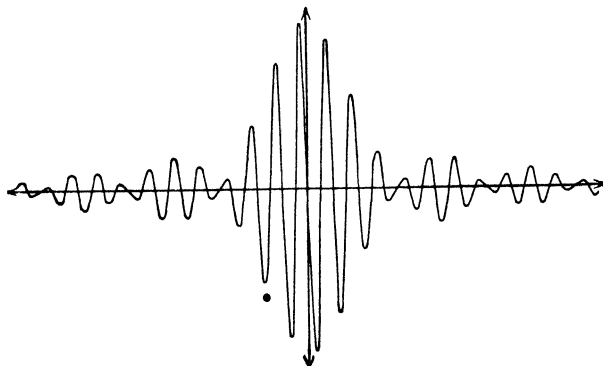


FIG. 2.—A wave-group.

So the velocity v with which the material particle allocated to the vibration event moves is nothing other than the group-velocity of the wave-group. We might therefore identify the moving particle with a travelling wave-group, which belongs to a narrow spectral region of mean frequency $\nu = \frac{c^2}{\mu h}$. The wavelength of the phase-waves comes out as

$$\lambda = \frac{u}{\nu} = \frac{c^2}{v} \frac{h}{\mu c^2} = \frac{h}{\mu v}. \quad (60)$$

^{*} This figure has been taken from F. Flamm, see reference (11).

[†] See, for example, de Broglie, "An Introduction to the Study of Wave Mechanics," p. 44 (transl. by H. T. Flint), Messrs. Methuen & Co. Ltd. A still more detailed treatment is given by Sommerfeld, "Wave Mechanics."

For electrons which move with a velocity about one quarter of that of light this gives a wave-length of roughly 10^{-9} cm.* The relationship between the wave-length and the momentum G of the particle is given by

$$G = \mu v = \frac{h\nu}{c^2}v = \frac{h\nu}{u} = \frac{h}{\lambda}. \quad (61)$$

So far the only noteworthy feature about the preceding argument is the fact that we have encountered the equation $v = c^2/u$ twice; firstly as that relation between particle-velocity and phase-velocity which enables us to ascribe an invariant frequency to a material particle; secondly, as the expression for the group-velocity of the wave-group. The circumstances which imbue these speculations with a certain physical truth present themselves subsequently in the light of the facts now to be enumerated.

20. The Optical-Mechanical Analogy

According to Fermat's Principle the path of a ray of light between two given points is such that less time is taken in traversing the actual path than in traversing any other path. If we call the differential of the path ds and the wave-velocity u , then

$$\int_A^B \frac{ds}{u} = \text{a minimum.} \quad (62)$$

Now the motion of a material point takes place according to the Principle of Least Action given in § 3 in such a way that the path integral of its velocity is a minimum, thus

$$\int_A^B \mu v ds = \text{a minimum.} \quad (11)$$

If we make use of (61) we see that (11) assumes the form of (62). So the conception of phase-waves allows us to reduce the fundamental principles of wave propagation to those of the motion of material particles, or vice versa.

* These waves are, of course, not to be confused with Röntgen rays of the same wave-length, for the phase-velocity of the material waves (phase-waves) is greater than that of light. In passing we may remark that this circumstance does not conflict with the law of the theory of relativity which states that no effect can be propagated with a velocity greater than that of light. For the transmission of the phase of a wave of infinite extent does not represent a transference of energy. The energy of the wave-group is transported with the group-velocity v , which is less than c .

21. The Interpretation of Bohr's Quantum Conditions

The conception of phase-waves further makes it possible to form a convincing picture for interpreting Bohr's quantum conditions for the circular orbits of the hydrogen atom. The general quantum conditions (29a) give in the special case of circular orbits

$$\oint p_{\phi} d\phi = 2\pi r\mu v = n\hbar. \quad (63)$$

Here r denotes the orbital radius. By using the relation (61) for the wave-length of the phase-waves (63) becomes

$$2\pi r = n\lambda. \quad (64)$$

The circular orbits allowed by the quantum conditions are thus distinguished by the fact that their circumferences are integral multiples of the wave-length of the phase-waves assigned to the revolving electrons.

22. The Occurrence of Interference Phenomena in the Case of Electrons

Quite independently of the purely theoretical considerations of de Broglie certain recent experimental researches have given rise to the suspicion that under certain circumstances cathode rays which have hitherto been regarded as purely corpuscular rays, also assume the character of a wave-radiation. In this connexion we must mention above all the important experiments of Davisson and Germer.²¹ They allowed a small pencil of electrons to fall on a nickel crystal and then determined by measurement how the intensity of the rays reflected in different directions depended on the velocity of the electrons. Certain maxima occur which may be interpreted, similarly to those that occur in the interferences of Röntgen rays in the case of Laue's and Bragg's experiments, as a diffraction phenomenon of wave-radiation at the space-lattice of the crystal. A knowledge of the angles of reflection and the lattice constant enable us to calculate, by a well-known method, the wave-length of the radiation. It was found that the apparent wave-length of the electron-radiation is inversely proportional to its velocity, as is required by the formula (60) of de Broglie's theory.

A phenomenon which dates back further than that of the

reflection of electrons by crystals was discovered almost simultaneously * by Ramsauer³⁶ (method of circular electronic beams), and Townsend and Bailey (diffusion method), and consists essentially in the fact that atoms offer less resistance to electrons of low velocity (of the order of one volt, that is, $5.95 \cdot 10^7$ cms./sec.) than to electrons moving with higher speeds. In the case of the inert gases this decrease of the effective cross-section of the atoms is particularly marked. There appears to be a definite low velocity of the electrons for which each type of atom has a definite minimum † cross-section. It is possible to consider this phenomenon, too, as a diffraction effect of the phase-waves of the electrons.

Finally, we may mention that the conception of material waves has been made use of by Einstein (23) in dealing with the statistics of the ideal gas, but a discussion of these interesting ideas would lead us too far here.

23. The Conception of Phase-Waves as a Guide to Wave-Mechanics

The interpretation of the quantum conditions as an integral relation entering into the phase-waves, which was given in § 21, tended to inspire the hope that with the help of the concept of waves it would be possible to arrive at an understanding of quantum conditions by a less artificial route than before. The discontinuous element introduced into atomic physics by the quantum conditions (29) does not harmonize well with the remaining ideas, according to which the atom is to be regarded as a sort of planetary system. The singling out of a discrete series of orbits, the traversing of the electronic paths without the emission of radiation, these are features which have been impressed somewhat violently on the otherwise classical picture of the electron. On the other hand, the occurrence of discrete values for the frequencies is quite familiar and natural even within the framework of classical physics, namely, in vibration problems of elastic continua. In the case of vibrating strings or of vibrating membranes and so forth there are associated with every system of nodes or nodal lines quite definite vibration frequencies, which come directly out of the very nature of the

* For a full report of the history and results of this "Ramsauer-Townsend effect" see Brose and Saayman, *Ann. d. Phys.*, 5, pp. 797-852, 1930. The theoretical aspect is treated by Holtsmark, *Zeitsch. f. Phys.*, 1930.

† It is important to emphasize this point, as many text-books and reference books make the quite unjustified and erroneous statement that electrons of vanishing small velocity pass freely through the atoms of the inert gases. —(Transl.)

problem without further supplementary conditions. It is known that the differential equations of the elastic vibrations of such systems give solutions that are of use for the problem in question only if certain parameters that occur in them (which are connected with the vibration frequency) assume definite discrete values, the so-called *proper values*. It readily suggested itself to try and apply these ideas, which are familiar in the theory of elasticity, to de Broglie's phase-waves and to interpret the quantized elements of atoms and their radiation as proper values of vibration problems. This attempt has been carried out with great success by Schrödinger, without, however, one of the greatest hopes that had been raised by de Broglie's theory being fulfilled, in that the conception of phase-waves was obliged to take refuge in multi-dimensional spaces.

Recapitulating, we may say that de Broglie's theory seeks to resolve the conception of the point-mass into that of a fictitious wave-field; the motion of a point-mass takes place analogously to that of a group of waves that obey Fermat's principle. The phase-waves of de Broglie differ from electromagnetic waves in having no direct physical reality. In the case of electromagnetic waves there is a real subject of the vibration event; the elements which execute the vibrations are the components of the electric and magnetic intensity of field, that is, quantities which are accessible in principle to direct physical measurement. The vibrating something of de Broglie's theory is, on the other hand, the "field-scalar" ψ which is linked up with physically observable quantities through a rather long chain of abstractions.

CHAPTER VI

THE WAVE-MECHANICS OF SCHRÖDINGER

24. The Wave-Equation for the One-Electron Problem

As was shown in § 21, de Broglie's wave-conceptions offered an interpretation for the simplest case of quantum orbits. But beyond this they were unable to account for the finer details of spectra. The position was to some extent comparable with that which came about in the realm of classical optics at the time when the knowledge of the wave nature of light first took sway. It was possible to explain the simplest cases of interference and diffraction merely from the fact that light is a wave event without entering further into its laws. To treat any rather more complicated problems it is necessary to have a knowledge of the partial differential equations according to which the vibration occurs.

In the case of de Broglie's material waves Schrödinger succeeded in setting up an equation of this kind, which at once showed itself to be an extraordinarily effective instrument for treating quantum problems. We shall first make our acquaintance with it for the case of the one-electron problem, that is, for that wave-field that is to replace in wave-mechanics the picture of a single point-mass (electron) moving in a field of force. We shall outline the derivation of this equation along the line of thought pursued by Schrödinger in his second communication.* The allocation of a frequency ν and a phase-velocity u to a point-mass moving with a velocity v is carried out by Schrödinger, without the use of the relativity theory, on the basis of the quantum postulate (28) and classical dynamics. Hence in place of the relations (48) and (57) of de Broglie's theory the following transcriptive relations hold according to Schrödinger :

$$h\nu = E, \quad (65)$$

$$uv = \frac{E}{\mu}. \quad (66)$$

* The collected papers of Schrödinger on Wave-mechanics have been issued in English in one volume by Messrs. Blackie & Co. Ltd. We here retain the references to the original German papers, as they are easily traced in the English volume.—(Transl.)

Here, as always, E denotes the total energy (sum of the kinetic and the potential energy) of the point-mass. If we set $E = \mu c^2$, we of course arrive at equations (48) and (57) again. Dividing (65) by (66) likewise leads, as in de Broglie's theory, to the relation between the wave-length of the phase-wave and the momentum of the point-mass

$$\frac{1}{\lambda} = \frac{\mu v}{h}. \quad (61)$$

The vibrations of a mechanical and electromagnetic nature that occur in macroscopic physics in general obey a partial differential equation of the form

$$\Delta\psi - K\psi = 0. \quad (67)$$

Here the physical meaning of the quantity K is the square of the reciprocal of the phase-velocity of the waves. This is a constant which is characteristic of the material ("material constant") for electromagnetic and elastic waves in isotropic bodies. On the other hand, the phase-velocity of the fictitious wave-field here under consideration is by no means a constant; rather, it depends in accordance with (66) on the incidental value of the speed v of matter, and so is a function of position.

In (67) we imagine the quantity K , which is now no longer constant, to be replaced by $1/u^2$ and attack the equation with the usual assumption, according to which ψ is represented as the product of a function of the co-ordinates and an exponential of the time

$$\psi = f(x, y, z)e^{2\pi i \nu t}. \quad (68)$$

This gives

$$\ddot{\psi} = -4\pi^2 \nu^2 \psi,$$

and hence

$$\Delta\psi + \frac{4\pi^2 \nu^2}{u^2} \psi = 0. \quad (69)$$

The quantity ν/u that occurs in (69) is nothing else than the reciprocal of the wave-length, for which we above deduced the relation (61). Hence if we base our calculation on the classical expression for the kinetic energy T we have

$$\left(\frac{\nu}{u}\right)^2 = \left(\frac{1}{\lambda}\right)^2 = \left(\frac{\mu v}{h}\right)^2 = \frac{2\mu}{h^2} \cdot T = \frac{2\mu}{h^2} (E - V). \quad (70)$$

If we substitute this in (69) we get

$$\Delta\psi + \frac{8\pi^2 \mu}{h^2} (E - V) \psi = 0, \quad (71)$$

which is Schrödinger's *wave-equation*.

The picture of an electron moving in a field of force is thus replaced by that of an infinitely extended wave-field that satisfies equation (71). On account of the term $-V$ in the bracket the coefficient of ψ in this equation is variable from place to place; moreover, it assumes different values according to the total amount of energy we consider assigned to the motion of the electron.

25. The Idea of Proper Values

The one-dimensional equation

$$\frac{\partial^2 \psi}{\partial x^2} - K \frac{\partial^2 \psi}{\partial t^2} = 0 \quad (72)$$

that corresponds to (67) represents, as we know, the equation of a vibrating string. It offers the simplest example of the occurrence of proper values in this class of problems. With the assumption

$$\psi(x, t) = f(x)g(t)$$

it becomes

$$gf'' - Kfg'' = 0 \quad \text{or} \quad \frac{f''(x)}{f(x)} = K \frac{g''(t)}{g(t)},$$

and this equation is identically satisfied in x and t , if both sides assume a value independent of x and t , that is, assume a constant value, which we shall call λ . (This λ has nothing to do with the wave-length λ that occurs in (60) and (61).) This leads to the equation in $f(x)$:

$$f''(x) + \lambda f(x) = 0. \quad (73)$$

If the string has the length l and is firmly held at both ends, the following boundary conditions become added:

$$f(0) = 0 \quad \text{and} \quad f(l) = 0. \quad (74)$$

A solution of equation (73) that satisfies the first boundary condition is given by

$$f(x) = A \sin \sqrt{\lambda} x. \quad (75)$$

In order that the second boundary condition of (74) should be fulfilled it is necessary to have

$$\sqrt{\lambda} = n \frac{\pi}{l}; \quad n = 1, 2, 3 \dots \quad (76)$$

It is easy to show that the equation (73) has solutions, which fulfil both boundary conditions and do not vanish identically,

only for the values of λ given by (76). The values of the parameter λ given by (76) are called the *proper values* of the problem represented by the differential equation (73) and the boundary conditions (74). The solutions (75) are called the *proper functions* of the problem.

Schrödinger's theory hinges on the following very remarkable circumstance. The wave-equation (71) has solutions that are unique, continuous and finite for all values of the co-ordinates * only if the parameter E that occurs in it assumes either any positive values whatsoever or a discrete series of negative values. Now the negative values of the energy exactly correspond to the orbits of the electron that lie in finite regions; further, the discrete proper values of E agree exactly with the energy values of Bohr's quantum orbits.

26. Application to the Hydrogen Atom

If the field of force that acts on the electron is the Coulombian field of a nucleus with charge $+e$ we have

$$V = -\frac{e^2}{r}, \quad (77)$$

and the equation (71) becomes

$$\Delta\psi + \frac{8\pi^2\mu}{h^2}\left(E + \frac{e^2}{r}\right)\psi = 0. \quad (78)$$

The additive constant of the potential energy has been chosen so that $V = 0$ for $r = \infty$. This signifies for the mechanical aspect of the problem that the Kepler ellipses situated entirely in finite regions have negative energy-values. The value $E = 0$ corresponds to parabolic orbits; positive energy-values correspond to hyperbolic orbits.

To integrate equation (78) we transform it, in conformity with the character of the problem, to polar co-ordinates and make the following assumption for ψ :

$$\psi = f(r, \theta, \phi)e^{2\pi i vt} = \chi(r)g(\theta, \phi)e^{2\pi i vt}. \quad (79)$$

Precisely as in the above example of the vibrating string the equation (78) may now be split up into one in $g(\theta, \phi)$ and another in $\chi(r)$. The solutions of the equation in $g(\theta, \phi)$ come out as a series of functions which are called surface harmonics. If we denote the zero points of one of these functions by $\theta_1, \theta_2, \dots$ and $\phi_1, \phi_2, \dots$, respectively, the equations

$$\theta = \theta_1, \theta = \theta_2, \dots \quad \text{and} \quad \phi = \phi_1, \phi = \phi_2, \dots$$

* This postulate is called the natural boundary condition of the problem.

represent cones and meridian planes respectively, on which $g(\theta, \phi)$ and hence also ψ is always equal to zero. They are the nodal planes of the vibration represented by the solution in question. If we denote the number of nodal planes belonging to a definite solution $g(\theta, \phi)$ by $l + 1$, ($l = 0, 1, \dots$), then the corresponding function $\chi(r)$ must satisfy the equation

$$\frac{d^2\chi}{dr^2} + \frac{1}{r} \frac{d\chi}{dr} + \left[\frac{8\pi^2\mu}{h^2} \left(E + \frac{e^2}{r} \right) - \frac{l(l+1)}{r^2} \right] \chi = 0. \quad (80)$$

The surface harmonics that occur as solutions of the wave-equation in themselves already form a discrete series of functions given by the integral number l with its successive values. But only equation (80) together with the postulate that $\chi(r)$ must be unique and finite leads to the result that if E is to be negative, it can assume only definite discrete values (the proper values of the problem). We find for these proper values, according to Schrödinger ³⁷

$$E_n = - \frac{2\pi^2\mu e^4}{h^2} \cdot \frac{1}{n^2}; \quad n = l + 1, l + 2, l + 3 \dots \quad (81)$$

This expression agrees perfectly with that which resulted from Bohr's frequency conditions for the case of the hydrogen atom (equation (31), § 9). The difference between this and the original quantum theory is that now the introduction of quantum conditions is not necessary at all. Rather, the discrete energy values E_n come out of the vibration problem under consideration just as naturally as the system of harmonic overtones in the case of the vibrating string. This result must be regarded as a decided step forward.

27. The Interpretation of the Process of Radiation by means of Wave-Mechanics

If, retaining Bohr's frequency condition, we restrict ourselves to calculating only the frequencies of the lines, the finding of the proper values in itself completes the task; the frequencies may then be calculated with the help of (28). Like Heisenberg's mechanics, however, wave-mechanics goes further than this in also allowing us to calculate the components of intensity. We shall deal with this question later. First we wish to consider the pictorial (*anschaulich*) interpretation of the phenomenon of radiation by Schrödinger's theory.

The frequency condition (28) is doubtless to be regarded as one of the most securely founded pillars of the quantum theory. There are, however, certain difficulties in the ideas of quantum transitions between the individual stationary orbits, which

become essentially diminished in wave-mechanics. The picture of the hydrogen atom is drawn in all its details in Schrödinger's theory in contradistinction to Heisenberg's theory: it, of course, looks quite different from the Rutherford-Bohr planetary model. The atomic nucleus is now the centre of a field of stationary waves, and the number of nodal planes of this field differs according to the "state of excitation" of the atom. As already mentioned above, there exist meridional nodal planes $\phi = \text{const.}$ and nodal cones $\theta = \text{const.}$ To these there are also to be added nodal spheres $r = \text{const.}$ corresponding to the zero points of the solutions of (80). The number and position of the nodal surfaces is dependent on the two quantum numbers l and n . The geometrical form of the system of stationary waves is quite irrelevant for the sequel and the nodal planes are mentioned only as showing a similarity with the well-known pictures of stationary vibrations in membranes and similar objects, in which similar integral relations occur. But the way in which ψ -waves depend on the time has a direct physical meaning. In accordance with (65) and (68) a partial vibration belonging to the proper value E_n is given by

$$\psi = f_n(r, \theta, \phi) e^{\frac{2\pi i E_n t}{h}}. \quad (82)$$

Now if the atom is excited to perform two partial vibrations with the energy-values E_m and E_n simultaneously, the "chord" of the two vibrations will give rise to beats or difference tones whose frequency is given in agreement with (28) by

$$\nu(nm) = \frac{1}{h} (E_m - E_n).$$

Thus wave-mechanics gives a sort of interpretation of Bohr's frequency condition. The picture of the quantum switch of the electron from one stationary path to another becomes replaced by the picture that a partial vibration of energy E_m gradually passes over into one of energy E_n in which process, during the dying away of the one partial vibration and the gathering of strength of the other, the energy difference is radiated out in the form of an electromagnetic wave whose frequency agrees with the *beat frequency* of the two ψ -waves.*

* The picture of the mechanism of the emission of radiation that we have here sketched has been regarded as unacceptable by some theoretical physicists on account of the difficulties that occur in the many-electrons problem which will be discussed in § 30. Cf. in this connexion Born¹⁷ who clings essentially to the pictures of Bohr's stationary states and brings the field-scalar ψ (contrary to Schrödinger's interpretation which will be discussed in the next paragraph) into relationship with the probability of the atom being in a definite stationary state.

28. The Physical Meaning of the Field-Scalar ψ . Calculation of the Intensities of Spectral Lines

In the interpretation of the process of radiation outlined at the end of the last section the question naturally forces itself upon us: why is it just the beat-frequencies and not those of the primary ψ -waves themselves that are observed as electromagnetic waves? Schrödinger was able to give a quite plausible answer to this question by giving the ψ -waves, about whose physical character no assumption had yet been made at all, a definite physical meaning; they are to be vibrations of the electric charge of the electron continuously distributed over the whole wave-field. The density of charge ρ of the field is not to be given by the complex quantity ψ itself expressed in (82) but by the square of its absolute value. So if we designate the complex quantity * conjugate to ψ by $\bar{\psi}$, then, according to Schrödinger,

$$\rho = \psi\bar{\psi}. \quad (83)$$

As a consequence of this assumption it is found that a single partial vibration of ψ does not generate a variable electric moment of the atom, but that, on the other hand, when several partial vibrations occur simultaneously the electric charges oscillate with the frequency of the combination-tone or beat-tone. Suppose the two partial vibrations corresponding to the two proper values E_m and E_n are excited. For the sake of brevity we shall again introduce the angular velocity in terms of the number of radians per second (cf. p. 11):

$$\frac{2\pi E_n}{h} = \omega_n, \quad \frac{2\pi E_m}{h} = \omega_m.$$

Then

$$\begin{aligned} \rho = \psi\bar{\psi} &= (f_n e^{i\omega_n t} + f_m e^{i\omega_m t})(\bar{f}_n e^{-i\omega_n t} + \bar{f}_m e^{-i\omega_m t}) \\ &= f_n \bar{f}_n + f_m \bar{f}_m + f_n \bar{f}_m e^{i(\omega_m - \omega_n)t} + f_n \bar{f}_m e^{i(\omega_n - \omega_m)t}. \end{aligned} \quad (84)$$

The term of $\psi\bar{\psi}$ that changes with the time therefore represents a vibration whose frequency is given by

$$\nu(nm) = \frac{1}{2\pi}(\omega_m - \omega_n) = \frac{1}{h}(E_m - E_n). \quad (85)$$

According to the laws of classical electrodynamics such a vibration of electric charges must excite an alternating electromagnetic field of the same frequency.

* The square of the absolute value of a complex quantity is given by its product with the corresponding conjugate complex number:

$$(a + ib)(a - ib) = a^2 + b^2.$$

The picture of the physical character of the quantity ψ as given by the equation (83) not only serves to give us an understanding of Bohr's frequency condition, but also enables us to calculate the intensities of the spectral lines. As explained in § 4 the intensity components of the field of radiation are proportional to the square of the second derivatives, with respect to the time, of the components in question of the electric moment of the charges generating the fields. The word "component" is here to be understood in the two-fold sense mentioned in § 4, except that now we are not dealing with "genuine" Fourier components, since it is not now a question of the harmonic vibrations $n\nu$ of a fundamental frequency ν but rather of the individual frequencies of emission $\nu(nm)$. The formulæ of Hertz (21), (22) and (27) that were adduced to calculate the intensities refer, as their form shows, to a vibrating point-charge, whereas we are here dealing with vibrations of continuously distributed charges. The ψ -field of the atom extends, strictly speaking, over the whole of space. The part, however, within which the density of charge is *sensibly* different from zero is limited to atomic dimensions. For calculating the field of radiation at macroscopic distances from the centre of the atom we may therefore take our stand without misgivings on Hertz's formulæ, taking careful note, however, that in the present case the electric moment is not simply the product of the charge by the elongation of a vibrating corpuscle but is calculated by integrating over the whole ψ -field. Hence the x -component of the electric moment of the atom is given by

$$p_x = \int \rho x d\tau = \int x \psi \bar{\psi} d\tau, \quad (86)$$

where $d\tau$ signifies the three-dimensional element of volume and the integration is to be performed over the whole of space; x denotes the ordinary Cartesian co-ordinate. For wave-motion we depend, in conformity with (21) and (22), only on the part of the electric moment that varies with time; and this part occurs, as we explained above, only when several partial vibrations occur simultaneously and then in the form of beats. The partial vibration of the electric moment that belongs to the difference tone $\nu(nm)$ is given, according to (84), by

$$p_x(nm) = \int x \psi_n \bar{\psi}_m d\tau. \quad (87)$$

Further, it follows from (84) that we obtain the second time derivative of the moment by multiplying by $(2\pi\nu(nm))^2$; if we then square the product once again, then we get, according to

(27), the corresponding intensity component of the line with the frequency $\nu(nm)$.

Hence Schrödinger's theory, like Heisenberg's, yields unique rules for calculating the intensity components of spectral lines. The application of his theory to the Stark effect (which may be treated by a method of perturbations as shown by Schrödinger in his third paper⁴⁰) led to results that agree better with the results of observation than those of Kramers³¹ obtained on the basis of the Correspondence Principle.

29. Generalization of the Wave-Equation Connexion with Heisenberg's Theory

The wave-equation (71) of the one-electron problem, when multiplied by $-\frac{h^2}{8\pi^2\mu}$, assumes the form

$$\frac{1}{2\mu}\left(\frac{h}{2\pi i}\right)^2\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + V\psi - E\psi = 0, \quad (88)$$

or, expressed a little more simply, symbolically,

$$\left\{\frac{1}{2\mu}\left(\frac{h}{2\pi i}\right)^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V\right\}\psi - E\psi = 0. \quad (89)$$

Now the expression in the curly brackets is constructed analogously to the Hamiltonian function $H(p, q)$, given in equation (5), for a single point-mass. We have only to substitute the operational symbols

$$\frac{h}{2\pi i} \frac{\partial}{\partial x}, \quad \frac{h}{2\pi i} \frac{\partial}{\partial y}, \quad \frac{h}{2\pi i} \frac{\partial}{\partial z},$$

respectively, for p_x, p_y, p_z to get the bracketed expression on the left-hand side of (89). In this way the function $H(p, q)$ becomes an *operator*, the product of which with any quantity simply denotes the carrying out of the operation in question. If any operator contains, as in the above case, besides the operational symbols, actual numbers as summands (as here, for example, the quantity V), then these quantities are to be multiplied in just the ordinary way when the operation is executed. Let a be a pure number, then $\frac{\partial}{\partial x} + a$ is an operator which when applied to a function f gives

$$\left(\frac{\partial}{\partial x} + a\right)f = \frac{\partial f}{\partial x} + af.$$

The square of the operator $\frac{\partial}{\partial x}$ of course denotes two successive differentiations. With the help of this symbolic method the rules for setting up equation (71) may now be put into the following form. We form the operator corresponding to the Hamiltonian function $H(p, q)$ by making the symbolic substitution

$$p_k \rightarrow \frac{h}{2\pi i} \frac{\partial}{\partial q_k}. \quad (90)$$

The symbolic product of the operator that results in this way with the wave-function ψ must be equal to the ordinary product $E\psi$. Expressed in a formula this means

$$\left\{ H\left(\frac{h}{2\pi i} \frac{\partial}{\partial q}, q\right) - E \right\} \psi = 0. \quad (91)$$

(We must here remark that before making the symbolic substitution (90) the expression in the curly bracket of (91) would vanish. For H denotes the energy function and E the numerical value of the energy. Since we are dealing with conservative systems, we always have $H(p, q) = E$. We must also point out that a transformation to other than Cartesian co-ordinates, for example, to polar co-ordinates, may be performed only after the change (90) has been effected.)

The form of the wave-equation given by (91) is now no longer tied to the one-electron problem and it therefore readily suggests itself to regard equation (91) as the perfectly general wave-equation. This assumption has, indeed, actually been made by Schrödinger and, with its help, very remarkable relationships with Heisenberg's theory have come to light. The theorems given in the sequel hold independently of the form of the function $H(p, q)$, that is, they apply to any arbitrary atomic problem.

(a) The rather arbitrarily introduced commutation laws (38) and (45)-(47) of Heisenberg's theory become, if we replace the momenta by partial differential quotients according to (90), a quite natural operational rule. For if f is a function of the quantities q_k and also any other variables, then we have, as can easily be verified,

$$\frac{\partial}{\partial q_n}(q_m f) - q_m \frac{\partial f}{\partial q_n} = \begin{cases} f & \text{for } n = m, \\ 0 & \text{for } n \neq m. \end{cases} \quad (92)$$

In the symbolism of operators this may be written

$$\frac{\partial}{\partial q_n} q_m - q_m \frac{\partial}{\partial q_n} = \begin{cases} 1 & \text{for } n = m, \\ 0 & \text{for } n \neq m. \end{cases} \quad (93)$$

If we multiply the equation (93) by $\hbar/2\pi i$ and then carry out the substitution given in (90) in the reverse direction we immediately arrive at Heisenberg's relations.*

(b) The proper values E_n of equation (91) agree exactly with the diagonal terms $H(nn)$ of the solution matrix, that is with the energy-levels which are calculated on the basis of Heisenberg's calculus for the problem given by the same function $H(p, q)$.

(c) If from the proper functions of the equation (91) we form the definite integrals given in (87) and used for calculating the intensity components, then these integrals agree, except for a normalizing factor, with the matrix-elements $q(m, n)$, which would come out of the Heisenberg theory:

$$q_k(nm) \sim \int q_k \psi_n \bar{\psi}_m d\tau; \quad (94)$$

here $d\tau$ denotes the product of all rectangular co-ordinate differentials: the integral is to be taken over the whole of configuration space.

(d) An analogous relation holds for the elements of the momentum matrix p :

$$p_k(nm) = \frac{\hbar}{2\pi i} \int \psi_n \frac{\partial \psi_m}{\partial q_k} d\tau. \quad (95)$$

30. Criticism of Wave-Mechanics

As is evident from the relationships just described between the two theories, wave-mechanics accomplishes in practice the same as the matrix theory. With the help of a system of solutions of the wave-equation (91) we can calculate the matrix components of Heisenberg's theory, and hence answer the relevant questions concerning frequencies and intensity components of spectral lines. In view of the great difficulties that attach to the calculations with infinite matrices this tracing back of the problem to the realm of partial differential equations, which has been well developed by mathematicians, must be regarded as a great advance in method. Further, since wave-mechanics and quantum mechanics lead in principle to the same final results as regards frequencies and intensities, the results enumerated in § 15 which speak in favour of quantum mechanics may equally well be considered as evidence in favour of Schrödinger's theory.

The generalization of wave-mechanics given in (91) and its

* It must be observed that the numbers n and m that occur in the equations (92) and (93) have the significance of co-ordinate indices and are not to be confused with the similarly named current numbers of the matrix-elements.

connexion with Heisenberg's theory has been acclaimed on the one hand as a triumph for wave-mechanics; on the other, however, it finally dashes the hopes raised by de Broglie's reflections that field-physics would become reinstated. "Field-physics" is taken as standing for the sum total of those theories which describe physical events causally in space and time by means of partial differential equations. Classical hydrodynamics, the theory of elasticity and the theory of the magnetic field given by Maxwell's equations are typical field theories. Classical physics had become completely wrecked on the quantum problems. Bohr's form of the quantum theory (and still more Heisenberg's quantum mechanics) signify a complete departure from field-physics. The birth of de Broglie's conceptions of phase-waves and still more Schrödinger's treatment of the one-electron problem led physicists to expect that it would be possible to describe quantum phenomena by a theory of the type of electrodynamics. But the present position shows this hope to be illusory. For in all problems in which several mutually reacting electrons occur, the ψ -waves are no longer vibrations in physical space of three dimensions or in Minkowski's four-dimensional world, but rather vibrations in the "configuration space" of the system. This configuration space is a pure abstraction. It is a manifold of just as many dimensions as the system in question has degrees of freedom. For example, for a system of four mass-points it has twelve dimensions. Thus although de Broglie's waves had no direct meaning, yet the vibration of the unknown something did take place in the physical space-time continuum, in Minkowski's "world." Now Schrödinger, by means of his interpretation of the field-scalar ψ given in equation (83) infused a physical "soul" into de Broglie's waves, but in doing so he drove them off the scene of action of the physical world and banished them into the spirit regions of multi-dimensional configuration space. Hence, although wave-mechanics is governed by a partial differential equation (71) or (91), respectively, there can be no question of a resuscitation of a field-physics in the sense of Maxwell's electrodynamics.

Such a return to a pure field-physics is perhaps impossible altogether. The opinions of physicists differ on this point. Whereas Einstein ²⁴ still nurses this hope, Bohr ³ and Heisenberg believe that they have arguments that, on grounds of principle, make a causal description of atomic happenings in space and time impossible.

Some of the defects that were held up against Heisenberg's theory in § 16 also apply to wave-mechanics. For this theory, too, takes over in the form of the Hamiltonian function $H(p, q)$

both the atomic model and the elementary laws that act between the particles unchanged from Bohr's theory. The interpretation of the field-scalar ψ given by (83), which endows it with a very clear and direct physical meaning, seems even to introduce an additional difficulty. For the wave-equation (78) is based, among other things, on Coulomb's elementary law in the form of the potential expression $-\epsilon^2/r$. The fluctuating density of charge $\rho = \psi\bar{\psi}$, however, behaves as if this law does not in the least concern it.

PART II

THE FURTHER DEVELOPMENTS SINCE 1926

31. Introduction

THE task of describing quantum-mechanics and wave-mechanics in the simplest possible language in a report of limited length is an almost impossible undertaking, and the authors are by no means convinced that they have come even within a reasonable distance of achieving this object in their present effort. Even in the first decade of its existence the quantum theory had developed into a science of considerable dimensions. After the appearance of the epoch-making works of Heisenberg and Schrodinger, a particularly rapid rate of development set in, with the result that within hardly more than four years a new branch of knowledge emerged out of thin air, so to speak, and grew into a subject which to be adequately treated would require about five hours of lectures per week throughout the academic year. Even at the present time the quantum theory is a more extensive branch of knowledge than was, for example, Maxwell's theory at the time when M. Abraham wrote his well-known textbook on electrical theory. This new science is distinguished, however, not only by its wide range, but also by its depth. The study of it is rendered so difficult by the unusual concepts and abstractions it involves and the elaborate mathematical language in which it is expressed that the majority of physicists even now are far from following the newer developments of the theory with real understanding. It is not to be wondered at that the non-professional physicist experiences still greater difficulty in penetrating into the channels of thought of the new theory. Nevertheless, an excursion into these realms should well repay the trouble for those who have sufficient mathematical training. The æsthetic sense of the abstract investigator cannot fail to experience continual pleasure in the beauty and harmony of spectra and their fine structure, as has been justly emphasized by Sommerfeld, while the investigator with practical aims and interests may hope to be enlightened on the question of the constitution of matter and the nature of chemical bonds. This circumstance justifies the great number of new accounts of quantum

and wave-mechanics among which the present report is to be counted.

In this second part of our report Section VII deals with the extension of the general foundations of wave-mechanics, namely, Schrödinger's time-dependent wave-equation and the relativistic equation, as well as the more recent generalization of the theory by Dirac. Before the Dirac theory is discussed a digression on the lines of reasoning and empirical facts of the older quantum theory that lead to the hypothesis of the spinning electron by Uhlenbeck and Goudsmit is interpolated.

Sections VIII and IX give a survey of the achievements of the theory up to 1929. The subjects include the theory of the Zeeman and Stark effects, the scattering of α -rays, the emission of electrons from cold metals, the Auger effect, radio-activity, dispersion, the photo-electric effect, the Compton effect, the many-bodies problem, the solution of the helium problem, the formation of molecules. Further, the relationship of the theory to the statistics of systems of many degrees of freedom (Bose-Fermi) is discussed.

The last section, X, gives a discussion of the attempts hitherto made to interpret the theory and contains comments on the concepts involved. It deals with wave-packets, the statistical interpretation of wave-mechanics, the two-fold nature of matter and radiation (corpuscular and undulatory), uncertainty relations and so forth. The present position may be characterized as follows. A mathematical formalism has been found which accounts in an astonishingly accurate way not only for certain details of spectra, but also for a number of other physical phenomena and the chemical behaviour of many elements; but this success was gained only by introducing very unfamiliar conceptions and ideas, which demanded an extensive revision of our view of the fundamental principles of physical events and of causality.

A few words may be added about the range of the material discussed. Although the authors have by no means fought shy of mathematical developments, having, indeed, given the method used in many calculations in considerable detail, they have been obliged to refrain from giving those parts of the theory that required the use of heavier mathematical apparatus. These parts include—we mention only the most important subjects—investigations that involve the theory of groups, the "quantizing of amplitudes," and the transformation theory in its exact form. The omission of these subjects is not, however, to be regarded as indicating an unfavourable judgment—unless, indeed, it be of the ability of the authors to give even a partially intelligent account of them at the present early stage.

CHAPTER VII

RELATIVITY AND ELECTRON SPIN

32. Recapitulation. The Application of the Wave-Equation to Conservative Systems

To help the reader to understand the sequel more easily we shall begin with a short recapitulation. The form of wave-mechanics given in Chapter VI, which is its simplest expression, relates to conservative systems, that is, it deals in particular with those atomic problems in which one or more electrons move in static fields of force, the action of the electrons on one another being negligible. The most general wave-equation at that stage was given by equation (91) :

$$\left\{ H\left(\frac{h}{2\pi i} \frac{\partial}{\partial q}, q\right) - E \right\} \psi = 0. \quad (91)$$

This equation is to be interpreted as follows. In the Hamiltonian function of the system, $H(p, q)$, the momenta p_k are to be replaced by the operators $\frac{h}{2\pi i} \frac{\partial}{\partial q_k}$, that is,

$$p_k \rightarrow \frac{h}{2\pi i} \frac{\partial}{\partial q_k}, \quad (90)$$

which causes the expression in the outer brackets of (91) to become a differential operator of the second order, which is to act on the "wave-function" ψ . The energy-parameter E here denotes the constant total energy of the conservative system, which is equal to the numerical value of the Hamiltonian function of the system :

$$E = H(p, q). \quad (96)$$

If the substitution of (90) were not carried out beforehand, the bracketed expression in (91) would vanish identically. As an example of the application of the symbolic expression used in (91) we shall consider the case of the motion of a single electron of mass μ in a potential field $V(q)$. We have

$$H = \frac{1}{2\mu} \sum_{\infty} p_i^2 + V(q).$$

The wave-equation corresponding with this, when multiplied by $-\frac{8\pi^2\mu}{h^2}$, is

$$\Delta\psi + \frac{8\pi^2\mu}{h^2}(E - V)\psi = 0. \quad (71)$$

In contradistinction to the objective of the traditional physics of fields the chief problem here is not to integrate equation (71), that is, not to determine the function $\psi(q)$, which is its solution, but rather to determine those proper values of the parameter E , for which (71) possesses solutions that do not vanish identically and which satisfy the natural boundary conditions (that is the solutions must be finite and continuous and vanish at infinity). The discrete proper values E of the problem represent the energy-levels of the system, that is, the energies of the stationary orbits in the language of the older quantum theory. The frequencies of the spectrum emitted from the system in question are given by Bohr's frequency condition,

$$h\nu = E_m - E_n.$$

Thus the frequencies of the spectrum are obtained merely by determining the proper values E_n of the problem. To calculate the intensities of the lines (and the states of polarization given by the intensities) we further require the solutions ψ_n that belong to the proper values E_n ; these solutions are called the proper functions of the equations (71) or (91), respectively. It has been shown in §§ 28 and 29 how we get from the proper functions to the matrices $q(nm)$ and $p(nm)$. The intensities of the lines are then calculated by means of equation (27), § 4.

The above formulæ thus enable us to calculate the frequencies and intensities of line-spectra of those atomic systems in which the electron or electrons are situated in a statical field of force. This assumption is accurately realized only in the case of the one-electron problem and the corresponding perturbation effects (for example, the Stark effect). To these restrictions on the range of application of the formulæ there is to be added another of a formal nature due to the mathematical scheme that is used. The replacement of the momenta by differential operators, as given by (90), has a sense only for the p_k 's themselves and for positive integral powers of them. For example, if n is an integer the corresponding rule of substitution runs

$$p_k^n = \left(\frac{h}{2\pi i}\right)^n \frac{\partial^n}{\partial q_k^n}. \quad (90a)$$

So a wave-equation can be formed on the model of a Hamiltonian function only if H can be expressed by integral powers and

products of powers of the p_k 's. This circumstance has played a part in the setting up of the relativistic wave-equation (§ 34).

33. The Wave-Equation Depending on the Time

In atomic problems concerned with the motion of electrons in fields that vary with the time, for example, in the problem of dispersion, equations of the type (71) or (91) are no longer sufficient. In these equations the time-dependence of the ψ -function was eliminated in accordance with § 24 by the assumption usually made in the theory of vibrations, namely, by

$$\psi = f(x, y, z)e^{2\pi i\nu t}. \quad (68)$$

The equations in question therefore contain only the space-co-ordinates as independent variables; instead of wave-equations, they should more correctly be called "vibration" or "amplitude equations."

We read in Schrödinger's fourth communication that "the urgent need was felt to extend the theory to non-conservative systems, as it is only in this way that we can study the behaviour of the system under the action of given external forces, for example, of a light wave or of a passing outside atom." * Now if the potential energy contains the time explicitly, it is in general impossible to satisfy the wave-equation by means of an assumption of the form (68), just as, analogously, the equation for the forced vibrations of an oscillator,

$$m\ddot{x} + kx = f(t),$$

cannot be fulfilled by a condition that is assumed to vary periodically with the proper frequency of the freely vibrating system. In such cases, then, the amplitude equation is of insufficient range and we must have recourse to the real wave-equation.

This reversion to the real wave-equation cannot be undertaken by reversing the procedure of § 24 which led from the classical wave-equation

$$\Delta\psi - \frac{1}{v^2}\ddot{\psi} = 0 \quad (67)$$

to Schrödinger's equation (71). For this method uses the relation $\ddot{\psi} = -4\pi^2\nu^2\psi$, corresponding to assumption (68), which, as above mentioned, is *not* to be used for the general case of non-conservative systems.

To set up the time-dependent Schrödinger equation we shall

* The extension of the theory to non-conservative systems had been carried out earlier by Heisenberg within the framework of his matrix theory.

therefore not link up with the wave-equation (67) of field-physics at all. Rather, we shall obtain the desired equation by an appropriate transposition of the formalism by which the partial differential equation of Hamilton and Jacobi (to be referred to as H.P. in the sequel) is formed from the energy expression assumed or Hamilton's function, respectively. It must be stated that in the arguments outlined below we are dealing with a purely formal method for setting up a time-dependent wave-equation, a method which by no means has the character of inevitability. We shall outline the derivation of the equation in two steps, namely (a) for conservative systems, (b) for non-conservative systems.

(a) *Conservative Systems*.—Let the energy equation be given by

$$H(q_1 \dots q_f, p_1 \dots p_f) - E = 0. \quad (97)$$

The integration of the corresponding system of $2f$ equations of motion (12) can, as we know, be reduced to the integration of a single partial differential equation of the first order (H.P. equation), which is obtained by making the substitution (97) :

$$p_k = \frac{\partial S}{\partial q_k}, \quad (98)$$

where S is the unknown function (of action) which is to be determined from the resulting partial differential equation of the first order. The effect of the substitution (98) is to convert the equation (97) into the H.P. equation :

$$H\left(q_1 \dots q_f, \frac{\partial S}{\partial q_1} \dots \frac{\partial S}{\partial q_f}\right) - E = 0. \quad (99)$$

The application of this equation to the classical treatment of a problem of motion, that is, as if we are dealing with particle mechanics, depends on the following circumstance. If we have found a complete integral

$$S(q_1 \dots q_f, E, \alpha_1 \dots \alpha_f) \quad (100)$$

of equation (99), that is, an integral which depends on the f independent variables q_k , the energy parameter E and, in addition, f arbitrary integration constants α_k , then the integrals of the equations of motion are given implicitly by the system of equations :

$$\frac{\partial S}{\partial q_k} = p_k, \quad \frac{\partial S}{\partial \alpha_k} = \beta_k, \quad \frac{\partial S}{\partial E} = t - t_0, \quad (101)$$

in which the β_k 's are f further arbitrary constants. After E has been eliminated the equations (101) represent a system of

2*f* equations, which give us the unknowns q_k and p_k as functions of the time and the integration constants. These relationships which have been known to mathematicians for nearly a hundred years were of importance for the train of thought that led Schrödinger to create wave-mechanics. But they play no further part in the method of presentation which we are about to give.

From this and the preceding paragraphs we see that we can arrive at both the H.P. equation as well as Schrödinger's equation by starting out from the energy equation $H(q, p) = E$ and replacing the momenta by the partial derivatives with respect to the corresponding canonically conjugate co-ordinates. What difference is there in the routes by which the two equations (H.P. and Schrödinger's) are reached? In the case of the H.P. equation the differential operator $\frac{\partial}{\partial q_k}$ is *first* performed on the function S , and then the result of this operation, namely the partial derivative $\frac{\partial S}{\partial q_k}$, is introduced as the argument into the function H , after which the operations prescribed by the functional structure of H , for example, powers of $\frac{\partial S}{\partial q_k}$, are performed. As an illustration take a point-mass in a potential field $V(q)$. The energy-law gives

$$H = \frac{1}{2\mu}(p_1^2 + p_2^2 + p_3^2) + V(q) = E. \quad (102)$$

The Hamilton-Jacobi partial differential equation (H.P.) is

$$\frac{1}{2\mu} \left\{ \left(\frac{\partial S}{\partial q_1} \right)^2 + \left(\frac{\partial S}{\partial q_2} \right)^2 + \left(\frac{\partial S}{\partial q_3} \right)^2 + V(q) \right\} - E = 0. \quad (103)$$

In contrast with this the transition from the energy-law to Schrödinger's equation is effected as follows. The p_k 's in H are replaced by the operators $\frac{\hbar}{2\pi i} \frac{\partial}{\partial q_k}$ and the operations prescribed by the function H , such as raising to powers and so forth, are performed symbolically on these operators *before* the differentiations are performed on the ψ -function. This gives rise to a composite differential operator $(H - E)$ which is to act on the wave-function ψ only *after* the operator has been completed. For example, the Schrödinger equation corresponding to the energy-law (97) runs:

$$\left\{ -\frac{\hbar^2}{8\pi^2\mu} \left(\frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} + \frac{\partial^2}{\partial q_3^2} \right) + V(q) - E \right\} \psi = 0. \quad (104)$$

(b) *Non-conservative Systems*.—In general we are concerned in particle mechanics with a time-dependent function of action, which is related to the time-independent part S of the action-function discussed under (a) by the equation

$$W = E - St, \quad (105)$$

which is valid only for conservative systems. The Hamiltonian function is defined for non-conservative systems by

$$H = \sum p_k \dot{q}_k - T + V, \quad (106)$$

in which T represents the kinetic energy of the system. In the general case H must *not* be identical with the total energy E . If in (106) we replace the $\dot{q}_k$'s by the p_k 's, H again becomes a function of the q 's and p 's and the time t :

$$H(q_1 \dots q_f, p_1 \dots p_f, t). \quad (107)$$

The corresponding H.P. equation then runs

$$H\left(q_1 \dots q_f, \frac{\partial W}{\partial q_1} - \frac{\partial W}{\partial q_f}, t\right) + \frac{\partial W}{\partial t} = 0. \quad (108)$$

Thus we arrive at the time-dependent H.P. equation formally by replacing the energy (taken as negative) in the relation $H(q, p, t) - E = 0$ (which, we must remark in passing, is no longer fulfilled for a constant E in the case of non-conservative systems) by the differential quotient $\frac{\partial W}{\partial t}$. The corresponding transition to the time-dependent Schrödinger equation will then naturally have to be effected in such a way that the difference $H - E$ becomes an operator; this is done by first replacing the momenta in accordance with (90) by the corresponding derivatives with respect to the co-ordinates and then making the substitution

$$(-E) \rightarrow \frac{h}{2\pi i} \frac{\partial}{\partial t}. \quad (109)$$

The desired time-dependent Schrödinger equation therefore runs

$$\left\{ H\left(q, \frac{h}{2\pi i} \frac{\partial}{\partial q}, t\right) + \frac{h}{2\pi i} \frac{\partial}{\partial t} \right\} \psi = 0. \quad (110)$$

The function $\bar{\psi}$, which is the conjugate complex function to ψ , must then of course satisfy the equation which is obtained from (110) by replacing i by $-i$:

$$\left\{ H\left(q, -\frac{h}{2\pi i} \frac{\partial}{\partial q}, t\right) - \frac{h}{2\pi i} \frac{\partial}{\partial t} \right\} \bar{\psi} = 0. \quad (110)^*$$

For the case where H has the form

$$H = \frac{1}{2\mu} \sum p_k^2 + V(q, t)$$

(110) becomes

$$\Delta\psi - \frac{8\pi^2\mu}{h^2} V\psi \pm \frac{4\pi i\mu}{h} \frac{\partial\psi}{\partial t} = 0, \quad (111)$$

where the upper sign holds for ψ and the lower for $\bar{\psi}$. The striking feature about this equation is that it is of the second order in the space-co-ordinates and of the first in t . Thus it is not of the type of the wave-equation (67) of field-physics but of the type of the equation which expresses the conduction of heat. The fact that the solutions of (111) are periodic functions of the time and do not depend exponentially on it as in the solutions of the equation of the conduction of heat is due to the coefficient of $\frac{\partial\psi}{\partial t}$ in (111) being imaginary.

The application of the time-dependent wave-equation (111) to the Compton effect and the problem of dispersion will be given in §§ 47-49.

34. The Relativistic Wave Equation according to Schrödinger

The discussion so far has referred to the translation of the "classical," that is, non-relativistic equations of motion into the language of wave-mechanics. The need was next felt to extend these arguments to relativistic mechanics. The reasons for this are obvious: according to the principle of relativity all physical laws must be invariant with respect to the Lorentz-Einstein transformation. This postulate has repeatedly shown itself to be of great heuristic value; we need only mention the law of the inertia of energy which was first deduced quite generally from the principle of relativity. The introduction of relativity mechanics into the realm of the quantum theory of spectra had, in particular, given us the first explanation of the fine-structure of hydrogen-like spectra. Thus wave-mechanics would be incomplete if it could not be brought into a form which is invariant with respect to the Lorentz-Einstein equations of transformation, as demanded by the theory of relativity.

A further weakness of wave-mechanics, as expressed in the form so far given, namely, in equations (71), (91) and (111), is that it refers only to atomic mechanical systems in which the forces may be derived from a scalar potential $V(q)$. Now one of the most important tests of spectral theory is given by the

Zeeman effect, the resolution of the energy-levels in an external magnetic field. The forces that act on the electron in the presence of magnetic fields cannot be derived, as we know, from a scalar potential V . Besides depending on the co-ordinates they also depend on the velocity of the particles. The ponderomotive forces that an electromagnetic field given by the field-vectors $\mathbf{E}$ and $\mathbf{H}$ exerts on a charge e moving with velocity $\mathbf{v}$ is given by

$$\mathbf{f} = e(\mathbf{E} + \frac{1}{c}[\mathbf{v}\mathbf{H}]). \quad (112)$$

Here the electric intensity of field $\mathbf{E}$ and the magnetic intensity of field $\mathbf{H}$ may be derived, in the general case of an electromagnetic field varying with the time, from a scalar potential V and a vector potential $\mathbf{A}$ in the following way :

$$\mathbf{E} = -\text{grad } V - \frac{1}{c}\dot{\mathbf{A}} \quad \mathbf{H} = \text{curl } \mathbf{A}, \quad (113)$$

where V and $\mathbf{A}$ themselves are solutions of the generalized Poisson equations :

$$\Delta\mathbf{A} - \frac{1}{c^2}\ddot{\mathbf{A}} = -4\pi\frac{\rho\mathbf{v}}{c} \quad \Delta V - \frac{1}{c^2}\ddot{V} = -4\pi\rho, \quad (114)$$

where ρ and $\mathbf{v}$ denote the density and velocity of the field-producing charges, while $\mathbf{v}$ is the velocity of the point-charge e which is influenced by the field.

Our task, then, is to bring the wave-equation into a relativistically invariant form and at the same time to generalize it so that it will allow us also to treat problems in which the forces acting on electrons are given by the equations (112) to (114).

Schrödinger,⁴¹ as well as Gordon⁸⁶ and Klein⁸⁹ succeeded in achieving both these objects by applying the formal method (explained in the previous section) of translating the wave-equation of the H.P. equation into that form of it which describes in a relativistically invariant form the motion of an electron under the influence of an electromagnetic field.

The equation in question runs

$$\begin{aligned} \left(\frac{1}{c}\frac{\partial W}{\partial t} + \frac{e}{c}V\right)^2 - \left(\frac{\partial W}{\partial x} - \frac{e}{c}\mathbf{A}_x\right)^2 - \left(\frac{\partial W}{\partial y} - \frac{e}{c}\mathbf{A}_y\right)^2 \\ - \left(\frac{\partial W}{\partial z} - \frac{e}{c}\mathbf{A}_z\right)^2 - \mu^2 c^2 = 0. \end{aligned} \quad (115)$$

Now the rule for converting the H.P. equation into a wave-equation is given by the symbolic substitution

$$\frac{\partial W}{\partial t} \rightarrow \frac{h}{2\pi i} \frac{\partial}{\partial t}, \quad \frac{\partial W}{\partial x} \rightarrow \frac{h}{2\pi i} \frac{\partial}{\partial x}, \text{ etc.} \quad (116)$$

In view of the remarks in the previous section the differentiations denoted by the operators (116) must be performed *after* the squares have been taken, and in the present case the operators must act only on the wave-function ψ .*

For example, the term of the resulting wave-equation which is due to the first bracket in (115) is

$$-\left(\frac{h}{2\pi c}\right)^2 \frac{\partial^2 \psi}{\partial t^2} + \frac{2e}{c} V \frac{h}{2\pi i} \frac{\partial \psi}{\partial t} + \left(\frac{e}{c}\right)^2 V^2 \psi.$$

The complete relativistic-magnetic equation which results when the substitution (116) has been effected and after we have multiplied by $-4\frac{\pi^2}{h^2}$ is

$$\Delta\psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} \mp \frac{4\pi ie}{hc} \left(\frac{V}{c} \frac{\partial \psi}{\partial t} + \mathbf{A} \text{ grad } \psi \right) + \frac{4\pi^2 e^2}{h^2 c^2} \left(V^2 - \mathbf{A}^2 - \frac{\mu^2 c^4}{c^2} \right) \psi = 0. \quad (117)$$

By using the customary nomenclature of relativistic electrodynamics equation (117) may be written in a form which brings out more clearly its symmetry with respect to the space-time co-ordinates. We set

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict, \quad (118)$$

$$\Phi_1 = \mathbf{A}_x, \quad \Phi_2 = \mathbf{A}_y, \quad \Phi_3 = \mathbf{A}_z, \quad \Phi_4 = iV, \quad (119)$$

$$p_k = \frac{h}{2\pi i} \frac{\partial}{\partial x_k}, \quad (120)$$

where p_k now no longer denotes a momentum component, but rather is already that operator which is to replace the momenta in passing over to the wave-equation. Lastly, let μ_0 denote the rest-mass of the electron. Then the resulting symmetrical form of the wave-equation is

$$\left\{ \sum_1^4 \left(p_k + \frac{e}{c} \Phi_k \right)^2 + \mu_0^2 c^2 \right\} \psi = 0. \quad (121)$$

Again the expression in the brackets is to mean that first the operations within the brackets are to be performed and only

* According to Sommerfeld (§ 9, "Wave Mechanics," Messrs. Methuen & Co. Ltd., 1930) the operators (116) have also to act on the potential functions V and $\mathbf{A}$. But, owing to the condition $\text{div } \mathbf{A} + \frac{1}{c} \dot{V} = 0$, in summing over the four co-ordinates those terms which contain the derivatives of the potentials vanish. Hence in the present case it is simpler to agree to allow the operators to act only on the wave-function ψ itself.

then are the operators p_k or their powers, respectively, to act on ψ .

In equation (121) the wave-equation of the one-body problem has the most general and most complete form possible at the stage of development so far reached in this report of the theory. It is natural to enquire whether it is also to be regarded as the final form. The answer depends on the result of investigating with what success the equation may be applied to problems outside the region originally envisaged by the wave-equation. These problems are the Zeeman effect and the fine-structure of the hydrogen lines.

To calculate the Zeeman effect we simplify equation (121) by cancelling the relativistic terms (higher powers of $1/c$) and by retaining only the first powers of the components of the magnetic field strength. This gives us the approximate equation *

$$\Delta u + \frac{4\pi i}{h} \left\{ \frac{e}{c} (\mathbf{A} \cdot \text{grad } u) - \mu_0 \frac{\partial u}{\partial t} \right\} - \frac{8\pi^2 \mu_0}{h^2} e V u = 0, \quad (122)$$

which in fact leads to exactly the same resolutions in the case of the normal Zeeman effect as the original Bohr theory. The progress made in the wave-mechanical deduction consists in the fact that by forming the Heisenberg matrices in accordance with § 28 we arrive at unique selection and polarization rules for the Zeeman components (Sommerfeld, "Wave Mechanics," § 10).

To treat the problem of relativistic fine-structure we specialize the general equation (121) by setting $\mathbf{A} = 0$, and, in view of the static character of the problem, we again introduce the time-independent wave-function called u by making the assumption

$$\psi = u(x, y, z) e^{\frac{2\pi i}{h} Et}. \quad (123)$$

We then get the relativistic equation for static fields

$$\Delta u + \frac{4\pi^2}{h^2 c^2} \{ (E + \mu_0 c^2 - eV)^2 - \mu_0^2 c^4 \} u = 0, \quad (124)$$

which, when the terms containing c^2 in the denominator are struck out, gives the original Schrödinger equation (71). If

* The variable u that occurs in (122) is not directly identical with the wave-function ψ that occurs in (121), but rather results from it by separating out a time factor; this is done by assuming for the wave-function

$$\psi = u e^{\mp \frac{2\pi i}{h} \mu_0 c^2 t}.$$

In contradistinction to Sommerfeld we call the time-dependent wave-function ψ and the wave-function independent of the time u .

we work out the proper values of this equation we get the ordinary hydrogen terms of the form R/n^2 , every term as in the older theory being resolved n times, with the difference that now the separation is much wider than before. Thus whereas Sommerfeld's original theory was able to give the fine-structure of the hydrogen spectrum correctly, Schrödinger's theory (and hence also Heisenberg's—since it is in practice equivalent to that of wave-mechanics) would yield a fine-structure of the hydrogen spectrum that would be appreciably coarser than that given by the actually observed values. The disclosure of this circumstance was not able, however, to shake the confidence felt in the new theory, as it was surmised even at that time with considerable certainty that the fine-structure of hydrogen is also influenced by a second factor which, when combined with the relativistic differences of levels, would give the true final picture of the resolution. This second factor is given by the electron spin, introduced by Uhlenbeck and Goudsmit, which we shall discuss in the following sections.

35. A Survey of the Development of the Theory of Multiple Quantizing *

To set out the arguments which led to the enunciation of the hypothesis of the spinning electron we must go back a little way and recount the history of the development of Bohr's atomic theory. In doing so we shall use throughout §§ 35-37 the ideas of electronic orbits of the older quantum theory and shall subsequently trace out the avenues leading to the ideas of quantum and wave-mechanics.

The first great advance made in the quantum theory of spectra after the pioneer work of Bohr was due to Sommerfeld who in 1915 succeeded in giving on the basis of the theory of relativity an explanation of the fine-structure of the hydrogen spectrum. This resulted in the important discovery that the energy-levels of the stationary states of one-electron systems form not a one-dimensional but a two-dimensional manifold and so are to be ordered according to two current numbers, since the stationary states themselves are also correspondingly characterized by two quantum numbers. The stationary orbits of the hydrogen electron are Kepler ellipses whose major axis

* The phrase "multiple quantizing" is to be understood here in the sense of the older quantum theory; it refers to the introduction of *several* quantum numbers (Sommerfeld, Planck, Schwarzschild) and has nothing to do with the same term as applied to the quantizing of amplitudes in the scheme of quantum electrodynamics recently formulated.

is proportional to the square of the "principal quantum number" n and whose semi-latus-rectum is proportional to the square of the "subsidiary quantum number" k (also called azimuthal quantum number). According to the general form (29) of the quantum conditions the product of the quantum numbers with h are equal to the "action variables" I_k of the problem. The relations *

$$I_1 = nh, \quad I_2 = kh$$

hold, in which, according to what was said above, I_1 is proportional to the square root of the major axis and $\frac{I_2}{2\pi}$ represents the "angular momentum of the orbit," that is, the moment of momentum of the revolving electron with respect to the nucleus. If we choose $\frac{h}{2\pi}$ as the unit of moment of momentum, the orbital angular momentum becomes simply equal to k .

According to classical mechanics the major axis of the orbital ellipse is fixed in space. The motion of the electron may be expanded in accordance with (20) in a Fourier series with respect to a single fundamental frequency as argument, namely, the frequency of revolution of the electron. According to the fundamental formula

$$\nu_k = \frac{\partial H(I_1 \dots I_f)}{\partial I_k} \quad (17a)$$

we see that since $\nu_2 = 0$ the energy H depends only on the one action variable I_1 , and hence only on the quantum number n : orbits having the same major axis and but different eccentricities have the same energy. Multiply periodic systems of this type (also called conditionally periodic), in which the degree of periodicity, that is, the number of mutually independent fundamental (or intrinsic) frequencies, is less than the number of degrees of freedom, are called *degenerate*. Corresponding to the classical characteristic of degeneracy, namely, the coincidence of certain fundamental frequencies (or the occurrence of linear integral relationships between them) we have in the quantum theory, in accordance with (17a) or with the Bohr frequency condition that is related to it, the coincidence of energy-levels that belong to different quantum states. Thus in the quantum theory a system is degenerate if certain stationary states with different

* In § 8 the various quantum numbers were denoted by n_k (see (29)), where the letter k was used as an index and is not to be confused with Bohr's subsidiary or azimuthal quantum number k . The connexion between the notations of § 8 and the present § is given by: $n_1 = n$, $n_2 = k$.

quantum numbers have the same energy or, translated into the newer language of wave-mechanics, if one and the same proper value has several different proper functions belonging to it. Now we know that the degeneracy may be overcome by the agency of small perturbing forces: through them new independent fundamental frequencies occur in the motion and the energy-levels that are coincident in the degenerate system separate. A perturbing force capable of bringing about this separation is given by the relativistic change of mass of the electron, which acts like a central force that does not obey Coulomb's law exactly. Under its influence the electron describes a Kepler orbit whose perihelion performs a slow rotation in which the angle between the major axis and a fixed direction in space is precisely the angle variable $w_2 = \nu_2 t$ canonically conjugate to I_2 . By (17a) the orbital energy will then depend on the major axis and, in addition, the orbital moment of momentum and hence, on account of (29), on the azimuthal quantum number k . The frequency of revolution ν_2 of the rotation of the perihelion due to the variation of mass is of a much smaller order of magnitude than the frequency of revolution ν_1 of the electron itself. Consequently, according to (17a), the energy depends only slightly on I_2 and so on the azimuthal quantum number k , that is, the energy-levels belonging to different values of k differ only very little from one another. Hence the resulting fine-structure of the individual hydrogen lines can be shown experimentally only with difficulty.

The idea of multiple quantizing was introduced into spectral theory by Sommerfeld's theory of fine-structure and almost simultaneously by the treatment of the Stark and the Zeeman effects. Its fruitfulness became particularly manifest when Bohr succeeded in extending corresponding arguments to the theory of alkali-spectra. Among the optical-spectra of many-electrons systems the spectra of the alkalis are most closely related to the hydrogen spectrum because they are produced by quantum transitions of one electron, the so-called "radiating electron" (*Leuchtelektron*), which in the mean is at a greater distance from the nucleus than the remaining electrons of the atom, so that its motion can be described approximately by that of a point-charge in a central field. The atomic core, that is, the system which consists of the nucleus and the remaining electrons of the atom, generates a field that approximates to spherical symmetry, but such that the force no longer *exactly* depends on the distance in accordance with Coulomb's law, since the atomic core acts like a point-charge only at great distances. The departures from Coulomb's law are equivalent to the effect

of perturbing forces, which like the relativity corrections overcome the degeneracy of the system and produce a separation of the energy-levels that belong to the same n but to different k 's. The difference, as compared with hydrogen, consists in the fact that the perturbing forces of the core-electrons are much more effective than the relativistic change of mass, so that the separation of the energy-levels is considerably greater than in the case of hydrogen. This means that we are no longer dealing with a fine-structure but with a sub-division of the spectrum into different rows of spectral series.

Earlier than this the empirical research of spectroscopists had disclosed that the spectral terms of the alkalis are to be

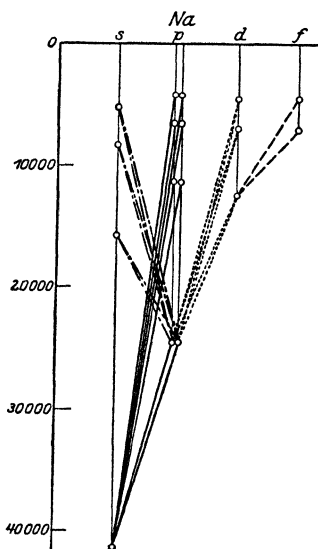


FIG. 1.—Scheme of the terms of the sodium spectrum. (From F. Hund, "Linienspektren," Berlin, Springer, 1927.)

ordered according to two current numbers (*Laufzahlen*). These current numbers are now to be identified, except for an additive constant, with the quantum numbers. Fig. 1 depicts the well-known scheme of the sodium terms. The heights of the individual points above a zero-level (not shown in the figure) represents the energy-value of the terms. The terms in the same vertical column belong to the same values of k . For reasons to be given later, it has been found expedient to introduce the subsidiary quantum number $l = k - 1$ and to abandon k .* The series of terms called $s, p, d, f \dots$ by spectroscopists have the following values of the subsidiary quantum number assigned to them:

$$\begin{array}{cccccc} s & p & d & f & \dots \\ l = 0 & 1 & 2 & 3 & \dots \end{array}$$

If the terms of the hydrogen spectrum were to be cast into a scheme corresponding to that of Fig. 1 we should obtain a figure which differs from Fig. 1 by having the levels belonging to the same values of n at nearly equal depths from the top horizontal

* This means that we should now regard l and not k as the orbital moment of momentum. In the original Bohr theory the orbits belonging to the s -terms and having $l = 0$ raised a difficulty in that they would have a vanishingly small moment of momentum, that is, they would be "pendulum orbits" which would lead to a collision between the nucleus and the electron. This difficulty disappears in the newer theory (cf. § 38).

line; in the case of the alkalis the levels for the same values of n lie at different depths. Hence in the case of hydrogen several spectral series combine to form a single series with fine-structure. For example, the Balmer series would be composed, in the terminology of the alkali and other spectra, of the series: $2p - ms$ (sharp subsidiary series); $2p - md$ (diffuse subsidiary series), and $2s - mp$ (principal series).*

The subsequent development of spectral theory had now to take into account the empirical fact that the individual energy-levels of the alkali-spectra (and also of other optical-spectra) which had already been ordered by means of two quantum numbers, themselves exhibited a fine-structure. In the case of the alkalis all terms except the s -term are double; in the case of the alkaline earths, which stand in the next column of the periodic system, triplets occur; in the next column there are again doublets and, in addition, quartets; and so forth. Long before the quantum theory came into being Rydberg set up a spectroscopic "Alternation Law" which he discovered empirically; this states that in the first, third, fifth, columns of the periodic system the terms separate into an even number of components, whereas in the second, fourth and sixth rows the term-separations are odd.

Superficial reflection may suggest that the fine-structure of alkali-spectra is, like that of hydrogen, due to relativity, but closer inspection soon shows that this view is untenable. For relativistic change of mass effects a separation of the energy-levels having different k -values (or, rather, l -values in our new terminology) but one and the same n -value. In the case of the alkali-spectra this separation has moreover been performed so thoroughly by the disturbing influence of the core-electrons that the resolution into the individual series above described results. So these already separated alkali energy-levels will, on account of the relativistic change of mass, experience a further but quite unimportant displacement; not, however, a further resolution.

In 1920 Sommerfeld, therefore, felt constrained to introduce in addition to the quantum numbers n and k a third quantum number j , which he called the "inner quantum number," to explain the complex structure of spectra. At that time no link presented itself for the interpretation of this inner quantum number in terms of elements of the Bohr theory (which would allow the energy-levels to be calculated from the model). This interpretation was only to be obtained by proceeding inductively

* In "The Origin of Spectra," by Foote and Mohler, the term "*subordinate series*" is used instead of "*subsidiary series*".

step by step. It was first found that the observed number of fine-structure components of the lines could be accounted for on the following assumptions:

(a) To each doublet of terms we must ascribe a pair of values to j that differ by unity; as we pass from one series to the next the values of these pairs increase by unity. The numbering of the j 's for the alkalis that is regarded as correct at the present time is:

$$p: \frac{1}{2} \text{ and } \frac{3}{2}; \quad d: \frac{3}{2} \text{ and } \frac{5}{2}; \quad f: \frac{5}{2} \text{ and } \frac{7}{2}; \quad \cdot$$

and so forth.

(b) For transitions between lines the following selection rule holds:

$$\Delta j = \begin{cases} +1 \\ 0 \\ -1. \end{cases} \quad (125)$$

This selection rule, which has been derived from experimental results, gives the key to the interpretation of j on the Bohr model. We shall only briefly outline the underlying ideas. We compare the selection rule (125) with that which holds for the azimuthal quantum number k or, more conveniently now, the subsidiary quantum number l : this rule is $\Delta l = \pm 1$. On the basis of the Correspondence Principle this is to be interpreted as follows: the quantum numbers multiplied by h are, according to (29), equal to the corresponding action variables of the system. If the motion of the associated canonically conjugate variables takes place "free of overtones," then the spectrum ascribed to it by the quantum theory can also contain only the "fundamental frequencies," that is, only those transitions are connected with radiation for which the quantum number in question (here l) changes by at most 1. From the validity of the selection rule (125) for j we may therefore conversely deduce that hj is an action variable, whose corresponding angle variable executes a uniform rotation, just as is the case with the perihelion distance, whose canonically conjugate action variable, the orbital moment of momentum, is proportional to l .

Besides the orbital moment of momentum the angular momentum of the whole atomic system is an action variable whose associated angle variable performs rotations "free of overtones." Hence by the Correspondence Principle the quantum number belonging to the total moment of momentum must, like that belonging to the orbital moment of momentum, obey a selection rule of the form (125). In this way the conjecture arose that the inner quantum number j represents the total moment of

momentum of the atom. The source of the difference between the orbital moment of momentum and the total moment of momentum will be discussed in the next paragraph. Before entering on this question we shall further consider the question of multiple quantizing by discussing the introduction of the fourth quantum number m .

Even in the case of the relativistic one-electron problem the motion attains the degree of periodicity 3 as soon as a perturbation, such as in the form of an external electric or magnetic field, occurs. The atom is then no longer degenerate (even when regarded as a three-dimensional system). Corresponding to the three independent classical fundamental frequencies, three quantum numbers n , l and m occur. The additional motion that occurs in the magnetic field consists of a precession of the orbital plane about an axis given by the direction of the magnetic field (Larmor precession). The frequency of the Larmor precession is proportional to the magnetic intensity of field and depends further only on the position and mass of the revolving electrons, but not on the other elements of the orbit. It amounts to

$$\omega = \frac{1}{2\pi} \frac{e|\mathbf{H}|}{2\mu c}. \quad (126)$$

The action variable canonically conjugate to the angle of precession, which advances regularly, is the component of the moment of momentum in the direction of the magnetic field. If we denote it by I_3 , then the "magnetic quantum number" m is to be introduced by means of the quantum condition

$$I_3 = \frac{mh}{2\pi}. \quad (127)$$

If we again choose $\frac{h}{2\pi}$ as the unit of angular momentum, then m becomes directly equal to the component of the moment of momentum in the direction of the field. The condition (127) then simply asserts that the component of the moment of momentum in the direction of the field can assume only integral values. This "direction quantizing" has been strikingly confirmed by the well-known experiments of Gerlach and Stern.

Since there is also a magnetic moment (cf. § 36) corresponding to the mechanical moment of momentum of moving charges, and since the energy of a magnet in the field depends on the orientation of the axis of its moment with respect to the direction of the field, the different orientations permitted by the quantum condition (127) have different energies in the magnetic field.

Thus the separation of terms known as the Zeeman effect occurs in the magnetic field.

In the course of time it became clear that even in the case of the one-electron problem the orbital moment of momentum and the total angular momentum were not identical (cf. § 38). We have then to regard m as the component of the total angular momentum in the direction of the field. Since this component can at most attain the absolute value j , which occurs in the parallel or anti-parallel position, the range of the magnetic quantum number is given by

$$m = -j, -j + 1, \dots, j - 1, j. \quad (128)$$

Pauli, by making use of the knowledge of the four-fold quantizing of electronic orbits, gradually acquired in the manner above sketched, was able to follow up Stoner's extension of Bohr's theory of the periodic system by laying down a principle which has shown itself to be very fruitful. It states: no electron in an atom can have exactly the same values for each of the four quantum numbers n, l, j, m as any other electron in the atom. The maximum number of electrons possible in the individual atomic shells (K - L - M -shells, and so forth) may be derived from *Pauli's Principle*, as also the related lengths of the periods in the periodic system.

36. The Hypothesis of the Spinning Electron

Since, according to the arguments of the preceding paragraph, there is a difference between the orbital moment of momentum and the total moment of momentum the question arises as to the interpretation of the additional moment of momentum $j - l$ in the Bohr model. The first idea that suggested itself was to ascribe the additional moment of momentum to the atomic core. Consider the simplest case, that of an alkali-atom. In the case of lithium the radiating electron describes a two-quantum orbit about the atomic core, which consists of the nucleus and the two one-quantum electrons of the K -shell. If both the K -electrons revolve around the nucleus in the same direction they would give rise to a resultant moment of momentum differing from zero. This moment of momentum of the atomic core would in combination with the orbital moment of momentum l produce the resultant moment of momentum j .

But later developments proved that the view that the residual moment of momentum was to be ascribed to the atomic core was untenable. The detailed arguments are given in "Linien-spektren und periodisches System der Elemente," §§ 14, 16, 18,

by F. Hund (Berlin, Springer, 1927). Of the different reasons which speak against this assumption (they include the anomalous Zeeman effect, the diamagnetic behaviour of the atomic core, the dependence of the magnitude of the alkali doublets on the atomic number, the spectra of the alkaline earths), we shall here single out only one for discussion as it involves arguments which have an important bearing on the sequel. If an electrically charged point-mass of mass μ and carrying a charge e describes an orbit so that its moment of momentum is p , then the system at the same time has a magnetic moment of amount $\frac{e}{2\mu c}p$. For example, the electron of the hydrogen atom in its ground state produces a magnetic moment of amount

$$\frac{e}{2\mu c} \frac{h}{2\pi} \quad (\text{Bohr magneton}). \quad (129)$$

In the same way the atomic core, whose revolving charges all have the same sign, would have to have a magnetic moment $\frac{e}{2\mu c}$ times as great as its moment of momentum.

But the empirical system of series spectra and of the anomalous Zeeman-effect leads to the following conclusions: the additional moment of momentum of unknown origin (difference between l and j) must amount to $\pm \frac{1}{2}$, expressed in terms of $\frac{h}{2\pi}$ as the unit; on the other hand, the corresponding magnetic moment, whose magnitude can be deduced from the Zeeman separations, must be a *whole* Bohr magneton. Hence the ratio of the magnetic moment to the moment of momentum in the case of the source of the additional angular momentum must be $\frac{e}{\mu c}$ instead of $\frac{e}{2\mu c}$, as would have to be expected for a system of revolving point-charges. This circumstance together with others that spoke against the assumption of a moment of momentum of the core led Pauli as early as 1924 to the idea of ascribing the difference between the orbital angular momentum and the total angular momentum to the electron itself. Towards the end of 1925 Uhlenbeck and Goudsmit⁸⁷ put forward the radical hypothesis of the spinning electron.

The electron is to be regarded as having a moment of momentum of its own of magnitude $\frac{1}{2} \cdot \frac{h}{2\pi}$ and a magnetic moment $\frac{e}{2\mu c} \frac{h}{2\pi}$. The electron behaves simultaneously as a top and a

magnet, which would indicate that it is of finite extent and rotates about an axis which passes through its centre. The ratio of the magnetic moment to the angular momentum is $\frac{e}{\mu c}$, that is, twice as great as in the case of a system consisting of revolving point-masses with charges of the same sign. Uhlenbeck and Goudsmit have, in fact, shown that the ratio $\frac{e}{\mu c}$ presumed by them

results for a massive rotating sphere of constant density of mass and provided with a surface charge. But it is probably premature to form a detailed picture of the way in which the mechanical and the magnetic moment of the electron comes about. The assumption that the electron has the moments given above is sufficient in itself to give the necessary data in terms of the Bohr model which will furnish the theory of the complex structure of spectra and of the anomalous Zeeman-effect.

Thus the spinning electron, like the atom as a whole, is subject to directional quantizing. In consequence of its revolution around the nucleus it is subjected to a magnetic field that is generated by the nucleus, which may be regarded as moving relatively to it. Its axis of rotation in the absence of a magnetic field or in the presence of only a weak external field can only take up positions parallel or anti-parallel to the axis of the orbital moment of momentum. In strong magnetic fields, on the other hand, the orientation is directed by the external field (Paschen-Back effect, cf. Hund, loc. cit., § 15).

Expressed in units of $h/2\pi$ the angular momentum of the spinning electron, which we shall call the "spin momentum" in the sequel, has the value $\frac{1}{2}$. Thus we ascribe to the electron the "spin quantum number" $s = \pm \frac{1}{2}$. The total moment of momentum for systems with one radiating electron is then given by

$$j = l + s = l \pm \frac{1}{2}, \quad (130)$$

which leads exactly to the values of j for the alkali doublets, given on p. 74. The doublet separation of the alkali levels is of course to be regarded as due to the magnetic energy of the mutual action between the magnetic field produced by the relative motion of the nucleus at the point where the electron happens to be and the electron itself acting as a magnet. If $\mathbf{H}$ is the intensity of the magnetic field mentioned and $\mathbf{M}$ the magnetic moment of the electron, then the additional magnetic energy is given by

$$\Delta E = - (\mathbf{H}\mathbf{M}), \quad (131)$$

where the scalar product (HM) is positive for the parallel position and negative for the anti-parallel position.

From the point of view of electronic orbits it then only remained to explain why the s -levels were not separated into doublets. In the empirical scheme of terms this fact had been taken into account by setting $l = 0$ for the s -levels and by making the quantum number j denote the absolute value of the total angular momentum. The s -terms then only have the value $j = \frac{1}{2}$. In terms of the Bohr model this means that the orbits with $l = 0$ must be assumed to have neither angular momentum nor magnetic moment. On the older view this requirement could have been fulfilled only by the "pendulum paths" and these were regarded as impossible as they would involve a collision between the nucleus and the electron. According to wave-mechanics, however, the angular momentum and the magnetic moment vanish quite naturally for $l = 0$ and the associated energy-levels must also necessarily be single and not double (cf. § 38).

Lastly, we shall briefly indicate in what way Pauli's principle, formulated in the preceding section, allows us to determine in conjunction with the hypothesis of the spinning electron the maximum number of electrons in the K -, L -, M -shells of the atom. We first observe that the quantum numbers n , l and j refer to absolute values and are therefore essentially positive. On the other hand, s denotes the component of the spin momentum in the direction of the orbital moment of momentum and m the component of the total moment of momentum in the direction of a definite axis (given, say, by an external magnetic field). Hence s and m can assume positive and negative values equally well.

Following Pauli we call electrons that belong to the same values of n and l , or that in Bohr's original terminology revolve in given n_k -orbits, *equivalent* electrons. We next enquire what is the maximum number of equivalent electrons that can occur in an atom. For a given value of l we have, by (130), the following two values of j :

$$j = l + \frac{1}{2}, \quad j = l - \frac{1}{2}.$$

The number of possible values of m for these j -values may be seen from the scheme:

$j = l - \frac{1}{2} : m = -l + \frac{1}{2}, -l + \frac{3}{2}, \dots, l - \frac{3}{2}, l - \frac{1}{2}$	Total: $2l$
$j = l + \frac{1}{2} : m = -l - \frac{1}{2}, -l + \frac{1}{2}, \dots, l - \frac{1}{2}, l + \frac{1}{2}$	Total: $2l + 2$
	In all: <u><u>$4l + 2$</u></u>

Thus the total number of equivalent electrons for a given value of the quantum number l is $z = 2(2l + 1)$. Now for a given value of the principal quantum number n there are n possible values for l , namely, $l = 0, 1, \dots, n - 1$. (As we know, the earlier azimuthal quantum number $k = l + 1$ used by Bohr had values running from 1 to n .) This leads to the following scheme for the maximum number of electrons possible in each shell :

<i>K</i> -shell	.	.	$n = 1$	$l = 0$	$z = 2$
<i>L</i> -shell	.	.	$n = 2$	$l = 0$	$z = 2$
				$l = 1$	$z = 6$
Maximum number of <i>L</i> -electrons = 8					
<i>M</i> -shell	.	.	$n = 3$	$l = 0$	$z = 2$
				$l = 1$	$z = 6$
				$l = 2$	$z = 10$
Maximum number of <i>M</i> -electrons = 18					
<i>N</i> -shell	.	.	$n = 4$	$l = 0$	$z = 2$
				$l = 1$	$z = 6$
				$l = 2$	$z = 10$
				$l = 3$	$z = 14$
Maximum number of <i>N</i> electrons = 32					

In this way we arrive exactly at the numbers 2, 8, 18, 32 . . . characteristic of the lengths of the rows of the periodic table.

The distribution of the values of the individual quantum numbers among the eight electrons of the *L*-shells (neon-configuration), for example, is to be pictured in the following way on the basis of Pauli's hypothesis and the spin-hypothesis. For the sake of clearness we have added in the table the values of the spin quantum number s , which are, however, redundant as their values are already given by the other four numbers :

n	l	(s)	j	m
2	0	$(\frac{1}{2})$	$\frac{1}{2}$	$\frac{1}{2}$
2	0	$(-\frac{1}{2})$	$\frac{1}{2}$	$-\frac{1}{2}$
2	1	$(\frac{1}{2})$	$\frac{3}{2}$	$\frac{3}{2}$
2	1	$(\frac{1}{2})$	$\frac{3}{2}$	$\frac{1}{2}$
2	1	$(\frac{1}{2})$	$\frac{3}{2}$	$-\frac{1}{2}$
2	1	$(\frac{1}{2})$	$\frac{3}{2}$	$-\frac{3}{2}$
2	1	$(-\frac{1}{2})$	$\frac{1}{2}$	$\frac{1}{2}$
2	1	$(-\frac{1}{2})$	$\frac{1}{2}$	$-\frac{1}{2}$

It is surprising that the application of Pauli's principle to the inner electronic configurations should give the correct results so readily. For the quantum numbers to which this principle was originally applied referred to one-electron systems, whereas

in the case of the completed shells we are dealing with systems in which an appreciable interaction of the electrons might have been expected.

37. The more recent View of the Fine-Structure of the Hydrogen Spectrum

The discussion in § 35 shows that from the point of view of the older theory there is a fundamental difference between the fine-structure of hydrogen and of that of the alkalis: in the case of hydrogen the energy-levels having the same values for n and different values for l form a fine-structure group, whereas in the case of the alkalis the corresponding levels belong to the different series of terms $s, p, d, f, \dots$. On the other hand, the fine-structure components of the alkali-terms have the same values for l : they differ in their values for the inner quantum number j or, according to the interpretation based on the spin-hypothesis, owing to the magnetic moment of the radiating electron being able to assume two different positions. Apart from the interpretation of the quantum theory there is, in fact, a fundamental empirical difference between the two types of fine-structure. In the series of hydrogen terms the number of the fine-structure components is equal to the principal quantum number n . For the ground-orbit of the Balmer series this number is 2, for the initial orbits of this series, it is, however, 3, 4 . . . *ad infin.* On the other hand, in the case of the $p, d, f, \dots$ series of the alkali-terms the fine-structure consists only of doublets.

In spite of this fundamental difference evidence was forthcoming in the course of time which favoured the view that there was a common origin to the two types of fine-structure. Further, the Röntgen spectra, which, as we have long known, exhibit a hydrogen-like character, also showed a definite similarity to alkali-spectra, according to theoretical investigations by Landé and experimental researches by Millikan and Bowen. This led Sommerfeld to remark incidentally that the equation

$$\text{hydrogen spectra} = \text{Röntgen spectra} = \text{alkali spectra}$$

compels us to assume a relationship between the natures of the spectra and fine-structures of hydrogen and the alkalis. Faith in this conjecture was further strengthened by the spin-hypothesis and its results. So long as the additional moment of momentum responsible for the difference between l and j was ascribed to the atomic core, it had necessarily to be assumed that it vanished in the case of hydrogen, whose atomic core consisted of the proton alone. As soon, however, as a moment of momentum and a magnetic moment are ascribed to the electron itself, two

different orientations of the spin moments must be allowed in the case of hydrogen, which correspond to different amounts of magnetic energy. As a consequence of this we should expect a further doublet resolution of the individual fine-structure components. From the present point of view, then, the theoretical treatment of the problem of the fine-structure of hydrogen would have to be distinguished in two points from the original theory of Sommerfeld: (a) the calculation of the energy-levels for the n_l -orbits must be made not according to particle-dynamics but according to wave-mechanics with the use of the relativistic Schrödinger equation (124), § 34; (b) the doublet separation of the individual n_l -terms caused by the magnetic energy of the spinning electron must be taken into account.

The carrying out of these calculations brought to light the following highly remarkable result, which reminds us of the dramatic climaxes of the stories by Jules Verne. The alterations in the theory due to (a) and (b) compensate each other in such a way that the completed theory (to the degree of approximation applied) leads precisely to the fine-structure formula of Sommerfeld once again! In treating the fine-structure problem on classical lines and in ignorance of electron spin and wave-mechanics Sommerfeld had committed a double error. But these errors compensate themselves so that the fine-structure measurements of Paschen give exactly the results predicted by theory. Thus Sommerfeld's result had strengthened the confidence of theoretical physicists in multiple quantizing and had given added impetus to the development of the theory. If calculations had been made before electron spin had been recognized, but with Schrödinger's equation, which is presumably more correct than the earlier Bohr assumptions, results would have been obtained which would have been immediately contradicted by experiment, so that confidence in the new theory would have been considerably shaken.

The fact that in spite of the doublet separation due to spin the number of energy-levels is not increased as compared with those that arise in the classical treatment, is due to a special feature of degeneracy that presents itself in this problem and that leads to the coincidence of certain energy-levels. The ground-term of hydrogen, like the s -terms of the alkalis, remains single (cf. § 35). For the case of the two next highest terms, whose difference as we know yields the H_α -line the table below gives the term-values and their allocation to the quantum numbers according to Sommerfeld's original theory as well as according to the present view. Here $\alpha = \frac{2\pi e^2}{hc}$ denotes Sommerfeld's

fine-structure constant and $\Delta\nu = \frac{R\alpha^2}{2^4}$ denotes the doublet separation of the ground-term of the Balmer series.

Term Value.	Values Assigned to the Quantum Numbers			
	Old View.		New View.	
	n	k	n	l s f
$R\left(\frac{1}{2^2} + \frac{1}{4} \frac{\alpha^2}{2^4}\right) + \Delta\nu$	2	1	$\left\{ \begin{array}{l} 2 \\ 2 \end{array} \right.$	$\left\{ \begin{array}{l} 0 \\ 1 \end{array} \right.$ $\left\{ \begin{array}{l} \frac{1}{2} \\ -\frac{1}{2} \end{array} \right.$ $\left\{ \begin{array}{l} \frac{1}{2} \\ \frac{1}{2} \end{array} \right.$
$R\left(\frac{1}{2^2} + \frac{1}{4} \frac{\alpha^2}{2^4}\right)$	2	2	2	1 $\frac{1}{2}$ $\frac{3}{2}$
$R\left(\frac{1}{3^2} + \frac{1}{4} \frac{\alpha^2}{3^4}\right) + 2\left(\frac{2}{3}\right)^4 \Delta\nu$	3	1	$\left\{ \begin{array}{l} 3 \\ 3 \end{array} \right.$	$\left\{ \begin{array}{l} 0 \\ 1 \end{array} \right.$ $\left\{ \begin{array}{l} \frac{1}{2} \\ -\frac{1}{2} \end{array} \right.$ $\left\{ \begin{array}{l} \frac{1}{2} \\ \frac{1}{2} \end{array} \right.$
$R\left(\frac{1}{3^2} + \frac{1}{4} \frac{\alpha^2}{3^4}\right) + \frac{1}{2}\left(\frac{2}{3}\right)^4 \Delta\nu$	3	2	$\left\{ \begin{array}{l} 3 \\ 3 \end{array} \right.$	$\left\{ \begin{array}{l} 1 \\ 2 \end{array} \right.$ $\left\{ \begin{array}{l} \frac{1}{2} \\ -\frac{1}{2} \end{array} \right.$ $\left\{ \begin{array}{l} \frac{3}{2} \\ \frac{3}{2} \end{array} \right.$
$R\left(\frac{1}{3^2} + \frac{1}{4} \frac{\alpha^2}{3^4}\right)$	3	3	3	2 $\frac{1}{2}$ $\frac{5}{2}$

The corresponding graphical representation is given in Fig. 2, which is taken from Sommerfeld's "Wave Mechanics," § 9E (p. 117 of the Engl. edition).

However satisfactory it may have been from one point of view that the electron spin corrected the term values that appeared to emerge incorrectly from the application of the relativistic wave-equation, the state of our knowledge in 1926 could not fail to make it appear a weakness in the theory that the spin-hypothesis had to be postulated *ad hoc* and that it constituted a foreign element in the theory. The need was doubtless felt of formulating the fundamental equations of wave-mechanics from the outset in such a way that the electron spin would result directly from them. This object was attained in a very important development of the theory by Dirac, on which we shall report in § 39.

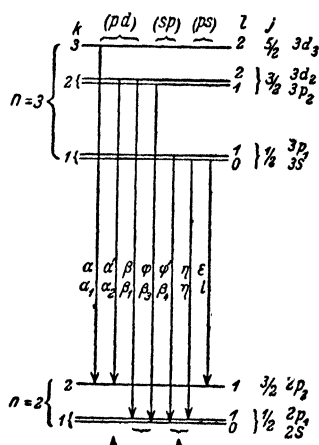


FIG. 2.—Initial and final levels of the H_{α} -line. On the left are the old figures of the quantum numbers, on the right the new. (From Sommerfeld's "Wave Mechanics," Ch. I, p. 117, Engl. edn.)

38. The Wave Mechanical Four-Current

Before entering on Dirac's theory we shall follow on the general relativistic wave-equation (121) § 34 by introducing an idea that will play an important part in the applications of the theory to be discussed in VIII and IX.

The calculation of the intensities, polarizations and the associated derivation of the selection rules from the wave-equation are made possible by means of the relationship, given in equations (94) and (95), between the proper functions of the wave-equation and the Heisenberg matrices (cf. §§ 14, 28 and 29) :

$$q_k(nm) = \int q_k \psi_n \bar{\psi}_m d\tau, \quad p_k(nm) = \frac{h}{2\pi i} \int \psi_n \frac{\partial \bar{\psi}_m}{\partial q_k} d\tau.$$

These relations, which are very important for the sequel, are based on the fact that the proper functions ψ_n of the wave-equation represent an orthogonal system of functions ; that is

$$\int \psi_n \bar{\psi}_m d\tau = 0, \text{ for } n \neq m. \quad (132)$$

The orthogonal relation (132) may be deduced quite easily for the original form of Schrödinger's equation (71) with the help of Green's theorem. This gives us for two functions f and g :

$$\int (f \Delta g - g \Delta f) d\tau = \oint \left(f \frac{dg}{dn} - g \frac{df}{dn} \right) d\sigma, \quad (133)$$

in which the surface integral on the right-hand side is to be taken over the boundary of the integrated space on the left-hand side. If we now set $f = \psi_n$, $g = \psi_m$ and integrate over the whole of space, the right-hand side vanishes, because the integrand of the surface integral being of higher order than the second vanishes.* The left-hand side is transformed by substituting for $\Delta \psi$ its value from Schrödinger's equation (71). We then get

$$(E_n - E_m) \int \psi_n \bar{\psi}_m d\tau = 0,$$

which gives for $E_n \neq E_m$ the relation of orthogonality (132). Thus the proper functions form an orthogonal system of functions only when there is no degeneracy. There is also an analogous orthogonal relation for the product of a proper function ψ_n with the conjugate complex value $\bar{\psi}_m$ of another proper function

$$\int \psi_n \bar{\psi}_m d\tau = 0, \text{ for } n \neq m. \quad (132a)$$

This is arrived at in the same way as above by setting $f = \psi_n$, $g = \bar{\psi}_m$ in (133).

* The ψ 's vanish at least like $\frac{1}{r}$ and therefore $\frac{d\psi}{dn}$ at least like $\frac{1}{r^2}$.

The derivation of (132) and (132a) from Green's theorem (133), which is here given, relates to the proper solutions of wave-equations of the form $\Delta\psi \sim \psi$. By means of a generalization of Green's theorem it may further be shown that the orthogonality of the proper functions also holds for the solutions of those differential equations which may be written in the form

$$L(v) = 0, \quad (134)$$

where $L(v)$ is a so-called self-adjoint differential expression.* Now the left-hand side of the most general (relativistic-magnetic) form of the wave-equation is a self-adjoint differential expression of this kind, hence the orthogonal relation (132) also holds for the solutions of the general wave-equation.

On account of the homogeneity of the wave-equation its functional solutions are given except for an arbitrary constant factor. Therefore in order that the matrices above given, which were formed from the proper functions, may have a definite numerical value, it is necessary to normalize the proper functions. This is done in the usual way by means of the rule

$$\int |\psi_n|^2 d\tau = \int \psi_n \bar{\psi}_m d\tau = 1. \quad (135)$$

This last equation may be combined with the orthogonal relation (132a) in the one expression

$$\int \psi_n \bar{\psi}_m d\tau = \delta_{nm}. \quad (132b)$$

Now the solutions v_n of the time-dependent wave-equations (111) and (121) respectively are themselves functions of t . Thus if the normalizing condition (135) is fulfilled for a definite moment of time, it will be fulfilled at any moment only if

$$\frac{d}{dt} \int \psi_n \bar{\psi}_m d\tau = 0, \text{ for } n \neq m \text{ and } n = m, \quad (136)$$

which, on account of the dependence of the proper functions v on the time, is by no means obvious *from the outset*. But further investigation shows that the relation (136) is actually fulfilled also in the case of the proper functions v that vary with the time; this likewise is connected with the circumstance that the left-hand side of the wave-equation is a self-adjoint differential expression. For it is always possible to form out of the solutions of such equations a definite vectorial differential expression whose generalized divergence (sum of the partial

* See Courant and Hilbert, "Methoden der mathematischen Physik," Vol. I, p. 200, Springer, Berlin. Sommerfeld, "Wave Mechanics," pp. 105 and 127.

derivatives of the individual components with respect to corresponding co-ordinates) vanishes. For simplicity we trace the line of reasoning used in the derivation for the case of the non-relativistic wave-equation (III), restricting ourselves to the case with no magnetic field, that is $\mathbf{A} = 0$. From the above-mentioned generalization of Green's theorem it may be deduced that the solutions of equation (III) fulfil the following condition :

$$\frac{\partial}{\partial t}(\psi\bar{\psi}) + \frac{\hbar}{4\pi i\mu} \operatorname{div}(\psi \operatorname{grad} \bar{\psi} - \bar{\psi} \operatorname{grad} \psi) = 0. \quad (137)$$

The equation (137) also holds if we insert any arbitrary proper function for ψ and $\bar{\psi}$ in it ; thus generally for $n = m$ or $n \neq m$:

$$\frac{\partial}{\partial t}(\psi_n \bar{\psi}_m) + \frac{\hbar}{4\pi i\mu} \operatorname{div}(\psi_n \operatorname{grad} \bar{\psi}_m - \bar{\psi}_m \operatorname{grad} \psi_n) = 0.$$

If we integrate (137) over the whole of space the second term containing the space divergence may be transformed by means of Gauss's theorem into a surface-integral which vanishes for the same reason as the surface-integral in (133). Hence there remains

$$\frac{d}{dt} \int (\psi_n \bar{\psi}_m) d\tau = 0$$

in which m and n have arbitrary values. This ensures the " persistence of orthogonality " for $n \neq m$ and the " persistence of normalization "

$$\int \psi_n \bar{\psi}_m d\tau = \delta_{nm}$$

for $n = m$.

In using the proper functions " normalized to unity " according to (132b) the equation (83) of § 28 for the density of Schrödinger's " charge-cloud " is to be modified by explicitly adding the factor whose value was left open in the un-normalized ψ -function that was there used. For the charge-cloud * corresponding to the individual electron we must have

$$\rho = e\psi\bar{\psi} \quad (83a)$$

since if we use this expression in integrating over the whole of space and also the relation (132b) the value e actually results for the total charge. Thus the equation (83a) allocates to the wave-function a fictitious density of distribution of the electron regarded as diffused or " smudged " (*verschmiert*) over the whole of space (cf. § 68 on this point).

* Concerning the physical interpretation of the charge-cloud cf. §§ 68 and 69.

The relativistic generalization of this allocation is obtained by using equation (121). The generalization of the relation (137) corresponding to the transition from the non-relativistic wave-equation (III) to the relativistic wave-equation (137) is *

$$\sum_{k=1}^4 \frac{\partial S}{\partial x_k} = 0 \quad (138)$$

with

$$S_k = \frac{e}{\mu_0} \left\{ \frac{\hbar}{4\pi i} \left(\psi \frac{\partial \bar{\psi}}{\partial x_k} - \bar{\psi} \frac{\partial \psi}{\partial x_k} \right) - \frac{e}{c} \Phi_k \psi \bar{\psi} \right\}. \quad (139)$$

In this the dependence of the wave-function ψ on the time is expressed by

$$\psi = v(x, y, z) e^{\frac{2\pi i}{\hbar} \mu_0 c^2 t}. \quad (140)$$

Equation (138) is to be regarded as the wave-mechanical counterpart to the laws of conservation of matter and electricity. We know that in the mechanics of continua the equation of continuity holds; this states that the increase of density per unit of time at any point of the body is equal to the excess of the amount of matter flowing into the element of volume over the amount flowing out. Thus if we denote the density of a material fluid by ρ_m , the density of electric charge by ρ and the flux-vector in both cases by $\mathbf{v}$, then the following two equations hold: the hydrodynamic equation of continuity

$$\frac{\partial \rho_m}{\partial t} + \text{div}(\rho_m \mathbf{v}) = 0 \quad (141)$$

and the similarly formed law of conservation of electricity

$$\frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{v}) = 0. \quad (141a)$$

Now we know that in relativistic electrodynamics the electric density of charge is combined with three components of current density to form a tensor of the first degree, the so-called four-current:

$$s_1 = \frac{\rho \mathbf{v}_x}{c}, \quad s_2 = \frac{\rho \mathbf{v}_y}{c}, \quad s_3 = \frac{\rho \mathbf{v}_z}{c}, \quad s_4 = i\rho.$$

In the language of the theory of relativity the law of the conservation of electricity then assumes the form

$$\frac{\partial s_k}{\partial x_k} = 0,$$

* Cf. Sommerfeld, "Wave Mechanics," p. 107, equation (25), which also explains the reason for the value of the normalizing factor used in (139).

which is in perfect formal agreement with (138). To this circumstance there is to be added the fact that the time component s_4 of the expression (139) becomes, if we use (140), the expression (83a) for the charge density, but multiplied by i (see Sommerfeld, "Wave Mechanics," equation 24, p. 107). In this way we are led to regard the expression (139) in dynamic problems as the four-current of the "fictitious fluctuating cloud-charge" in just the same way as we earlier regarded the expression (83a) as the charge-density of this cloud.

Corresponding to the "smudging" of the charge of the electron over the whole of space there is a similar smudging of its mass. We may then speak, in the same sense and with the same reservation (§ 68) as was done in the case of a charge-cloud, of a corresponding "mass-cloud" of the electron, whose material density is given by

$$\frac{\mu_0}{e} S_k = \frac{h}{4\pi i} \left(\psi \frac{\partial \bar{\psi}}{\partial x_k} - \bar{\psi} \frac{\partial \psi}{\partial x_k} \right) - \frac{e}{c} \Phi_k \psi \bar{\psi}. \quad (139a)$$

With the help of these conceptions of a fictitious density of electricity and matter or of the associated four-current, certain difficulties may now be cleared away which presented themselves in the older theory of electronic orbits. Before dealing with them we must interpolate a few explanatory remarks. Hitherto in our wave-mechanical discussion we have treated the proper functions and the proper values for the sake of simplicity as a one-dimensional manifold and hence have provided them with only one current number n or m , respectively. Now in the case of multi-dimensional problems multiple quantizing presents itself in the present case just as much as in the older quantum theory as set out in § 35; the proper values and proper functions form a manifold that is to be ordered according to various current-numbers (*Laufzahlen*). We designate the quantum numbers of the one-electron problem (disregarding the electron spin for the moment) in the usual manner by n, l, m again (the letter m in the expressions used above, for example, $\psi_n \bar{\psi}_m$ signified only a different value of the one quantum number n , whereas in the sequel it is again to be regarded as the "magnetic quantum number" in the sense of § 35). We should therefore have to write in a more detailed manner

$$\begin{array}{l} \psi_{n, l, m} \quad \text{instead of} \quad \psi_n \\ \text{and} \quad \psi_{n', l', m'} \quad \text{instead of} \quad \psi_m \end{array}$$

where n', l', m' denote any other numerical values of the three quantum numbers n, l, m . If no external fields are present and if the relativity corrections are neglected it is known that

double degeneracy occurs, which consists in the proper values being dependent on n but not on l and m , so that the one-dimensional scheme of the proper values $E_1, E_2 \dots E_n$ is justified when the approximation mentioned is applied.

We next follow along the general line of argument of § 35. We stated there that for certain reasons it had been found expedient to introduce the quantity $l = k - 1$ as the azimuthal quantum number and that accordingly l was also to be regarded, instead of k , as the moment of momentum of the orbit (in terms of the unit $\hbar/2\pi$). The s -orbits therefore have zero moment of momentum which is also in agreement with the fact which emerges from the magnetic properties, since these orbits have no magnetic moment. On the basis of the idea of electronic orbits one could only imagine an orbit of vanishingly small moment of momentum and magnetic moment to be realized by having the electron swinging to and fro on a straight line through the nucleus. If the electrons and nuclei are of finite dimensions these fictitious pendulum orbits would lead to collisions between these corpuscles; for this reason pendulum orbits were simply excluded without further reason in the beginnings of the quantum theory. The above-mentioned normalization of the s -terms, however, compelled the re-introduction of these pendulum orbits into the theory by a back door, as it were, without the difficulties of collision remarked on above being produced.

In the light of wave-mechanics the position now presents itself as follows. In place of the individual electronic orbits we have the wave-field of the proper functions $\psi_{n, l, m}$. Just as in classical field-physics the electromagnetic field transports energy the ψ -waves carry electric charge or matter with them in accordance with (139) and (139a) respectively. We shall designate the space-part of the four-current given by (139), that is, the ordinary three-dimensional electric current-density, by s , and the corresponding material current given by the space-part of (139a) by s_m . If, further, we designate the radius vector from the origin of co-ordinates (atomic nucleus) to the volume-element $d\tau$ by r , we get for the moment of momentum $\mathbf{J}$ and the magnetic moment $\mathbf{M}$ of the system

$$\mathbf{J} = \int [\mathbf{r} s_m] d\tau \quad (142)$$

$$\mathbf{M} = \frac{1}{2c} \int [\mathbf{r} s] d\tau. \quad (142a)$$

The quantities s and s_m depend on the proper functions $\psi_{n, l, m}$ in accordance with (139) and (139a). If we now form the

momenta (142) and (142a) for the proper functions belonging to the s -terms with $l = 0$ we get $\mathbf{J} = \mathbf{M} = 0$.

The form itself of wave-mechanics thus leads to the vanishing of the mechanical moment of momentum and of the magnetic moment of the s -terms without any mention being made of electron orbits at all, still less of pendulum orbits through the nucleus.

Moreover, it may be shown that in the case of many-electron systems the momenta of the completed electron shells (inert gas configuration) calculated according to (142a) likewise vanish (Sommerfeld, "Wave Mechanics," Ch. I, § 8c). These results are in agreement with the empirical fact that hydrogen and the inert gases are diamagnetic.

For $l \neq 0$ the result obtained differs somewhat from that of the older quantum mechanics, since the momenta are not simply proportional to l , but to $\sqrt{l(l+1)}$. The orbital moment of momentum that is associated with a state having the azimuthal quantum number l is thus given by

$$s = \frac{h}{2\pi} \sqrt{l(l+1)}. \quad (143)$$

From equation (139) for the four-current it is further possible to derive the boundary conditions for the ψ -function at any surfaces of discontinuity. Since the four-current of matter and electricity must everywhere be finite and continuous, it follows that both the function ψ and the derivative $\frac{d\psi}{dn}$ (where n is not a quantum number, but denotes the direction of the normal) must have the same values on both sides of a surface of discontinuity. If, for example, the electric intensity of field at the bounding surface of a body suddenly decreases from a high value to zero then the boundary conditions for ψ and $\frac{d\psi}{dn}$ just mentioned hold at this "discontinuous surface of the field." We shall make use of this fact in treating the problem of the escape of electrons from metallic surface (§ 46).

39. Dirac's Theory of the Spinning Electron in the Absence of External Fields *

The formal development of the quantum mechanics so far developed is crowned by Dirac's theory of the spinning electron.

* Dirac. See reference (72).

In expounding it we must report in even broader outline than in many other parts of this report, and we cannot give the details of the calculations. Indeed, we cannot even undertake to set forth all the arguments that led to its enunciation since this would require a knowledge of certain branches of quantum-mechanics (involving the theory of transformations), into which we can later enter only cursorily owing to the difficulty of the mathematics.

We shall begin by enumerating the achievements that are to be credited to Dirac's theory of the spinning electron. It represents a relativistically invariant theory which at the same time yields the effects of spin without the addition of further assumptions. In its application to atomic problems its successes include the correct fine-structure formula as well as the Zeeman effect of the doublet atoms. Apart from its having certain weaknesses of principle, which are to be discussed below, it solves the problems associated with electron spin in their entirety.

The fundamental idea underlying the enunciation of the new equation of Dirac is essentially the following. The relativistic wave-equation (I2I) had two weaknesses. It did not embrace the spin phenomena and it was of the second order in time, whereas the classical wave-equation (III) (or rather, the equation of heat conduction) was of the first order in time. The fact that this intrusion of the higher order is a weakness can be exactly realized only in connexion with reasoning involving the theory of transformations (cf. § 68). Here we shall let the simple assertion suffice that it is found to be mathematically impossible to apply the general theory of transformations to differential equations that are of higher order than the first in time.*

Perhaps the following remark will help to make the position clearer. The relativistic Hamiltonian function is, as we know,

$$H = \mu c^2 \sqrt{1 + \frac{p_x^2 + p_y^2 + p_z^2}{\mu^2 c^2}} + V$$

* We shall here anticipate the account of the method of variation of constants given in § 47. If we apply the process of calculation there given to a differential equation which is of the second instead of the first order in time as in (226), we see at once that the problem is sometimes indeterminate. Besides knowing the a_k 's at the time $t = 0$ we should then also have to know the $\dot{a}_k$'s at the time $t = 0$, which, for example, for the application of the theory of perturbations lies outside the range of our physical knowledge.

The assumption

$$a_m = 1, \quad a_n = 0 \quad (n \neq m)$$

would then by no means suffice, rather we should require to know all the $\dot{a}_n$'s at the time $t = 0$. It is difficulties similar to these that occur in applying the theory of transformations.

and the corresponding equation of Hamilton and Jacobi

$$\mu c^2 \sqrt{1 + \frac{p_x^2 + p_y^2 + p_z^2}{\mu^2 c^2}} + V = \frac{\partial W}{\partial t}. \quad (144)$$

It is of the *first order* in the momenta, but *not linear*. Now, in § 32 (cf. 90a) we emphasized the necessity of taking up only whole functions of p into Hamilton's function. We avoided trouble by forming the following equation which is certainly correct in classical physics :

$$\left(\frac{\partial W}{\partial t} - V\right)^2 = \mu^2 c^4 \left(1 + \frac{p_x^2 + p_y^2 + p_z^2}{\mu^2 c^2}\right). \quad (145)$$

The quantum-mechanical substitution (116) was then applied to this equation, which is now quadratic in the momenta. Actually this detour, undertaken on formal grounds, raised the order of the equation unjustifiably, and this is the very step that is retracted by Dirac in his new formulation. Of course, other means must then be adduced to overcome the difficulties of the function H which is irrational in p . We shall now turn our attention to these means.

If we write down the equation of the first order in time it must on account of the symmetry of space and time demanded by the theory of relativity also be of the first order in the space co-ordinates. Now Schrödinger's original equation, which is of the second order in the space co-ordinates, has been excellently confirmed by experiment, if we except spin effects. It is not possible for a differential equation of the first order with respect to a function ψ to be fully equivalent to a differential equation of the second order in ψ . But we may certainly imagine a system for two or more functions ψ_i from which, by elimination, differential equations of the second order f in the ψ_i 's arise. These equations of elimination must then closely approximate to Schrödinger's original equations.*

After these introductory remarks we shall write down Dirac's assumption without giving further reasons. For the case of the force-free electron it runs :

$$(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 \mu c) \psi = 0 \quad (146)$$

* We may illustrate what has been said in the text by the relation of Newton's equations of motion of classical mechanics for a point-mass with one degree of freedom to Hamilton's equation of motion. Newton's equation of motion $\mu \ddot{x} = -\frac{\partial V}{\partial x}$ is of the second order in the time t . We have equivalent to it the system of two differential equations of the first order (1) $\mu \dot{x} = p$; (2) $\frac{dp}{dt} = -\frac{\partial V}{\partial x}$, in which p denotes the momentum of the point-mass. Substituting (2) in (1) at once leads to Newton's equation of motion.

where $p_0 = -\frac{\hbar}{2\pi i} \frac{1}{c} \frac{\partial}{\partial t}$, $p_1 = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x}$ and so forth. The quantities α_k are coefficients whose laws of calculation we leave open for the present. We can only say that they are certainly not the laws used in calculating with ordinary numbers. For in that case only the one function ψ would occur in the equation (146), whereas we had already asserted earlier that we required to set up a system of equations for *several* unknown functions. Thus the quantities α_k would represent something like a scheme of coefficients in two rows in a system of linear equations; the scheme has to be fixed by means of additional postulates.

For this purpose we allow the operator

$$(-p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 \mu c)$$

to act once more on the expression (146),

$$\left\{ \begin{aligned} &(-p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 \mu c) \\ &(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 \mu c) \psi = 0 \end{aligned} \right\} \quad (147)$$

We must next assume that the unknown quantities α_i do not depend on the co-ordinates and may therefore be commuted with the momenta. For in the case of a force-free electron the points of space are equivalent, so that we may set

$$\alpha_k p_i \psi = p_i \alpha_k \psi \dots \quad (148)$$

Equation (147), written out completely, will then appear as follows:

$$I + II + III = 0 \dots \quad (147a)$$

where the symbols I , II , III are

$$\begin{aligned} I &\equiv \left[\frac{\hbar^2}{4\pi^2} \left\{ \frac{\partial^2}{c^2 \partial t^2} - \left(\alpha_1^2 \frac{\partial^2}{\partial x_1^2} + \alpha_2^2 \frac{\partial^2}{\partial x_2^2} + \alpha_3^2 \frac{\partial^2}{\partial x_3^2} \right) \right\} + \alpha_4^2 \mu^2 c^2 \right] \psi. \\ II &\equiv -\frac{\hbar^2}{4\pi^2} \left\{ (\alpha_1 \alpha_2 + \alpha_2 \alpha_1) \frac{\partial^2}{\partial x_1 \partial x_2} + (\alpha_1 \alpha_3 + \alpha_3 \alpha_1) \frac{\partial^2}{\partial x_1 \partial x_3} \right. \\ &\quad \left. + (\alpha_2 \alpha_3 + \alpha_3 \alpha_2) \frac{\partial^2}{\partial x_2 \partial x_3} \right\} \psi. \\ III &\equiv -\frac{\hbar \mu c}{2\pi i} \left\{ (\alpha_1 \alpha_4 + \alpha_4 \alpha_1) \frac{\partial}{\partial x_1} + (\alpha_2 \alpha_4 + \alpha_4 \alpha_2) \frac{\partial}{\partial x_2} \right. \\ &\quad \left. + (\alpha_3 \alpha_4 + \alpha_4 \alpha_3) \frac{\partial}{\partial x_3} \right\} \psi. \end{aligned}$$

It is not permissible, of course, to write $\alpha_i \alpha_k$ in place of $\alpha_k \alpha_i$, since after what has been said the quantities α will certainly not be commutative with respect to multiplication.

We now *postulate* that in the *force-free case* the *differential equation of the second order with respect to the quantities ψ_i* obtained by a process of elimination agrees with the relativistic equation (121). This postulate is actually easily fulfilled by suitably splitting up (147a). We satisfy it by assuming

$$\left. \begin{aligned} I = II = III = 0 \\ \alpha_i^2 = 1 \dots \\ \alpha_i \alpha_k + \alpha_k \alpha_i = 0 \quad (i \neq k) \end{aligned} \right\} \quad (149)$$

It is the relations (149) that serve to determine the quantities α_k . If we know the quantities α_i we shall also know how many functions ψ_i we are compelled to introduce. We shall not enter here into a discussion of the general and systematic method of solving (149) but shall simply give the result. For this purpose we shall follow Dirac and consider six schemes of coefficients of four rows and four columns, each of which we shall denote in turn by the letters $\sigma_1, \sigma_2, \sigma_3, \rho_1, \rho_2, \rho_3$. These schemes of coefficients are :

$$\begin{aligned} \sigma_1 &= \begin{Bmatrix} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ 0 & 0 & I & 0 \end{Bmatrix} & \sigma_2 &= \begin{Bmatrix} 0 & -I & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{Bmatrix} \\ \sigma_3 &= \begin{Bmatrix} I & 0 & 0 & 0 \\ 0 & -I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & -I \end{Bmatrix} & \rho_1 &= \begin{Bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \end{Bmatrix} \\ \rho_2 &= \begin{Bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{Bmatrix} & \rho_3 &= \begin{Bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & -I & 0 \\ 0 & 0 & 0 & -I \end{Bmatrix} \end{aligned}$$

The σ 's and ρ 's are thus matrices of four rows, which are to obey the same laws of calculation as those in § 13 (36, 37) to which the Heisenberg matrices are subject. For example, for ρ_1 we have that all its elements ρ_1^{ik} equal zero except $\rho_1^{13}, \rho_1^{24}, \rho_1^{31}, \rho_1^{42}$, which are all equal to one. The lower index of the matrices ρ_i, σ_i always serve to characterize the matrix, the upper indices refer to the corresponding element of the matrix denoted by the lower index. In the same way we have for σ_2 : $\sigma_2^{ik} = 0$ except $\sigma_2^{12} = \sigma_2^{34} = -i, \sigma_2^{21} = \sigma_2^{43} = i$. For the multiplication of two matrices we then have the law $(\sigma_i \rho_k)^{ln} = \sum_m \sigma_i^{lm} \rho_k^{mn}$ fully analogous to (37).

It may now be shown that these matrices σ_i and ρ_k satisfy the relations

$$\sigma_i^2 = \rho_i^2 = I, \quad (149a)$$

$$\sigma_i \sigma_k + \sigma_k \sigma_i = 0, \quad (149b)$$

$$\rho_i \rho_k + \rho_k \rho_i = 0. \quad (149c)$$

$\sigma_i^2 = I$ of course denotes that the matrix $\sigma_i \sigma_i$ obtained by multiplication is a unit matrix [cf. § 12, (41)]:

$$(\sigma_i \sigma_i)^{lm} = \delta_{lm} \begin{cases} \delta_{lm} = 1, & l = m \\ \delta_{lm} = 0, & l \neq m \end{cases}.$$

As an illustration let us calculate σ_2^2 :

$$(\sigma_2^2)^{lm} = \sum_k \sigma_2^{lm} \sigma_2^{km}.$$

This sum over k contains at the most one term, since σ_2 has only one term in each row and in each column. For $l = m$ this term is actually present, for example,

$$(\sigma_2^2)^{11} = \sum_k \sigma_2^{1k} \sigma_2^{k1} = \sigma_2^{12} \cdot \sigma_2^{21} = (-i)i = I.$$

But for $l \neq n$ this term does not exist; for example,

$$(\sigma_2^2)^{12} = \sum_k \sigma_2^{1k} \sigma_2^{k2} = \sigma_2^{12} \sigma_2^{22} = 0, \text{ since } \sigma_2^{22} = 0.$$

Thus σ_2^2 is actually a diagonal matrix.

In the same way we may confirm (149b). For we have

$$(\sigma_1 \sigma_2)^{lm} = \sum_k \sigma_1^{lk} \sigma_2^{km},$$

$$(\sigma_2 \sigma_1)^{lm} = \sum_k \sigma_2^{lk} \sigma_1^{km}.$$

Thus, say for $l = 3$,

$$(\sigma_1 \sigma_2)^{3m} = \sigma_1^{34} \sigma_2^{4m} = i \delta_{3m},$$

$$(\sigma_2 \sigma_1)^{3m} = \sigma_2^{34} \sigma_1^{4m} = -i \delta_{3m}.$$

By addition (149b) immediately follows.

Now we might already take the ρ_i 's and the σ_i 's for the α_i 's, since they satisfy the same relations (149). But on account of (146) we require four and not three quantities. This fourth quantity is obtained by using the relation

$$\sigma_i \rho_k = \rho_k \sigma_i. \quad (149d)$$

(149d) is verified by direct calculation; we shall confirm it for the product of σ_3 by ρ_1 :

$$(\sigma_3 \rho_1)^{lm} = \sum_k \sigma_3^{lk} \rho_1^{km},$$

$$(\rho_1 \rho_3)^{lm} = \sum_k \rho_1^{lk} \sigma_3^{km},$$

say $l = 4$, then

$$\begin{aligned}(\sigma_3 \rho_1)^{4m} &= \sigma_3^{44} \rho_1^{4m} = -\delta_{2m}, \\ (\rho_1 \sigma_3)^{4m} &= \rho_1^{42} \sigma_3^{2m} = -\delta_{2m}.\end{aligned}$$

In the same way it is easy to prove the correctness of the relations

$$\left. \begin{aligned}\rho_3 \rho_1 &= i \rho_2 \\ \rho_2 \rho_3 &= i \rho_1 \\ \rho_1 \rho_2 &= i \rho_3\end{aligned} \right\} \quad (149e)$$

and corresponding relations for σ .

If we now set

$$\alpha_1 = \rho_1 \sigma_1 = \sigma_1 \rho_1, \quad (150a)$$

$$\alpha_2 = \rho_1 \sigma_2 = \sigma_2 \rho_1, \quad (150b)$$

$$\alpha_3 = \rho_1 \sigma_3 = \sigma_3 \rho_1, \quad (150c)$$

$$\alpha_4 = \rho_3, \quad (150d)$$

all the relations (149) are satisfied. For example, we get

$$\alpha_2^2 = \rho_1 \sigma_2 \rho_1 \sigma_2 = \rho_1 \rho_1 \sigma_2 \sigma_2 = I$$

$$\alpha_3 \alpha_2 = \rho_1 \sigma_3 \rho_1 \sigma_2 = \rho_1 \rho_1 \sigma_3 \sigma_2 = -\rho_1 \rho_1 \sigma_2 \sigma_3 = -\rho_1 \sigma_2 \rho_1 \sigma_3 = -\alpha_2 \alpha_3.$$

With the quantities $\alpha_1 \dots \alpha_4$ that we have obtained we shall now write down the equations (146) explicitly:

$$(p_0 + \rho_1 \sigma_1 p_1 + \rho_1 \sigma_2 p_2 + \rho_1 \sigma_3 p_3 + \rho_3 \mu c) \psi = 0 \quad (151)$$

and from them, by using (149e),

$$(\rho_1 p_0 + \sigma_1 p_1 + \sigma_2 p_2 + \sigma_3 p_3 - i \rho_2 \mu c) \psi = 0. \quad (151a)$$

The equation (151) of course stands for a group of four equations. For example, the expression $\rho_3 \psi$ when written out in full is

$$(\rho_3 \psi)^i = \sum_k \rho_3^{ik} \psi^k \quad (i = 1, 2, 3, 4) \quad (151b)$$

and the sum over k here reduces on account of the matrix ρ (or σ , respectively) to one term. Thus the first equation (151a) runs

$$+ \frac{\hbar}{2\pi i} \left(-\rho_1^{13} \frac{1}{c} \frac{\partial \psi_3}{\partial t} + \sigma_1^{12} \frac{\partial \psi_2}{\partial x_1} + \sigma_2^{12} \frac{\partial \psi_2}{\partial x_2} + \sigma_2^{11} \frac{\partial \psi_1}{\partial x_3} - i \rho_2^{13} \mu c \psi_3 \right) = 0 \quad (152)$$

If we carry out this calculation for all four equations according to the scheme given we obtain the following system, which we shall write down explicitly for the sake of clearness and completeness:

$$(p_0 + \mu c) \psi_1 + (p_1 - i p_2) \psi_4 + p_3 \psi_3 = 0, \quad (152a)$$

$$(p_0 + \mu c) \psi_2 + (p_1 + i p_2) \psi_3 - p_3 \psi_4 = 0, \quad (152b)$$

$$(p_0 - \mu c) \psi_3 + (p_1 - i p_2) \psi_2 + p_3 \psi_1 = 0, \quad (152c)$$

$$(p_0 - \mu c) \psi_4 + (p_1 + i p_2) \psi_1 - p_3 \psi_2 = 0, \quad (152d)$$

40. Dirac's Theory of the Spinning Electron in the Presence of External Fields

For the case where an electromagnetic field is present we use the generalization previously given (cf. § 34 (118) to (121)) and write in place of p_0 and p_1

$$p_0 + \frac{e}{c}\phi_0, \quad p_1 + \frac{e}{c}\phi_1$$

and so forth. In place of (151) we then get

$$\left\{ p_0 + \frac{e}{c}\phi_0 + \rho_1 \left[\sigma_1 \left(p_1 + \frac{e}{c}\phi_1 \right) + \sigma_2 \left(p_2 + \frac{e}{c}\phi_2 \right) + \sigma_3 \left(p_3 + \frac{e}{c}\phi_3 \right) \right] + \rho_3 \mu c \right\} \psi = 0. \quad (153)$$

In the case where an electromagnetic field is present we already see an essential difference between Dirac's system of equations and those hitherto used (117) and (121), respectively, whereas the equation of the second order obtained by elimination in the *force-free case* was made to agree exactly with the old relativistic equation (121). If we apply the following operator (cf. (147)) as before to (53),

$$\left[- \left(p_0 + \frac{e}{c}\phi_0 \right) + \rho_1 \left\{ \sigma_1 \left(p_1 + \frac{e}{c}\phi_1 \right) + \sigma_2 \left(p_2 + \frac{e}{c}\phi_2 \right) + \sigma_3 \left(p_3 + \frac{e}{c}\phi_3 \right) \right\} + \rho_3 \mu c \right],$$

we get

$$\begin{aligned} & \left[- \left(p_0 + \frac{e}{c}\phi_0 \right)^2 + \left\{ \sum_k \sigma_k \left(p_k + \frac{e}{c}\phi_k \right)^2 + \mu^2 c^2 \right\} \psi \right. \\ & \quad + \rho \left[\left\{ \sum_k \sigma_k \left(p_k + \frac{e}{c}\phi_k \right) \right\} \left(p_0 + \frac{e}{c}\phi_0 \right) \right. \\ & \quad \left. \left. - \left(p_0 + \frac{e}{c}\phi_0 \right) \left\{ \sum_k \sigma_k \left(p_k + \frac{e}{c}\phi_k \right) \right\} \right] \right] \psi = 0. \end{aligned} \quad (154)$$

We then carry out in part the differentiations

$$\left(p_k = \frac{\hbar}{2\pi i} \frac{\partial}{\partial x_k} \right),$$

indicated in (154) so far as they relate to the potentials ϕ . After a few simple intermediate calculations, in which all we have to

remember is that the p 's are not commutative, we get the following system of equations:

$$\left\{ -\left(p_0 + \frac{e}{c}\phi_0\right)^2 + \sum_k \left(p_k + \frac{e}{c}\phi_k\right)^2 + \mu^2 c^2 + \frac{eh}{2\pi c} \sum_k \sigma_k \mathbf{H}_k + \frac{ieh}{2\pi c} \rho_1 \sum_k \sigma_k \mathbf{E}_k \right\} \psi = 0. \quad (155)$$

Here $\mathbf{E}_k$ and $\mathbf{H}_k$ denote the corresponding component of the electric and magnetic field strength which may be derived by (113) from the ϕ_i 's.

Apart from the terms $\sum_k \sigma_k \mathbf{H}_k$ or $\rho_1 \sum_k \sigma_k \mathbf{E}_k$ respectively, (155) agrees exactly with (121). But the two additional members lead to a two-fold difference with respect to the earlier relativistic equation.

In the first place, (155) again represents a group of four equations, but now the equations for the ψ_i 's are *no longer separated*; in general *each equation will contain several ψ 's*, cf. (151b). (We shall see below that in the special case of a static central field such a separation can be carried out successfully.)

The second essential difference compared with (121) consists in the form of the factors that are associated with the newly added ψ 's in (155).

Let us take as an illustration the special case $\mathbf{H}_2 = \mathbf{H}_3 = \mathbf{E}_1 = \mathbf{E}_2 = \mathbf{E}_3 = 0$. Then the first equation of (155) acquires the additional term

$$\frac{eh\mathbf{H}_1}{2\pi c} \sigma_1^{12} \psi_2.$$

We may characterize this (in the language of the older theory which is, of course, no longer fully adequate) by saying that the electron apparently has a magnetic moment of amount $\frac{eh}{4\pi\mu c}$ or, respectively, an analogous electric moment.* But this tallies exactly with the one part of Uhlenbeck and Goudsmit's hypothesis discussed in § 36; it here follows as a consequence of Dirac's general assumption. Of course a direct general comparison is not possible on account of the fact mentioned, namely, because the ψ_i 's are not separable. We shall get to know the "spin-term" more exactly below when we deal with the case of a central field. In § 43 we shall make clear why an electric moment, which here presents itself quite formally, also occurs.

To prove that Dirac's assumptions are perfectly correct we

* It is necessary to divide by 2μ in order that the formal analogy with (111) may be made perfect.

should also prove that (146) is invariant with respect to the Lorentz-Einstein transformation. This is not immediately obvious since we do not know how the $\alpha_1 \dots \alpha_4$ behave towards Lorentz-Einstein transformations. Dirac has shown that the necessary invariance of the system of equations (146) is actually preserved.

We have yet to enquire into the second part of Uhlenbeck and Goudsmit's hypothesis according to which the electron also has a moment of momentum of magnitude $\frac{1}{2} \frac{h}{2\pi}$. To the degree of approximation that is admissible in a new theory which applies the conceptions of the old theory this likewise follows from Dirac's equations without further additions being necessary. We shall briefly indicate the ideas that lead to this result.

In § 38 *definite* values were given for the orbital moment of momentum, such as $j = 0$ for the s-shell. From this we infer that the orbital moment of momentum represents an integral of the quantum-mechanical equations. This is, however, no longer possible in Dirac's theory. Rather, calculation shows that what is constant is not the orbital moment of momentum alone but that quantity increased by an additional term of form similar to those in (155). The amount of this additional term is exactly $\frac{1}{2} \frac{h}{2\pi}$. This is in agreement with the spectroscopic facts that led to the introduction of the inner quantum number j (cf. §§ 35-37).

For spherically symmetrical fields in which the vector potential vanishes the equations (153) may be considerably simplified and transformed. After rather tedious calculations, which we shall not repeat here, we obtain as the result of elimination an equation of the second order of the form :

$$\frac{\partial^2 \chi}{\partial r^2} + \frac{2}{r} \frac{\partial \chi}{\partial r} + \left[\frac{\left\{ \frac{E}{c} + V(r) \right\}^2 - \mu^2 c^2}{h^2} - \frac{j(j+1)}{4\pi^2 r^2} \right] 4\pi^2 \chi - \frac{1}{\frac{E}{c} + V + \mu c} \frac{\partial V}{\partial r} \left(\frac{\partial \chi}{\partial r} + \frac{j+1}{r} \chi \right) = 0, \quad (156)$$

where

$$\chi = \frac{1}{r} \psi_3.$$

Then ψ_1 is given by the equation

$$\frac{\partial \psi_3}{\partial r} + \frac{j}{r} \psi_3 = \frac{2\pi \left(\frac{E}{c} + V + \mu c \right)}{h} \psi_1, \quad (157)$$

The same equations hold for ψ_2 and ψ_4 so that instead of four different solutions we get only two. This has a general significance which we shall discuss later in greater detail. In (156) and (157) E denotes the desired proper value parameter, V is the potential energy which is assumed to be a function only of r , j corresponds in a certain sense to our earlier l .

We shall postpone the comparison of (156) with the results of other theories and experiment until § 43.

After this account of the achievements of Dirac's theory it seems appropriate to say in a few words why its author still does not consider that it can be regarded as a satisfactory solution of the electron problem. For this purpose we revert to the equation (153) for a particle carrying the charge e :

$$\left[p_0 + \frac{e}{c}\phi_0 + \rho_1 \left\{ \sum_k \sigma_k \left(p_k + \frac{e}{c}\phi_k \right) \right\} + \rho_3 \mu c \right] \psi = 0. \quad (153)$$

If we write down the equation for a particle having a charge $-e$ and the same mass μ moving in the same field, we have

$$\left[p_0 - \frac{e}{c}\phi_0 + \rho_1 \left\{ \sum_k \sigma_k \left(p_k - \frac{e}{c}\phi_k \right) \right\} + \rho_3 \mu c \right] \psi = 0. \quad (153a)$$

Now we are perfectly free to alter our definition of the momenta by writing i in place of $-i$, that is

$$p_0 = + \frac{h}{2\pi i} \frac{\partial}{\partial t}, \quad p_k = - \frac{h}{2\pi i} \frac{\partial}{\partial x_k}.$$

This can alter nothing in the final result but simply causes some of the quantities ψ to occur instead of their complex conjugates. (153a) then runs

$$\left[p_0 + \frac{e}{c}\phi_0 + \rho_1 \left\{ \sum_k \sigma_k \left(p_k + \frac{e}{c}\phi_k \right) \right\} - \rho_3 \mu c \right] \psi = 0. \quad (153b)$$

But we may at any time replace ρ_3 by $-\rho_3$, for this does not conflict * with any of the relations (149a to d), (150a to d).

But with this substitution (153b) becomes

$$\left[p_0 + \frac{e}{c}\phi_0 + \rho_1 \left\{ \sum_k \sigma_k \left(p_k + \frac{e}{c}\phi_k \right) \right\} + \rho_3 \mu c \right] \psi = 0. \quad (153c)$$

The equation (153c) for the particle with the reverse sign $-e$ has thus become fully identical with the original equation (153)

* This example shows us that our solution for a_k is by no means the only one; nor is this required for our purpose.

for the particle with the charge $+e$ and having the same mass. We may describe the position as follows: Dirac's system of equations (153) refers to particles of charge $+e$ as well as to those of charge $-e$; but this means that it has too many solutions for the charge $+e$. As neither $+e$ nor $-e$ has an advantage over the other, exactly half of the solutions refer to a positively charged particle and the other half to a negatively charged particle. Since in general two different solutions may be combined together, that is, a particle can pass from the state characterized by a proper function belonging to $+e$ to the state characterized by a proper function belonging to $-e$, this signifies that according to Dirac's theory the electrons can change their sign. This is an effect for which no experimental evidence has ever been observed. It can also be shown that in the same process, namely, in the transition of the charge $+e$ to $-e$, the sign of the total energy E must change. This results from the same process of substitution that earlier led to the equations (153a to c). Since E here denotes the total energy, thus, for example, in the field-free case

$$E = mc^2 \text{ and kinetic energy,}$$

Dirac's theory leads to the quite unintelligible result that particles with *negative mass* must exist.

We have here given the objections of principle that may be urged against Dirac's theory. It must be emphasized that they do not apply to Dirac's theory alone but just as much, for example, to the earlier relativistic equation (121), as can be shown in exactly the same way. In practice, therefore, until these difficulties are removed we shall merely retain one half of the solutions which correspond to positive mass and subject them to comparison with experiment. That this comparison comes out satisfactorily in every respect was mentioned in the introduction and will be dealt with in somewhat greater detail in §§ 43, 44.

For historical reasons it seems appropriate to us to point out that the idea of describing the behaviour of the spinning electron by means of several proper functions and not by one alone was used before Dirac by Pauli³⁵ and Darwin.⁶⁹ Their theory, which involved only two functions, had the disadvantage, compared with Dirac's, that it was not relativistically invariant.

CHAPTER VIII
ACHIEVEMENTS OF THE THEORY
ONE-BODY PROBLEMS

41. Oscillator and Rotator. Remarks Concerning Method

As a first application of the theoretical points of view developed in Chapters IV, VI and VII we shall deal with the quantum-mechanical properties of the oscillator and the rotator. The treatment of these systems which are mathematically the simplest will at the same time give us an opportunity of illustrating methods that are constantly being applied in quantum mechanics for determining proper values.

Schrödinger's equation for the linear oscillator whose potential energy is proportional to the square of the elongation runs

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2\mu}{h^2}\left(E - \frac{a^2}{2}x^2\right)\psi = 0 \dots \left(2\pi\nu = \frac{a}{\sqrt{\mu}}\right). \quad (158)$$

In looking for those solutions of Schrödinger's equation that satisfy certain conditions of regularity (cf. §§ 25, 26) we must concern ourselves with the singular points of the differential equation. These singular points are those values of the coordinates for which one coefficient of the desired function or its derivative becomes infinite. We also consider the equation divided by the factor multiplying the highest derivative, so that the term $\frac{d^2\psi}{dx^2}$ has the coefficient unity. Besides these singular points the behaviour "at the point infinity" must always be investigated separately. By this we mean the following: we replace the independent variable by means of the substitution

$$\xi = \frac{1}{x},$$

and then enquire whether the differential equation

$$\frac{d^2\psi}{d\xi^2} + f(\xi)\frac{d\psi}{d\xi} + g(\xi)\psi = 0$$

that results from the original equation in this way is regular or singular at the point $\xi = 0$. According to this we speak of the original differential equation as being regular or singular "at the point infinity."

If we assume as a solution of the differential equation a power series

$$\psi = c_0 + c_1(x - x_0) + c_2(x - x_0)^2 + \dots$$

whose coefficients c_i are determined by substitution in it, this solution is convergent, according to general theorems, in a circular region around x_0 as far as the next singular point. The behaviour of the power series in the neighbourhood of x_0 , if x_0 is itself a singular point, must then be specially investigated according to well-established methods of the theory of differential equations. (We have to deal below with a case of this kind when discussing hydrogen, where a power series that begins at the point 0 must be assumed after the behaviour in the neighbourhood of $r = 0$ has been specially investigated.)

The differential equation (158) has no singular point in finite regions. On the other hand, however, the point at infinity is singular, as is seen by making the substitution $x = \frac{1}{\xi}$:

$$\frac{d^2\psi}{d\xi^2} + \frac{2}{\xi} \frac{d\psi}{d\xi} + \frac{8\pi^2 m}{h} \left(\frac{E}{\xi^4} - \frac{a^2}{2\xi^6} \right) \psi = 0. \quad (158a)$$

So a power series in x ,

$$\psi = c_0 + c_1 x + c_2 x^2 + \dots,$$

assumed to start at the origin $x = 0$, will in general not remain limited for great values of x . We next introduce new dependent and independent variables by means of the following substitutions, inserting appropriate abbreviations for the coefficients of (158):

$$\begin{aligned} \frac{8\pi^2 \mu E}{h^2} &= \epsilon, & \frac{2\pi \sqrt{\mu} a}{h} &= b, \\ \psi &= e^{-\frac{2}{b} x^2} v(x), & x &= \frac{y}{\sqrt{b}}. \end{aligned}$$

Then (158), considered as a differential equation in the unknown function v , which is taken as a function of y , assumes the following form:

$$\frac{d^2 v}{dy^2} - 2y \frac{dv}{dy} + \left(\frac{\epsilon}{b} - 1 \right) v = 0. \quad (159)$$

(The substitution $\psi = e^{-\frac{b}{2}x^2}$ is suggested by a consideration of the approximate behaviour of ψ at the point $\xi = 0$.) We now endeavour to integrate (159) by means of a power series in y . A condition that must be postulated is that, for large and real values of y , v approaches infinity less rapidly than $e^{+\frac{y^2}{2}}$. We are then certain that $\psi = e^{+\frac{y^2}{2}} v$ vanishes for large values of y .

Substituting

$$v = c_0 + c_1 y + c_2 y^2 + \dots$$

in (159) gives the following recurrence formula for the c_k 's :

$$c_k \left(-\frac{\epsilon}{b} + 2k + 1 \right) = c_{k+2} (k+1)(k+2). \quad (160)$$

Two cases are now to be distinguished: (a) a coefficient c_{n+2} vanishes; (b) no coefficient c_{n+2} vanishes.

In the first case, on account of (160), all the higher coefficients vanish, the series ceases and becomes a polynomial of the n th degree in y . Our condition of regularity is then certainly fulfilled, for every polynomial approaches infinity less rapidly than the exponential function. But c_{n+2} (when c_n differs from zero) can vanish only if

$$\epsilon = b(2n + 1). \quad (161)$$

(161) determines an *infinite discrete manifold* of values for ϵ , which allow the condition of regularity to be fulfilled. In the old terminology (161) runs

$$E_n = \left(n + \frac{1}{2} \right) h\nu. \quad (162)$$

If (161) is not fulfilled we are no longer dealing with a polynomial but with a whole transcendental function. But in that case our postulate of regularity can no longer be fulfilled: for a simple estimate of the coefficients shows that for great values of y v behaves like e^{y^2} and accordingly goes to infinity like $e^{\frac{y^2}{2}}$. We may therefore conclude by saying that a regular solution of (158) exists only for the values given by (161) or (162) of the proper value parameter ϵ or E respectively.

The polynomials here derived are called Hermitean polynomials* and are usually denoted by the abbreviation H_n . This

* For the account here given cf. Courant and Hilbert, "Methoden der mathematischen Physik," Vol. I, particularly Ch. V.

method of breaking off a series that has been assumed by giving an appropriate value to the proper function that enters into the recurrence formula, is the most important means of arriving at a rigorous solution of such proper value problems. The subsequent calculation is usually carried out by rigorously solving one problem by splitting off a part of the potential energy and calling it a small perturbation and then treating the complete problem according to the method of perturbations described in detail below.

The feature of the result (162), which is remarkable from the physical point of view, is that the oscillator is no longer quantized in integers as in the earlier theory but in half-integers. This does not show itself in the spectrum, since in forming the energy-differences the constant part $\frac{1}{2}h\nu$ cancels out. But this additional term plays a certain part in many considerations affecting zero point energy (scattering of Röntgen rays by atoms of crystals, which execute harmonic vibrations of the type in question), as well as in the case of many phenomena connected with band-spectra, into which we cannot enter more closely here. It appears that the experimental results, although they are not quite unambiguous, do speak in favour of the new quantizing in half-integral numbers.

The selection rule for the ideal oscillator of the old Bohr theory, according to which the quantum number n may change by only ± 1 , can easily be derived from (162). For it can be shown simply that

$$x_{nm} = \int_{-\infty}^{+\infty} x \psi_n \psi_m dx = \int_{-\infty}^{+\infty} x e^{-bx^2} H_n(\sqrt{b}x) H_m(\sqrt{b}x) dx \quad (163)$$

can differ from zero only for $m = n \pm 1$.

We next discuss the rotator. Schrödinger's equation for the rotator that turns about a fixed axis is (since no potential energy is present)

$$\frac{d^2\psi}{d\phi^2} + \frac{8\pi^2\mathcal{I}}{h^2} E \psi = 0, \quad (164)$$

where $\mathcal{I}$ is the moment of inertia.

The proper functions are

$$\psi = \frac{\cos}{\sin} \left(\frac{\sqrt{8\pi^2 E \mathcal{I}}}{h} \phi \right). \quad (165)$$

Now we must demand that

$$\psi(\phi + 2\pi) = \psi(\phi)$$

since the angle ϕ is defined except for multiples of 2π and since ϕ must be a uniform function of ψ . So we obtain

$$\sqrt{\frac{8\pi^2 E \mathcal{F}}{h^2}} = l, \quad l = 0, 1, 2, \dots$$

Hence we get for the energy

$$E_l = \frac{l^2 h^2}{8\pi^2 \mathcal{F}} \quad (166)$$

in complete agreement with the value of the old Bohr theory.

The position is essentially different if we consider a rotator whose axis is not fixed in space. This is the only case that actually occurs in physical nature, being realized, for example, in the rotations of diatomic molecules. For them, on the assumption that the moments of inertia about both axes are equal, Schrödinger's equation is

$$\frac{1}{\sin \theta} \left(\frac{\partial}{\partial \theta} \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} + \frac{8\pi^2 \mathcal{F} E}{h^2} \psi = 0. \quad (167)$$

In (167) θ and ϕ denote the known polar angles that fix the direction. The equation corresponds fully to the motion of a point-mass which is at a constant distance from the origin, that is, to the spherical pendulum. This differential equation may be solved by the method of separation of variables. For this purpose we set

$$\psi = f(\theta)g(\phi).$$

Equation (167) then separates into two independent ordinary differential equations:

$$\frac{d^2 g}{d\phi^2} = -A^2 g(\phi), \quad (168)$$

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{df}{d\theta} \right) + \left(\frac{8\pi^2 \mathcal{F} E}{h^2} - \frac{A^2}{\sin^2 \theta} \right) f(\theta) = 0, \quad (169)$$

of which equation (168) has the solutions

$$\rho = \frac{\cos}{\sin} [A\phi]. \quad (170)$$

To make

$$g(\phi + 2\pi) = g(\phi)$$

A must have the values m , where $m = 0, 1, 2, \dots$. Equation (169) may be successfully solved by the same methods (using a power series) as were applied in the treatment of the oscillator. For this purpose we first set

$$\cos \theta = x$$

and then obtain in place of (169)

$$(1 - x^2) \frac{d^2 f}{dx^2} - 2x \frac{df}{dx} + f(x) \left(\frac{8\pi^2 \mathcal{F} E}{h^2} - \frac{m^2}{1 - x^2} \right) = 0. \quad (171)$$

Equation (171) is the differential equation of the associated spherical harmonics (cf. Courant and Hilbert, *loc. cit.*).

Its singular points are $x = \pm 1$, if we leave out of account the point "infinity" which does not at present interest us. If we again assume a power series of the form

$$f = a_0 + a_1 x + a_2 x^2 + \dots,$$

starting from the zero point, it will be convergent only as far as the points ± 1 , that is, within the circle of radius unity.

If we first set $m = 0$ it is again possible to get a recurrence formula with two terms. We again show, as in the case of the oscillator, that regularity can be obtained at the points $x = \pm 1$ only in the case where the series breaks off. The condition for this breaking off is now given similarly to (162), by fixing the coefficients of f ,

$$\frac{8\pi^2 \mathcal{F} E}{h^2} = l(l + 1); \quad l = 0, 1, 2, \dots \quad (172)$$

From this we obtain for the energy of the rotator with a free axis,

$$E = \frac{(l + \frac{1}{2})^2 h^2}{8\pi^2 \mathcal{F}} - \frac{1}{4} \frac{h^2}{8\pi^2 \mathcal{F}}. \quad (173)$$

The polynomials that result in this way are the well-known Legendre polynomials, usually denoted by $P_l(x)$, being simple spherical harmonics. From them we get the solutions of equation (171) for $m \neq 0$ in the formula

$$f_l^m(x) = (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_l(x). \quad (174)$$

Simple inspection of (171) shows, as earlier in the case of equation (159), that it does not change its form when the sign of the independent variable is reversed. From this it follows that the Hermitean and the Legendre polynomials always contain either only even or only odd powers of the independent variables.

These polynomials, like the Hermitean polynomials, satisfy the general relation of orthogonality (cf. § 38).

Concerning the physical interpretation of (173) we observe that the energy levels of a rotator with a fixed axis are not the same as those of a rotator with a free axis, as is shown by

equations (166) and (173). Apart from a constant term that is of no importance for spectra the rotator with a free axis is quantized in *half-integral* numbers, a fact that was observed in many cases of band-spectra before the new quantum theory had been proposed. For in this case simply by forming

$$h\nu_{nm} = E_n - E_m$$

it is possible to ascertain the difference between half-integral and integral quantizing, which was not possible in the case of the oscillator.

42. Perturbation Theory of Conservative Systems

In this section we are occupied with the achievements of the theory with respect to the *one-body problem*. This is understood to comprise those cases in which forces act on the particle in question, whereas the reaction of the particle need not be considered. The problems in this category are almost all of such a type that the corresponding mechanical equations or, respectively, the Schrödinger wave-equation, cannot be rigorously integrated; nor can the proper values in general be rigorously determined. The simple methods of § 42 generally fail for complicated systems. To overcome these difficulties a number of methods of approximation have been worked out which are summarized under the heading of *the theory of perturbations of quantum mechanics*. In applying these perturbation methods it is found most expedient to link up with Schrödinger's wave-equation, for which methods have been devised by Schrödinger⁴⁰ himself, Born¹⁷ and Dirac.⁷³ We shall now discuss the perturbation theory of closed conservative systems but without applying the method particularly rigorously or generally.

We first illustrate the process for the simple case involving one degree of freedom. Let the corresponding wave-equation be

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2\mu}{h^2}(E - V)\psi = 0. \quad (71)$$

Here V denotes the total potential energy, which we suppose separable into the two portions,

$$V = V_0 + \lambda V_1.$$

We now assume that we have integrated the equation rigorously for $V = V_0$, that is, that we have determined the proper functions and the corresponding proper values ψ_k^0 , E_k^0 . We next enquire what influence the additional term λV_1 in the wave-equation has

on the definite proper functions and proper values. (The factor λ is used here, as usual in perturbation theory, only to indicate that we are dealing with a term of different order of magnitude, as we shall naturally neglect terms of higher order in a method of approximation, so that in the final result λ must simply be set equal to 1.) The characteristic feature of the calculation is that we enquire into the change caused in a definite proper value E_k and a definite proper value ψ_k by the perturbation. Hence we tentatively assume as the solution of the differential equation containing the perturbation term

$$\begin{aligned}\psi_k &= \psi_k^0 + \lambda\psi_k', \\ E_k &= E_k^0 + \lambda E_k',\end{aligned}$$

and seek to determine the ψ_k' 's and E_k' 's in such a way that (71) is satisfied except for terms of second and higher order in λ .

Equation (71) then separates into the two equations :

$$\begin{aligned}\psi_k^{0''} + \frac{8\pi^2\mu}{h^2}(E_k^0 - V_0)\psi_k^0 &= 0, \\ \psi_k^{1''} + \frac{8\pi^2\mu}{h^2}(E_k^0 - V_0)\psi_k' &= -\frac{8\pi^2\mu}{h^2}(E_k' - V_1)\psi_k^0.\end{aligned}\quad (71b)$$

Here (71b) is the same as the wave-equation for the unperturbed case except that it is *not homogeneous* in ψ_k' .

To perform the integration we make use of the well-known theorem which states that a function satisfying certain conditions of regularity may be expanded in a series of proper functions that are mutually orthogonal :

$$f(x) = \sum_n a_n \psi_n(x). \quad (175)$$

The coefficients a_n are here determined by the relation

$$a_n = \int f(x) \psi_n(x) dx. \quad (176)$$

This is a direct consequence of (132), as may be seen immediately by multiplying the right- and the left-hand side of (175) by ψ_n .

We perform this expansion for the right- and left-hand side of (71b). We set

$$\psi_k' = \sum_i a_i \psi_i^0. \quad (177)$$

The coefficients of the unknown function ψ_k' are determined by substituting from (177) in the differential equation (71b),

$$\psi_i^{0''} - \frac{8\pi^2\mu}{h^2}V\psi_i^0 = -\frac{8\pi^2\mu}{h^2}E_i^0\psi_i^0.$$

By inserting this relation in the left-hand side of (71) we get

$$\sum_i a_i (E_i^0 - E_k^0) \psi_i^0 = (E_k' - V_1) \psi_k^0. \quad (178)$$

The right-hand side may likewise be expanded in terms of proper functions

$$(E_k^0 - V_1) \psi_k^0 = \sum_i b_i \psi_i^0, \quad (179)$$

and it is to be emphasized that all the coefficients b_i are already known on account of the 0 th approximation. Now by well-known theorems, not only must the total expression on the left be equal to that on the right but each individual coefficient of ψ_i^0 on the left must equal the corresponding coefficient on the right, thus

$$a_i (E_i^0 - E_k^0) = b_i. \quad (180)$$

This is seen most readily by multiplying the right- and left-hand sides by ψ_i^0 and integrating over x . All the coefficients as far as the l th then vanish. In this way we get as our new proper function

$$\psi_k^1 = \sum_i \frac{b_i}{E_i^0 - E_k^0} \psi_i^0. \quad (181)$$

This solution is complete except for the k th coefficient, which remains indeterminate on account of the vanishing denominator. We determine it by making the additional condition that

$$b_k = 0. \quad (182)$$

This eliminates all discontinuities from the series (181).

The condition (182) may always be satisfied since we still have at our disposal the proper value perturbation E_k' in order to make the proper solution regular. From (179) and (176) it then results that

$$b_k = \int dx (\psi_k^0)^2 (E_k' - V_1) = 0, \quad (183)$$

and hence, on account of (135), we get for the E_k' 's,

$$E_k' = \int dx V_1 (\psi_k^0)^2. \quad (184)$$

This completes the proof of the proper value problem to a first degree of approximation. We shall write down the final formulæ once again, grouped together. They are

$$\psi_k' = \sum_l \frac{b_l}{E_l^0 - E_k^0} \psi_l^0, \quad (181)$$

$$b_k = 0, \quad (182)$$

$$E_k' = \int V_1(\psi_k^0)^2 dx, \quad (184)$$

$$\psi_k = \psi_k^0 + \psi_k', \quad E_k = E_k^0 + E_k'.$$

In practical cases in which we are often concerned only with the calculation of the proper value perturbation, that is, with the change of energy, and not with a knowledge of the proper function, we may make use of the relation (184) by determining the perturbation proper values from the knowledge of the proper functions of zero order without* having to know the proper functions of the first order.

In applying the process to systems of several degrees of freedom nothing is changed in the case of the so-called non-degenerate systems. We take the terms degenerate and non-degenerate systems as denoting in the usual way such as have several or only one proper function, respectively, belonging to one proper-value (cf. § 35). By general mathematical considerations we have that for one degree of freedom only one proper function belongs to each proper value in the differential equations that come into question. But in general this is no longer the case with differential equations involving several degrees of freedom. If we follow out the application of the scheme so far used to the case of several degrees of freedom with degeneracy we see immediately at what point of the calculation certain inferences are made impossible by the degeneracy. We have again to determine the a_{k_1} 's and a_{k_2} 's for those proper functions which have the same proper value. For example, when two proper functions k_1, k_2 have the same proper value E_k we get the meaningless equations

$$a_{k_1} = \frac{b_{k_1}}{E_k - E_k} \quad (180a)$$

$$a_{k_2} = \frac{b_{k_2}}{E_k - E_k} \quad (180b)$$

* It is to be recalled that there is an exact analogy to this perturbation process in the mechanical perturbation process of the old Bohr theory; in this, too, it was found sufficient for determining the energy to the first order to know the motion to zero order. The addition term to the energy was then simply equal to the mean value of the perturbation potential during the unperturbed motion, quite analogously to (184).

fully analogously to the earlier equation (180). But earlier we could satisfy the single equation (182) with the one proper value parameter E' ; this is, however, no longer possible now, since we have two equations (180a, b) and only one proper value parameter E_k' . This mathematical consideration suggests that the influence of the perturbation consists in splitting up the energy-value E_k in question into two energy-values:

$$E_k^I = E_k^0 + \lambda E_k', \quad E_k^{II} = E_k^0 + \lambda E_k''.$$

We then have just enough coefficients to satisfy both equations (180a, b).

But a slight complication enters into the problem here. For just as in the case of any particular proper value E_k under consideration each of the two proper functions $^*\psi_{k_1}$ and ψ_{k_2} separately satisfy the equation

$$L(\psi_k) + \alpha(E_k - V)\psi_k = 0,$$

so also any arbitrary linear combination satisfies

$$\psi = c_1\psi_{k_1} + c_2\psi_{k_2},$$

where, on account of

$$\int \psi \bar{\psi} dx = 1,$$

we must have

$$c_1\bar{c}_1 + c_2\bar{c}_2 = 1. \quad (185)$$

We cannot straightway say, however, which coefficients are the correct ones in the perturbed system, but must decide this from our perturbation calculus. Recapitulating then, we may say that the difference as compared with the case of non-degenerate systems consists in the following two points. Firstly, a term to which several proper functions belong generally splits up under the influence of the perturbation into several terms, and secondly, the initial system of the ψ_k 's is not yet known to us but must first be fixed if the problem is to be solved to the first approximation.

We shall illustrate the course of the calculation by taking the case of a two-fold degenerate system. But the formulæ may be applied without difficulty to degeneracy of any order except that they become a little more extended. Let, then, E_k be a proper value of the unperturbed differential equation

$$L(\psi) + (E - V_0)\psi = 0,$$

* ψ_{k_1} and ψ_{k_2} may also be assumed here as orthogonal and normalized. If they were not so originally, it is easy to form two linear combinations that satisfy these conditions.

and let the two proper functions ψ_{k_1} and ψ_{k_2} belong to E_k . From these two proper functions we form the linear combination

$$\begin{cases} \psi_y^{I^0} = c_{11}\psi_{k_1} + c_{12}\psi_{k_2} \\ \psi_k^{II^0} = c_{21}\psi_{k_1} + c_{22}\psi_{k_2} \end{cases} \quad (186)$$

and again assume as the solution of the general equation

$$\begin{cases} \psi_k^I = \psi_k^{I^0} + \phi_k^I \\ \psi_k^{II} = \psi_k^{II^0} + \phi_k^{II} \end{cases} \quad (187)$$

$$E_k = E_k^0 + \epsilon_k. \quad (188)$$

Substituting (187) and (188) in the differential equation and arranging terms in powers of λ , we get

$$L(\phi_k^I) + (E_k^0 - V_0)\phi_k^I = -(\epsilon_k - V_1)\phi_k^{I^0}. \quad (189)$$

Equation (189) is fully analogous to (71b). To obtain the two necessary conditions (cf. 180a, b) we now require, analogously to (182), that

$$\int dx (\epsilon - V_1) \psi_k^{I^0} \psi_{k_1} = 0, \quad (190a)$$

$$\int dx (\epsilon - V_1) \psi_k^{I^0} \psi_{k_2} = 0. \quad (190b)$$

Substituting the expressions (186) and taking account of orthogonality, we get the two linear equations

$$c_{11}(\epsilon_{11} - \epsilon) + c_{12}\epsilon_{12} = 0, \quad (191a)$$

$$c_{12}\epsilon_{12} + c_{22}(\epsilon_{22} - \epsilon) = 0, \quad (191b)$$

where ϵ_{lm} , for example, is an abbreviation for

$$\int dx V_1 \psi_{k_1} \psi_{k_2} = \xi_{lm}.$$

Equations (191a, b) form a system of homogeneous linear equations for determining the c_{ik} 's. We know that this system has a solution that differs from zero only if the determinant of its coefficients vanishes:

$$\begin{vmatrix} \epsilon_{11} - \epsilon & \epsilon_{12} \\ \epsilon_{12} & \epsilon_{22} - \epsilon \end{vmatrix} = 0. \quad (192)$$

This determinant is a quadratic function of ϵ . By choosing a suitable value for ϵ we can thus always make it vanish. The different solutions for ϵ then give the *new proper values*. After the ϵ_1 's and ϵ_2 's have been determined we can use (191a, b) to calculate the quantities c_{ik} ; and by substituting these values in (186) and (189), as earlier, we can arrive at the ϕ_k^I 's and ϕ_k^{II} 's.

The same holds here that has already been emphasized in the non-degenerate case, namely, the *proper values can be determined to a first approximation* with the help of (192) *before the proper functions are known to a first approximation*. Indeed, we need not even have calculated the correct scheme of coefficients c_{ik} ; it suffices to know the proper functions ψ_{k1}, ψ_{k2} of zero order with arbitrary coefficients.*

We cannot here discuss the treatment of perturbations of higher orders (that is, of involving terms proportional to λ^2 and so forth), nor questions of convergence. On the other hand, it is important for many practical cases to observe the following points closely. In systems which, like the hydrogen atom, for example, have a discrete as well as a continuous spectrum † it is not sufficient to assume in the expansion the proper functions ψ_k simply as a series of discrete proper functions; rather we must also add the continuous spectrum

$$\psi_k^I = \sum_l a_{lk} \psi_l + \int dE a(E) \psi(E).$$

The effect of taking into account the continuous proper value spectrum is highly important, particularly in non-stationary problems.

As our present account is rather in the nature of a report we should have refrained from giving this perturbation method, which is exceedingly elegant mathematically, were it not that the questions that occur in considering many-body problems had not suggested to us to give a more detailed account of the treatment of degenerate systems. As we shall see in § 52 the determinant (192) is actually the mathematical expression for that great group of "exchange phenomena" which are so characteristic of the new quantum theory and which have been verified in such a surprising way.

In concluding our perturbation calculations we wish to remark that the processes have been described by Schrödinger⁴⁰ himself, who illustrated them in numerous applications.

43. The Unperturbed Hydrogen Atom

The successive achievements of quantum mechanics in dealing with the hydrogen atom and the singly ionized helium atom, which differs from it essentially only in having twice its nuclear

* A perturbation process due to Wentzel and Brillouin, which brings out the relationship with the earlier Sommerfeld-Wilson quantum condition, is dealt with in Sommerfeld's "Wave Mechanics," p. 158.

† In the earlier quantum theory the discrete spectrum corresponds to elliptic orbits, the continuous spectrum to hyperbolic orbits.

charge, have been described in detail in §§ 26, 35-37, so that it only remains to tabulate the separate formulæ and to discuss the influence of external fields. The original wave-equation of Schrödinger (or, respectively, the matrix problems equivalent to it) led in the case of the hydrogen problem to the equation

$$\Delta\psi + \frac{8\pi^2\mu}{h^2}\left(E + \frac{Ze^2}{r}\right)\psi = 0. \quad (193)$$

Here Z denotes the nuclear charge, which for He, for example, is to be set equal to 2. The other symbols have the usual significance.

Equation (193) may be solved expediently in spherical coordinates by means of the method of separation of variables: $\psi = \chi(r)g(\theta)h(\phi)$. We then get for $g(\theta)h(\phi)$ the equation (167) in which the coefficient of $g(\theta)h(\phi)$ again becomes $l(l+1)$. (This is quite comprehensible, for $g(\theta)h(\phi)$ describe the motion for $r = \text{const.}$, that is, the motion of a "spherical pendulum" or a rotator with a free axis.) The solutions are the corresponding associated Legendre functions:

$$g(\theta)h(\phi) = e^{\pm im\phi} P_l^m(\cos \theta), \quad (194a)$$

$$l = 0, 1, 2 \dots; \quad m = 0, 1, 2 \dots; \quad m \leq l.$$

For $\chi(r)$ we get the equation

$$\frac{d^2\chi}{dr^2} + \frac{2}{r} \frac{d\chi}{dr} + \left(\frac{8\pi^2\mu E}{h^2} + \frac{8\pi^2\mu Ze^2}{h^2 r} - \frac{l(l+1)}{r^2} \right) \chi = 0. \quad (194b)$$

This equation (194b) with its singular points $r = 0$, $r = \infty$ may be discussed, for example, along the lines of the method of § 41. We shall here merely quote the results.

1. For $E < 0$ there are no solutions that remain finite in the whole of space except when

$$E_n = - \frac{2\pi^2 e^4 \mu Z^2}{n^2 h^2}, \quad n = 1, 2 \dots \quad (195)$$

In that case

$$\nu_{mn} = \frac{E_m - E_n}{h} = \frac{2\pi^2 e^4 \mu Z^2}{h^3} \left(\frac{1}{n^2} - \frac{1}{m^2} \right). \quad (196)$$

(196) is Balmer's formula without the fine-structure.*

$\chi(r)$ then has the form $L(r)e^{-\frac{r}{an}}$; here $L(r)$ is a polynomial

* It can easily be shown that the influence of the motion of the nucleus comes out exactly as in the earlier theory of Bohr; this is done by writing down the wave-equation for 6 degrees of freedom and by treating the motion of the centre of gravity by means of the separation method.

in r , whose form and degree depend on n and l (Laguerre polynomial). α stands for $\frac{h^2}{4\pi^2 e^2 m Z}$ and is equal to the radius of the ground orbit in the old theory ($\alpha = 0.53 \cdot 10^{-8}$). The figures 3 and 4 give the succession of values of χ and $r^2 \chi^2$ for the lowest quantum numbers.

We see from them that, when ψ has values of the order of magnitude of this radius, a is appreciably different from zero, but that for greater values it descends very rapidly.

2. For $E > 0$ we have for any arbitrary values of E solutions

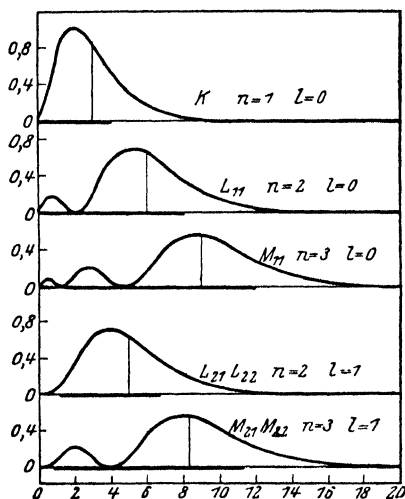


FIG. 3.

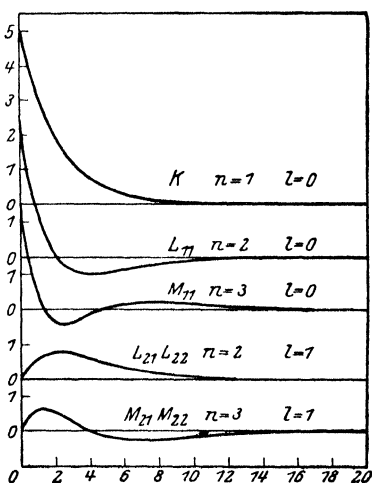


FIG. 4.

Charge density and proper function for the case of the H-atom.

that are regular in the whole of space; accordingly we speak in this case of a continuous proper value spectrum. The solutions cannot be given in a compact form; only their asymptotic expression for great values of r is known, which runs

$$\chi(r) \sim \frac{1}{r} e^{-\frac{2\pi i \sqrt{E_\mu}}{\hbar} r}.$$

To this degree of approximation the hydrogen atom is a multiple degenerate system. To each value of l and m there belongs one and the same value of the energy, provided n is the same. By simply counting up the spherical harmonics we get the degree of degeneracy n^2 .

The relativistic wave-equation (117),

$$\frac{d^2\chi}{dr^2} + \frac{2}{r} \frac{d\chi}{dr} + 4\pi^2\chi \left[\frac{\left(\frac{E}{c} + \frac{Ze^2}{c^2}\right)^2 - \mu^2c^2}{h^2} - \frac{l(l+1)}{4\pi^2r^2} \right] = 0,$$

gives, when $\phi = \frac{Ze}{r}$, a fine-structure, that is, different values for the energy when the l 's are different but n is the same. But the numerical value is not in agreement with the results of experiment.

The relativity correction shows itself in (117) in the same way, except for a constant term, as the effect of an additional force proportional to r^{-3} .

As explained in § 37 the electron spin has also an influence on the calculation of the energy-values. It is obtained as follows: the electron is equivalent to a magnet of magnetic moment $\mathbf{m} = \frac{eh}{4\pi\mu c}$ equal to a Bohr magneton. It moves with the velocity $\mathbf{v}$ relative to the nucleus. We may also, instead, regard the electron as being at rest and the nucleus moving relatively to it with the velocity $-\mathbf{v}$. On account of its motion the charge on the nucleus excites at the point where the electron is situated a magnetic field* of magnitude $\mathbf{H} = \frac{e}{cr^3}[\mathbf{r}\mathbf{v}]$ or, if we express $\mathbf{v}$ in terms of the angular momentum,

$$\mathbf{H} = \frac{e}{\mu cr^3}N \quad (\text{where } N = \text{angular momentum of the orbit}). \quad (197)$$

The potential energy of the electron is then given according to (131) by

$$V_{\text{spin}} = \mathbf{H}\mathbf{m}. \quad (198)$$

A more detailed discussion of the conditions that result in the individual transformations, as worked out by Thomas, show that the expression (198) for calculating the energy must be reduced by the factor $\frac{1}{2}$:

$$V_{\text{spin}} = \frac{e^2}{2\mu^2c^2r^3}N. \quad (199)$$

Thus a moving magnet experiences a force in the electric field. The term-values are now calculated by substituting the two perturbation terms,

$$V_{\text{rel}} \sim \frac{1}{r^2}, \quad V_{\text{spin}} \sim \frac{1}{r^3}$$

* This is evident, since the rotating nucleus is equivalent to a circular current, for example, by Biot and Savart's law.

in the wave-equation and by using the method described for degenerate systems to calculate the energy.

In this way we get as the complete wave-equation

$$\Delta\psi + \frac{8\pi^2m}{h^2} \left(E + \frac{e^2}{r} - V_{\text{rel}} - V_{\text{spin}} \right) \psi = 0.$$

We obtain as our final result, using the nomenclature of §§ 35-37,

$$E_{nlj} = Rh \left[\frac{Z^2}{n^2} + \frac{4\pi^2 e^4 Z^4}{h^2 c^2 n^3} \left\{ \frac{j(j+1) - l(l+1) - s(s+1)}{2l(l+\frac{1}{2})(l+1)} + \frac{3}{4n} - \frac{1}{l+\frac{1}{2}} \right\} \right], \quad (s = \frac{1}{2}). \quad (200)$$

The expression (200) shows us immediately the essential difference in the old and the new view of the fine-structure: whereas earlier, in the absence of the first term which is proportional to Z^4 , a separation occurred only for *different* values of l (orbital moment of momentum); this is now no longer the case, since the values $j = l \pm \frac{1}{2}$ can now also be distinguished from each other. Doublets having the same values for l but different values for j are called spin or, better, relativity doublets, whereas doublets with the same j but having different values of l are called screening doublets.

In the case of hydrogen or He^+ we have

$$E_{nl_1j} - E_{nl_2j} = 0, \quad (201)$$

$$l_1 = j + \frac{1}{2}, \quad l_2 = j - \frac{1}{2}, \quad (202)$$

that is, screening doublets do not exist at all in this case. But they *do* occur in Röntgen spectra in which the screening action of the other electrons has to be taken into consideration besides the nuclear charge.

Dirac's theory of the spinning electron gives the described effects without additional assumptions. This is easily seen if we approximately transform the last term of (156); since this term expresses the difference between Dirac's equation for χ and the usual relativistic equation. It must be observed that on account of the large numerator μc this difference represents only a small correction. We then get exactly the additional term (199). Gordon⁸⁴ and Darwin⁷⁰ have shown that the exact integration of Dirac's equation leads to the following expression for the energy:

$$E = \frac{\mu c^2}{\sqrt{1 + \frac{\alpha^2}{(n + \sqrt{j^2 - \alpha^2})^2}}}, \quad (203)$$

$$\sim - \frac{2\pi e^2 Z}{n}$$

where

which is precisely the expression of Sommerfeld's fine-structure formula as obtained on the older quantum theory of Bohr. The expression (200) represents the first approximation to (203).

We must add a short remark about the quantity j (the total moment of momentum) of Dirac's theory. In (156) j occurs in two forms: as $j(j+1)$ and $j+1$. The first term assumes the same values 0, 2, 6 . . ., no matter whether j has positive values including zero, or positive and negative values excluding zero. But for the values $j = k$, $j = -k - 1$, which lead to the *same value*, $k(k+1)$, for $j(j+1)$, we get for $j+1$ the two different values $k+1$ and $-k$; that is, the two associated energy-values are different. So Dirac numbers j thus:

$$j = -1, +1, -2, +2 \dots \quad (204)$$

and hence obtains the earlier selection rule for l . We cannot, however, pursue these questions further here.

44. Influence of External Fields on the Hydrogen Atom

The influence of an external electrical field on the hydrogen atom is obtained, according to wave-mechanics, only as far as members that already contain the square or higher powers of the field-intensity, just as in the Bohr theory. If we assume a homogeneous electrical field in the z -direction the wave-equation of the system runs

$$\Delta\psi + \frac{8\pi^2\mu}{h^2} \left(E + \frac{e^2Z}{r} - eFz \right) \psi = 0. \quad (205)$$

As in mechanics we may also separate this equation by introducing so-called parabolic co-ordinates, defined by

$$x = \sqrt{\lambda\sigma} \cos \phi, \quad y = \sqrt{\lambda\sigma} \sin \phi, \quad z = \frac{1}{2}(\lambda - \sigma).$$

In spherical co-ordinates λ and σ have the graphical meaning $r \pm z$. Equation (205) then becomes

$$\begin{aligned} \frac{\partial}{\partial \lambda} \left(\lambda \frac{\partial \psi}{\partial \lambda} \right) + \frac{\partial}{\partial \sigma} \left(\sigma \frac{\partial \psi}{\partial \sigma} \right) + \frac{1}{4} \left(\frac{1}{\lambda} + \frac{1}{\sigma} \right) \frac{\partial^2 \psi}{\partial \phi^2} + \frac{2\pi^2\mu}{h^2} \left[E(\lambda + \sigma) \right. \\ \left. + 2Ze^2 - \frac{1}{2}e^2(\lambda^2 - \sigma^2) \right] \psi = 0. \quad (206) \end{aligned}$$

Equation (206) may then be separated into three equations by making the assumption

$$\psi = f(\lambda)g(\sigma)h(\phi).$$

These equations are of the same type as the original equations (194a, b) of Schrödinger and may easily be integrated approximately. We get for the energy in the electrical field

$$E = E_0 + \left. \begin{aligned} & \frac{3h^2}{8\pi^2\mu Ze^2} Fnn' - \frac{h^6 n^4 F^2}{16\mu^3} \\ & [17n^2 - 3n'^2 - 9m^2 + 19] \\ & \pm n', m = 0, 1, 2, 3 \dots \end{aligned} \right\}. \quad (207)$$

In (207) n' and m signify new quantum numbers. We see that if the term $\sim F$ is taken into account the degeneracy involves only one quantum number, whereas according to the complete formula (207) the energy depends on all three quantum numbers.

The term linear in F agrees with the well-known formula of Schwarzschild and Epstein for the resolution of the hydrogen lines in moderately intense electrical fields. For very intense fields the quadratic term must be taken into account; this term is identical with the old expression except for the summand 19 and the different numbering of m . A method of testing the quantum-mechanical formula also consisted in evaluating the quadratic term for the ground-state in which we must set $n = 1$, $n' = m = 0$. We then obtain for the electrical energy of the hydrogen atom in the ground-state, referred to the field-free state $F = 0$,

$$\Delta E = - F^2 \frac{9h^6}{256\pi^6 Z^4 \mu^3 e^6}, \quad (208)$$

whereas the value in the older theory was $\frac{2}{3}$ of this amount. Now on the classical theory we have for the dielectric constant of a gas with N atoms in unit volume

$$\epsilon = 1 - \frac{8\pi \Delta E \cdot N}{F^2}. \quad (209)$$

Substituting (208) in (209) we obtain for the dielectric constant of atomic hydrogen

$$\epsilon = 1 + \frac{9Nh^6}{32\pi^5 \mu^3 e^6}.$$

This value agrees with that obtained by experiment considerably better than that calculated on the basis of the older theory.¹⁰¹

The relative intensities of the individual components, into which each spectral line resolves in the Stark effect, has been calculated by Schrödinger⁴⁰ and found to be in satisfactory agree-

ment with experimental results. As a measure of the intensity of a line he used the corresponding dipole-moment

$$x_{n_1 n_1'; n_2 n_2'} = \int x \psi_{n_1 n_1'} \psi_{n_2 n_2'} d\tau,$$

which corresponds to the old assumption (27a) of Heisenberg.

After what has been said (cf. §§ 35-37) it is evident that the Zeeman effect of the hydrogen atom is of particular interest, since we have to regard the hydrogen atom as resembling an alkali atom. We add the following details about the anomalous Zeeman effect to what has already been said in § 37. The energy of a particle of magnetic moment $\mathbf{m}$ in a field $\mathbf{H}$ is, by (131),

$$\Delta E = - (\mathbf{Hm}). \quad (131)$$

The magnetic moment which, in the language of electron spin, is composed of the orbital moment of momentum of the electron and its own moment of momentum is not directly known in the case of a weak magnetic field, since both the spin and the orbital moment of momentum are not integrals of the equations of motion but only the resultant total moment of momentum. A knowledge of the total moment of momentum alone or, respectively, of the magnetic quantum number m does not immediately allow us to specify the magnetic moment. Before the advent of quantum mechanics Landé had already given a formula, obtained partly empirically and partly with the help of the principle of correspondence, which allowed the magnetic resolution to be calculated from a knowledge of the multiplet structure (doublet, triplet, . . .), of the nature of the term ($s, p, d, \dots$) and of the resulting moment of momentum. According to him the magnetic energy, say, of a doublet atom is

$$\Delta E = \frac{eh\mathbf{H}}{4\pi\mu c} mg, \quad (210)$$

in which g stands for the Landé factor,

$$g = \frac{j + \frac{1}{2}}{l + \frac{1}{2}}. \quad (211)$$

This formula, which has been well confirmed by experiment, holds only for weak fields, that is, for those in which the separation of the individual terms due to intra-atomic forces is great compared with the resolution of the term caused by the external magnetic field. In the case of very intense magnetic fields, that is those in which the relative strengths are reversed, a very general effect occurs which is called the Paschen-Back effect, after its discoverers. This consists in the components of a multiplet

closing together, as it were, into one line, which then itself shows the normal Lorentz resolution.

Hydrogen satisfies the conditions here described perfectly. In the magnetic field we have besides the perturbation terms involving $\frac{1}{r^2}$ and $\frac{1}{r^3}$, the perturbation due to the field itself. If **N** denotes the orbital moment of momentum and **S** the spin of the electron, then on account of (131) we have

$$-\Delta E = \frac{eh}{4\pi\mu c} \{(\mathbf{HN}) + (\mathbf{HS})\}. \quad (212)$$

Here (**HN**) denotes, in a certain sense, the perturbation of the orbit by the magnetic field, whereas (**HS**) gives a measure of the interaction between the spin and the magnetic field. The more exact calculation, first given in an important paper by Heisenberg and Jordan,⁹⁰ leads to the following expression for the change of energy of the hydrogen atom :

$$\Delta E = h\omega_m m - \frac{h\omega_D}{4} \pm \frac{h}{2} \sqrt{\omega_m^2 + 2\omega_D m + \omega_D^2 (l + \frac{1}{2})^2}. \quad (213)$$

In (213) ω_m stands for $\frac{e\mathbf{H}}{4\pi\mu c}$, that is the change of frequency in the normal Zeeman effect, while ω_D gives the amount of the doublet separation due to inner atomic forces. This expression contains as a limiting case Landé's *g*-formula on the one hand, and the Paschen-Back effect on the other. For if we set

$$\omega_m < \omega_D,$$

then (213) becomes

$$\Delta E = h\omega_m m g,$$

which is the same as the earlier expression (210). Conversely we get for

$$\omega_m \gg \omega_D,$$

a value for the energy-difference equal to the Lorentz value.

Besides yielding these energy-levels quantum mechanics has also made it possible to derive a further series of formulæ, partly obtained by means of the principle of correspondence and partly empirically, dealing with the intensities of the components of multiplet structures and of their Zeeman effects which agree excellently with the results of experiment.

But for further details we must refer the reader to the fundamental paper by Heisenberg and Jordan⁹⁰ already mentioned. Finally we wish to state without giving calculations

that Dirac's theory of the spinning electron leads to Sommerfeld's fine-structure formula, the separations in the Zeeman effect and the intensities of the components, all in complete agreement with the elementary theory of the spinning electron and with the results of experiment.

45. Problems of the Diffraction of Matter

The similarity between the form of Schrödinger's equation and the equations of optics allow us to predict certain phenomena qualitatively and then to confirm them quantitatively by translating the corresponding phenomena in the realm of radiation into terms of matter. There are two examples, in particular, that have become of importance for the essential questions involved. We are referring to the scattering of α -particles by atomic nuclei and the selective reflection of electrons by crystals. The first phenomenon was already known to the classical theory and could be explained by it. If an α -particle flies past a centre of attraction, which for simplicity we assume to be at rest, it will be deflected by amounts depending on the distance of nearest approach to this centre. Consider a stream of α -particles moving originally in the same direction towards the scattering layer of matter. Let us assume that in consequence of the great distances between the scattering centres one α -particle will as a rule be appreciably influenced by only *one* nucleus, if at all; the following well-known formula of Rutherford then holds for the intensity of the scattering i at an angle θ to the direction of incidence :

$$i = \frac{\pi i_0 e^2 E^2}{2\mu^2 v^4 \sin^4 \frac{\theta}{2}}, \quad (214)$$

where e , E are the charges of the α -particle and the nucleus, i_0 is the intensity of the incident stream, v denotes the initial velocity and μ the mass of the α -particles.

We shall deal exhaustively in §§ 70-72 with the fundamental difference between the view of the classical and the quantum theory, which consists in the fact that in the classical theory we can in principle still specify the shortest distance p ("impact parameter") to within which an α -particle moving with velocity v approaches the nucleus, whereas in the quantum theory it is no longer possible to determine p and v simultaneously. Here we shall merely deal with the problem as to how we can account for such a scattering of the α -particles by means of Schrödinger's equation and the expressions (139) for the four-current.

In our present case Schrödinger's differential equation reads

$$\Delta\psi + \frac{8\pi^2\mu}{h^2}(W - V(r))\psi = 0. \quad (215)$$

Here $V(r)$ denotes the potential energy of the centre of attraction ; and we consider $V(r)$ as replaced by the potential of the nucleus $\frac{eE}{r}$ for the close distances $V(r)$ that come into question, whereas for great distances we should in addition have to take into account the screening of the electronic envelope. We may interpret this equation optically as if we were dealing with the diffraction of a light-wave in a medium in which the index of refraction is a function of the radius. In the case of a diffraction of this kind the light is scattered in all directions according to well-known laws, and our present aim is to determine the scattered intensity as a function of the angle of scattering. In optics the intensity of a wave is proportional to the square of its amplitude ; and quite similarly our expressions (139) for the four-current are proportional to the square of ψ and its derivative.

The integration of the equation is best performed in practice by taking account of the influence of the perturbation by successive approximations. Before outlining the method of calculation we shall state the result, which is that the quantum-mechanical expression for the intensity of the scattering agrees *exactly* with that of Rutherford. This applies not only to the approximate but also to the rigorous solution of the problem (see references 60 and 85).

The calculation of such problems proceeds along the following lines. We first consider the unperturbed motion : this is the case where the action of the attracting centre is left out of account. Schrödinger's equation then reads

$$\left. \begin{aligned} \Delta\psi_0 + k^2\psi_0 &= 0 \\ \left(k^2 = \frac{8\pi^2\mu W}{h^2} = \frac{4\pi^2}{\lambda^2} \right. \\ \lambda &= \frac{h}{\mu v} = \text{the de Broglie wave-length} \\ \psi_0 &= e^{ikz} \end{aligned} \right\} \quad (216a)$$

if we chose the co-ordinate system so that the unperturbed wave advances parallel to the z -axis.

As a next approximation we set

$$\psi = \psi_0 + \lambda\psi_1$$

and determine ψ_1 by substituting in (215), neglecting terms of the form λ^2 ,

$$\Delta\psi_1 + k^2\psi_1 = \frac{8\pi^2\mu}{\hbar^2}V\psi_0 = \frac{8\pi^2\mu}{\hbar^2}\frac{eE}{r}e^{ikz}. \quad (216b)$$

This non-homogeneous equation may then be integrated according to well-known methods of potential theory and gives as a result for large values of r , which alone come within experimental observation, an expression of the form

$$\psi_1 = \frac{1}{r}f(\theta)e^{ikr}, \quad (216b)$$

where θ denotes the angle between the angle of incidence, which lies in the z -direction, and the angle of scattering.

According to (139) there then corresponds to this form of the proper function (216b) a four-current which depends on the angle and whose radial component is given by

$$I_r \sim \frac{1}{r^2}f^2(\theta). \quad (217)$$

The number of particles that pass through the surface of a sphere of radius r is given by

$$i \sim r^2 I_r \sim f^2(\theta). \quad (218)$$

Calculation shows that (218) agrees exactly with the expression (214). The account here given is a special application of the general theory of M. Born¹⁷ for the collision of material particles. In his paper Born was also successful in deriving the fundamental result of the experiments of Franck and Hertz, according to which an electron colliding with an atom is either scattered, the energy remaining constant (elastic collision),—such as also occurred in the case of the α -particle discussed above,—or else passes on with a loss of energy (*inelastic collision*). Born also investigated the convergence of this process of approximation and found it to be confirmed in general.

Rutherford's formula (214) has been well confirmed, except for certain slight deviations, by experimental results. After what has been said it is clear that these deviations can no longer be placed to the account of quantum mechanics but must be interpreted by means of assumptions about the structure of the atomic nucleus.

The second "diffraction phenomenon" is the selective reflection of electrons by crystals, which perhaps furnishes the

strongest evidence for the wave-nature * of matter. If a monochromatic Röntgen ray of wave-length λ impinges on the surface of a regular crystal, then, as is well known, intensive diffraction phenomena occur in special directions if the following relations due to Laue apply

$$d(\alpha - \alpha_0) = n_1\lambda, \quad (218a)$$

$$d(\beta - \beta_0) = n_2\lambda, \quad (218b)$$

$$d(\gamma - \gamma_0) = n_3\lambda, \quad (218c)$$

in which

α, β, γ denote the direction cosines of the diffracted ray,
 $\alpha_0, \beta_0, \gamma_0$ denote the direction cosines of the incident ray,
 d = the lattice constant,
 n_1, n_2, n_3 = integers.

Analogous experiments, in which the Röntgen ray was replaced by a beam of electrons moving with uniform velocity, led to exactly the same phenomena; in this case the quantity

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

must, by (60), be substituted in Laue's relations. Just as in the case of Röntgen rays a series of differently arranged experiments were devised to demonstrate diffraction phenomena, so these same phenomena could be shown to occur in the case of electron rays. The first discoverers of this phenomenon were the two American physicists, Davisson and Germer,²¹ who worked with what was essentially the same as the original arrangement of Laue, Friedrich and Knipping, the only difference being that they worked with reflected instead of transmitted light. This was necessary because the crystal has an incomparably greater absorptive power for electrons than for Röntgen rays. G. P. Thomson¹¹¹ and Rupp^{106, 107}, who used the method of Debye and Scherrer for Röntgen rays, investigated a great number of substances by means of cathode rays over a wide range of velocities and obtained regular Laue photographs in the transmitted beam. Kikuchi⁹⁸ took photographs through thin laminæ of mica and obtained besides the well-known space-lattice interference pattern, also plane-lattice patterns.

In all experiments in which relatively slow cathode rays

* The expression wave-nature is by no means intended to anticipate the discussion given below, in which we shall have to decide in favour of the point of view of particles, but rather to emphasize the formal analogy that exists between optical and material phenomena in questions of intensity and diffraction.

were used (energy of the primary rays being of the order of 100 volts), the Laue relations were not exactly satisfied, the agreement becoming better for the faster electrons. The cause of the deviation from Laue's relations could be traced back to an index of refraction of the crystal (in some cases even good quantitative values for the deviation could be obtained in this way). For just as a medium which is differently constituted optically from a vacuum *changes the velocity of transmission of light or its wave-length*, so a medium in which a potential exists different from that in empty space *changes the corresponding wave-length of the electrons*. If we imagine two contiguous space-regions, in one of which the potential is zero while it has the constant value Φ in the other, then the following wave-equation holds for electrons which are supposed to have the same total energy E in the two regions :

$$\frac{d^2\psi_1}{dx^2} + \frac{8\pi^2\mu}{h^2}E\psi_1 = 0, \quad (219a)$$

$$\frac{d^2\psi_2}{dx^2} + \frac{8\pi^2\mu}{h^2}(E - \Phi)\psi_2 = 0. \quad (219b)$$

From (219a) and (219b) we get

$$\psi_1 = e^{ik_1x}, \quad k_1 = \sqrt{\frac{8\pi^2\mu E}{h^2}} = \frac{2\pi\mu v}{h} = \frac{2\pi}{\lambda_1},$$

$$\psi_2 = e^{ik_2x}, \quad k_2 = \sqrt{\frac{8\pi^2\mu(E - \Phi)}{h^2}},$$

$$\frac{k_2}{k_1} = \sqrt{\frac{E - \Phi}{E}} = \sqrt{1 - \frac{\Phi}{E}}.$$

The factor of x in ψ_1 and ψ_2 is inversely proportional to the wave-length. Thus the wave-length differs in the two space-regions by the factor

$$\lambda_1 = \lambda_2 \sqrt{1 - \frac{\Phi}{E}}. \quad (220)$$

From the optical relation $\lambda_1 = n\lambda_2$ we have thus to regard

$$n = \sqrt{1 - \frac{\Phi}{E}}$$

as the index of refraction of the medium 2 with regard to the medium 1. The introduction of an index of refraction to explain the deviations from Laue's relation was first suggested by Eckart⁷⁸ and later its relationship with the energy properties of a crystal was exhaustively discussed particularly by Bethe.⁶²

Qualitative support for the introduction of the refractive index is given by the fact that for great values of E the deviations from Laue's relation vanish, since then $\frac{\Phi}{E} \rightarrow 0$. By comparing results of observation with the relation (220) it is possible to determine for a great number of metals the mean potentials Φ that exist in them. In this way it is found that Φ is always less than zero and is of the order of magnitude of 10 volts on the average. The sign of Φ is in good agreement with the fact that the electrons are rigidly held in the interiors of crystals. We can here do no more than call attention to the relationship of these quantities with the new statistics and with the Richardson Effect (cf. § 62), especially as the implications involved seem as yet by no means clear.

46. Crossing a Potential Barrier

In this section we shall discuss a problem which is commonly designated as the phenomenon of crossing a threshold potential (*Potentialschwelle*). The principle involved in this phenomenon is essentially as follows.

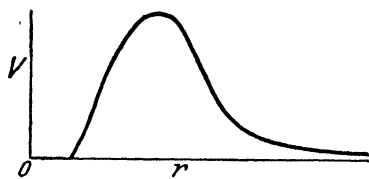


FIG. 5.—Diagrammatic representation of a potential barrier.

Suppose the variation of potential to be represented, say, as a function of the radius purely diagrammatically by Fig. 5. In the language of classical mechanics if

a particle happens to be in the region of small values of r , then it will be able to pass over into the region of large values of r only if its kinetic energy is great enough to overcome the potential "hill" or barrier (*Potentialberg*) indicated in the figure. If its kinetic energy is less it will remain permanently confined within the region of small values of r .

In quantum mechanics this type of description is for a number of reasons insufficient, indeed inadmissible. We shall leave over the general significance which the example here discussed has for the interpretation of the quantum-mechanical equations, as we wish to discuss it fully later in § 72 in connexion with Heisenberg's Uncertainty Relation. For the present we shall restrict ourselves to a formal discussion of the wave-equation involved.

From this wave-equation we see that we must attack the problem by finding a proper function with an associated proper

value, which is valid in the whole region and we must insert in the differential equation the value of the potential as a function of the radius.

We shall now take as a concrete example of our method of treatment a case which is of great practical importance, namely, the emission of electrons from cold metals. This refers to the observation that intense electric fields (of the order of 10^6 volts per cm.) are able to make electrons escape from metals even at room temperatures. Since according to the laws of electrostatics the electric field does not penetrate into the metal this effect is not immediately intelligible. In §§ 61, 62 we shall have to deal more fully on the basis of the new quantum statistics with the structure of metals and the electrons contained in them. Here it will suffice to state the fact, which has been confirmed by a series of experimental observations (Richardson effect, photo-electric effect, and so forth), that the electrons in

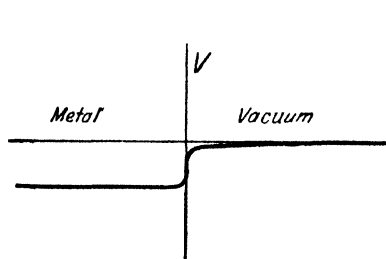


FIG. 6.—Potential barrier at a metal-vacuum boundary.

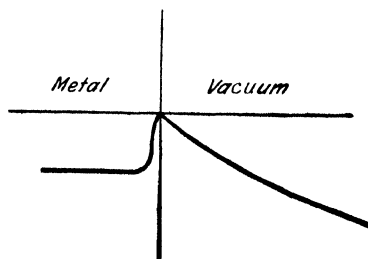


FIG. 7.—Metal vacuum. Potential when an external field is applied.

the metal are held back by a potential difference of the order of a few volts compared with the outside space.* The variation of potential has then the following diagrammatic form.

At the surface of the metal we have the above-mentioned sudden change to a potential, which remains unchanged for infinite distances if external fields are not acting (Fig. 6). If we now suppose an external field to be impressed acting in such a direction that electrons would be driven away from the metal, then the potential is of the form illustrated in Fig. 7. According to classical mechanics an electron can escape, even when an external field is acting, only if its kinetic energy is greater than the potential drop between the metal and vacuum. According to wave-mechanics we have to find, as stated above, a solution of Schrödinger's equation for the whole of space. In practice

* The reader is recommended to read the present section again if possible in conjunction with §§ 61 and 62.

we form it by treating the two regions, metal and outer space, separately and compound the two solutions of ψ_M and ψ_A by means of the boundary condition (cf. § 38) that holds at the surface.

In this way we obtain a complete solution that is valid for the whole of space. We have

$$\Delta\psi_M + \frac{8\pi^2\mu}{h^2}(E - V_M)\psi_M = 0, \quad (221a)$$

$$\Delta\psi_A + \frac{8\pi^2\mu}{h^2}(E - V_A)\psi_A = 0, \quad (221b)$$

$$\psi_M(o) = \psi_A(o), \quad (221c)$$

$$\frac{d\psi_M}{dx}(o) = \frac{d\psi_A}{dx}(o). \quad (221d)$$

In external space the electron has a continuous proper value spectrum, as is easily seen. If we then build up the electric current from ψ_A according to (139), we obtain the number of electrons that escape per second. This furnishes a wave-mechanical explanation of the phenomenon of the emission of electrons from cold metals.⁸²

Our calculation assumes, of course, that the number of electrons that escape from the metal is so small in comparison with the total quantity present that the circumstances are not appreciably affected by the current; this condition is always fulfilled experimentally.

A comparison of the results given by the detailed theory with those obtained by experiment are very satisfactory, at any rate qualitatively.

A case which is fully analogous to that of electron emission is that of the interpretation of the emission of α -rays by radioactive bodies, which has been given by Gurney and Condon⁸⁸ as well as by Gamow⁸³. In our ignorance of the laws of nuclear structure, the fact that α -particles are held captive in atomic nuclei is explained, at present, as follows. According to Coulomb's law an α -particle would always have to be repelled by the nucleus and could never approach infinitely near to it. The holding of an α -particle in the nucleus is thus interpreted purely schematically by assuming that the Coulombian potential near the nucleus alters its form (cf. the diagram in Fig. 5).

This reversal of the value of the potential is to make it possible for the α -particle to be bound to the nucleus. By this means of representation we have again reduced the problem to the above-mentioned crossing of a potential barrier. In the interior of the nucleus the α -particle is under the action of forces that

are unknown to us and about which we make the following simplifying assumptions. We treat the α -particle in the interior of the nucleus as if it were free up to the distance $r = r_0$. From r_0 to infinity we then again assume the Coulombian form of potential (Fig. 8). This makes the problem exactly the same as the extraction of electrons out of cold metals. We must now again find a solution of Schrödinger's equation which is valid throughout the whole of space and which we again find it expedient to build up from two parts which link together at the point r_0 . The frequency with which an α -particle escapes from the nucleus is then calculated exactly as in the case of electronic emission, and the surprising result is obtained that the empirical relation of Geiger and Nuttall comes out directly by calculation from the theoretical assumptions. This relation asserts that the decay constant of an element and the kinetic energy of the α -particle that is emitted during decay obey the relation



FIG. 8.—Diagrammatic representation of the potential in the interior of and near an atomic nucleus.

$$\log \lambda = A + B \log E. \quad (222)$$

Here A and B are constants for one and the same radioactive family.

47. Perturbation Theory of Non-Stationary Systems

Schrödinger's equation (71) is, as we know, correct only if we are justified in speaking of a constant energy of the system. In all cases in which time-dependent forces exist [for example, $V = V(t)$] this equation no longer holds, and in place of (71) we must adduce the complete equation

$$\Delta\psi - \frac{4\pi i\mu}{h} \frac{\partial\psi}{\partial t} - \frac{8\pi^2\mu}{h^2} V\psi = 0 \quad (111)$$

for purposes of discussion.

Since the exact calculation can hardly ever be carried out practically in these cases, a very striking perturbation method has been worked out which allows us in many cases to predict the effect of perturbing forces without performing the integrations in detail.

If we write (111) in the form

$$\Delta\psi - \frac{4\pi i\mu}{h} \frac{\partial\psi}{\partial t} - \frac{8\pi^2\mu}{h^2} (V_0 + V_1)\psi = 0, \quad (223)$$

where

$$V(x, y, z, t) = V_0(x, y, z) + V_1(x, y, z, t),$$

then we may write down the general solution of

$$\Delta\psi - \frac{4\pi i\mu}{h} \frac{\partial\psi}{\partial t} - \frac{8\pi^2\mu}{h^2} V_0\psi = 0 \quad (223a)$$

in the form

$$\psi = \sum a_n e^{\frac{2\pi i E_n t}{h}} v_n + \int dE a(E) e^{\frac{2\pi i Et}{h}} v(x, E). \quad (224)$$

Here the a_n 's denote constant factors, the v_n 's denote the *known proper functions* of the unperturbed system, and our manner of writing the expression shows that we are taking into account not only the discrete but also the continuous proper value spectrum. For the sake of convenience of notation we shall in the sequel write down only the discrete terms, but it must be remembered that the formulæ are to be supplemented by the proper functions of the continuous spectrum, the summation signs then being supplanted by integrals. On account of

$$\int d\tau \psi \bar{\psi} = \int d\tau \left[\sum_n a_n \bar{a}_n \psi_n \bar{\psi}_n + \sum_k \sum_l a_k \bar{a}_l \psi_k \bar{\psi}_l e^{\frac{2\pi i (E_k - E_l) t}{h}} \right] = 1$$

we have

$$\sum_n a_n \bar{a}_n = 1.$$

And, indeed, this relation holds for all times, not only initially. We may now also try to satisfy the complete equation (223) with the assumption (224), but the a_n 's will then no longer be constants but functions of the time, which we shall determine from (223). We assume the a_n 's to be known at the time $t = 0$. Substituting (224) in (223) then gives

$$V_1 \sum_k a_k e^{\frac{2\pi i E_k t}{h}} v_k - \frac{h}{2\pi i} \sum_k \dot{a}_k e^{\frac{2\pi i E_k t}{h}} v_k = 0. \quad (225)$$

If we set $V_{km}(t) = \int a_k v_m V_1 d\tau$, multiply (225) by one v_k and integrate, we get

$$\frac{da_k}{dt} = \frac{2\pi i}{h} \sum_l V_{kl}(t) e^{\frac{-2\pi i (E_k - E_l) t}{h}} a_l. \quad (226)$$

This is a system of total differential equations in the unknowns $a_k(t)$, in which the coefficients $V_{kl}(t)$ are *known functions of the time*. Up to this point the argument has been absolutely rigorous. To perform the integration, however, it is necessary

to make approximative assumptions. But we must first prove that if $\sum a_n \bar{a}_n = 1$ holds at any time then it must hold at all times even if the perturbing forces depend on the time. For this purpose we form

$$\dot{a}_k a_k + \dot{\bar{a}}_k \bar{a}_k$$

with the help of (226); and since $\bar{V}_{km} = \bar{V}_{mk}$ we obtain the desired result,

$$\frac{d}{dt} \sum_n a_n \bar{a}_n = 0. \quad (227)$$

This relation (227) is fundamentally important for the interpretation that is to be discussed in § 68. For it allows us to set the quantities $\bar{a}_n \bar{a}_n$ proportional to the number of atoms that exist in the state n , since the total number of atoms does not change, as we see by (227).

We now perform the approximative integration by inserting for the quantities a_k on the right-hand side their values a_k^0 at the time $t = 0$.

We then obtain as the solution of (226)

$$a_k(t) = a_k(0) + \frac{2\pi i}{h} \sum_l \int_0^t V_{kl} e^{\frac{2\pi i(E_k - E_l)t}{h}} a_l dt. \quad (228)$$

This solution is of course only a first approximation, but it may be improved by further approximations and it is satisfactory in practice for all cases that occur (except for Dirac's dispersion theory).

The physical meaning of the time-integrals in (228) may easily be given. If we assume that for $t = 0$ all the a_l 's except a_m are zero, then

$$a_l(t) = \frac{2\pi i}{h} \int_0^t V_{lm}(t) e^{\frac{2\pi i(E_l - E_m)t}{h}} a_m dt. \quad (229)$$

So these integrals are a measure of the probability of a transition from m to l within the time $0 \leq t \leq t_1$, the initial state $a_m = 1$, $a_n = 0$ ($n \neq m$) being assumed.

The form of (229) becomes particularly interesting if the perturbing forces and so also V_{km} are periodic functions of the time. By performing the integration we then get from (228) that

$$a_l = a_m \left\{ c_{mn} \frac{e^{\frac{2\pi i(E_n - E_l + h\nu)t}{h}}}{E_n - E_l + h\nu} + \bar{c}_{nm} \frac{e^{\frac{2\pi i(E_n - E_l - h\nu)t}{h}}}{E_n - E_l - h\nu} \right\}. \quad (230)$$

This represents a_l as a Fourier series with respect to the time, in which the amplitudes have characteristic denominators:

$$E_n - E_l \pm h\nu.$$

In this Fourier series those terms will predominate, for which a numerator becomes very small or even vanishes. But even when the denominator vanishes,

$$E_k - E_l \pm h\nu = 0, \quad (231)$$

the corresponding expression in (229) still remains finite. If we imagine the numerator of the expression (229), namely, $e^{\frac{2\pi i}{h}(E_n - E_l \pm h\nu)t}$ to be expanded for very small values of t , we obtain for the amplitude

$$a_l \sim t. \quad (232)$$

Thus the calculation of time-dependent perturbation problems will in general depend on our finding out whether small or vanishing denominators occur. These "resonance" denominators, as they are called, then make the corresponding term predominate over all other possibilities of transition that may come into question.

We are dealing here with what may be regarded as an extension of the conception "degeneracy." Whereas in the case of the stationary solution degeneracy is definable so that two proper functions *have the same energy*, that is, depend on the time in the same way, namely $e^{\frac{2\pi i E}{h}t}$, now, in the time-dependent case, we can only speak of *the two proper functions depending on the time in the same way*, without ascribing to them the same energy on this account. Indeed, the energy may be allocated uniquely only if the proper function is a simply periodic function of the time. We shall learn the deeper reason for this later when we discuss Heisenberg's Uncertainty Relation.

48. Theory of Dispersion

We treat as our first time-dependent perturbation problem the phenomena of dispersion. For this purpose we take over from the classical theory the following theorem which must certainly also hold in quantum electro-dynamics: the index of refraction of a medium is determined by the electric moment that the particles acquire under the influence of the incident wave. Suppose unit volume to contain N vibration elements

which are independent of one another. If we call the perturbing force $\mathbf{E}$, then we have for the index of refraction of the medium

$$n^2 - 1 = 4\pi N \frac{\mathbf{m}}{\mathbf{E}}, \quad (233)$$

where $\mathbf{m}$ denotes the electric moment of an atom under the influence of the perturbing force.

The calculation is fully analogous to the scheme given in § 47. Let the complete equation be

$$\Delta\psi - \frac{4\pi i\mu}{h} \frac{\partial\psi}{\partial t} - \frac{8\pi^2\mu}{h^2} (V_0 + Az \cos 2\pi\nu t) \psi = 0, \quad (234)$$

where $Az \cos 2\pi\nu t$ denotes the potential from which the electric intensity of field is derived.

In the following discussion we assume that the incident frequency ν coincides with *no* absorption frequency $\nu_{kl} = \frac{E_k - E_l}{h}$, so that none of the denominators in (238) vanishes. The treatment of the absorption process as well as of the reaction of the atom to the field of radiation are necessarily excluded by this assumption. The assumption is necessary to assure that there is convergence in our approximative calculation.

Let us consider the dispersion of a system, in which before the incidence of the wave all atoms are in the state denoted by the energy E_n . We may then write

$$\psi_n = \psi_n^0 + \lambda\phi_n(x, y, z, t)$$

and we thus again obtain for ϕ_n a non-homogeneous equation

$$\Delta\phi_n - \frac{4\pi i\mu}{h} \frac{\partial\phi_n}{\partial t} - \frac{8\pi^2\mu}{h^2} V_0\phi_n = \frac{8\pi^2\mu Az\psi_n^0}{h^2} \cos 2\pi\nu t. \quad (235)$$

By considerations that are exactly analogous to those of §§ 42 and 47 we then obtain

$$\psi_n^0 = v_n^0 e^{\frac{2\pi i E_n t}{h}} \quad (236a)$$

$$\phi_n = \omega_n^+ e^{\frac{2\pi i (E_n + h\nu)t}{h}} + \omega_n^- e^{-\frac{2\pi i (E_n - h\nu)t}{h}} \quad (236b)$$

$$\Delta\omega_n^\pm + \frac{8\pi^2\mu}{h^2} (E_n + h\nu - V_0) \omega_n^\pm = \frac{4\pi^2\mu}{h^2} Az v_n^0. \quad (237)$$

From the solution for $\omega_n^\pm$ or ϕ_n we can, by (139), then obtain the electric density, which comes out as

$$\rho_{nn} = v_n \bar{v}_n + 2 \cos 2\pi\nu t \sum_l \frac{(E_n - E_l) Az v_{nl} v_n^0 \bar{v}_l^0}{(E_n - E_l)^2 - h^2 \nu^2}, \quad (238)$$

and correspondingly the electric moment in the x -direction, say, as

$$\mathbf{m}_x^{nn} = \int x \rho_{nn} d\tau.$$

This completes the mathematical part of the problem and we have now to give the interpretation of (238).

In the classical theory the dispersion formula of a system, which had several proper frequencies (in the case of ideal linear oscillators), was given by

$$n^2 - 1 = \sum_l \frac{S_l}{\nu_{nl} - \nu^2}. \quad (239)$$

Here we have the expression analogous to (238). In the quantum theory as in the classical theory *the observed frequencies of emission* essentially determine the form of the dispersion formula. In the classical theory these frequencies ν_{nl} coincided with the frequencies of the motion. In the quantum theory this is no longer the case; nevertheless only observable optical frequencies enter into the dispersion formula, as is demanded by the Correspondence Principle. The quantities S_l in (239) give a measure of the intensity of the oscillators. The quantity

$$(E_n - E_l) A z_{nl} \nu_n \nu_l$$

in (238) corresponds to them.

A further refinement is given by the dispersion formula just obtained,—a refinement which Kramers had deduced earlier from considerations of correspondence. Using the language of the original Bohr theory we suppose an atom to be in an excited state with energy $E_n > E_l$. Then the oscillators that enter into the dispersion formula and play a predominant part are not only those terms of the sum (238) that correspond to the transitions $E_n - E_l$ ($l > n$), which we may in a certain sense call absorption terms, but also the terms $(E_n - E_l)$ where $l < n$.*

Finally, the quantum formula of dispersion contains possibilities which we have already encountered in the classical theory, although in another field. We are referring to the occurrence of the so-called combination-vibrations (sum- and difference-tones). By this we mean the following. If a mechanical system, say a linear oscillator, in which the restoring force is proportional to the elongation, is caused to vibrate by some periodic external force, only two different frequencies occur in its motion, namely,

* We here meet with the behaviour which plays an important rôle in Einstein's well-known derivation of Planck's radiation formula in the guise of positive and negative induced absorption.

its proper frequency ν_0 and the frequency of the external perturbation ν . The position is quite different when we are dealing with a non-harmonic oscillator, that is one in which the restoring force contains also higher powers of the elongation. Then the forced vibration contains, besides the two frequencies just mentioned, also new frequencies of the form

$$\nu_0 \pm \nu$$

which have been called, owing to their structure, combination vibrations. This phenomenon has long been known in acoustics and has been the subject of much research. In the dispersion phenomena of optics it has hitherto sufficed in practice to assume the configurations in question to be represented very approximately by harmonic oscillators. It is only in recent years that experiments, particularly those of the Indian physicist Raman ¹⁰⁵ have shown that this is not always sufficient. Certain molecules, which have several proper frequencies (nuclear vibration, rotation, electronic motion) exhibit in the light scattered from them not only the frequency of the incident ray but also this frequency augmented or diminished by an amount equal to the proper frequency of the vibration or rotation.

This is then a purely classical effect which only mere chance has brought to light so late and it fits excellently into the scheme of the quantum theory. The determination of the electric density ρ_{nn} of the state n given in (238) does not alone completely determine the behaviour of an atom. There must also be added the electric moment which is assigned to the transitions

$$\rho_{nk}d\tau = \psi_n \bar{\psi}_k d\tau.$$

This electric moment then has the factor $e \frac{2\pi i(E_n - E_k \pm h\nu)t}{h}$ which depends on the time. This agrees exactly with what we found earlier about combination tones.

For the sake of completeness we add that in making calculations about dispersion phenomena the continuous proper value spectrum must be taken specially into account, as if we base our arguments on a model it is possible to derive values for the intensities of the oscillators which are in agreement with experimental results only with the help of the continuous proper value spectrum.

49. The Photo-Electric Effect and the Compton Effect

The photo-electric effect and the Compton effect have always been regarded as the strongest supports of the quantum

theory of radiation. Expressed in terms of this theory these phenomena manifest themselves as follows :

1. A light-quantum (photon) of frequency ν vanishes while an electron, which we suppose attached to the atom by an amount of energy E_0 , escapes with a kinetic energy W , so that

$$W = h\nu - E_0. \quad (240)$$

This is Einstein's Photo-electric Equation. E_0 is the negative term-value of the electron (it must be remembered that the energy of an electron bound to an atom is always negative).

2. A photon of frequency ν which propagates itself along the Z -axis encounters a free electron at rest; the frequency ν of the photon becomes changed into ν' and its direction is turned through an angle θ , while the electron moves off with the velocity v in the direction making the angle ϑ with the Z -axis. The

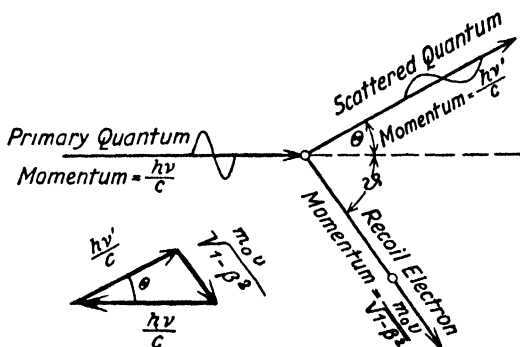


FIG. 9.—To illustrate the laws of conservation of energy and momentum in the Compton effect.

following relations then hold between the quantities ν , ν' , θ , v and ϑ (see Fig. 9) :

$$h\nu = h\nu' + \mu c^2 \left(\frac{1}{\sqrt{1 - v^2/c^2}} - 1 \right); \quad (241)$$

$$\frac{h\nu}{c} = \frac{h\nu'}{c} \cos \theta + \frac{\mu v}{\sqrt{1 - v^2/c^2}} \cos \vartheta, \quad (242a)$$

$$0 = -\frac{h\nu'}{c} \sin \theta + \frac{\mu v}{\sqrt{1 - v^2/c^2}} \sin \vartheta. \quad (242b)$$

Elimination immediately shows that the wave-lengths of the primary and the scattered ray of light are connected by the following relation :

$$\lambda^1 = \lambda + \frac{h}{\mu c} \sin^2 \frac{\theta}{2}. \quad (243)$$

The relations (241), (242a), (242b) represent the laws of conservation of energy and momentum, an amount of energy $h\nu$ and of moment $\frac{h\nu}{c}$ being ascribed to the light-quantum of frequency ν . The process of scattering by free electrons thus entails, as (241) shows, a *reduction of the frequency* of the incident light: this is an obvious consequence of the energy law, since the energy of the electron is obtained at the expense of the primary radiant energy.

In principle those electrons may be called free which require only very little work to release them in comparison with the energy quantum $h\nu$ of the incident radiation. This holds sufficiently approximately for the scattering of Röntgen rays by elements of low atomic number. As the atomic number

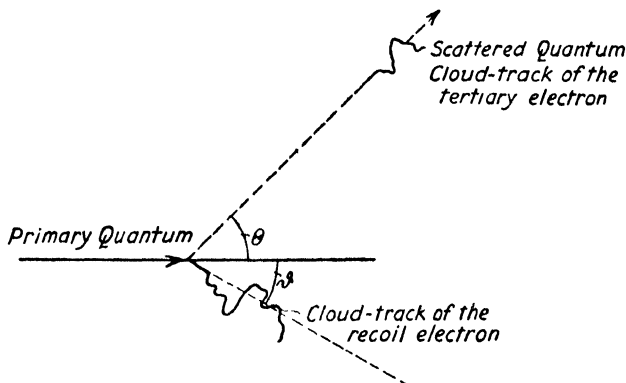


FIG. 10.—The experiment of Compton and Simon.

increases (or the frequency decreases) the conditions can no longer be represented so simply and finally; so far as the release of electrons is concerned, they pass over into the laws of the photo-electric effect, for which only the question of the equating of the energy is of importance. We need not trouble about the equating of the momentum in the case of the photo-electric effect because the electrons are bound with considerable energy in comparison with $h\nu$ and react with the atomic core, which takes up corresponding quantities of momentum.

The laws of the photo-electric and the Compton effect given in (240) and (241, 242) have been confirmed experimentally over a very wide range. Besides the observations made of the energy of the scattered electrons and the changes of frequency of the scattered light there are two other experiments, one by Bothe and Geiger,⁶⁶ the other by Compton and Simon,⁶⁷ which are of

great interest. The first of these two experiments showed that the scattering of the light and the release of the electrons in the Compton effect occurred *simultaneously* so far as could be determined within the limits of experimental error. The experiments of Compton and Simon proved that the law of the conservation of momentum holds not only statistically but also for every individual process. By using the Wilson camera (see Fig. 10) to photograph a great number of individual processes they showed that within the limits of experimental error the angles θ and ϑ actually satisfy the relationships (241), (242a, b). Apart from the questions of intensity to be discussed below this was the position of the theory before the advent of quantum mechanics. We see that the explanation appears necessarily to involve corpuscular assumptions about the nature of light, the connexion with the undulatory properties of light being left in abeyance. After what has been said it seems justifiable to regard it as a particularly striking success of quantum mechanics that it can represent the effects described without making particular corpuscular assumptions about the nature of light. It becomes particularly easy to explain formally the photoelectric effect, which is really contained in the general formulæ of § 47. If, as the facts demand, we regard the unperturbed atom as being in the state denoted by the proper function v_m , and if we allow the periodic perturbation $V_1 = Az \cos 2\pi\nu t$ to act on it, then according to (230), new proper functions present themselves, whose coefficients have the denominators

$$E_l - E_m - h\nu. \quad (244)$$

The decisive terms are those with vanishingly small denominators. Thus we immediately obtain Einstein's photoelectric law:

$$E_l = h\nu + E_m \quad (E_l > 0, E_m < 0).$$

Expressed in the language of quantum mechanics the photoelectric effect takes place as follows: an electron with the proper function $v_m e^{\frac{2\pi i E_m t}{h}}$ is excited by the incident frequency ν to make transitions to the state with the proper function $v_l e^{\frac{2\pi i}{h}(h\nu + E_m)t}$. In this expression we must always have

$$E_l > 0$$

in order that the electron may actually be able to escape. Thus we are here dealing with a combination of a discrete proper value with a continuous one.

In principle the laws of the Compton effect may be deduced

in exactly the same way as those of the photo-electric effect, except that now a special condition is added. For, while in the case of the photo-electric effect the initial term may be combined with any arbitrary final term, in the Compton effect this is possible only for such final terms for which the momentum of the electron together with the momentum of the emitted radiation satisfies the particular relations (242*a*, *b*). This is the result of a calculation which we cannot here repeat in full.^{76, 88}

The surprising results of this purely corpuscular theory of light may thus be obtained as consequences of quantum mechanics, the expression for the four-potential of light being taken over from Maxwell's theory, just as was done earlier with the expression for the potential of the atomic nucleus. Hence at the present stage of our knowledge there is no imperative reason for assuming the existence of "atoms of light." To what extent certain general views about the parallelism of waves and corpuscles may lead us to assume light-corpuscles will not be discussed here; this question stretches over into the region of quantum electrodynamics, which is at present only in process of evolution.

So far we have not yet spoken of the frequency with which the scattering processes occur nor of the directional distribution. The purely corpuscular theory gives no basis for determining these quantities; indeed, even the old quantum theory had to seek the help of the correspondence principle. The new quantum theory is able to answer these questions unambiguously, since, as we know, the four-current is a measure of the rate of happening of elementary processes. Now since in general the proper functions will depend on the direction we have a possibility of interpreting the relevant questions. Without giving details we can here state that the theoretical predictions of the way in which intensity and angle of emission is involved in individual scattering processes are in good agreement with experimental results.*

* With reference to the special refinements of quantum-mechanical treatment we have yet to mention the investigation of Wentzel into the Compton effect in the case of bound electrons, which are to make possible the transition above described to the pure photo-electric effect, as well as the rigorous calculation of the Compton effect according to Dirac's theory in place of the relativistic wave-equation of Schrödinger and Gordon, which has been given by Klein and Nishina.¹⁰⁰ The latter have found that for the regions of frequency that come into question there is almost complete agreement with the researches described in the text; derivations in intensity and angular distribution were to be expected only in the case of very high frequencies, such as those of cosmic radiation.

CHAPTER IX

MANY-BODY PROBLEMS

50. Exchange Degeneracy. A Particular Example

IN the preceding pages we have occupied ourselves exclusively with so-called one-body problems, and have pictured the mutual action between the corpuscle in question and other atoms by means of external fields which for their part undergo no changes through the events in the atom under consideration. The following sections are to discuss in detail the interaction of atomic configurations. In the case of these many-body problems certain new phenomena, for which we can find no analogy in classical physics, present themselves. It is thanks to these new phenomena that we are now able quantitatively to interpret a whole series of phenomena which were mysterious even qualitatively, such as multiplet structure and related phenomena. To illustrate the state of affairs as strikingly as possible we shall begin by taking an example due to Heisenberg,²⁷ which exhibits all the characteristic features.

Let us consider two linear oscillators which are initially to perform oscillations along a straight line without interaction. The wave-equation of the system then runs

$$\frac{\partial^2 \psi}{\partial x_1^2} + \frac{\partial^2 \psi}{\partial x_2^2} + \frac{8\pi^2\mu}{h^2} \left[E - \frac{a^2}{2}(x_1^2 + x_2^2) \right] \psi = 0. \quad (245)$$

The solution ψ of this wave-equation is simply a product of the two wave-functions of the individual oscillators, or, expressed more generally, a sum of linear functions of these products. We cannot at this stage distinguish whether for a given total energy E the oscillator 1 has the energy ϵ_1 and the oscillator 2 the energy ϵ_2 , or vice versa, so long as we know only that $\epsilon_1 + \epsilon_2 = E$.

The general solution may be expressed by the sum

$$\begin{aligned} \psi &= c_1 \psi_1 + c_2 \psi_2, \\ \psi_1 &= \psi_{\epsilon_1}(x_1) \psi_{\epsilon_2}(x_2), \quad \psi_2 = \psi_{\epsilon_2}(x_1) \psi_{\epsilon_1}(x_2), \end{aligned} \quad (246)$$

in which, on account of the condition of orthogonality,

$$\int \psi_i \bar{\psi} dx_1 dx_2 = 1, \quad \int \psi_i \bar{\psi}_k dx_1 = \delta_{ik},$$

the coefficients c_i must satisfy the equation $\sum c_i^2 = 1$ but are otherwise quite arbitrary. Thus for each proper value we have not one but several proper functions; we may say, therefore, that the system is degenerate.

As is well known, such degenerate systems are treated as if subject to a slight perturbation, and we have to investigate whether the proper values are then split up so that ultimately only one proper function belongs to each proper value. The present case differs from those of energy degeneracy considered earlier inasmuch as in the case of energy degeneracy the fact of the particle being in the state 1 or 2 made a physical difference, whereas in our case, on account of our having assumed the two oscillators to be absolutely identical, we cannot distinguish whether we are dealing with the state ϵ_1, ϵ_2 or ϵ_2, ϵ_1 . This type of degeneracy, which is caused by the parts of a mechanical system being exactly similar (interchangeable), is called *Exchange Degeneracy*.

We now assume a slight perturbation of the potential energy, say, of the form $\lambda x_1 x_2$. This perturbation function is *symmetrical** in the co-ordinates of the particle. This is not due to chance but rather is a physical necessity on account of the fact mentioned above, namely, that the particles may be exchanged. Also, every other perturbation function (electromagnetic field) must be symmetrical with respect to the co-ordinates of the particles. The wave-equation in the corresponding co-ordinates then runs

$$\frac{\partial^2 \psi}{\partial x_1^2} + \frac{\partial^2 \psi}{\partial x_2^2} + (E - V_0 - \lambda x_1 x_2) \psi = 0. \quad (247)$$

It may be separated by means of the substitution

$$X_1 = \frac{1}{\sqrt{2}}(x_1 + x_2), \quad X_2 = \frac{1}{\sqrt{2}}(x_1 - x_2). \quad (248)$$

As before we then get the solution as a product of two proper functions of oscillators, but this time they are of different frequencies:

$$\begin{aligned} \nu_1^2 &= \nu_0^2 + \lambda^2, & \nu_2^2 &= \nu_0^2 - \lambda^2, \\ E &= (n_1 + \tfrac{1}{2})h\nu_1 + (n_2 + \tfrac{1}{2})h\nu_2. \end{aligned} \quad (249)$$

* A function $f(x_1, x_2)$ is called symmetrical in x_1 and x_2 if $f(x_1, x_2) = f(x_2, x_1)$; anti-symmetrical functions in x_1 and x_2 are those for which $f(x_1, x_2) = -f(x_2, x_1)$.

This does away with the degeneracy; the terms are no longer equidistant but lie at different distances.

We are next concerned with the question as to what transitions are possible in the coupled system, firstly for when there are no external perturbations. In § 14 it was shown that according to classical theory the emission of radiation from an electromagnetic system with an arbitrary distribution of charges is determined according to Hertz's theory by its dipole-moment. This assumption was transferred in § 14 with the help of the correspondence principle to quantum mechanics, the matrix-element of the dipole-moment, ex_{mn} being assigned to the transition $m \rightarrow n$ (Cartesian co-ordinates being used).

These reflections now need to be carried a little further. Even in the classical theory we find that not only the dipole-moment but also the so-called multipoles determine the emission. In the case of two linear oscillators, which we are considering, the dipole-moment is given by

$$e(x_1 + x_2) = e\sqrt{2}X_1,$$

the higher poles are given by expressions of the type

$$x_1^2 + x_2^2,$$

and so forth. If the dipole happens to vanish in consequence of some mechanical property of the system, then the multipoles determine the emission. This holds, as we have said, even in the classical theory. In quantum mechanics the matrix-elements of the multipole-elements

$$(x_1^2 + x_2^2)_{mn}$$

and so forth, are then to be formed in an exactly corresponding way; these matrix-elements allow us to calculate the probability for a transition between m and n .

In the special case when (248) and (249) hold the matrix of the separation co-ordinate X_1 depends only on the quantum number n_1 , that of the separation co-ordinate X_2 depends only on X_2 . By (248) the matrix of the dipole-moment $e(x_1 + x_2)_{mn}$ comes out now as

$$eX_{1n_1n_1'}.$$

From this we see that the dipole-moment to be allocated to the transitions depends solely on n_1 , that is, on account of the dipole-moment transitions only from $n_1 \rightarrow n_1 \pm 1$ may occur, as in the case of the linear oscillator. Thus spontaneous changes of the quantum number n_2 are not possible in this order of approximation.

To determine the exact values of the transition probabilities we must now also consider the influence of the multiple moments. Of these we may say at once without further calculation that they are symmetrical in the co-ordinates x_1 and x_2 , a property of considerable importance. It is obvious, because there is no distinction between the two oscillators. But a symmetrical function of x_1 and x_2 , when expressed in the co-ordinates X_1 and X_2 , must contain only even powers, since on account of (248) it changes sign when x_1 is exchanged with x_2 . The matrix-elements of the linear oscillator are now (cf. § 41) of the form

$$x_{ik} \neq 0 \quad \text{for} \quad i = k \pm 1.$$

In the same way the matrix-elements of X_2 are equal to

$$X_{2n_1, n_2'} \neq 0 \quad \text{for} \quad n_2 = n_2' \pm 1.$$

In calculating the matrix-elements of an even power of X_2 we immediately get, by applying the rule of matrix multiplication (37), that only those matrix-elements (ik) can differ from zero for which i and k differ by even numbers. From this it follows that the radiation, even when all multipoles are taken into account, can only cause n_2 to change by even multiples. The system cannot spontaneously pass from transitions from a state with an odd value for n_2 to one with an even value for n_2 . The totality of all term-values thus resolves itself into two systems which do not combine with each other. If we imagine the one system to be realized initially in physical nature then the second system cannot be derived from it by spontaneous emission.

This statement is now to be extended to include external perturbations. Such perturbations likewise alter nothing in the inferences made above. For since they are certainly symmetrical in x_1 and x_2 , and hence will certainly involve only X_2 as an even power, they similarly yield only such matrix-elements for which X_2 changes only by even numbers. So the resolution into two non-combining term-systems also persists in the most general case in which emission of radiation and external forces are taken into account.

These far-reaching conclusions were made possible solely because the particles considered were exactly similar; this is the characteristic feature of all atomic theories. We may therefore expect, even in the case of general mechanical systems, to find the same behaviour, namely, the removal of degeneracy by coupling two similar systems and the occurrence of non-combining term-systems.

51. Particle Degeneracy. The General Proof for the Problem of n -Bodies

To arrive at a general proof of the occurrence of non-combining term-systems we consider the time-dependent Schrödinger equation of the many-body problem :

$$\frac{h}{2\pi i} \cdot \frac{\partial \psi}{\partial t} + (H_1 + H_2 + \dots H_f)\psi = 0. \quad (250)$$

Here H_i is the operator of the i th electron. We shall disregard the question of mutual action for the present. A solution of this wave-equation is then given by

$$\psi_I = e^{\frac{2\pi i E t}{h}} \psi(E_1, x_1) \psi(E_2, x_2) \dots \psi(E_f, x_f). \quad (251)$$

The following relation holds :

$$\Sigma E_i = E. \quad (251a)$$

From this solution we obtain just as before any arbitrary number of other solutions by commutating the indices and compounding the resulting expressions by means of arbitrary coefficients to form linear functions :

$$\psi_{II} = e^{\frac{2\pi i E t}{h}} \psi(E_1 x_k) \dots \psi(E_k x_1), \quad (252)$$

$$\psi_{III} = c_I \psi_I + c_{II} \psi_{II}, \quad (252a)$$

and so forth.

In the absence of mutual action and perturbing forces each electron will then remain in its own state or will spontaneously emit radiation. If we imagine mutual action to be added, then by the formulæ of the general theory of perturbations (cf. § 47) energy will pass from one electron to other electrons, that is, the initial state will not remain constant. We now enunciate the following general theorem : if at the time $t = 0$ the solution in question is symmetrical (or anti-symmetrical) in the co-ordinates x_1, x_2 , it also remains symmetrical (or anti-symmetrical) for arbitrary perturbing forces for all times.

Thus at the time $t = 0$ let $\psi(1, 2, 3, \dots) = \psi(2, 1, 3)$, then we also have $H[\psi(1, 2, 3)] = H[\psi(2, 1, 3 \dots)]$.

Since ψ_I, ψ_{II} each alone satisfies Schrödinger's equation, so also does their difference. Thus when $t = 0$ we have

$$H[\psi_I - \psi_{II}] - \frac{h}{2\pi i} \frac{\partial}{\partial t} [\psi_I - \psi_{II}] = 0. \quad (253)$$

This means that if the difference vanishes at a certain moment of time, it vanishes for all moments of time, that is, *always*. An existing symmetry can therefore never be disturbed. If

a system of definite symmetry is present initially, then it can never pass over into one of different symmetry.

The fact that $\psi_I - \psi_{II}$ also satisfy Schrödinger's equation can be asserted immediately only because from physical reasons of symmetry every perturbation function must be symmetrical in all co-ordinates. This shows that there are certainly non-combining systems of solutions of Schrödinger's equation. The question as to the number of non-combining systems that occur in an N -body problem is an algebraic problem which Wigner was the first to solve.

We now select from all the non-combining term-systems two which can combine neither with each other nor with any other term-system. In the case of the first we assume that the proper function is symmetrical in the *Cartesian co-ordinates of all its particles*, that is, that its form does not change for any permutation of the indices. We call this the *symmetrical* solution. The second solution is that for which, when two arbitrary particles are exchanged the proper function changes its sign. We call this the *anti-symmetrical* solution. According to what has already been said neither solution can combine with any other, since every other solution is less symmetrical than the symmetrical and more symmetrical than the anti-symmetrical solution. We obtain the symmetrical or, respectively, the anti-symmetrical solution in the simple case in which the solution is given as the product of the proper functions of the individual systems, as follows. Let

$$\psi_{\text{sym}} = \sum \psi_1 \psi_2 \dots \psi_f, \quad (254)$$

$$\text{and} \quad \psi_{\text{anti}} = \begin{vmatrix} \psi_1(x_1) & \dots & \psi_f(x_1) \\ \psi_1(x_2) & \dots & \\ \dots & \dots & \\ \psi_1(x_f) & \dots & \psi_f(x_f) \end{vmatrix}. \quad (255)$$

The summation sign in (254) denotes that summation is to be performed for all permutations of the indices.

The fact that two symmetrical or anti-symmetrical proper functions can combine only with those of similar symmetry can also be seen easily as follows. Let V' be the perturbation function (mutual action) which, as has been repeatedly emphasized, is symmetrical in the co-ordinates of the particles. The probability of a transition $m \rightarrow$ is then proportional to

$$\int dx_1 \dots dx_f \dots \psi_n V' \psi_n \dots \quad (256)$$

(cf. § 47) and is, of course, independent of the way in which the indices are assigned to the variables.

If we exchange say x_i with x_k then V' and the element of integration remains unchanged, but ψ_m, ψ_n change their signs if they have not the same symmetry in x_i and x_k . Since, on the one hand, the integral is not allowed to change its value when the change in indices is effected, and, on the other hand, is multiplied by -1 , it must vanish. This holds for every pair of indices i and k . Thus a function which is symmetrical (or anti-symmetrical) in all its indices can combine only with symmetrical (or anti-symmetrical) functions.

The present state of the *theory* does not enable us to discover what term-system is realized in nature, whether this term-system depends on the nature of the particle, or whether several term-systems can exist for one particle. On the other hand, *experiment* teaches us that for electrons only *one term-system*, namely, the *anti-symmetrical type*, is *actually realized*. For this is the content of Pauli's Exclusion Principle (*Verbot*) which was formulated earlier and is here expressed in the language of wave-functions. If we consider, for example, two electrons in an atom, then if their mutual action is disregarded, their anti-symmetrical solution is given by

$$\psi_{\text{anti}} = \psi_1(x_1)\psi_2(x_2) - \psi_1(x_2)\psi_2(x_1).$$

We see that it vanishes identically if the wave-functions of the individual electrons are the same. This is the content of Pauli's *Verbot* extended in the sense that even in all those cases in which the electrons occur in different states only the anti-symmetrical solution can be realized. For if in the case of electrons in different states another solution, for example, the symmetrical solution, could also occur in nature, then a transition to another state could occur, in which both electrons would have the same quantum numbers while having symmetrical proper functions, and this is in contradiction to Pauli's Principle.

In the case of protons experiments consisting of observations of band-spectra and the specific heat of the hydrogen molecule have shown very recently that for them, too, the anti-symmetrical system alone is realized. We shall develop this point further below.

52. The Helium Atom

The most important application of the theory to many-body problems consists in the exact analysis of the helium atom. As described in Part I of our Report it was primarily this problem which brought the Bohr theory to its fall. After the above remarks we may express this by saying that the helium problem

defined solution because it had been found impossible in principle to give grounds for the existence of non-combining term-systems.

In helium two electrons associate themselves together under the influence of a nucleus to form an atom: in the argument that follows we assume this nucleus to be permanently at rest. If we neglect the mutual action at first, then we again have the typical case of degeneracy; for a particular energy E there are several proper functions $\psi_1(\epsilon_1)\psi_2(\epsilon_2)$ and $\psi_1(\epsilon_2)\psi_2(\epsilon_1)$ respectively, if we are given only that $\epsilon_1 + \epsilon_2 = E$.

If we next take the mutual electrical action into account each term resolves itself and we then get two different proper functions.

If the mutual electrical action is taken into account this degeneracy lapses. It is at this stage that the great importance of the perturbation theory of degenerate systems given in § 42 becomes manifest. Applying the formulæ (187)-(191) literally, we have to determine both the ϵ 's as the roots of the determinantal equation as well as the c_{ik} 's. Here again the exact similarity of both electrons plays a decisive part. Owing to this we have not only the relation $\epsilon_{12} = \epsilon_{21}$ which holds for *all* systems, but also

$$\epsilon_{11} = \epsilon_{22}. \quad (257)$$

[This follows from the meaning of ϵ_{ik} , cf. (191).] So we get as the solution of (190)

$$\epsilon = \epsilon_{11} \pm \epsilon_{12}, \quad (258)$$

and correspondingly for c_{ik}

$$c_{11} = \pm c_{12}, \quad (259a)$$

$$c_{12} = c_{22}. \quad (259b)$$

Further, on account of the normalizing condition, we have

$$c_{11} = \frac{1}{\sqrt{2}}. \quad (259c)$$

Thus the desired solutions are

$$\psi_S = \frac{1}{\sqrt{2}}[\psi_{\epsilon_1}(x_1)\psi_{\epsilon_2}(x_2) + \psi_{\epsilon_2}(x_1)\psi_{\epsilon_1}(x_2)], \quad (260a)$$

$$\psi_A = \frac{1}{\sqrt{2}}[\psi_{\epsilon_1}(x_1)\psi_{\epsilon_2}(x_2) - \psi_{\epsilon_2}(x_1)\psi_{\epsilon_1}(x_2)]. \quad (260b)$$

(260a) is the symmetrical solution; (260b) the anti-symmetrical solution. In the case of the two-electron problem no other solutions exist.

This does not yet completely represent the data given by the helium atom. In helium two term-systems are known to occur, the ortho- and the para-system, of which the former is a triplet system, the latter a singlet system. The ground-state belongs to the *P*-system (para-). We must therefore refine our model a little further since so far we have been concerned only with two term-systems that absolutely do not combine, whereas the *O*- and the *P*-systems *do* combine with each other although only feebly. The feature that has been left out of account in our treatment is the electron spin. The two magnetic moments can be parallel as well as anti-parallel to each other. In exchanging two electrons we must take care to observe whether their magnets point in the same direction, as it is only in this case that the two electrons can be called fully equivalent and then their exchange signifies no change in the physical state.

We shall now correct our account by including electron spin. We shall call the term which is symmetrical in the co-ordinates of the electrons, except for the spin, the *P*-term, and that which is anti-symmetrical in the co-ordinates of the electrons, except for the spin, the *O*-term. We shall agree that the sign of the proper function that characterizes the spin is to change if the direction of spin changes. (This assumption is found to be verified when the conditions are investigated by the theory of groups.) In the case of the *P*-terms the electron magnets must be anti-parallel, in that of the *O*-terms they must be parallel. This assumes that the total proper function depending on co-ordinates and spin changes its sign when two electrons are exchanged.

When we speak of the "direction of the electron magnets," this has a meaning only if we distinguish some direction in space physically. For this purpose we imagine a magnetic field acting; in it the electron moments can assume the following values. Here m = total moment, m_i = individual moment :

$$\begin{array}{lll}
 m = m_1 + m_2, & & \\
 m_1 = \pm \frac{1}{2}, & m_2 = \pm \frac{1}{2}, & \\
 m = 1, & m_1 = m_2 = \frac{1}{2}, & (261\alpha) \\
 m = 0, & m_1 = \frac{1}{2}, & m_2 = -\frac{1}{2}, \quad (261\beta) \\
 m = 0, & m_1 = -\frac{1}{2}, & m_2 = \frac{1}{2}, \quad (261\gamma) \\
 m_1 = -1, & m_1 = -\frac{1}{2}, & m_2 = -\frac{1}{2}. \quad (261\delta)
 \end{array}$$

With every value of m there is associated, by (131), a definite value of the energy. But the energy value belonging to $m = 0$ may be realized in two ways, as our table shows; that is, it is degenerate.

Hitherto in discussing the directions of spin we have argued

as if there were no mutual action between the electrons. This must now be allowed for if we are to arrive at the correct final result. The mutual *electrical* action of the electrons causes no alteration in the above tabulation since it is independent of the direction of the moments. But there also exists a mutual *magnetic* action which *does* depend on the relative orientation. So we must now supplement our perturbation calculation, in which hitherto only Coulombian forces have been considered, with the forces due to spin, and we must take into account the circumstance mentioned above, namely, that the initial system is also degenerate with respect to the *orientation of spin*.

The mutual action of the spins is shown by calculation to do away with this degeneracy in the orientation just as the Coulombian forces did by resolving the terms into *P*- and *O*-terms. The two degenerate configurations denoted by (261β) and (261γ) resolve into a symmetrical and an anti-symmetrical configuration (cf. 259). On the whole we thus obtain *three* solutions that are symmetrical in the electron moments and *one* that is anti-symmetrical.

If we denote the solution that is symmetrical and anti-symmetrical in the electron *co-ordinates* by Φ_s and Φ_a , respectively, the three proper functions that are symmetrical in the moments by χ_{s1} , χ_{s2} , χ_{s3} , and the one which is anti-symmetrical in the moments by χ_a , the complete scheme of all the solutions runs

$$\Phi_s \cdot \chi_{s1}, \quad \Phi_s \cdot \chi_{s2}, \quad \Phi_s \cdot \chi_{s3}, \quad \Phi_a \cdot \chi_a, \quad (262a)$$

$$\Phi_s \cdot \chi_a, \quad \Phi_a \cdot \chi_{s1}, \quad \Phi_a \cdot \chi_{s2}, \quad \Phi_a \cdot \chi_{s3}. \quad (262b)$$

The system $(262a)$ is symmetrical in the electrons, the system $(262b)$ anti-symmetrical. Only the latter is realized in nature.

We now recapitulate the scheme of the conditions that hold in the case of the helium atom. First approximation: no mutual action between the electrons. Second approximation: the mutual electrical action is taken into account. Each term resolves into two terms; the associated proper function is symmetrical and anti-symmetrical, respectively, in the electron co-ordinates. Third approximation: the mutual magnetic action of the electrons is taken into consideration. Every term (already resolved previously into two terms) is now resolved into four times two terms, of which half are symmetrical in the electron co-ordinates, the other half being unsymmetrical, while six terms are symmetrical in the electron moments and two are anti-symmetrical. From these eight terms we obtain the complete anti-symmetrical system containing four series of terms by compounding the solutions that are symmetrical in the

co-ordinates with the anti-parallel magnets and conversely. In accordance with Pauli's Principle the other four symmetrical solutions are not realized in nature.

This scheme of the terms of the helium atom already contains all the essential features of the helium spectrum: there are two term-systems which combine feebly with each other (only in consequence of the mutual magnetic action); the one system is simple (*P*-terms) and the ground-state belongs to it. This is immediately obvious since in the ground-state the magnets must be anti-parallel when the electrons are in the same "orbits."

This accounts for all the qualitative features of the helium spectrum and shows even more than is given by observation. For originally the ortho-helium spectrum was considered to be a doublet spectrum and this contradicted Rydberg's Law of Alternation. Only recent observations have shown that this spectrum is a blurred triplet. This view that it is a triplet is supported by intensity measurements which gives an anomalous value, from the point of view of the rules of Burges and Dorgelo, for the intensity ratio in the case of a doublet but the correct value for blurred triplets. The dia-magnetism of helium is also a consequence of these qualitative assumptions. For in the ground-state the orbital impulse $s = 0$ and the magnets are anti-parallel, so that there is no magnetic moment.

There are now two quantitative problems to be solved: the lowest term which is equal to the ionization potential must be calculated, and the deviation of the terms from the hydrogen terms for excited states must be determined. The former problem was first solved satisfactorily by Kellner,⁹⁷ who used a method of approximation due to Ritz. The excited terms were calculated by means of the theory of perturbations by Heisenberg himself, and satisfactory agreement was obtained at any rate for the terms having greatest orbital momentum. Kellner's value for the ionization potential came very near the observed value. More recent investigations by Hylleraas have improved this agreement to within a few thousandths.

53. Theory of Band-Spectra and Nuclear Spin

The theoretical treatment of molecules consisting of like atoms has led to important disclosures concerning nuclear spin and the statistics that hold for protons (cf. § 64). This is most easily seen by discussing the following simplified example. Let us consider a molecule consisting of two like atoms (for example, H_2 , N_2 , O_2). To a first approximation we can resolve the proper functions into factors that correspond with the individual degrees

of freedom of the molecules. These degrees of freedom are : (1) The motion of the centre of gravity of the molecule ; this motion is certainly independent of the internal motions of nuclei and electrons, and its proper function is certainly symmetrical in the co-ordinates of both nuclei and hence is no longer of importance for the subsequent discussion. (2) The vibrations of the nuclei relative to each other, which express themselves in the so-called molecular vibration spectra. The co-ordinates of both nuclei also occur quite symmetrically here, so that we may exclude this proper function just like that of translation in judging the character of the symmetry. (3) The motion of electrons within the molecule. There is a paper by Hund^{92a} on this point ; he has shown that the associated proper function is symmetrical or anti-symmetrical in the co-ordinates of the nuclei according to the character of the motion ; in the ground-state it is always symmetrical. (4) The rotation of the molecules about an axis through their centre of gravity, which expresses itself in the rotation spectrum. If the angle about this axis is called ϕ and the angular momentum $p_\phi = \frac{nh}{2\pi}$, then the corresponding proper function is proportional to $e^{im\phi}$, cf. (165) and (165a).

If ϕ increases by the amount π this denotes that the nuclei are exchanged. In this exchange the proper function changes its sign or remains unchanged according as n is odd or even. Thus we can draw up the following scheme of the symmetrical and the anti-symmetrical solutions of a molecule : at the same time we again call attention to the fact that *the symmetry is being considered with reference to the nuclei and not with reference to the electron co-ordinates* :

Anti-symmetrical System.	Symmetrical System.
Electron proper function, symmetrical, $p = \frac{h}{2\pi}(1, 3, 5, 7 \dots)$	Electron proper function, symmetrical, $p = (0, 2, 4 \dots) \frac{h}{2\pi}$
Electron proper function, anti- symmetrical, $p = \frac{h}{2\pi}(0, 2, 4 \dots)$	Electron proper function, anti- symmetrical, $p = (1, 3, 5, 7 \dots) \frac{h}{2\pi}$

We cannot yet specify which of these two systems, which do *not*, as we know, combine with each other, is realized in nature. On the other hand, we *do* know the following from the above scheme. Suppose only the anti-symmetrical system is realized. Then on account of the selection principle for the rotator (cf. § 41) a change in p can occur only if an electron transition occurs simultaneously, that is, a change of the electron term from a form symmetrical in the nuclear co-ordinates to one that is

anti-symmetrical, and vice versa. Let us consider as an illustration the transition from a symmetrical to an anti-symmetrical electron term. Then transitions only, say, from 1 to 2, 3 to 4, and so forth are possible but not those from 2 to 3, 4 to 5, and so forth; every alternate line which we should ordinarily expect according to the selection principle drops out. This has actually been observed in the case of oxygen bands. The argument is not affected by the question as to whether the symmetrical or the anti-symmetrical term-system is realized.

This theoretical view has recently been confirmed by observations of the absorption-spectrum of oxygen.* All these considerations depend essentially on the complete similarity of both nuclei. If we imagine a nucleus replaced by an isotope, all the arguments based on exchange become void and *all* the spectral lines must appear, not merely alternate lines. This reflection has been confirmed by extremely precise observations of the absorption spectrum of oxygen in which nothing was previously known of the existence of isotopes. If we consider only the strong lines, every second line is found missing, quite in agreement with the above scheme. Between the strong lines, however, that occur we observe very weak absorption lines which apparently correspond to forbidden transitions. But these absorption lines do not occur at the place which is allocated to oxygen of molecular weight 32 and the corresponding moment of momentum but are slightly displaced as if oxygen of molecular weight 33 and 34 were present. This obviously points to the existence of oxygen isotopes to which such selection rules no longer apply. The proportion in which these isotopes are mixed is of the order of magnitude 1 per cent., so that it has been able to elude observers who determined the atomic weight chemically.

Besides this type of bands, in which alternate lines are missing, other bands also occur in which all lines are present but show varying intensity. Every second line is weaker than its predecessor. This fact may be interpreted very elegantly if we again ascribe a spin to both nuclei, just as we have done in the case of the helium electrons. We must then again consider the various possibilities that arise through the adjustment of the nuclear moments. The reasoning is completely analogous to that for the helium electrons. We again have three cases of parallel and one of anti-parallel moments. We can again complete our system by forming the following scheme :

* Glaue and Johnston, *Journ. Amer. Chem. Soc.*, **51**, 1436 (1929).

Anti-symmetrical System.

1. Electron proper function, symmetrical,

$$(a) \psi = \frac{h}{2\pi}(0, 2, 4 \dots)$$

(nuclear magnets anti-symmetrical).

$$(b) \psi = \frac{h}{2\pi}(1, 3, 5 \dots)$$

(nuclear magnets symmetrical).

2. Electron proper function, anti-symmetrical,

$$(a) \psi = \frac{h}{2\pi}(0, 2, 4 \dots)$$

(nuclear magnets symmetrical).

$$(b) \psi = \frac{h}{2\pi}(1, 3, 5 \dots)$$

(nuclear magnets anti-symmetrical).

Symmetrical System.

1. Electron proper function, symmetrical,

$$(a) \psi = \frac{h}{2\pi}(0, 2, 4 \dots)$$

(nuclear magnets symmetrical).

$$(b) \psi = \frac{h}{2\pi}(1, 3, 5 \dots)$$

(nuclear magnets anti-symmetrical).

2. Electron proper function, anti-symmetrical,

$$(a) \psi = \frac{h}{2\pi}(0, 2, 4 \dots)$$

(nuclear magnets anti-symmetrical).

$$(b) \psi = \frac{h}{2\pi}(1, 3, 5 \dots)$$

(nuclear magnets symmetrical).

It is quite immaterial whether the symmetrical or the anti-symmetrical system is realized in nature: at any rate in both schemes the group with parallel nuclear moments occurs three times as frequently as those with anti-parallel nuclear moments.

The variation of intensity is directly due to this. It is again thanks to the nuclear moments that every line can occur but that every second line occurs three times as often as its predecessor. The argument is completely analogous to that in the case of helium, in which three ortho-terms and only one para-term occurs. To find which term-system occurs in nature we should have to ascertain for the electron ground-state whether the *a priori* frequency of occurrence of the even or the odd rotation numbers is the greater. Spectral observation has disclosed that in the case of hydrogen the alternation of intensity is in the ratio 1 : 3.

The question as to the term-system realized in nature, which is also the *anti-symmetrical system for protons*, was first answered by observations of the specific heat of hydrogen which we shall discuss fully below (cf. § 64).

54. The Calculation of the Hydrogen Molecule

The old quantum theory made no useful contribution to the problem of the formation of molecules. This is easily understood if we recall that even in the simpler example of the helium atom it met with no success. The new quantum mechanics at first, indeed, seemed in a certain respect to lessen the prospect of explaining atomic attraction. As we know, neutral atoms attract each other if they possess dipoles or higher poles. If they are completely spherically symmetrical they exert no attractive forces on each other. Now the hydrogen atom in its

ground-state is spherically symmetrical. This spherical symmetry is not, indeed, preserved when two atoms approach each other, for then a perturbation effect presents itself. From elementary electrostatic considerations unperturbed atoms would not act on each other at all.

Nevertheless Heitler and London had considerable success in quantitatively interpreting the formation of a hydrogen molecule, without assuming mutual polarization, on the basis of the reciprocal action of unperturbed atoms. This is owing to the circumstance that we are here again dealing with a degenerate system in which new possibilities are created by the particle degeneracy.

What has been said may be expressed quantitatively as follows: Schrödinger's equation for a hydrogen molecule with immovable nuclei separated by a distance a_0 is

$$\Delta_1\psi + \Delta_2\psi + \frac{8\pi^2\mu}{h^2}\left(E - \frac{e^2}{r_{ab}} - \frac{e^2}{r_{12}} + \frac{e^2}{r_{a1}} + \frac{e^2}{r_{a2}} + \frac{e^2}{r_{b1}} + \frac{e^2}{r_{b2}}\right)\psi = 0. \quad (263)$$

Here a, b denote the two nuclei; 1, 2 refer to the electrons, and ra_1 is the distance between the nucleus a and the electron 1. This problem is treated to a first degree of approximation by means of perturbation theory. We assume as the foundation of our solution that the electron 1 moves round the nucleus a , the electron 2 about the nucleus b , any mutual action between the two atoms being disregarded for the present.

We then take as our perturbation energy the influence of the nucleus a and the electron 1 on the nucleus b and the electron 2, and vice versa. Now the initial state under consideration is degenerate. For the electron 2 might equally well be attached to the nucleus a and the electron 1 to the nucleus b . This degeneracy gives two proper functions for the ground-state, and if the mutual action is taken into account the proper values then split up and we again obtain a symmetrical and an anti-symmetrical solution. Only the anti-symmetrical solution gives a negative value and gives in fact a definite minimum for the energy for $a_0 = 0.75 \cdot 10^{-8}$ cm.

The anti-symmetrical solution always has an energy that lies above zero, so that for it we find no combination of atoms into a molecule. But since by Pauli's Principle the definitive solution must be anti-symmetrical in the co-ordinates, we get for the solution of the molecule problem the symmetrical solution with anti-parallel electron magnets.

According to the calculations of Heitler and London and

later work by Wang the quantitative values obtained for the work of dissociation and the nuclear distance are

$$D = 3.8 \text{ eV}, \quad \alpha_0 = 0.75 \cdot 10^{-8} \text{ cm.}$$

whereas the observed values are

$$D = 4.3 \text{ eV}, \quad \alpha_0 = 0.76 \cdot 10^{-8} \text{ cm.}$$

Following on this, Heitler and London have used analogous methods to investigate the problem of the formation of the helium molecule. In this case, too, the symmetrical solution would lead to molecule formation; but now we no longer pay due regard to Pauli's Principle by imagining the electron magnets to be in anti-parallel adjustment. Even in the case of the helium atoms we have to do with two electrons and anti-parallel spin and so have no further possibility of assuming two other electrons in this same state. Pauli's Principle thus yields the observed fact that the helium atom in the ground-state is unable to form a molecule.

This reason, however, based on exclusion, no longer holds if the helium atoms are excited. The electrons are then in different quantum states, and in principle it becomes possible to bring about molecule formation by suitably disposing of the position of the electron magnets. Molecules of excited helium have actually been observed.

55. The Prospects Opened up by the Theory of Groups and Chemistry

Right at the beginning we emphasized the reasons for not dealing in our account with the treatment of problems of the theory of groups. Nevertheless we find it desirable at this stage at least to indicate roughly what far-reaching importance the theory of groups has for the questions just outlined, which occur in many-body mechanics. In all problems of particle degeneracy we are always concerned with the multiplicity of terms which is produced by the complete similarity of the elementary particles. These degenerate terms can be split up by means of external fields or mutual action. Since these perturbing forces also always have certain elements of symmetry which are in turn conditioned by the equality of all elementary particles, it is easily understood that here that branch of mathematical theory is destined to play a decisive part which deals with symmetry when similar elements are exchanged and with the ordering of such exchanges into definitely distinguished classes. In mathematical language we are in all cases dealing

with the theory of an equation whose order is equal to the degree of degeneracy of the term in question.

The degeneracy phenomena just treated then become physically observable both in the case of multiplet spectra and their changes in the electric and magnetic field, as well as in the important energy question of the resolution of energy-states under the mutual action of atoms, that is, chemical problems. In discussing the problem of the helium model we have already seen that molecule-formation in the ground-state is rendered impossible by Pauli's Principle. London¹⁰² succeeded in showing quite generally that the occurrence of homopolar bonds depends on whether the two atoms can place together one or more electrons whose magnetic moments can adjust themselves to the anti-parallel position without transgressing Pauli's Principle. For example, chlorine possesses seven electrons in the 3-quantum state. It is able to enter into a homopolar bond either by retaining one electron as a valency electron, the six others placing themselves anti-parallel in pairs, or by having three electrons as valency electrons and two pairs, or five valency electrons and one pair, or seven valency electrons and no pair. Inert gases with a complete shell in which the magnets are anti-parallel in pairs exhibit no homopolar valency in the ground-state.

In view of the considerable mathematical difficulties involved we cannot enter further into these questions here.

56. The Auger Effect

An interesting phenomenon that often occurs in conjunction with the ordinary photo-electric effect is the so-called Auger effect. This refers to the observation that in a photo-electric process, besides the electrons whose energy satisfy the relation (240), others also often occur with the energy

$$E = P_1 - 2P_2. \quad (264)$$

This effect is to be interpreted as follows. Using the language of the light-quantum theory let us suppose a light-quantum of energy $h\nu$ removes from the K -shell an electron for which the relation (240) holds. The ionized atom then has a gap in the K -shell which is filled by the transition of an electron from a state of higher energy (L -, M -shell and so forth) to the K -shell and so forth until the atom is fully reorganized again. In these transitions the difference of energy can be emitted either in the form of radiation, and then Röntgen rays of frequency

$$h\nu' = P_1 - P_2. \quad (265)$$

should be observable in this process; or else—and this is the Auger effect—it may happen that the energy $P_1 - P_2$ that is set free is not emitted as radiation but is transferred to another electron, say one from the L -shell, which then escapes with the energy

$$E = P_1 - 2P_2.$$

In the language of quantum mechanics we are in this case concerned with an intensity question: according as to whether the probability of an emission of energy or of a mechanical process (transference of energy) is greater an emission of Röntgen ray frequency or an Auger effect takes place. It is of course possible in principle to calculate these two probabilities, and this has been done approximatively by Wentzel⁶⁰ and Fues.

We may express the position in a language already familiar to us as follows: Between the proper functions of the initial-state, which lie in the discrete spectrum, and those of the end-state, which partly lie in the continuous spectrum, resonance occurs, in virtue of (264) which gives rise to transitions.

57. Preliminary Remark about Classical Statistics

To work out the correct formulations of the new quantum statistics correctly we shall find it useful briefly to recall the principles of the statistical theory of heat which has hitherto prevailed. In the applications to be considered we deal essentially with systems that consist of many individual particles, of which each may assume the terms allowed by the quantum theory independently of the state of all the other particles. Classical statistics and the statistics of the Bohr quantum theory—which produced no change in the statistical method—led to the following rule for calculating mean values. If we number the term values and the atoms consecutively a state is characterized by specifying the atoms by means of their numbers, which denote that they have the first, second . . . term. The probability of a distribution of which we know only that n_1 atoms have the first term, n_2 the second term, and so forth, is then given by the well-known expression due to Boltzmann:

$$W = \frac{N!}{n_1! n_2! \dots n_k!}, \quad (\sum n_i = N) \quad (266)$$

This expression is obtained as follows. The value of the numerator is the number of all possible permutations of the numbers

of atoms that have the different terms. But we must divide this number by all those permutations in which two particles having the same term-values are exchanged. This explains the presence of the individual factors of the denominator. The statistical problem consisted in finding that distribution of the n_i 's for which the expression (266) becomes a maximum under the following conditions :

$$\begin{aligned}\Sigma n_i &= N, & (267a) \\ \Sigma n_i \epsilon_i &= E. & (267b)\end{aligned}$$

The first condition states that the total number of particles must remain constant, the second that the total energy remains constant during every permutation. The most probable distribution that results is then

$$n_i = N \frac{e^{-\frac{\epsilon_i}{kT}}}{\sum_l e^{-\frac{\epsilon_l}{kT}}} \quad \begin{matrix} (k = \text{Boltzmann's constant,} \\ T = \text{absolute temperature}). \end{matrix} \quad (268)$$

By means of Boltzmann's principle, which connects probability with entropy, or by an equivalent method, expressions for the characteristic thermodynamic functions are then obtained. The free energy comes out as

$$F = -kT \log \sum_l e^{-\frac{\epsilon_l}{kT}} \quad (269)$$

and the following well-known thermodynamic relations then hold :

$$\frac{\partial F}{\partial T} = -S \quad (S = \text{entropy}) \quad (270)$$

$$F = E + T \frac{\partial F}{\partial T}. \quad (271)$$

The essential feature, then, of the hitherto prevailing statistics is: a state is uniquely defined if we specify for *each individual atom* which term it possesses.*

In the whole of this discussion we have assumed that every term-value can be realized in only one way, that is, we have assumed that degenerate states do not occur. If a term-value can be realized in various ways, for example, the l th term in g_l ways, the formulæ (266), (268), (269) must be supplemented

* In principle this definition agrees with that given for micro-state in some text-books.

by factors, so-called *weight-factors*, which denote the number of possible ways of realizing a state. We then obtain, for example,

$$N_i = \frac{N g_i e^{-\frac{\epsilon_i}{kT}}}{\sum_i g_i e^{-\frac{\epsilon_i}{kT}}}, \quad (268a)$$

$$F = -kT \log \sum_i g_i e^{-\frac{\epsilon_i}{kT}}. \quad (269)$$

58. The New Statistics

The preceding remarks about classical statistics are now taken over in an appropriate form into quantum statistics but become definitely modified for two reasons. The individual particle still has its state characterized by its single proper value or its corresponding proper function, respectively. But the state of the total system is now no longer given when we know the proper value of every individual particle. This would be so if the proper values were single in the whole system; but they are not, on account of the particle-degeneracy discussed earlier. Let us consider, for example, a system of two particles of which we suppose each to have, say, three different terms (proper functions), mutual action being left out of account. We shall consecutively number the terms 1, 2, 3 and denote the particles by means of the indices a and b . We then assume in the sequel that the proper values of the particles are not appreciably altered by the presence of the second particle.

If we had only to deal with classical statistics there would now be nine different states which we may characterize as follows. The particle a can assume any of the states 1 to 3 and, independently of it, the second particle may do the same. The nine proper functions that we should then obtain if we counted classically, would be,

$$\left. \begin{array}{lll} \psi_a(1) \psi_b(1) & \psi_a(1) \psi_b(2) & \psi_a(1) \psi_b(3) \\ \psi_a(2) \psi_b(1) & \psi_a(2) \psi_b(2) & \psi_a(2) \psi_b(3) \\ \psi_a(3) \psi_b(1) & \psi_a(3) \psi_b(2) & \psi_a(3) \psi_b(3) \end{array} \right\}. \quad (270)$$

But by the theory of many-body problems we find that in quantum mechanics the totality of the solutions here written down is not realized but only a term-system which does not combine with the others. In our case we have, as in that of helium, only two non-combining term-systems, namely, the

symmetrical and the anti-symmetrical system. The symmetrical system has six different states : *

$$\left. \begin{aligned} &\psi_a(1)\psi_b(1), \quad \psi_a(1)\psi_b(2) + \psi_a(2)\psi_b(1), \quad \psi_a(1)\psi_b(3) \\ &\quad \quad \quad + \psi_a(3)\psi_b(1) \\ &\psi_a(2)\psi_b(2), \quad \psi_a(2)\psi_b(3) + \psi_a(3)\psi_b(2), \quad \psi_a(3)\psi_b(3) \end{aligned} \right\} \quad (271)$$

whereas the anti-symmetrical system has only three :

$$\left. \begin{aligned} &\psi_a(1)\psi_b(2) - \psi_a(2)\psi_b(1), \quad \psi_a(1)\psi_b(3) - \psi_a(3)\psi_b(1), \\ &\quad \quad \quad \psi_a(2)\psi_b(3) - \psi_a(3)\psi_b(2) \end{aligned} \right\} \quad (272)$$

So we see that in quantum mechanics all the different states are not possible but only a smaller manifold selected from this total number, the proportion being dependent on the choice of the term-system. *Thus a state is characterized in quantum mechanics by its proper function, which refers to all elements of the system.* Exchanging two particles can at the most effect a change in the proper function. For the sequel we shall adopt the convention that proper functions that differ only in sign are not to be counted as different. This is reasonable physically ; for since all physical statements (cf. § 38) are determined by the square of the proper function, two proper functions that differ only in sign give rise to no observable differences.

The remarks just made essentially prescribe the quantum-statistical treatment. Let us recapitulate : on reasonable grounds we regard as different states of a quantum-mechanical system only such whose proper functions are different (apart from their sign). All such states are regarded as being equally probable. The calculation of statistical mean values is thus reduced to counting up all the proper functions which the quantum-mechanical system possesses. But we must know at the outset which statistics, that is, which choice of the non-combining system (symmetrical, anti-symmetrical) is valid for the particle in question.

59. Statistical Methods of Calculation. The Statistics of Bose and Fermi

The counting up of all the proper functions to solve the statistical problem, as mentioned in the previous section, is a quite correct procedure but is inconvenient for practical purposes. In the case of statistical problems of mechanical systems we are dealing with an enormous number of elementary particles. For example, 2 grms. of H_2 contain 6.10^{23} molecules. Each

* We disregard constant normalizing factors.

proper function which defines a state depends on $3N$ variables so that a direct calculation can hardly be carried out. For purposes of calculation it is more convenient to give formulæ that refer exclusively to mechanical properties of the individual particle.

We revert again to our earlier example of two particles, of which each can assume three states, and we assume, to begin with, that in nature only the symmetrical term-system is realized. The above-mentioned difference between the classical and the new statistics, namely, that in the former, to determine the state fully it is necessary to know the proper function of the individual particle, whereas in quantum mechanics it is necessary to know only the proper function of the total system, suggests to us to enquire whether by changing the statistics it is possible to arrive at a simple method of counting up the proper functions of the total system. We therefore no longer enquire in the sequel: what is the state of the individual atom 1, 2, 3 . . ., but rather, how many of the g_i -states that occur contain n_i -atoms? This reversed procedure is reasonable, for in it we have already taken into account the essential feature of quantum mechanics. It is a matter of indifference to us whether the atom a is in the first state and the atom b in the second state or vice versa. These apparently different states are no longer different from the point of view of quantum mechanics, for the proper function that characterizes the state is symmetrical in the particle co-ordinates and consequently remains unchanged when the particles a and b are exchanged.

If we carry this process out in our particular case the position is as follows. We enquire: what is the probability that both atoms are in one and the same state? This probability, which is obtained by counting up the proper functions (271), is exactly $\frac{1}{2}$, whereas the probability that both atoms are in different states is also $\frac{1}{2}$. According to Boltzmann's method of counting [cf. (270)] the probability that the atoms are in one and the same state is only $\frac{1}{3}$.

The argument is fully analogous in the case of anti-symmetrical proper functions. In this case again we enquire: how great is the number of the states which are occupied by no atom at all or by only one atom? One completely specified state can never be occupied by more than one atom. In our earlier example we saw that of the three possible states two are always occupied. The probability that all atoms are in one state is zero.

This simple method of counting for three possible states and two atoms may now be formally extended very simply. Let

the system in question have the energy-levels $\epsilon_1, \epsilon_2, \dots, \epsilon_k$, of which the l th is g_l -times degenerate. We must again emphasize that these data concerning terms and their multiplicity refer to the *individual atom* and not to the total system. We now assert that we can obtain the possible proper functions of the total system to their correct degree of multiplicity by answering the following statistical questions which deal only with terms of the individual atom.

First let the whole system that is realized be symmetrical. We enquire which is the most probable distribution n_l of the atoms under the subsidiary conditions (267a, 267b). The answer now differs from (266) thus:

Let us determine the maximum of the product

$$\overline{|Z|} \frac{g_l!}{g_l^{(0)}! g_l^{(1)}! g_l^{(2)}! \dots} \quad (g_l = \text{the weight factor}), \quad (273)$$

$$(0 \leq g_l^{(k)} \leq g_l; \quad g_l^k = 0, 1, 2, \dots) \quad \sum_i g_l^{(i)} = g_l,$$

with the subsidiary conditions

$$\sum_i \sum_l g_l^{(i)} = N, \quad (274)$$

$$\sum_i \sum_l g_l^{(i)} \epsilon_l = E, \quad (275)$$

where $g_l^{(i)}$ denotes the number of states with the energy ϵ_l which i atoms contain. Then the most probable values of n_l are given by

$$n_l = \sum_i g_l^{(i)}. \quad (276)$$

If n_l is known we easily arrive at the thermodynamic functions [cf. (269)].

To prove our statement let us first consider only exchanges which concern the ϵ_l term. If we consider two atoms a and b exchanged, as was done, for example, in dealing with Boltzmann's statistics, then no new proper functions of the total system result, for it is symmetrical in all atoms. But we do obtain a new proper function if we select two out of the g_l states with energy ϵ_l and exchange their occupation numbers (*Besetzungszahlen*). This exchange leaves the conditions (274) and (275) intact. In the scheme (271) it corresponds to the exchange of two proper functions. We can effect only $g_l!$ of these exchanges; this gives the numerator of one member of (273). But if the two states with energy ϵ_l contained the *same number* of atoms, their exchange leads to nothing new. So we must subtract all those possibilities that correspond to this exchange. Now there are $g_l^{(i)}$ states with energy ϵ_l , which i atoms contain; an exchange

of two such states may not therefore be counted as a new possibility, we must divide by $g_i^i!$. This holds for every i and so gives the denominator of (273).

The line of reasoning is quite similar but a little simpler for the anti-symmetrical system of proper functions. These considerations hold literally, except that g_i^i is different from zero only for g_i^0 and g_i^1 . Since no more than one atom can exist in one state this limitation is obvious. So we obtain in place of (273)

$$\overline{|l|} \frac{g_i!}{g_i^0! g_i^1} = \overline{|l|} \frac{g_i!}{g_i^0! (g_i - g_i^0)!} = \overline{|l|} \frac{g_i!}{(g_i - g_i^0)! g_i^1} \quad (277)$$

and in place of (274), (275), (276) we have

$$\sum_i g_i^I = N, \quad (278)$$

$$\sum_i \epsilon_i g_i^I = E, \quad (279)$$

$$n_i = g_i^I. \quad (280)$$

Let us recapitulate the meaning of the new statistical methods of calculation. Just as in Boltzmann's statistics we are concerned only with the multiplicity of the terms of the individual atom. But whereas in the former case a difference arose according as to whether the atom a or b had a certain term, it is immaterial in the present case. The only matter of importance is how many states of definite energy occur (g_i) and how many of them contain no atom or several atoms (g_i^0, g_i^I).

These new methods of statistical calculation are designated by the names of their protagonists, Bose⁸⁵ (symmetrical proper functions), and Fermi⁸¹ (anti-symmetrical proper functions). They were evolved before quantum mechanics came into being. We shall return below (cf. § 60) to the line of argument that led to this new type of statistics.

60. Theory of the Ideal Gas

In quantum mechanics as in classical theory we take the ideal gas to mean a system of particles which have no appreciable potential energy with respect to each other. The difference, as compared with classical theory, of course now consists in the fact that even in the case of those systems in which the influence of the potential energy on the proper values present is very small, the conditions of symmetry of the total system must not be lost sight of. We carry out the calculations, according to the formulæ given in the previous section, both for symmetrical as well as for

anti-symmetrical proper functions, and shall consider special cases only when we get to the final result.

The wave-equation of a system which consists of N particles which are particularly independent of one another then runs

$$\sum_1^N i\Delta_i \psi + \frac{8\pi^2\mu}{h^2} E\psi = 0. \quad (278)$$

We can separate this equation into n equations of which each one refers only to a single particle,

$$\Delta_i \psi_i + \frac{8\pi^2\mu}{h^2} \epsilon_i \psi_i = 0, \quad (278a)$$

and the sum of the proper values satisfy the condition

$$\sum_i \epsilon_i = E. \quad (278b)$$

Let these N particles be enclosed in a vessel, whose sides are given by

$$0 \leq x \leq a, \quad 0 \leq y \leq b, \quad 0 \leq z \leq c.$$

We write down as the condition for this enclosing of the particles

$$\left. \begin{aligned} \psi_i(0, y, z) &= \psi_i(x, 0, z) = \psi_i(x, y, 0) = \\ \psi_i(a, y, z) &= \psi_i(x, b, z) = \psi_i(x, y, c) = 0 \end{aligned} \right\}. \quad (279)$$

These conditions (279) make it certain that at the boundary both the density and the current vanish (cf. § 38).

By separating the variables x, y, z we obtain as the solution for one particle

$$\psi_k = A \sin k_1 x \sin k_2 y \sin k_3 z. \quad (280)$$

The conditions (279) then give for k_1, k_2, k_3 the relations

$$k_1 = \frac{n_1 \pi}{a}, \quad (281a)$$

$$k_2 = \frac{n_2 \pi}{b} \quad (n_1, n_2, n_3 = 0, 1, 2), \quad (281b)$$

$$k_3 = \frac{n_3 \pi}{c}, \quad (281c)$$

and for ϵ_i , by (278a),

$$\epsilon_{n_1, n_2, n_3} = \frac{h^2}{8\mu} \left(\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} + \frac{n_3^2}{c^2} \right). \quad (282)$$

Thus according to (282) only discrete energy-values are allowable for each particle. But the possible proper values lie so close together that we can represent the number of proper values in a definite interval to a sufficient degree of approximation by means of a continuous function. This continuous function tells us how many proper values lie in the interval ϵ , $\epsilon + d\epsilon$. The procedure is identical with the calculation of the proper vibrations of a cavity or a crystal * in the old quantum theory. In this way we obtain

$$g(\epsilon)d\epsilon = \frac{2\pi abc}{h^3}(2m)^{\frac{3}{2}}\epsilon^{\frac{1}{2}}d\epsilon = \frac{2\pi V}{h^3}(2m)^{\frac{3}{2}}\epsilon^{\frac{1}{2}}d\epsilon. \quad (283)$$

By making use of this value we can now evaluate the formulæ (273) to (276) and (277) to (280). The required most probable distribution of the g_s 's and n_s 's results immediately by the ordinary method of calculating the maximum of (277) and (273), respectively. The calculation runs parallel in the case of Boltzmann, Bose and Fermi. We obtain as our result :

$$\text{Bose :} \quad n(\epsilon) = \frac{g(\epsilon)}{e^{\alpha + \frac{\epsilon}{kT}} - 1}, \quad (284)$$

$$\text{Fermi :} \quad n(\epsilon) = \frac{g(\epsilon)}{e^{\alpha + \frac{\epsilon}{kT}} + 1}. \quad (285)$$

The constant α is determined, as above, by the condition

$$\Sigma n_s = \int n(\epsilon)d\epsilon = N.$$

It is instructive to compare (268), (284) and (285) with one another. For this purpose we write (268a) in the form

$$n_i = g_i e^{\alpha + \frac{\epsilon_i}{kT}}, \quad (268b)$$

where α is determined by (267a). We see at once that if $e^{\alpha + \frac{\epsilon}{kT}} \gg \gg 1$ both the new types of statistics pass over into that of Boltzmann. The statistics of Bose and Fermi differ only in the sign of 1 in the denominator. The meaning of α is, just as in classical theory, equal to $(+kT)$ times the free energy. The value of α determines the amount of the deviation from classical behaviour, the so-called degeneracy.

We now discuss the case of the symmetrical proper function. This is of importance for the case of light-quanta. Following Einstein we imagine the radiation to be composed of atoms of

* See, for example, F. Reiche, "The Quantum Theory," pp. 31 and 38.

light for which a system of statistics holds which was given in a paper by Bose published before the advent of quantum mechanics; this system of statistics is now recognized as that of symmetrical proper functions. By Einstein's fundamental hypothesis of light-quanta we must then set

$$\epsilon = h\nu.$$

The subsidiary condition (274) does not occur in this case ($\alpha = 0$). For it is quite possible for a light-quantum, say of energy $h\nu$, to be split up into two light-quanta of energy $\frac{h\nu}{2}$.

No limit is set to the number of particles in the case of light-quanta. Only the expression for $g(\epsilon)$ is to be corrected here, as we are in the present instance not calculating with the classical but with the relativistic wave-equation and we must take into account the different possibilities of polarization of the light. The expression of (283) is then replaced by

$$g(\epsilon) = g(h\nu) = \frac{8\pi V \nu^2}{c^3}. \quad (283a)$$

In this way we get for the number of light-quanta of frequency ν in the volume V

$$n(\nu)d\nu = \frac{8\pi \nu^2 V d\nu}{\frac{h\nu}{c^3(e^{h\nu/kT} - 1)}}. \quad (286)$$

If we multiply the number of light-quanta by their individual amounts of energy $h\nu$, we obtain for the energy-density

$$\rho(\nu) = \frac{8\pi h \nu^3}{c^3} \frac{1}{\frac{h\nu}{e^{h\nu/kT} - 1}} \quad (287)$$

in exact agreement with Planck's radiation law, which was initially obtained from totally different considerations.

The reason why the statistics of symmetrical proper functions (Bose's statistics) holds for light-quanta is not given *a priori*; rather our method of calculation is to be considered as being justified by its success.

In the development of the quantum theory the study of fluctuations of radiation has played an important part. This was understood as signifying the theoretical calculation of the deviations of the energy-distribution from its mean value, which was itself determined by Planck's law. These calculations, calculated thermodynamically on the basis of Planck's law, gave rise to considerable difficulty as they could be understood

from no known model. If the radiation was considered to be composed of waves, such as may be demanded by the classical theory of light, only one part of the fluctuation phenomena could be understood, whereas the other part remained unintelligible. If, conversely, the radiation was regarded, so to speak, as a gas consisting of light-atoms, a view which was favoured by a great number of rather recent experimental researches, such as the above described photo-electric effect, Compton effect and so forth, then only the above unexplained part of the fluctuation formula could be obtained; at that time, of course, owing to the state of the theory the statistical calculation was based on Boltzmann's formula. But if we calculate the fluctuations according to Bose's statistics then we arrive at the law of fluctuation of radiation which had been derived by purely thermodynamic reasoning on the basis of Planck's law of radiation. So that in this point, which cannot be tested experimentally, but is of great theoretical importance, the more recent view seems to show marked superiority.

The application of Bose's statistics to material systems has so far been unsuccessful in practice. This is due to the fact that e^α is great compared with unity for all the gases, temperatures and pressures that come into question, so that the Maxwellian distribution is valid. If we evaluate the subsidiary condition (274) we get the following approximate value for e^α :

$$e^\alpha = \frac{V}{N h^3} (2\pi m k T)^{\frac{3}{2}}. \quad (288)$$

Inserting the numerical values we find * that even for the gases of lowest molecular weight (excepting electrons) the value of (288) becomes comparable with unity only in the vicinity of the zero point and at so high a pressure that it is already far from possible to neglect the intra-molecular forces (van der Waals).

So Maxwell's distribution is valid for material gases to a high degree of approximation even when, theoretically, Bose's statistics are correct.

We now consider the special case of anti-symmetrical proper functions which is important for electrons and for which (285) holds. We again see that for $e^\alpha \gg 1$ (285) becomes transformed into the Maxwellian distribution. This is equivalent to

$$\frac{V}{N h^3} (2\pi m k T)^{\frac{3}{2}} \gg 1. \quad (288a)$$

* For He the value of e^α at ordinary temperatures and at atmospheric pressure is of the order 10^6 .

As in the case of Bose's statistics here, too, classical statistics hold to a high degree of approximation for all gases *except for free electrons*.

Those cases for which the condition (288a) is no longer valid are called degenerate cases, and we speak of a slightly degenerate gas as one which is only approximately and not completely described by the classical state, and of complete degeneracy when

$$e^\alpha \sim 1 \quad \text{or} \quad e^\alpha < 1,$$

that is, when classical statistics is not even approximately correct.

A remarkable property of the statistics of anti-symmetrical proper functions, which are also called Fermi statistics after their originator, consists in the occurrence of a *zero-point energy* and a *zero-point pressure*; that is, even at the absolute zero the Fermi gas still possesses kinetic energy. This can be seen qualitatively without any calculation from the fundamental principles of Fermi's statistics. The number of proper functions

below the energy ϵ_0 is, by (283), proportional to $\int_0^{\epsilon_0} \epsilon^{\frac{1}{2}} d\epsilon$. This

is certainly a finite number. Since not more than one atom can belong to any one proper function the energy of the whole gas must at least be so great that the number of proper functions below this energy is greater than or equal to the number of atoms present. An analogous argument can be applied to the momentum. This is represented quantitatively as follows.

The lowest amount of energy that a gas can have is determined by having the number of proper functions present, starting from the energy 0 of an individual atom up to a limiting energy ϵ_0 of an atom

$$\frac{2\pi V}{h^3} (2m)^{\frac{3}{2}} \int_0^{\epsilon_0} \epsilon^{\frac{1}{2}} d\epsilon = \frac{4\pi V}{3h^3} (2m)^{\frac{3}{2}} \epsilon_0^{\frac{3}{2}} = N, \quad (289)$$

exactly equal to the number of particles present. Then the total energy of the gas is

$$\begin{aligned} U_0 &= \frac{2\pi V}{h^3} (2m)^{\frac{3}{2}} \int_0^{\epsilon_0} \epsilon^{\frac{3}{2}} d\epsilon = \frac{4\pi V}{5h^3} (2m)^{\frac{3}{2}} \epsilon_0^{\frac{5}{2}} = \frac{3}{40} \left(\frac{6}{\pi}\right)^{\frac{1}{2}} \frac{h^2}{m} \left(\frac{N}{V}\right)^{\frac{5}{2}} \cdot N \\ &= \frac{3}{5} \epsilon_0 N. \end{aligned} \quad (290)$$

This total energy of the gas at the absolute zero is dependent on the volume. This in itself requires the existence of a pressure

at the absolute zero in accordance with the thermodynamical equation

$$\frac{\partial v}{\partial V} + p = T \frac{\partial p}{\partial T}.$$

This zero-point pressure amounts to

$$p = \frac{1}{20} \left(\frac{6}{\pi} \right)^{\frac{1}{3}} \frac{\hbar^2 (N)^{\frac{5}{3}}}{m (\bar{V})}. \quad (291)$$

There is another property of Bose's as well as of Fermi's statistics of the ideal gas which is of importance. Whereas in classical statistics the degrees of freedom of an atom which do not interact with one another, as, for example, the motion of translation of the centre of gravity and the motion of rotation, were treated as independent, this is not so in the statistics of Bose and Fermi. This can be understood most simply as follows. Two particles with the same proper function of translation but different states of rotation are not atoms in the same state and consequently are not interchangeable. On the other hand, these two states of motion recover their independence if the condition (288a) is fulfilled, so that we can carry out these statistics independently of one another for all material atoms (except electrons, which are to be discussed below).

61. Metallic Electrons as a Degenerate Fermi-gas

The classical theory of electrons had early attempted to interpret the conductivity of metals by assuming the presence of freely-moving electrons in them, and had achieved some notable successes. We recall that the Wiedemann-Franz Law was derived; this states that the quotient of the thermal conductivity by the electrical conductivity is proportional to the absolute temperature. The classical theory even gave an approximately correct value for the factor of proportionality. The development of the theory, however, led to a great number of inconsistencies in itself and also contradictions to experimental results, so that in general it was necessary to reject the picture of free metallic electrons. Perhaps the most important objection consisted in the observation that the specific heat of the metal at normal temperatures is only as great as it would be if calculated on the basis of Dulong and Petit's Law, so that the supposed free electrons, assumed to behave like gaseous atoms, appeared to have no specific heat at all.

Since the advent of the quantum theory attempts have

repeatedly been made, by means of certain changes in the statistics used, to regard these metallic electrons as a gas, but not a gas in the ideal state, but rather as degenerate. The application of Fermi's statistics to the electron gas by Sommerfeld first brought success in this field.

Sommerfeld pictures the electrons in the following way. Every atom in the metal parts with one or more of its outer electrons (valency electrons), so that it is left positively charged. The electrons can then move about inside the metal as if on the average no force is acting on them, since the free ionic charge is in the mean compensated by the electrons. This picture is, of course, only sketchy, yet it represents a very practical approximation for many questions and has served as the starting-point for valuable refinements of the theory of electrical conduction, which have been carried out by Bloch and Peierls with special attention to the potential conditions in the crystal lattice. In the present book we restrict ourselves, partly for mathematical reasons, to Sommerfeld's approximation.

If we apply our criterion (288) to this electron "gas" at room temperatures, we get

$$\epsilon^{\alpha} \ll 1.$$

It is only when T is of the order of magnitude 10^4 that ϵ^{α} approaches the value unity. We are therefore not dealing here with a case in which the "gas" behaves almost classically but rather with the opposite case of extreme degeneracy. Approximate formulæ for this case have already been set up by Fermi. They run

$$U = U_0 + \left(\frac{2\pi^8}{9}\right)^{\frac{1}{2}} \left(\frac{V}{N}\right)^{\frac{1}{2}} \frac{mk^2 T^2}{h^2} N, \quad (292)$$

$$p = p_0 + \left(\frac{16\pi^8 N}{24^3 V}\right)^{\frac{1}{2}} \frac{mk^2 T^2}{h^2}. \quad (293)$$

These equations not only inform us of the existence of zero-point energy and zero-point pressure, as has already been mentioned, but also tell us how the energy and the other thermodynamic functions behave at low temperatures and at a high degree of degeneracy. In particular, we see that the specific heat of a degenerate gas is much smaller than the value $3k/2$ which corresponds to classical statistics, and that as the temperature decreases this specific heat approaches zero. The essential difficulty involved in the classical assumption that there are electrons in the metal which move practically freely is thus eliminated.

It is also interesting to know the order of magnitude of the mean kinetic energy which a metallic electron has at room temperature. From (290) this comes out at about 10^{-11} erg. (6 volts). The mean kinetic energy on the classical theory would have amounted to about 10^{-14} erg. It must be noted that the dependence of these large amounts of energy (that is, the specific heat of the electrons) on the temperature is very small.

On the basis of these ideas and using methods analogous to those of the classical theory Sommerfeld has investigated the electrical and thermal conductivities of metals and has even succeeded in arriving at a quantitative explanation of some groups of the essential phenomena.

We shall not enter into a detailed treatment of non-equilibrium states here and shall give an account in the next two paragraphs only of thermionic phenomena and the paramagnetism of metals.

62. Photo-electric Limit and Richardson Effect

The emission of electrons by hot metals is governed by the well-known equation of Richardson. This states that the number of electrons per cubic centimetre which are in equilibrium with a hot metal at the surface of the metal is

$$\rho = AT^{\frac{3}{2}}e^{-\frac{\chi}{kT}}. \quad (294)$$

Here χ denotes the work which must be performed at the absolute zero of temperature to bring the electron from the interior of the metal to the surface. For many metals A is a constant which lies in the neighbourhood of $3 \cdot 10^{15}$. With our model of the metallic electrons this formula can be simply deduced in the following way. In thermodynamic equilibrium the difference of the free energy of the electrons in the metal and outside must be equal to the work required to carry out the transference. Now we have already pointed out earlier that the electrons in the metal have a very great mean kinetic energy even at the absolute zero. The fact that they do not escape from the metal is only due to the circumstance that there is an abrupt change of potential at the surface metal-vacuum, which tends to hold back the electrons in the metal. This change of potential must be so great that even when the kinetic energy of the electrons has been subtracted there is still some work necessary to extract the electrons from the metal. With our present theoretical knowledge we can only roughly calculate this abrupt change of potential, so that it must be regarded for the time being as an empirical constant characteristic of the metal.

We have yet to refine our statistical considerations in one respect. It is known that the electron is not a structureless point but rather has a mechanical (and a magnetic) moment. If we imagine a weak magnetic field to be imposed, whose influence on the energy values can be neglected—as has so often been done before in similar circumstances—then the electron can assume two positions, both of which are in themselves justifiable. Thus according to our concluding remark in § 60 we can carry out the whole statistical calculation as if we had two different kinds of electrons which are in thermodynamic equilibrium both with themselves and with outside space. For the free energy within the metal * the following relation then holds:

$$F_M = U + \frac{\partial F}{\partial T} = U_0, \quad (295)$$

and in outside space, where there are only very few electrons and for which we may consequently use classical statistics, we get for the value of the free energy,

$$F_G = kTN \left(\frac{3}{2} - \log \frac{2e^{\frac{3}{2}} V}{N h^3} (2\pi m k T)^{\frac{3}{2}} \right). \quad (296)$$

Our value deviates from the usual value by the term

$$- N k T \log 2,$$

which is explained by the fact that the electron spin can express itself in two ways. In equilibrium we must have

$$\frac{\partial F_M}{\partial N} - \frac{\partial F_G}{\partial N} = \chi_0 + kT,$$

where χ_0 denotes the abrupt change of potential. From this we get

$$\frac{N}{V} = \rho = 2 \frac{(2\pi m k T)^{\frac{3}{2}}}{h^3} e^{\frac{\chi_0 - \frac{\partial U_0}{\partial N}}{kT}}. \quad (297)$$

With Sommerfeld's model, then, the constant A of Richardson's equation (294) is universal and has the value

$$\frac{2}{h^3} (2\pi m k)^{\frac{3}{2}}.$$

* We are justified in calculating with the conditions that hold at the absolute zero, because the degeneracy is complete.

The work required at the absolute zero to remove an electron from the interior of the metal comes out as (cf. 290)

$$\chi = \chi_0 - \frac{\partial U_0}{\partial N} = \chi_0 - \frac{5}{3} \frac{U_0}{N} = \chi_0 - \epsilon_0. \quad (298)$$

Thus we see that this work is not equal to the whole change of potential, but, rather, equal to the potential change minus the *maximum kinetic energy* which the electron has at the absolute zero [cf. (290)]. But this work is also characteristic of the so-called long-wave limit of the photo-electric effect. If we suppose a light-quantum $h\nu$ to transfer its energy to an electron in a metal, then, from energy considerations, the electron can leave the metal only if

$$h\nu \geq \chi = \chi_0 - \epsilon_0.$$

Thus there is a long-wave limit of the photo-electric effect, for which the energy of a light-quantum coincides with Richardson's quantity χ .

63. Paramagnetism of Metals *

The fact that electrons have a magnetic moment makes it almost impossible on classical statistics to assume an electron gas in the interior of the metal. For in this case, the electrons would be constrained to point for the most part in parallel directions at low temperatures, and this would cause the metal to have a paramagnetic susceptibility which would depend strongly on the temperature. Nothing of the sort has been observed hitherto; rather the metals mostly exhibit diamagnetism or relatively weak paramagnetism which, moreover, is independent of the temperature. The theory of this phenomenon is given by Pauli ¹⁰³ as follows:

The electrons in the interior of the metal obey Fermi's statistics, hence the probability that an electron with kinetic energy ϵ_{kin} shall have a magnetic moment $+\frac{eh}{4\pi mc}H$ in the direction of the field is given by

$$n(\epsilon_{\text{kin}}, \mu) = \frac{g(\epsilon_{\text{kin}})}{e^{\alpha + \frac{\epsilon_{\text{kin}} + \mu H}{kT}} + 1}. \quad (299)$$

The mean magnetic moment of a metal is then given by

$$\int_0^\infty \mu n(\epsilon_1, +\mu) d\epsilon - \int_0^\infty \mu n(\epsilon_1, -\mu) d\epsilon.$$

* In § 63 we use m to denote the electronic mass and μ the magnetic moment for convenience. But this notation is not used in other paragraphs.

An approximate calculation then gives us for the magnetic moment per unit volume,

$$kH = \frac{e^2}{4\pi mc^2} \left(\frac{3N}{\pi V} \right)^{\frac{1}{2}} H. \quad (300)$$

We have here written down only the paramagnetic part which is independent of the temperature and which alone comes into consideration at room temperature; thus the weak part which depends on the temperature has been omitted. Of course, superimposed on this paramagnetic effect we have the diamagnetism due to the motion of the electrons inside the ions: this diamagnetism is not taken into account in our calculation. Pauli's formula gives very good agreement in order of magnitude for the case of the alkali metals, as can be seen from the following table:

		Na	K	Rb	Cs
K. 10^7 observed		5.9	5.1	0.9	— 0.5
K. 10^7 calculated		6.6	5.2	4.9	4.5

In the elements of higher atomic number the diamagnetism becomes predominant.

To help us understand this weak paramagnetism we may picture that the electrons in the metal crowd together in pairs in the states of smallest kinetic energy, just as in the ground-state of helium, and their magnets, owing to Pauli's Exclusion Principle, must be in anti-parallel directions. Thus the elementary magnets largely compensate each other owing to Pauli's Principle; not more than two electrons can remain in one state.

64. Heat of Rotation of the Hydrogen Molecule

At moderate temperatures the hydrogen molecule gives the value $\frac{5k}{2}$ for its specific heat; of this value the amount $\frac{3k}{2}$ may be regarded as due to translation and the remainder (k) as due to rotation about two axes through the centre of gravity with the moment of inertia $\mathcal{I}$. At low temperatures this specific heat diminishes, being $\frac{3k}{2}$ at about 70° . This decrease of the specific heat can be understood at once qualitatively as a result of quantizing the rotational energy. The older theories have been unable to furnish a quantitative reason for this decrease. (The fact that we have treated the heat of rotation according to quantum statistics quite independently of the motion of translation is justified because we are dealing with a gas of low density. Cf. the concluding remarks of § 60 in this connexion.) According

to quantum statistics the heat of rotation is given by the expression

$$\bar{\epsilon} = \frac{\sum_l g_l \epsilon_l e^{-\frac{\epsilon_l}{kT}}}{\sum_l g_l e^{-\frac{\epsilon_l}{kT}}}, \quad (300)$$

$$C_{\text{rot}} = \frac{d\bar{\epsilon}}{dT},$$

where g_l denotes the weight of the l th state. Now the determination of these weights depends entirely on what proper function is valid for the protons. It is at any rate clear from the remarks on the change of intensity in the band spectrum of hydrogen that the protons have a spin and that only one group of proper functions is realized in nature. The following discussion shows that it must be the asymmetrical proper function. For since, according to an investigation by Hund,^{92a} the electronic ground-state is always *symmetrical* in the *nuclear co-ordinates*, the rotation state 0 has the statistical weight 1 when the total proper function of the nuclei is symmetrical, whereas the rotation state 1 has the statistical value 3 (para- and ortho-term). By enumerating the spherical harmonics (cf. § 41) we find that the degree of degeneracy of the rotation of the ordinary rotator comes out as $2n + 1$. According to the above the states with even values for n have the factor 1 assigned to them, while those with odd values of n have the factor 3; these factors arise from the four different possibilities of orientation of the nuclear spin. If we correspondingly write down for the heat of rotation the expression

$$\bar{\epsilon} = \frac{\sum_0^{\infty} n g_{2n} \epsilon_{2n} e^{-\frac{\epsilon_{2n}}{kT}} + 3 \sum_0^{\infty} n g_{2n+1} \epsilon_{2n+1} e^{-\frac{\epsilon_{2n+1}}{kT}}}{\sum_0^{\infty} n g_{2n} e^{-\frac{\epsilon_{2n}}{kT}} + 3 \sum_0^{\infty} n g_{2n+1} e^{-\frac{\epsilon_{2n+1}}{kT}}}, \quad (301)$$

we get from it a formula for the specific heat which does not agree at all with the results of experiment. Dennison interprets this circumstance in the following way. If the nuclei had no spin transitions between even and odd rotation quantum numbers would be altogether impossible in the electronic ground-state. In consequence of the spin they are possible but occur very rarely, the number being proportional to the energy due to the interaction of the two magnets. Now it is quite reasonable to assume that during the relatively short time during which observations are made, which is of the order of several days, a combination of even with odd rotational quantum numbers

does not occur. We are then in a certain sense dealing with a mixture of two gases, one of which exhibits only even rotational quanta and the other only odd rotational quanta, and in which the ratio of the *a priori* probabilities is 1 : 3, as in the case of the para- and the ortho-term. Accordingly, we must assume for the free energy of a mixed gas of this kind, which does not set itself into equilibrium, that

$$\bar{\epsilon} = \frac{1}{4} \frac{\sum_0^{\infty} n g_{2n} \epsilon_{2n} e^{-\frac{\epsilon_{2n}}{kT}}}{\sum_0^{\infty} n g_{2n} e^{-\frac{\epsilon_n}{kT}}} + \frac{3}{4} \frac{\sum_0^{\infty} n g_{2n+1} \epsilon_{2n+1} e^{-\frac{\epsilon_{2n+1}}{kT}}}{\sum_0^{\infty} n g_{2n+1} e^{-\frac{\epsilon_{2n+1}}{kT}}}. \quad (302)$$

This leads to a value for the specific heat which agrees excellently with the observed values.

Dennison's hypothesis that we are here dealing with a mixture of two gases that rarely combine has recently received direct confirmation from experiments by Bonhoeffer and Harteck⁶⁴ and Eucken.⁶⁰ By experimenting with hydrogen under very high pressure it was possible to observe at low temperatures the transformation of the ortho-modification of greater energy and the para-modification of lesser energy. When this hydrogen was allowed to stand for a few days under high pressure its specific heat changed in perfect accord with theory, slowly assuming the value characteristic of para-hydrogen. The purpose of increasing the pressure was to provoke more frequent transitions between the two modifications by increasing the number of collisions.

Further, these experiments have furnished the proof that the anti-symmetrical proper functions also apply to protons. For since the ground-state is symmetrical in the nuclear coordinates it follows that if we choose the symmetrical proper functions the rotation quantum number 0 would have to have three times the *a priori* frequency of the rotation quantum number 1, which is in contradiction to observation, as this has yielded the inverse ratio as in the case of helium.

65. Theory of Ferro-magnetism

The quantum theory of the many-bodies problem has succeeded in making an unexpected contribution to the interpretation of ferro-magnetic phenomena. Although the mathematical apparatus that was required to deduce the results involved the most intricate parts of the theory of groups, and the calculation of the details encountered almost insuperable mathematical difficulties, we cannot refrain from giving a survey of the leading

physical ideas that entered into the investigation, although we cannot give the actual calculations.

As is well known, the difficulties of the classical theories in the field of ferro-magnetism were as follows: certain bodies (iron, cobalt, nickel, that is, ferro-magnetic substances) often exhibited spontaneous magnetization even when no external field was present. It was found in the case of some substances that this magnetization was little affected by external fields but depended particularly on the temperature. The ferro-magnetism vanished at higher temperatures above a critical point (Curie point) characteristic of the particular substance and passed over into paramagnetism which, of course, meant that the substance exhibited a magnetic moment different from zero only when an external exciting field was acting.

This complex of facts may be expressed in the language of statistics as follows. Starting from our knowledge of other cases, we assume that the atoms of ferro-magnetic substances carry elementary magnets. If a mutual action occurred in the interior of the metal with a tendency to direct the elementary magnets in parallel directions, this would make it possible to explain qualitatively the way in which ferro-magnetic phenomena arise. An interaction of this kind can be introduced purely formally, as was done by Weiss in his hypothesis of the internal field, and formulæ can then be obtained which account for a number of phenomena qualitatively and approximately quantitatively. It only remained to find the nature of these mutual forces of interaction in the interior of the metal.

Simple considerations about order of magnitude at once exclude mutual magnetic actions as they can produce only a zero effect or at most a very small effect. If we imagine electrical forces to act between the atoms such as to produce the directive effect in question, it is easy under certain assumptions to make calculations with simple models which show that ferro-magnetic phenomena cannot come about under these circumstances.

A very illuminating contribution to the nature of elementary magnets was made by the well-known experiments of Richardson, Barnett, Einstein, de Haas and others. In these effects it was shown that the ratio of the magnetic moment and the momentum of momentum behaves anomalously in the case of elementary magnets. When, later, the quantum theory of the Zeeman effect led to the same abnormal value of this ratio and finally led to the introduction of the spin hypothesis, it readily suggested itself to assume that in ferro-magnetism we were also dealing with a spin-effect. Heisenberg⁸⁹ combined this fact with the further fact that the forces between similar particles are

not characterized simply by their mutual electrical action but that the action of particle-exchange has also to be added. A system of N electrons, N being of the order of magnitude of the number of atoms in the mole, is degenerate if we disregard the interaction between the individual electrons.

If we take into account this interaction the system splits up into a great number of different terms, a definite part of which belongs to anti-symmetrical functions. (The argument runs along exactly the same lines as for the helium atom, except that now we have N instead of 2 electrons.) In a certain sense we treat the whole piece of metal with its N electrons as a uniform gigantic mechanical atom. Calculation shows that among these very many proper functions there will be a great number for which one part of the electrons (more than half of all the electrons) will have their elementary magnets pointing in *one* direction. Following exactly the statistical principles above described, we have now to determine the energy of each of such proper functions and further to calculate the resultant magnetic moment and finally the statistical mean value of the magnetic moment.

So long as some external field has not at some time or other directed the electronic magnets, even if only partially, the magnetic moment of a macroscopic piece of metal will have a mean value zero, since although the individual irregularly situated micro-crystals each have a magnetic moment of their own, their directions are entirely left to chance and so cancel out on the average. Even a weak magnetic field exerts a directive action, and when it is removed traces of the magnetization remain.

Thus ferro-magnetic phenomena represent a statistical effect and consequently, of course, depend on the temperature. Calculation also shows, in complete agreement with observation, that above the definite critical temperature the thermal motion of agitation becomes so intense that spontaneous magnetization can no longer come about.

In conclusion we shall once again emphasize the fundamental difference between this explanation and the classical theory—a difference which has led to the introduction of mutual electrical action.

In the classical theory the electrical interaction between two or more point-like atoms was the determining factor; in the quantum theory, on the other hand, the purely electrostatic mutual action of the two atoms supplemented by those terms which effect particle-exchange in consequence of degeneracy is the decisive element. These exchange terms enabled the quantum theory to give a successful interpretation.

CHAPTER X

INTERPRETATIONS OF THE THEORY

66. Enunciation of the Problem

IN the preceding paragraphs we have investigated a series of cases, as, for example, all non-stationary problems, in which the interpretation which we had given quantum mechanics in Part I of our report is no longer maintained. It was shown there how the stationary states of a system and the intensities in a transition are to be calculated by the method of matrices or with the help of the wave-equation. Nowhere was mention made of the position, velocity or other mechanical elements of reference of a particle, such as had been introduced in the Compton effect and so forth. Indeed, it was put forward as a particular advantage of Heisenberg's theory that the introduction of these "fundamentally unobservable" quantities was avoided in the theory. This may create the impression that we have made a retrogressive step in our discussion and puts us under the obligation of giving a comprehensive interpretation of the equations from our present point of view. In doing so we are fully aware that we are excluding from consideration the series of problems connected with so-called multiple quantizing, which plays a predominant part particularly in the papers by Jordan, Klein and Wigner^{95, 96}. In the researches just mentioned fundamental assumptions are involved the meaning of which cannot yet be clearly pictured and which will probably be elucidated only in connexion with the theory of quantum electrodynamics, which is in process of evolution at present.

We recall once again the two fundamental difficulties, at first sight apparently quite unrelated, which were overcome by the quantum theory. The first group comprises the whole complex of manifestations characterized by the term "light-quantum phenomena." Expressed crudely this means that the dominant facts can be formulated as if light, besides having the wave-nature which observation shows cannot be doubted, also has a corpuscular nature which is the governing factor in all questions connected with the laws of energy and momentum.

The second group of "incomprehensible manifestations in the classical sense" comprised all those phenomena associated with the words "stationary state and transitions." It appeared here as though the classical kinematical relations were no longer fully valid, as though, out of the totality of orbits possible according to dynamics and kinematics, for unaccountable reasons only certain orbits could be actually realized.

We must here add a general remark about a certain state of affairs which has frequently been discussed in the years preceding the advent of the new quantum mechanics. This concerns the fundamental question of measurement in atomic regions. It was a fundamental assumption of classical physics that the concepts with which it operated could be measured without introducing disturbances. This was to mean that for any individual measurement a test-body existed whose influence on the system to be measured could be made as small as was desired. This represented a *postulate* of classical physics. It had long been regarded as questionable whether this postulate would prove to be compatible with the existence of atoms, that is, with the smallest conceivable measuring bodies. But this problem had not attained to practical importance. Rather it was intended to be a critical objection to certain experiments which attempted to interpret the smallest particles, electrons and protons, in terms of the electromagnetic field.

67. Original Form of Heisenberg's Theory. Schrödinger's Wave-Mechanics in its Extended Sense

It is unnecessary to repeat here the ideas contained in Heisenberg's original investigations as they have been fully discussed in Part I of our report. The following remarks will serve to recall them sufficiently. Let a mechanical system be characterized by its Hamiltonian function, the co-ordinates and momenta, however, having nothing to do with space-time quantities but being simply mathematical quantities that obey certain rules of calculation. A unique rule of calculation leads for closed systems to a determination of the co-ordinate matrices and the diagonal matrix of the energy. This diagonal matrix is characterized as representing the totality of possible energy-levels of the system. The co-ordinate matrices are interpreted, in the special case where Hamilton's function is taken over into Cartesian co-ordinates from classical physics, as measures of the intensities of the spectral lines.

This theory is very nearly complete in itself. Its chief weakness, due, of course, to the immature state of development of

the theory, is that it makes no direct statement as to how the matrices which characterize the optical conditions are conceptually connected with the actual space-time co-ordinates and momenta. The theory is compelled to refer to the macroscopic conception of co-ordinates when it is a question of calculating intensities; not all matrices but only those formed from Cartesian co-ordinates are of direct use for this purpose. It was to be foreseen that the further development of the theory would furnish this link in some form or other.

Schrödinger's wave-mechanics may be discussed from two angles in this connexion. In the first place, it is completely equivalent as a mathematical process to Heisenberg's method of matrices. This unique correspondence, as we know, was shown to exist by Schrödinger³⁹ and Eckart⁷⁷. We are justified in describing any state of affairs in terms of proper functions and can then be certain that the corresponding theory of matrices can be set up.*

So far then there need have been no question of giving a special interpretation to Schrödinger's theory, and we might have remained satisfied with speaking of the different heuristic points of view that led to the two theories being proposed. In a second respect, however, the position is fundamentally different. Schrödinger has attempted in a very definite manner to establish a link with classical kinematics, indeed, more still, with classical field-physics. Although he himself¹⁰⁸ had given his support, some years before the new quantum mechanics had been proposed, to the possibility or rather the probability of a non-causal statistical theory of atomic processes, de Broglie's ideas appeared to him to offer the possibility of carrying through a more conservative programme with the help of his new operator calculus. His line of reasoning, on which we have already touched in § 28, runs somewhat as follows:

To make calculations about electrical radiation phenomena it is advantageous, as we know, to use the expression

$$\rho = \psi\bar{\psi}$$

for the electric density. Let us suppose that ψ is here represented by the most general expression (224). This value for the electric density makes it appear as though the electron, about whose ψ -function we are here concerned, is spread continuously over

* We feel that no error is being committed here in disregarding the special refinements that constitute the differences between non-commutative methods (matrices, q -numbers, on the one hand, and Wiener's, Schrödinger's, Jordan's and Dirac's operator calculus on the other hand) since nothing affecting the concepts fundamentally is involved in our simplified account.

the whole of space with a density that depends on its position. The proper function itself is represented by a trigonometrical series involving the time. The dependence of any term of this series on the time is of the form

$$e^{\frac{2\pi i E_n t}{h}}.$$

We then obtain as the time-dependence of the quantity $\psi\bar{\psi}$, in contrast to the time-dependence of the proper functions themselves, the expressions

$$\psi\bar{\psi} = \sum_k \sum_l a_k \bar{a}_l v_k \bar{v}_l e^{\frac{2\pi i}{h}(E_k - E_l)t}. \quad (303)$$

Now states, that is, proper functions, are not physically real elements of existence in Schrödinger's sense at all. These proper functions themselves and their dependence on time have not actually been physically observed. Only the frequencies and intensities of the radiation corresponding to a transition,

that is, which depend on the time according to $e^{\frac{2\pi i}{h}(E_k - E_l)t}$, are observable.

Another matter of importance pointed out by Schrödinger is the particular way in which such a "charge-cloud" is distributed over space (cf. § 43, Figs. 3 and 4). It is found that the charge is essentially restricted to a region which is of the order of magnitude of the usual atomic dimensions, so that no contradiction with experimental results arises from assuming a continuous distribution. Schrödinger further showed (103) by means of the example of linear oscillators that by appropriately superposing proper solutions

$$\Psi = \sum_k c_k \psi_k$$

it is possible to form a system whose centre of gravity behaves at all times like a point-mass vibrating harmonically. Such systems, which arise from the superposition of proper-solutions, are called *wave-packets*.

The transition from classical to quantum kinematics according to Schrödinger is now very simple in principle and easy to picture. *It is fully analogous to the transition from geometrical optics to wave-optics.* We can illustrate this most simply in the one-dimensional case. Let the wave-equation of light be given. From the wave-equation

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{n^2}{c^2} \frac{\partial^2 \phi}{\partial t^2},$$

where n is the index of refraction, we can, by substituting

$$\phi = e^{2\pi i \left(\mu - \frac{S}{\lambda n} \right)} (S = \text{phase of the light source}) \quad (304)$$

obtain the new following equation for S :

$$\frac{2\pi i}{\lambda n} \frac{d^2 S}{dx^2} + \frac{4\pi^2}{\lambda^2 n^2} \left(\frac{dS}{dx} \right)^2 = \frac{4\pi^2}{\lambda^2}. \quad (305)$$

We know that geometrical optics is valid for very short waves. If in (305) we neglect the first term we obtain the equation

$$\left(\frac{dS}{dx} \right)^2 = n^2,$$

which, as we know, governs the rectilinear propagation of light and the phenomena of reflection and refraction in the sense of geometrical optics.

In the same way, by means of the substitution

$$\psi = e^{\frac{2\pi i}{h} W} \quad (306)$$

we can pass from Schrödinger's equation to the equation

$$\frac{2\pi i}{h} W'' - \frac{4\pi^2}{h^2} W'^2 + \frac{8\pi^2 \mu}{h^2} (E - V) = 0, \quad (307)$$

which becomes, when $h \rightarrow 0$,

$$\frac{1}{2\mu} \left(\frac{dW}{dx} \right)^2 + V = E, \quad (308)$$

and this is the Hamilton-Jacobi differential equation of mechanics.

Schrödinger's interpretation may also be applied directly to the treatment of non-stationary problems. We are then, to use his language, dealing with a charge-configuration which no longer changes periodically.

Let us once again recapitulate this special form of Schrödinger's interpretation. According to it a continuous distribution of matter exists throughout space. Definite specified functions of the proper function characterize the density of the energy, momentum, charge and current at every place and for every moment of time. The function ψ is determined when we know its differential equation, that is, the external fields and the interaction of the particles as well as their initial state. From this initial state we then get rigorously the definite course of the function ψ for all times.

This view of Schrödinger's has not in general been able to

maintain itself. The great majority of physicists have adopted instead of it a view which bears a pronounced statistical character. Before we discuss in detail the grounds which primarily led to the departure from Schrödinger's interpretation, we shall give a short sketch of this third view.

68. The Statistical Interpretation of Quantum Mechanics. Theory of Transformations

The third interpretation of quantum mechanics is in essence purely statistical. In it the classical concepts of position, velocity, momentum, energy and so forth of the mechanical system are retained unaltered. The only restrictions it imposes relates to simultaneous observation of mechanical and kinematical quantities and the way in which they are linked together. Moreover, this interpretation retains the rigorous particle character which classical physics also ascribed to atoms. According to it, then, we can speak of an electron being at a perfectly sharply defined place. The essential part of the theory may be formulated somewhat as follows. Suppose a solution of Schrödinger's equation for a closed system is given in the form

$$\psi = e^{\frac{2\pi i}{h}Et} v_{n_1, n_2, \dots, n_f}.$$

If we know the values of these proper functions we also know the value of the energy and also $f - i$ further integrals of the system, where f denotes the number of degrees of freedom. This is to convey the following meaning. In the original theory of Bohr, when f quantum numbers were given, an equal number of functions of p and q was also given which, of course, had to be integrals of the mechanical equations of motion. Translated into the language of quantum mechanics this means that we can specify f quantities whose matrices must be diagonal matrices (for simplicity we leave out of account degeneracy). As an example let us take the quantities fixed for the hydrogen atom (cf. §§ 43 and 72). They are: (1) the energy $E(p, q)$; (2) the square of the angular momentum $\sum (xp_y - yp_x)^2$, which is determined by the quantum number l ; (3) the component of the angular momentum about the z -axis $(xp_y - yp_x)$, which is determined by the quantum number m .

We can then form from the proper functions in the manner already familiar to us such expressions as $\psi \frac{\partial \bar{\psi}}{\partial x_1} - \bar{\psi} \frac{\partial \psi}{\partial x_1}$, which we have already interpreted earlier as four-current and so forth. *In the new view these quantities then denote the probability that*

we shall experimentally find a particle in the region of co-ordinates $dx_1, \dots dx_f$, when we know that the system is in the state characterized by the proper functions $v_{n_1}, \dots v_f$.

This process can of course be extended quite formally to non-closed systems. The system in question, whose solution we shall suppose to be of the form (224), is then also characterized by the quantities a_n , which, on account of the normalizing condition, must obey the relation

$$\sum_n a_n \bar{a}_n = 1.$$

In the section on the variation of constants we showed that these relations if they once hold for one moment of time are valid for all times. This allows us to regard the quantities $a_k \bar{a}_k$ as a measure of the probability that the system in question will at a certain moment be in the state denoted by the proper function v_k .

For closed systems this interpretation leads to exactly the same laws as earlier.

We shall next discuss a number of details of the statistical interpretation here given, so that we may arrive at a better understanding of its general character. The first question that immediately suggests itself after what has been said is this: What relationship is there between those mechanical quantities about which we know something and those about which our present knowledge leads us to make probability assertions? Now every mechanical system of f degrees of freedom is characterized by $2f$ quantities, namely f co-ordinates and f generalized momenta. A proper function of a closed non-degenerate system fixes f and only f functions of p and q , which includes among them the energy. The quantities which are canonically conjugate to the f known quantities are not fixed and are precisely the quantities to which our probability statements refer. The process which we have here described is still somewhat specialized inasmuch as we have throughout assumed a solution of Schrödinger's equation as the basis of our calculation. Dirac⁷⁵ and Jordan⁹⁴ have extended the formal aspect in the following way, which is called the *transformation theory of quantum mechanics*. Just as Schrödinger's equation results by writing Hamilton's function as an operator and then seeking to determine its proper functions, so we can write any other arbitrary function of the co-ordinates and momenta in the form of an operator and determine by purely mathematical methods the proper values of this function just as we did earlier in the case of the Hamiltonian function. The essential result arrived at

by Dirac and Jordan may be stated as follows: To determine with what probability we may assign the value Q to a mechanical variable q in a system when we know that the value of another mechanical quantity $f(p, q)$ is F , we must solve a definite proper-value problem, which is formed with the function f as operator. The resulting proper-solutions then allow us to calculate the probability that the quantity q has the value Q . We see immediately that Schrödinger's case is a special case of this general process. Here, for example, the required quantity Q is the co-ordinate q itself, and the known quantity is the Hamiltonian operator. When substituted in the proper-value problem Hamilton's function gives precisely Schrödinger's equation, and the expression $\psi\bar{\psi}$ gives the probability that the system is in the position corresponding to the co-ordinate Q .

It is obvious that we cannot here enter more closely into the highly complicated calculations that lead to this theorem, although it has been possible to enunciate it simply. Without giving a proof we shall state the general result, which refers to the inter-dependence of canonically conjugate quantities and which we shall illustrate later in a great number of special cases.

If we can make a statement about the value of the quantity p —no matter whether we can assign to it the exact value P or whether we only know that it can assume *all possible* values whose mean deviation is $(\Delta p)^2$ —theory then leads to the result that the error (mean deviation) of the canonically conjugate quantity q satisfies the relation

$$(\Delta p)^2(\Delta q)^2 \sim h^2.$$

Thus if one of the two canonically conjugate quantities is exactly known the other is completely indeterminate. This formulation of Heisenberg's Uncertainty Relation has been included in our discussion of the transformation theory only for the sake of completeness; we shall discuss its fundamental importance in the light of various examples below.

69. Further Discussion of the Statistical Interpretation. Einstein's Imaginary Experiment

The statistical interpretation diverges so radically from the ideas hitherto accepted in physics that it appears appropriate to examine its conceptual foundations more closely.

Causal physics may be defined as follows: Let the state of a mechanical system be characterized by having $2f$ quantities specified, say the co-ordinates and the momenta. A knowledge

of these quantities at a definite time $t = t_0$, combined with a knowledge of the laws of mechanics (we disregard difficulties of calculation) allows us to determine the 2f quantities at any arbitrary time.

The statistical interpretation, on the other hand, states the following: At the time $t = t_0$ let all the quantities of the system that are simultaneously measurable be determined. A knowledge of physical laws then enables us to make probability statements about the frequency with which characteristic quantities can assume different values after the lapse of the same length of time in experiments which are repeated, the system being in the same state at the beginning of each experiment.

Both in classical and in quantum mechanics we must thus first specify what observable quantities characterize the state of a system. It may be the case from the very outset (and in § 70 we shall see that this is actually realized in nature) that in a mechanical system the quantities which are regarded as characteristic of the state of the system in classical and quantum mechanics are different. But we shall disregard this point here as it is unimportant *in the present connexion*.

The formulation of the statistical interpretation refers to the result of repeated series of experiments. We have left open the possibility that some of the quantities characteristic of the system change in the same way in every experiment, that is, that in a certain sense causality prevails for these quantities. For a statistical interpretation it is only necessary for *one* of the quantities characteristic of the mechanical system to change in a way which is not uniquely determined.

Let us turn to the example of the photo-electric effect (cf. § 49). We there made a statement both about the energy (that is, the absolute value of the velocity) and about the direction of the velocity of the emitted electrons. Now in that section we saw that two proper functions combined only if they were such that the energy of the electron satisfied the equation (240), whereas the angular distribution of the escaping electrons was characterized by the four-current. But this four-current did not enable us to predict unambiguously in what direction the electron would escape. Hence, regarding the way in which a photo-electric process takes place, we may say that the energy is uniquely determined, but it is possible to make only a probability statement about the direction of the velocity of the released electron.

We must emphasize a possibility which might hold for a statistical theory but, as we saw above, does *not* hold in the case of quantum mechanics. It is this: the energy of the

escaping electrons might have assumed all possible values and only their mean value for many experiments might have satisfied the equation (240). But the laws of conservation of energy and momentum (cf. § 49, Compton effect) show that this is not the case, but rather that exactly causal relations hold for these quantities.

These conditions are very clearly illustrated in an example proposed by Einstein. Let us consider the following experiment.

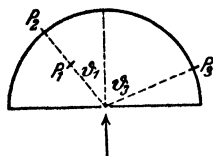


FIG. 11.—Einstein's imaginary experiment.

Suppose a beam of α -rays of equal energy and velocity fall on a small aperture in a screen and let it be scattered in all directions at the aperture. We wish to enquire into the fate of each individual particle after it has passed the aperture. We imagine a photographic film to be set up behind the screen (Fig. 11) on which we can fix the traces of the incident α -particle.

Then the quantity $\psi\bar{\psi} = f(\vartheta)$, which depends on the angle, gives us the probability that an α -particle moves off in the direction ϑ . Now we know from experiment, for example, from Wilson photographs, that an α -particle travels very nearly rectilinearly through a gas-filled space. Whereas, then, at first the probability of the α -particle arriving in the direction given by $+\vartheta$ or $-\vartheta$, respectively, was, for reasons of symmetry, equally great, we can assert with a high degree of probability that an α -particle of which we observe a trace (cloud-track) at P_1 , whose angle is ϑ_1 (cf. Fig. 11), in the gas-filled space between the screen and the film, will strike the film at a point whose angle is also nearly ϑ_1 . But we can also express this state of affairs as follows: Observing the α -particle at the angle ϑ_1 alters the probability of the α -particle being encountered at the angle ϑ_2 .

Up to this point there is nothing unusual in the argument. It gains special interest, however, through the circumstance, which has been pointed out by Einstein⁷⁹, that we are here dealing in a certain sense with "action at a distance." This is easily seen as follows. *Before* the α -particle was observed at P_1 there was a probability, say $w(\vartheta_3)$, that the α -particle would strike the film at the point P_3 . *After* the observation this probability is no longer $w(\vartheta_3)$ but 0 (or at any rate much smaller than $w(\vartheta_3)$). Now we can imagine the points P_1 and P_3 to be as far apart as we like, so that according to field-physics (relativity principle) a light-signal will take at least the time to pass from P_1 to P_3 :

$$\frac{\overline{P_1 P_3}}{c} = \Delta t.$$

In the case of quantum mechanics, however, the probability at P_3 has, on account of the observation at P_1 , changed *instantaneously*.

Although this remark is very interesting it does not seem to the present authors to disclose an inconsistency in quantum mechanics. It is possible to formulate parts of quantum mechanics so as to be relativistically invariant, and it is obvious that its results must not contradict the general theorems of the theory of relativity. Nor does this occur in the case in question. For the theory of relativity only forbids the transmission of a *signal with a speed greater than that of light*. In our case, however, we are dealing with the *change of probability of an experimental result*, and it is quite impossible to devise a real physical process which can be taken as equivalent to this behaviour of the wave-function. This example exposes very clearly the formal nature of the description given in terms of the wave-function. We also see clearly that in general it is impossible to make material events correspond directly with the behaviour of the proper functions, and that the attempt to do so easily leads to contradictions with the principle of relativity, which must hold for all physical theories.

Before ending this section we consider it advisable to warn the reader against an error which is often met with in discussions on the questions here treated. The view to which we object is the following: The statistical interpretation of quantum mechanics is not necessarily founded in the nature of things. Rather, the formal content of quantum mechanics, particularly in the Schrödinger version, suggests a causal interpretation. This causal interpretation consists in this: the proper function ψ is determined when we know it for the whole of space at the time $t = t_0$. It is then easily seen that Schrödinger's time-dependent equation (III) has only *one* solution which merges into the given initial state for $t = t_0$. [This already follows from the general theorems about the uniqueness of solutions of differential equations.] Thus the proper function ψ is completely determined "causally."

In our opinion a fallacy is here incurred, which arises through the concept necessary connexion (*Gesetzlichkeit*) being tacitly replaced by the related concept causality. The word *Gesetzlichkeit* is taken to stand quite generally for every connexion of a quantity depending on space and time with its values at any other point of space and time. "Causality" is the narrower concept. It refers to the necessary connexion (*Gesetzlichkeit*) of observable quantities, that is, those quantities whose value we can specify exactly even if only by means of

imaginary experiments. The formulation of the quantity ψ involves only a necessary connexion. For nowhere has the slightest attempt been made to devise an imaginary experiment which would allow the quantity ψ to be measured. Every system of physics whether it be causal or non-causal will have to furnish contents for laws. The quantities that enter into these laws, whose variations with time can be uniquely determined, will, however, be inaccessible to observation in statistical physics, while, at the same time, it is not possible to express the observable quantities as unique functions of the time.

70. Heisenberg's Uncertainty Relation. Purely Mechanical Problems

We have already briefly described the mathematical form of Heisenberg's Uncertainty Relation in connexion with the general theory of transformations. It runs as follows: If two quantum-mechanical quantities q and p are canonically conjugate to each other, and if we can make a statement about the quantity q or, respectively, its deviation $(\Delta q)^2$ from the mean value, then the relation

$$(\Delta q)^2 (\Delta p)^2 \sim h^2$$

holds for the deviation $(\Delta p)^2$ of the quantity p from the mean value.

In this section we shall chiefly deal with the physical interpretation of Heisenberg's law without laying stress on an exact mathematical formulation of the theorems used. This formulation can be performed in a rigorous and general way only with the help of the apparatus of the theory of transformations, which we cannot make use of here. But in spite of this we believe it possible, even without mathematical refinements, to give an exhaustive account of the essential content of Heisenberg's Uncertainty Relation.

We shall begin with some general remarks about "physical" quantities in classical and in quantum theory. We mean by a physical quantity one which requires for its quantitative determination at least an imaginary experiment; in other words, the quantity is *measurable*.

It was assumed in the classical theory, at first tacitly and then explicitly, that the quantities that enter into it could be measured without producing an appreciable change in the quantities to be measured; or, expressed more definitely and exactly, that a unique *measure* (*Messresultat*) could be assigned to a physical

quantity. It was further tacitly assumed in classical theory that *all the quantities that occur in it could be measured independently of one another.* To fulfil these assumptions the idea of a *test body* (*Probe-körper*) had been introduced, that is, a measuring instrument which could be chosen as small as one wished and whose reaction on the quantity to be measured or any other body that characterized the system could also be made as small as was desired.

Every atomistic view of matter lays the fiction of such a test-body open to doubt and may, indeed, rob it of all meaning. It is difficult to see how configurations of the smallest possible size are to be used to measure configurations of the same order of size without the process of measurement influencing the result in an unforeseen manner. The quantum theory, being a specifically atomistic theory, could not take over these classical assumptions without careful examination. The manner in which it modified them is clearly expressed in Heisenberg's Uncertainty Relation.

We must next point out the great complications involved in using a measuring-body; these are immediately evident. We can imagine configurations of arbitrary complexity with which our measurements can be carried out. Thus we can arrive at a general statement about the nature of quantum-mechanical quantities only by selecting a comparatively simple general property of all mechanical systems. Now no universal constants enter into the equations of quantum mechanics except the quantum of action. For *at the present stage of the theory* the charge and the mass of the particles are absolutely arbitrary, and without further provision it would be impossible for electrons or protons of a definite kind to enter into quantum mechanics. We may therefore expect that a general statement about quantum-mechanical measurements (at any rate at the present stage of the theory) will involve only the quantum of action.

To anticipate the result: quantum mechanics allows the classical assumption that a quantity in quantum mechanics can be measured to any degree of accuracy to hold good, but it shows that the second assumption is at fault, namely, the assumption that all physical quantities can be measured independently of one another to any degree of accuracy. The simultaneity of measurement restricts then to definite groups of physical quantities characterized by their mechanical simplicity.

To illustrate this we shall take several very simple examples from quantum mechanics. The solution of Schrödinger's equation for a free particle with constant energy is given by the

proper function * $\psi = e^{ikx}$. The momentum that must be ascribed to this particle is given by the expression

$$I_x \sim \psi \frac{d\bar{\psi}}{dx} - \bar{\psi} \frac{d\psi}{dx} \sim k.$$

It is therefore independent of x and everywhere proportional to k . Thus we are dealing with a mechanical system of *uniform momentum*, or, expressed in the usual way, *the momentum of our system is sharply defined*. Let us conversely enquire, in the sense of the statistical definition, into the probability of a particle with the above proper function being encountered in a definite region with the co-ordinate x_0 . This probability is given by

$$\psi\bar{\psi} \sim 1.$$

So it is independent of x ; that is, the position of the particle with this proper function is completely undetermined. The probability of finding it at any point of space is the same for all points. In the language of quantum mechanics, *the position of the particle is completely indeterminate*.

Now we may generalize the solution $\psi = e^{ikx}$ in a simple way so that the momentum and the co-ordinate exactly exchange their rôles as regards their degree of definiteness. For this purpose we form the expression

$$\psi = \int_{-\infty}^{+\infty} dk a(k) e^{\frac{k^2 \hbar i}{4\pi m} t} e^{ikx}. \quad (309)$$

ψ is likewise a solution of Schrödinger's equation (III), but it is now no longer *simply* dependent on the time but in the more general manner of (224).

We imagine the coefficients a_k to be determined in such a way (according to well-known mathematical theorems concerning the representation of indefinite functions by means of Fourier integrals) that for $t = 0$ ψ differs from zero only at *one point* of space and vanishes at all other points with the condition that $\int dx \psi\bar{\psi} = 1$.

If we form the momentum from (309) according to our ordinary rules, we see that it is now dependent on the co-ordinate x and can assume all possible values between $-\infty$ and $+\infty$ with different degrees of probability.

We can derive all the essential features from this simple

* We assume only plane waves in one direction and so for simplicity leave out of account the solution e^{-ikx} in our discussion.

example. The co-ordinate and the momentum are canonically conjugate to each other. A single term

$$e^{ikx} \quad \text{or} \quad e^{-ikx}$$

corresponds to a sharp value of the momentum. But this single term makes $\psi\bar{\psi}$ independent of x and therefore leads to indefiniteness in the co-ordinate. Conversely, fixing the co-ordinate in (309) requires an infinite number of terms of the form e^{ikx} . But corresponding to each of these terms there is a different momentum, so that we now have the indefiniteness in the quantity conjugate to the co-ordinate.

The position is the same, for example, with the two quantities energy and time, which are conjugate in classical mechanics. If the Hamiltonian function of the system is independent of the time the solution of Schrödinger's equation is, as we know,

$$\psi = v(x)e^{\frac{2\pi i}{h}Et}.$$

The quantity $\psi\bar{\psi}$ is then independent of the time and with it so are all the mean values

$$\int \psi f(p, q) \bar{\psi} dx$$

of the mechanical quantities. Or, expressed in our present terminology, *when the energy is sharply known the time-behaviour of the system is absolutely stationary.*

On the other hand, any arbitrary time-dependent solution (cf. (224)) can be represented by a series (or an integral) with terms of the form

$$\int dE a(E) v(x, E) e^{\frac{2\pi i}{h}Et}.$$

We can then determine the behaviour of the probability $\psi\bar{\psi}$ or the mechanical mean values

$$\int \psi f(p, q) \bar{\psi} dx.$$

These quantities will be dependent * on the time. On the other hand, it is no longer permissible to speak of a definite energy of the system. This energy will now be able to assume all

* For example, if for two terms

$$\psi = a_1 v_1 e^{\frac{2\pi i}{h}E_1 t} + a_2 v_2 e^{\frac{2\pi i}{h}E_2 t}$$

$\psi\bar{\psi}$ will be of the form

$$\psi\bar{\psi} = A + B e^{\frac{2\pi i}{h}(E_1 - E_2)t} + C e^{\frac{2\pi i}{h}(E_2 - E_1)t}.$$

possible values according to the moment of time at which it is measured.*

In our examples we have discussed a limiting case which is characterized by the fact that *one quantity is always sharply defined and the other always completely indeterminate*. The more general case will, of course, consist in this, that the proper function for each canonically conjugate quantity has an allowable interval within which the probability is in general variable. For this case the following simple discussion shows us how to obtain an estimate of the size of the interval or the range of the probability respectively, with which the quantum-mechanical quantity can assume the individual values in this interval.

Let p and q be canonically conjugate. That is, according to (90) and (93) the relation

$$(pq - qp) = \frac{h}{2\pi i}\psi$$

holds for the quantum-mechanical operators.

We now assume that †

$$\int \psi q \bar{\psi} d\tau = \bar{q} = 0, \quad \int \psi p \bar{\psi} d\tau = \bar{p} = 0. \ddagger \quad (310)$$

Thus the mean value of q and p vanishes. So our interest lies, in the sense of the theory of probability, in the mean square deviation from the mean-value :

$$\overline{(q - \bar{q})^2} = q^2 - \bar{q}^2 = \int \psi q^2 \bar{\psi} d\tau = (\overline{\Delta q})^2 \quad (311a)$$

$$\overline{(q - \bar{p})^2} = \bar{p}^2 - \bar{p}^2 = \int \psi p^2 \bar{\psi} d\tau = (\overline{\Delta p})^2. \quad (311b)$$

* Expressed in the language of matrices this means that the Hamiltonian operator will not now give only diagonal terms in the energy matrix, which would be independent of the time, but also terms that correspond to transi-

tions, that is, also depend on the time in the form given by $e^{\frac{2\pi i}{h}(E_m - E_n)t}$.

† In this section the bar over ψ always means "conjugate complex," but when placed over the mechanical variables it means the theoretical mean value of the probability.

‡ This does not restrict the generality of the case. For example, if

$$\bar{q} = \int \psi q \bar{\psi} d\tau = a,$$

$$p = \int \psi p \bar{\psi} d\tau = b,$$

we need only introduce the new quantities

$$\begin{aligned} p - a &= q' \\ p - b &= p' \end{aligned}$$

which, on account of

$$(p'q' - q'p')\psi = \frac{h}{2\pi i}\psi,$$

are then again canonically conjugate to each other and for which $\bar{q}' = \bar{p}' = 0$ then certainly holds.

is a measure of the probability that a particle will pass beyond the potential barrier. If we again use the above loose form of expression we may say: suppose a wave coming from the left with energy E falls on the potential barrier for which the potential jump is greater than the kinetic energy of the particle. Now whereas in the classical theory no particle at all would be able to cross this barrier, this is no longer the case in the quantum theory. Even if the potential rise is arbitrarily great there is a probability which differs from zero for a particle of lesser kinetic energy to cross it.

After the many examples of the last sections this method of formulation undoubtedly gives a paradoxical impression, for in a certain sense it denotes a transgression of the energy law. It means that kinetic energy even of the smallest amount suffices to overcome a potential difference of arbitrarily great amount, although only in rare cases. This paradox has arisen solely through an incorrect manner of expression. The proper functions $e^{\pm ikx}$ which have been used and to which a definite amount of energy can rightly be assigned, by no means exhibit the property that a wave falls on the potential barrier "from the left." Moreover, from what has been said this would be quite impossible, for the statement "from the left" restricts the position of the particle even if only to a small extent. On the other hand, the proper function $e^{\pm ikx}$ sharply fixes the momentum. On account of the Uncertainty Principle this simultaneous statement is impossible. Actually the proper function e^{-ikx} denotes only that a wave propagates itself from left to right.

We shall illuminate this question from another angle. The assumption for the solution in the two partial regions

$$\begin{aligned} x < 0, & \quad \psi = e^{-ikx}, \\ x > 0, & \quad \psi = 0, \end{aligned}$$

contradicts the boundary conditions, according to which ψ and $\frac{d\psi}{dx}$ must be continuous for

$$x = 0.$$

It would represent the case of a plane wave incident on the potential barrier from the left. If, conversely, we really wish to represent the case of a plane wave falling on the barrier we must assume the proper function for $x < 0$ in the form (224). But then all possible values of the energies and momenta occur in it and it is not surprising that a particle can surmount the barrier. Expressed simply this means that only those particles

would cross the barrier whose energy is greater than the potential difference.

If, conversely, we maintain the condition of constant momentum it is no longer possible to impose a restriction on the position, and we can now only say that the whole of space is already filled with ψ -waves; and in physical language this means that even at the beginning the probability of encountering a particle on the side $x > 0$ was different from zero.

As a further illustration let us take the discussion on the properties of the angular momentum and its components. It was shown in § 38 that the z -component of the angular momentum measured in terms of $\frac{h}{2\pi}$ is given by the value m , where m can assume all the values from $\pm 0, 1, 2, \dots l$. Here l is determined by the equation

$$M^2 = \frac{h^2}{4\pi^2} l(l+1),$$

where M^2 denotes the square of the angular momentum. Using pictorial language we would argue as follows: Let us suppose the atom to be directed, say, by an external field in such a way that its angular momentum falls in the direction of the field which we shall take as the z -axis. We then have

$$M_z^2 = M^2 = \frac{h^2}{4\pi^2} m^2 \leq \frac{h^2}{4\pi^2} l^2.$$

This value is less than M^2 .

To account for this contradiction we must consider the following point. The angular momentum, say, in the z -direction,

$$M_z = xp_y - yp_x,$$

corresponds to the quantum-mechanical operator

$$M_z\psi = \frac{h}{2\pi i} \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \psi. \quad (321a)$$

In the same way we have

$$M_x\psi = \frac{h}{2\pi i} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \psi, \quad (321b)$$

$$M_y\psi = \frac{h}{2\pi i} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \psi. \quad (321c)$$

These quantum-mechanical operators M_z, M_y, M_x cannot be commuted with one another. For example, if we determine the difference

$$M_x M_y \psi - M_y M_x \psi,$$

we get

$$(M_x M_y - M_y M_x)\psi = \frac{h}{2\pi i} M_z \psi. \quad (322)$$

It is easy to confirm (322) by simply performing the differentiation. Of course the corresponding relations that result by cyclic variation of x, y, z also hold.

Hence there is no meaning in saying that we can direct an atom by means of an external field in such a way that its angular momentum coincides with the direction of the field. We cannot simultaneously make measurements of M_x and M_y , as these two quantities cannot be commutated with one another, which is necessary if such an orientation of the atom is to be effected. For this orientation would signify that

$$\begin{aligned} M_x &= M_y = 0, \\ M_z &= \frac{lh}{2\pi}. \end{aligned}$$

But we *can* simultaneously measure M_x and M^2 , for M^2 is equal to the quantum-mechanical operator

$$(M_x^2 + M_y^2 + M_z^2)\psi = M^2\psi.$$

This simple calculation may be left to the reader who can easily convince himself by performing the differentiations that the quantum-mechanical operators M_x and M^2 can actually be commutated. Hence we cannot regard it as a paradox that the maximum value of M_x^2 does not coincide with M^2 but rather as a beautiful and striking confirmation of the line of reasoning on which Heisenberg's Uncertainty Relation is founded.

73. Concluding Remarks on Wave-Mechanics

We are now in a position to give the objections to Schrödinger's theory which have prevented its complete success. The first is that in his theory the particle character of the atom does not receive adequate recognition. Whereas according to Schrödinger the atoms are distributed throughout the whole of space so that we cannot really speak of its having a definite position, we know, for example, from Wilson's cloud-track experiment and others that it is possible to define the position of the electron very well. In answer to this Schrödinger points out that his wave-packets are also only of very small extent and hence do not contradict the results of observation. But the instructive example of the oscillator, to which we referred earlier, is by no means appropriate for characterizing this feature generally. For in the

case of an electron beam of appreciably uniform velocity there is no 'wave-packet' in the ordinary sense present at all, but rather a charge-density $\psi\bar{\psi}$ uniformly distributed over the whole of space. Now if we form a proper function ψ so that, say, at the time $t = 0$ the charge-density differs from zero only over a relatively small region, this packet will not remain together permanently but will in time become scattered and will finally assume arbitrarily great dimensions. It is this circumstance that makes statistics so significant since it only regards the charge-density as the probability that particles are at certain places. In Schrödinger's theory, however, which accepts the idea of charge-density literally, a dilemma arises in that a single observation of an electron by which its position is approximately determined, as in Wilson's cloud-chamber, would in a certain sense require that the wave-packet should suddenly contract to a single point (a very small region). This action would then really be a physical process that would have to occur with a speed greater than that of light, that is, it would be a signal and would be much more open to the objection raised above by Einstein than the probability packets. To maintain Schrödinger's view we would then have to argue somewhat as follows: It is not true at all that a single wave-packet scatters itself in the course of time. Every wave-packet in itself holds together and pursues according to laws at present unknown to us the path which we observe in Wilson's cloud-chamber. The gaps in our knowledge of atomic physics do not yet enable us to specify the exact path followed by every wave-packet, so we must remain satisfied with statistical data. This interpretation cannot be refuted. All we can say is that it goes beyond the limits of experiment at present and that at the present stage of our knowledge we have no evidence for believing that such parameters of atomic events which are to impose laws on the course of individual processes, will ever become known to us.

A second objection to Schrödinger's belief in the reality of the charge-cloud and his rejection of the concept of stationary states is given by the fact of spontaneous emission. We shall illustrate this case by a very simple example. Suppose an atom to be in the ground-state and let light be incident on it whose frequency is such that absorption processes (in the language of the older quantum theory) caused transitions to the state 3 which has the proper function ψ_3 . We can imagine the incident radiation to be as weak as we please (characterized, say, by its amplitude a). We then know from the calculus of perturbations that ψ_3 will also be proportional to this amplitude a . Moreover (cf. the theory of dispersion, § 48), excitation without

the absorption caused by the accompanying vibration of the other states will also be produced and will likewise throughout be proportional to a . According to the older quantum theory and also to the present statistical interpretation, the process must be regarded as consisting in the following: the atoms in the state 3 will pass spontaneously to the states 2 and 1, emitting the corresponding frequencies; and the radiated intensities will be proportional to the probabilities of transition. These probabilities of transition are completely independent of whether the state 2 is excited originally or not. According to Schrödinger's theory, on the other hand, the probabilities of transition would be proportional to

$$w_{3 \rightarrow 1} \sim a, \quad w_{3 \rightarrow 2} \sim a^2$$

because

$$\psi_1 \sim a^0, \quad \psi_2 \quad \text{and} \quad \psi_3 \sim a.$$

Thus we see that the transitions from 3 to 1 and 3 to 2 have different probabilities according to the intensity of the incident rays. By reducing the intensity of radiation we can reduce the probability of the transition $3 \rightarrow 2$ to as small a value as we please, which is contrary to accepted views and to observation.*

Recapitulating, therefore, we may express it as the opinion of the great majority of theoretical physicists who are working in this field that the present state of quantum mechanics does not permit an interpretation which involves a continuum, such as Schrödinger has attempted to put forward. There is of course nothing to prevent our conjecturing or hoping that further developments will yet allow us to formulate a causal theory. To what extent we are to set such hopes on the electrodynamics which is yet to be worked out cannot be discussed here. But we cannot refrain from remarking that the present statistical interpretation of quantum mechanics exhibits no logical weaknesses which do not also show themselves in other branches of physics.

* It must not be forgotten that the Schrödinger waves do not occur in three-dimensional actual space but in f -dimensional phase-space. If we take into account the interaction between electrons this problem of wave-packets becomes very complicated in general, and can by no means be formulated anything like as simply as in the case of the one-electron problem, which is due only to the accidental coincidence of actual space and phase-space in this particular instance. This whole complex of problems has been very convincingly treated from the statistical point of view by Jordan, Klein and Wigner. Owing to the complexity of the subject we cannot here enter into a discussion of their papers.

BIBLIOGRAPHY

A. MONOGRAPHS AND REPORTS

1. BIGGS, H. F., *Wave Mechanics, an Introductory Sketch*. Humphrey Milford, 1927.
2. BIRTWISTLE, G., *The New Quantum Mechanics*. Cambridge University Press, 1928.
3. BOHR, N., *Das Quantenpostulat und die neuere Entwicklung der Atomistik*. *Naturwiss.*, **16**, 245, 1928.
4. BORN, M., *Mechanics of the Atom*. Bell.
5. — *Probleme der Atomdynamik*. Springer, Berlin, 1926.
6. BRILLOUIN, L., *La nouvelle mécanique atomique*. *Journ. de Physique* (6), **7**, 135, 1926.
7. DE BROGLIE, L., *Les principes de la nouvelle mécanique ondulatoire*. *Journ. de Physique*, **7**, 321, 1926.
8. — *Collected Papers on Wave Mechanics*. Blackie.
9. — *Ondes et Mouvements*. Paris: Gauthier-Villars, 1926.
10. DARROW, K., *Introduction to Wave-Mechanics*, *The Bell System Technical Journal*, **6**, 653, 1927.
11. FLAMM, L., *Die neue Mechanik*. *Naturwiss.*, **15**, 569, 1927.
12. HAAS, A., *Wave-Mechanics and the New Quantum Theory*. Constable.
13. HEISENBERG, W., *Quantenmechanik*. *Naturwiss.*, **14**, 989, 1926.
14. JORDAN, P., *Die Entwicklung der neuen Quantenmechanik*. *Ibid.*, **15**, 614, 636, 1927.
15. LANDÉ, A., *Neue Wege der Quantenmechanik*. *Ibid.*, **14**, 543, 1926.
16. SCHRÖDINGER, E., *Collected Papers on Wave-Mechanics*. Blackie.

B. ORIGINAL PAPERS

17. BORN, M., *Zeitschr. f. Phys.*, **38**, 803, 1926 ; **40**, 167, 1926.
18. — and JORDAN, P., *Zeitschr. f. Phys.*, **34**, 858, 1925.
- 18a. — HEISENBERG, W., and JORDAN, P., *ibid.*, **35**, 557, 1925.
19. — and WIENER, N., *ibid.*, **43**, 172, 1927.
20. DE BROGLIE, L., *Thèses*. Paris, 1924 ; *Annales de Physique*, **10**, 22, 1925.
21. DAVISSON, C., and GERMER, L. H., *Nature*, **119**, 558, 1927 ; *Phys. Rev.* (2), **30**, 705, 1927.

22. DIRAC, P. A. M., *Proc. Roy. Soc.*, **109**, 642, 1925 ; **110**, 561, 1926 ;
' **112**, 661, 1926 ; **113**, 621, 1926 ; **114**, 243, 710, 1927 ; **117**,
610, 1928 ; **118**, 351, 1928.
23. EINSTEIN, A., *Berlin Ber.*, 1924, 261 ; 1925, 3, 18.
24. — *Ibid.*, 1923, 359.
25. ELSASSER, W., *Naturwiss.*, **13**, 711, 1925.
26. HEISENBERG, W., *Zeitschr. f. Phys.*, **33**, 879, 1925.
27. — *Ibid.*, **38**, 411, 1926.
28. — *Ibid.*, **39**, 499, 1926.
29. — *Ibid.*, **43**, 172, 1927.
30. — and JORDAN, P., *ibid.*, **35**, 812, 1926.
31. KRAMERS, H. A., *Dissertation. Kgl. Danske Vidensk. Selsk. Skrifter, Naturvidensk. Afd.*, **8**, Raekke, III, 3, 287, 1919.
32. — *Nature*, **113**, 673 ; **114**, 310, 1924.
33. — and HEISENBERG, W., *Zeitschr. f. Phys.*, **31**, 681, 1925.
34. LADENBURG, R., *ibid.*, **4**, 451, 1921.
35. PAULI, W., *ibid.*, **43**, 601, 1927.
36. RAMSAUER, C., *Ann. d. Physik*, **64**, 513, 1921 ; **66**, 546, 1921 ; **72**,
345, 1923.
TOWNSEND, J. S., and BAILEY, V. A., *Phil. Mag.*, **42**, 873, 1921.
BROSE, H. L., and SAAYMAN, E. H., *Ann. d. Phys.*, **5**, 797-852, 1930.
37. SCHRÖDINGER, E., *Ann. d. Phys.*, **79**, 361, 1926.
38. — *Ibid.*, **79**, 489, 1926.
39. — *Ibid.*, **79**, 734, 1926.
40. — *Ibid.*, **80**, 437, 1926.
41. — *Ibid.*, **81**, 109, 1926.
42. — *Ibid.*, **82**, 257, 1926.
43. SOMMERFELD, A., *Zeitschr. f. Phys.*, **47**, 1, 1928.
44. — *Naturwiss.*, **16**, 374, 1928.

C. RECENT BOOKS

45. SOMMERFELD, A., *Wave-Mechanics*. Methuen. (See also 58.)
46. DE BROGLIE, L., *Wave-Mechanics*. Methuen. (See also 52.)
47. DIRAC, P. A. M., *The New Quantum Mechanics*. Oxford.
48. HEISENBERG, W., *The Physical Principles of the Quantum Theory*.
Cambridge.
49. SCHRÖDINGER, *Four Lectures on Wave Mechanics*. Blackie.
50. MARCH, A., *Grundlagen der Quantenmechanik*. Barth, Leipzig.
51. BORN-JORDAN, *Elementare Quantenmechanik*. Berlin : Julius
Springer, 1929.
52. DE BROGLIE (übersetzt von R. PEIERLS), *Wellenmechanik*. Leipzig :
Akad. Verlagsges, 1929.
53. DARROW (übersetzt von RABINOWITSCH), *Einführung in die Wellen-
mechanik*. Leipzig : Hirzel, 1929.
54. FRENKEL, *Einführung in die Wellenmechanik*. Berlin : Julius
Springer, 1929.

55. GUTH, Quantenphysik. In the Handbuch der Physik of GEIGER-SCHEEL, **4**. Berlin : Julius Springer, 1929.
56. LANDÉ, Optik, Mechanik und Wellenmechanik, *ibid.*, **20**. Berlin : Julius Springer, 1928.
57. PLANCK-Heft of Naturwissenschaften. Heft 26. Berlin : Julius Springer, 1929.
58. SOMMERFELD, Atombau und Spektrallinien, Wellenmechanischer Ergänzungsband Braunschweig : Vieweg, 1929.
59. WEYL, Gruppentheorie und Quantenmechanik. Leipzig : Hirzel, 1928.
60. WENTZEL, Unperiodische Vorgänge in der Wellenmechanik. Comprehensive report in Phys. Z., **29**, 321, 1928.

D. ORIGINAL PAPERS

61. AUGER, Ann. de Phys., **6**, 183, 1926.
62. BETHE, *ibid.*, **87**, 55, 1928.
63. BOHR, Naturwissenschaften, **16**, 245, 1928.
64. BONHOEFFER and HARTECK, Berlin Ber., 103, 1929.
65. BOSE, S. N., Zeitschr. f. Physik, **27**, 384, 1924.
66. BOTHE-GEIGER, *ibid.*, **32**, 639, 1925.
67. COMPTON-SIMON, Phys. Rev., **26**, 289, 1925.
68. CONDON-GURNEY, *ibid.*, **33**, 127, 1929.
69. DARWIN, Proc. Lond. Phys. Soc., **116**, 227, 1927.
70. — *Ibid.*, **118**, 654, 1928.
71. DENNISON, *ibid.*, **115**, 483, 1927.
72. DIRAC, *ibid.*, **117**, 610 ; **118**, 351, 1928.
73. — *Ibid.*, **112**, 661, 1926.
74. — *Ibid.*, **114**, 243 u. 710, 1927.
75. — *Ibid.*, **113**, 621, 1927.
76. — Cambr. Phil. Soc., **23**, 500, 1926.
77. ECKART, Phys. Rev., **28**, 711, 1926.
78. — Proc. Nat. Acad., Washington, 1927.
79. EINSTEIN, Bericht vom Solvaykongress, 1927 (Diskussionsbemerkungen).
80. EUCKEN and HILLER, Zeitschr. f. phys. Chem. (B), **4**, 413, 1929.
81. FERMI, Zeitschr. f. Physik, **36**, 902, 1926.
82. FOWLER-NORDHEIM, Proc. Lond. Phys. Soc., **121**, 1928.
83. GAMOW, Zeitschr. f. Physik, **51**, 204, 1928.
84. GORDON, *ibid.*, **48**, 11, 1928.
85. — *Ibid.*, **48**, 180, 1928.
86. — *Ibid.*, **40**, 117, 1926.
87. GOUDSMIT-UHLENBECK, Naturw., **13**, 953, 1925.
88. HEISENBERG, Zeitschr. f. Physik, **41**, 239, 1927.
89. — *Ibid.*, **49**, 619, 1928.
90. — -JORDAN, *ibid.*, **37**, 263, 1926.
91. HEITLER, *ibid.*, **47**, 835, 1928.

92. HEITLER-LONDON, *Zeitschr. f. Physik*, **44**, 455, 1927.
- 92a. HUND, loc. cit., **42**, 93, 1927.
93. HYLLERAAS, *Zeitschr. f. Physik*, **54**, 347, 1929.
94. JORDAN, *ibid.*, **40**, 809, 1926 ; **44**, 1, 1927.
95. — -KLEIN, *ibid.*, **45**, 751, 1927.
96. — -WIGNER, *ibid.*, **47**, 631, 1928.
97. KELLNER, *ibid.*, **42**, 1927.
98. KIKUCHI, *Jap. Journ. Phys.*, **5**, 83, 1928.
99. KLEIN, *Zeitschr. f. Physik*, **37**, 895, 1926.
100. — -NISHINA, *ibid.*, **52**, 853, 1929.
101. LENNARD-JONES, *Proc. Lon. Phys. Soc.*, **105**, 640, 1924.
102. LONDON, *Zeitschr. f. Physik*, **46**, 455, 1927.
103. PAULI, *ibid.*, **35**, 765, 1925.
104. — *Ibid.*, **41**, 81, 1927.
105. RAMAN, cf. PRINGSHEIM, *Naturw.*, **16**, 597, 1928.
106. RUPP, *Ann. de Phys.*, **85**, 901, 1928.
107. — *Naturwissenschaften*, **17**, 174, 1929.
108. SCHRÖDINGER, *ibid.*, **11**, 1923.
109. — *Ibid.*, **14**, 664, 1926.
110. THOMAS, *Phil. Mag.*, **3**, 1, 1927.
111. THOMSON, G. P., *Proc. Lond. Phys. Soc.*, **117**, 600, 1928.
112. WIGNER, *Zeitschr. f. Physik*, **40**, 492 u. 883, 1927.

INDEX

- ABRAHAM, 57.
 Action at a distance, 190.
 Action variables, 7, 70.
 Alternating field, 19.
 Alternation law, Rydberg's, 73, 152.
 Amplitude equation, 61.
 Angle variables, 7.
 Anti-symmetrical functions, 143-148, 152, systems, 143, 152, 165.
 Atomic core, 76.
Aufpunkt (reference-point), 9.
 Auger effect, 158-159.
 Azimuthal quantum number, 17, 89, 100.
- BAILEY, 42.
 Balmer's series, 8, 73, 81, 115.
 Band-spectra, 152.
 Barnett, 179.
 Beat frequency, 49.
Besetzungszahlen, 164.
 Bethe, 127.
 Biot and Savart's law, 117.
 Bloch, 172.
 Bohr, 1, 4, 14, 17, 20, 33, 41, 48, 50, 51 69, 77, 79, 199.
 Bohr magneton, 77, 117.
 Boltzmann, 159, 160, 163.
 Bonhoeffer, 178.
 Born, 2, 6, 49, 108, 125.
 Bose, 163.
 Bothe and Geiger, 139.
 Brillouin, 114.
 Broglie de, wave-length, 35, 43.
 Brose, 42.
 Burgers and Dorgelo, 152.
- CANONICAL equations, 6, transformations, 6.
 Canonically conjugate quantities, 188.
 Causality, 191.
 Charge-cloud, 86, 184, -density, 116, 205.
 Classical statistics, 159.
 Components, Fourier and vector, 11, sp. intensity, 11.
 Compton and Simon, 139.
 Compton effect, 138, 198.
 Conditionally periodic motions, 7.
 Conjugate complex quantities, 50, 196.
 Correspondence principle, 20, 22, 23, 33, 34, 52, 74, 136.
- Courant and Hilbert, 85, 104, 107.
 Current numbers, 72, 88.
- DARWIN, 101, 118.
 Davisson and Germer, 41.
 Debye and Scherrer, 126.
 Degenerate systems, 79, 116, 134.
 Dennison, 177.
 Diagonal matrix, 26.
 Diamagnetic gases, 90.
 Diffraction of matter, 123.
 Dipole-moment, 144.
 Dirac, 2, 83, 90, 108, 118, 183, 187.
 Direction quantizing, 75.
 Dispersion, 21, 134.
 Doppler effect, 199.
 Doublets, 74.
 Dulong and Petit's law, 171.
- ECKART, 183.
 Einstein, 42, 136, 179, 190, 205, imaginary, experiment, 190, photo-electric equation, 138, 168.
 Electric moment, 12.
 Electrons, 101.
 Entropy, 160.
 Equivalent electrons, 79.
 Eucken, 178.
 Exchange degeneracy, 142.
- FERMAT'S Principle, 40, 43.
 Fermi, 163, 170, 171.
 Ferro-magnetism, 178-180.
 Field-scalar, 42, 50.
 Fine-structure, 69, 73, 81-83, 123.
 Forbidden transitions, 16.
 Forced vibrations, 61.
 Four-current, 84.
 Fourier series, 10, 22, 70.
 Frequency condition, 14, 26, 48, 50.
 Frequency, orbital and line, 20.
- GAMOW, 131.
 Gauss's theorem, 86.
 Geiger and Nuttall's relation, 131.
 Generalized co ordinates and momenta, 4.
 Gerlach, 75.
Gesetzlichkeit, 191.

Giauque and Johnston, 154.

Gordon, 66, 118.

Green's theorem, 85.

Group-velocity, 39.

Gurney and Condon, 130.

HAAS de, 179.

Half-integral numbers, 108.

Hamiltonian function, 4, 5, 31, 34, 52, 60, 91.

Harmonic components, 10.

Harteck, 178.

Heisenberg, 2, 11, 16, 20, 94, 122, 142, uncertainty relation, 128, 134, 188, 197, 201.

Hermitean matrix, 25, polynomials, 107.

Hertz's formulæ, 8, 9, 51.

Holtmark, 42.

Hund, 77, 177.

Hydrogen atom, charge-density and proper functions of, 116.

Hylleraas, 152.

IDEAL gas, 165.

Imaginary, 24.

Impact parameter, 123.

Indeterminacy, 194.

Infinite matrices, 24.

Inner quantum number, 73.

Interference, 41.

JORDAN, 2, 122, 181, 183, 187, 207.

KELLNER, 152.

Kepler orbits, 69, 71.

Kikuchi, 126.

Klein, 66, 181, 207.

Klein and Nishina, 141.

Kramers, 16, 33, 52, 136.

Kreisfrequenz, 8.

LAGUERRE polynomial, 116.

Landé, 81, factor of, 121, 122.

Larmor precession, 75.

Latenzzeit (retarded time), 2.

Laue, 126, 127.

Laufzahlen, 72, 88.

Legendre polynomial, 107, 115.

Leuchtelektron, 71.

Levels of sodium spectrum, 72.

Linear oscillators, 142.

Lissajou figures, 7.

London, 158.

Lorentz-Einstein transformation, 65, 99.

MAGNETIC quantum number, 75, 88.

Mass-cloud, 88.

Material constant, 45, waves, 35.

Matrices, 24-33, 94, 196.

Maupertius, 6.

Maxwell, 45, 141.

Maxwellian distribution, 169.

Metallic electrons, 171-173.

Millikan and Bowen, 81.

Minkowski, 55.

Multiple moments, 144, quantizing, 69, structure, 121.

Multiplication difference, 26.

NEGATIVE mass, 101.

Newton's equations of motion, 92.

Nodal lines, 42, planes, 49, surfaces, 49.

Non-conservative systems, 64.

ONDES de phase, 38.

Operators, 52, 53, 183.

Orbital moment of momentum, 118.

Ortho-system of helium, 150.

PARABOLIC orbits, 48.

Paramagnetism, 175.

Para-system of helium, 150.

Partial differential equation of Hamilton and Jacobi, 62, 185.

Paschen-Back effect, 33, 78, 121.

Pauli, 101, 175.

Pauli's exclusion principle, 76, 148, 158, 176.

Pendulum orbits, 72, 79, 89.

Pereils, 172.

Persistence of normalization, 86, of orthogonality, 86.

Perturbation, 108, 131, 145, functions, 143.

Phase-waves, 36, 38.

Photo-electric limit, 138, 168, 173.

Planck's constant, 14, radiation law, 168.

Polarization, 12.

Potential barrier, 128, 129, 203.

Poynting's vector, 9.

Principle of combination, 23, 26, of least action, 6, 40, of least time, 40.

Probabilities of transition, 15.

Probekörper, 193.

Proper functions, 47, values, 42, 46, 111-114, 132.

Proton, 155.

QUADRATIC matrices, 24.

Quantum conditions, 16, 41, numbers, 17.

Quellpunkt, 9.

RADIATING electron, 71.

Raman effect, 137.

- Ramsauer, 42.
 Reference-point, 9.
 Reiche, 167.
 Relativity, 36.
 Retarded time, 3, 9.
 Richardson, 128, 173, 179, effect, 128, 173.
 Ritz, 23, 152.
 Röntgen rays, scattering of, 105.
 Rotation, heat of, 177.
 Rotator, 107.
 Rupp, 126.
 Rutherford, 123, model, 13
 Rydberg's alternation law,
- SAAYMAN, 42.
 Scalar potential, 65.
 Scattering of α -particles, 123.
 Schrödinger, 2, 11, 16, 43, 44, 66, 92, 108,
 114, 124, 146, 183, 185, 194, 205, 207,
 wave-equation, 45, 60.
 Screening action, 118.
 Self-adjoint differential expression, 85.
 Sodium spectrum, 72.
 Sommerfeld, 9, 57, 67, 73, 81, 114, 123,
 172.
 Sommerfeld-Wilson condition, 17.
 Source-point, 9.
 Spectral series, 75, terms, 14.
 Spherical harmonics, 107, pendulum, 106.
 Spin momentum, 78.
 Spinning electron, 77, 151.
 Spin quantum number, 78, term, 98.
 Stark effect, 19.
 State of excitation, 49.
 Stationary states, 14.
- Stern and Gerlach, 75.
 Symmetrical functions, 143-148, 152, 165.
- TEST body, 193.
 Theory of transformations, 91.
 Thomas, 117.
 Thomson, G. P.,
 Thomson, J. J., 14.
 Townsend, 42.
 Information theory, 187.
 Transition probabilities, 145.
- SENBEEK and Goudsmit, 77.
 t matrix, 27.
- VERSCHMIERT, 86.
 Vibration equation, 61.
- WAVE equation, 45, 59, group, 38, 39,
 packet, 184, 205.
 Weight-factors, 161.
 Wentzel, 114, 159.
 Weyl, 197.
 Wiedemann-Franz law, 171.
 Wiener, 183.
 Wigner, 181, 207.
 Wilson, cloud track, 205, and Sommer-
 feld's condition, 17.
 World, Minkowski's, 55.
- ZEEMAN effect, 68, 78, 91, 121

