

COLLECTED PAPERS ON WAVE MECHANICS

ERWIN SCHRÖDINGER

AMS CHELSEA PUBLISHING
American Mathematical Society • Providence, Rhode Island



COLLECTED PAPERS
ON
WAVE MECHANICS

BY
E. SCHRÖDINGER
PROFESSOR OF THEORETICAL PHYSICS AT THE UNIVERSITY OF BERLIN

TOGETHER WITH HIS

FOUR LECTURES
ON
WAVE MECHANICS

CHELSEA PUBLISHING COMPANY
NEW YORK, N. Y.

THIRD (AUGMENTED) EDITION

FIRST GERMAN EDITION, LEIPZIG, 1927

SECOND GERMAN EDITION, LEIPZIG, 1928

FIRST ENGLISH EDITION, LONDON AND GLASGOW, 1928

SECOND (UNALTERED) ENGLISH EDITION, NEW YORK, 1978

THIRD (AUGMENTED) ENGLISH EDITION, NEW YORK, 1982

THIS THIRD ENGLISH EDITION INCORPORATES THE FULL TEXT OF A WORK ORIGINALLY PUBLISHED AS A SEPARATE WORK: 'FOUR LECTURES ON WAVE MECHANICS.' THIS WAS ORIGINALLY PUBLISHED AT GLASGOW IN 1928 AND REPRINTED AT GLASGOW IN 1929. THE PRESENT WORK IS PRINTED ON 'LONG-LIFE' ACID-FREE PAPER.

COPYRIGHT ©, 1982 BY CHELSEA PUBLISHING COMPANY
LIBRARY OF CONGRESS CATALOG CARD NUMBER 80-70108
INTERNATIONAL STANDARD BOOK NUMBER 0-8284-1302-9

PRINTED IN THE UNITED STATES OF AMERICA

Contents

PAPERS

	PAGE
ABSTRACT	ix
1. QUANTISATION AS A PROBLEM OF PROPER VALUES. PART I	1
2. QUANTISATION AS A PROBLEM OF PROPER VALUES. PART II	13
3. THE CONTINUOUS TRANSITION FROM MICRO- TO MACRO-MECHANICS	41
4. ON THE RELATION BETWEEN THE QUANTUM MECHANICS OF HEISENBERG, BORN, AND JORDAN, AND THAT OF SCHRÖDINGER	45
5. QUANTISATION AS A PROBLEM OF PROPER VALUES. PART III	62
6. QUANTISATION AS A PROBLEM OF PROPER VALUES. PART IV	102
7. THE COMPTON EFFECT	124
8. THE ENERGY-MOMENTUM THEOREM FOR MATERIAL WAVES	130
9. THE EXCHANGE OF ENERGY ACCORDING TO WAVE MECHANICS	137

LECTURES

1. DERIVATION OF THE FUNDAMENTAL IDEA OF WAVE MECHANICS FROM HAMILTON'S ANALOGY BETWEEN ORDINARY MECHANICS AND GEOMETRICAL OPTICS	155
2. ORDINARY MECHANICS ONLY AN APPROXIMATION, WHICH NO LONGER HOLDS FOR VERY SMALL SYSTEMS	160
3. BOHR'S STATIONARY ENERGY-LEVELS DERIVED AS THE FREQUENCIES OF PROPER VIBRATIONS OF THE WAVES.	163

4. ROUGH DESCRIPTION OF THE WAVE-SYSTEMS IN THE HYDROGEN ATOM. DEGENERACY. PERTURBATION	168
5. THE PHYSICAL MEANING OF THE WAVE FUNCTION. EXPLANATION OF THE SELECTION RULES AND OF THE RULES FOR THE POLARIZATION OF SPECTRAL LINES.....	170
6. DERIVATION OF THE WAVE EQUATION (PROPERLY SPEAKING) WHICH CONTAINS THE TIME	176
7. AN ATOM AS PERTURBED BY AN ALTERNATING ELECTRIC FIELD	177
8. THEORY OF SECONDARY RADIATION AND DISPERSION.....	181
9. THEORY OF RESONANCE RADIATION, AND OF CHANGES OF THE STATE OF THE ATOM PRODUCED BY INCIDENT RADIATION WHOSE FREQUENCY COINCIDES, OR NEARLY COINCIDES, WITH A NATURAL EMISSION FREQUENCY	185
10. EXTENSION OF WAVE MECHANICS TO SYSTEMS OTHER THAN A SINGLE MASS-POINT	189
11. EXAMPLES: THE OSCILLATOR, THE ROTATOR.....	192
12. CORRECTION FOR MOTION OF THE NUCLEUS IN THE HYDROGEN ATOM.....	197
13. PERTURBATION OF AN ARBITRARY SYSTEM	199
14. INTERACTION BETWEEN TWO ARBITRARY SYSTEMS.....	203
15. THE PHYSICAL MEANING OF THE GENERALIZED ψ -FUNCTION	205

Preface to the First (German) Edition

REFERRING to these six papers (the present reprint of which is solely due to the great demand for separate copies), a young lady friend recently remarked to the author: "When you began this work you had no idea that anything so clever would come out of it, had you?" This remark, with which I wholeheartedly agreed (with due qualification of the flattering adjective), may serve to call attention to the fact that the papers now combined in one volume were originally written *one by one* at different times. The results of the later sections were largely unknown to the writer of the earlier ones. Consequently, the material has unfortunately not always been set forth in as orderly and systematic a way as might be desired, and further, the papers exhibit a gradual development of ideas which (owing to the nature of the process of reproduction) could not be allowed for by any alteration or elaboration of the earlier sections. The *Abstract* which is prefixed to the text may help to make up for these deficiencies.

The fact that the papers have been reprinted without alteration in no way implies that I claim to have succeeded in establishing a theory which, though capable of (and indeed requiring) extension, is firmly based as regards its physical foundations and henceforth admits of no alteration in its fundamental ideas. On the contrary, this comparatively cheap method of issue seemed advisable on account of the impossibility at the present stage of giving a fresh exposition which would be really satisfactory or conclusive.

E. SCHRÖDINGER.

ZÜRICH, *November* 1926.

Publishers' Note

THIS translation has been prepared from the second edition of the author's *Abhandlungen zur Wellenmechanik*, published by Johann Ambrosius Barth, 1928. These papers include practically all that Professor Schrödinger has written on Wave Mechanics.

The translation has been made by J. F. Shearer, M.A., B.Sc., of the Department of Natural Philosophy in the University of Glasgow, and W. M. Deans, B.A., B.Sc., late of Newnham College, Cambridge.

The translators have tried to follow the original as closely as the English idiom would permit. The English version has been read by Professor Schrödinger. Throughout the book *Eigenfunktion* has been translated *proper function*, and *Eigenwert*, *proper value*. The phrase *eine stückweise stetige Funktion* has been translated *a sectionally continuous function*. These equivalents were decided upon after consultation with the author and with several English mathematicians of eminence.

Abstract

(The references are to pages.)

THE Hamiltonian analogy of mechanics to optics (pp. 13-18) is an analogy to *geometrical* optics, since to the path of the representative point in configuration space there corresponds on the optical side the *light ray*, which is only rigorously defined in terms of geometrical optics. The undulatory elaboration of the optical picture (pp. 19-30) leads to the surrender of the idea of the *path of the system*, as soon as the dimensions of the path are *not* great in comparison with the wave-length (pp. 25-26). Only when they are so does the idea of the path remain, and with it classical mechanics as an approximation (pp. 20-24, 41-44); whereas for "micro-mechanical" motions the fundamental equations of mechanics are just as useless as geometrical optics is for the treatment of diffraction problems. In analogy with the latter case, a *wave equation* in configuration space must replace the fundamental equations of mechanics. In the first instance, this equation is stated for purely periodic vibrations sinusoidal with respect to time (p. 27 *et seq.*); it may also be derived from a "Hamiltonian variation principle" (p. 1 *et seq.*, pp. 11-12). It contains a "proper value parameter" E , which corresponds to the mechanical energy in macroscopic problems, and which for a single time-sinusoidal vibration is equal to the frequency multiplied by Planck's quantum of action h . In general the wave or vibration equation possesses *no* solutions, which together with their derivatives are one-valued, finite, and continuous throughout configuration space, *except* for certain special values of E , the *proper values*. These values form the "proper value spectrum" which frequently includes continuous parts (the "band spectrum", not expressly considered in most formulae: for its treatment see p. 112 *et seq.*) as well as discrete points (the "line spectrum"). The proper values either turn out to be identical with the "energy levels" (=spectroscopic "term"-value multiplied by h) of the quantum theory as hitherto developed, or differ from them in a manner which is confirmed by experience. (Unperturbed Keplerian motion pp. 1-12; harmonic oscillator, pp. 30-34; rigid rotator, pp. 35-36; non-rigid rotator, pp. 36-40; Stark effect, pp. 76-82, 93-96.) Deviations of the kind mentioned are, *e.g.*, the appearance of non-integral quantum numbers

(viz. the halves of odd numbers) in the case of the oscillator and rotator, and further, the non-appearance of the "surplus" levels (viz. those with vanishing azimuthal or equatorial quantum number) in the Kepler problem. Even in these matters the agreement with Heisenberg's quantum mechanics is complete: this can be proved in general (see below and pp. 45-61). For the calculation of the proper values and the corresponding solutions of the vibration equation ("proper functions") in more complicated cases, there is developed a *theory of perturbations*, which enables a more difficult problem to be reduced by quadratures alone to a "neighbouring" but simpler one (pp. 64-76). To "degeneracy" corresponds the appearance of *multiple* proper values (p. 11, p. 33 *et seq.*). Especially important physically is the case where, as, *e.g.*, in the Zeeman and Stark effects, a multiple proper value is split up by the addition of perturbing forces (general case, pp. 69-76; Stark effect, pp. 93-96).

Up till now the function ψ has merely been defined in a purely formal way as obeying the above-mentioned wave equation, serving as its object, so to speak. It is necessary to ascribe to ψ a physical, namely an electromagnetic, meaning, in order to make the fact that a small mechanical system can emit *electromagnetic* waves of a frequency equal to a term-difference (difference of two proper values divided by \hbar) intelligible at all, and further, in order to obtain a theoretical statement for the intensity and polarisation of these electromagnetic waves. This meaning, for the general case of a system with an arbitrary number of degrees of freedom, is not clearly worked out until the end of the sixth paper (pp. 120-123; a preliminary attempt for the *one*-electron problem, on p. 60 *et seq.*, turned out incomplete). A definite ψ -distribution in configuration space is interpreted as a continuous distribution of electricity (and of electric current density) in actual space. If from this distribution of electricity we calculate the component of the electric moment of the whole system in any direction in the usual way, it appears as the sum of single terms, each of which is associated with a *couple* of proper vibrations, and vibrates in a purely sinusoidal manner with respect to the time with a frequency equal to the difference of the allied proper frequencies (p. 60 *et seq.*, where $\tilde{\psi}$ is to be replaced by ψ). This simplifies the calculation without essentially modifying it). If the wave-length of the *electromagnetic* waves, associated with this difference frequency, is large compared with the dimensions of the region to which the whole distribution of electricity is practically confined, then, according to the rules of ordinary electrodynamics, the amplitude of the partial moment in question (or, more accurately, the square of this amplitude multiplied by the fourth power of the frequency) is a measure of the intensity of the light radiated with *this* frequency, and with *this* direction of polarisation. The electrodynamic hypothesis concerning ψ , and the related purely classical calculation of the radiation, are verified by experience, in so far as they furnish the customary *selection and polarisation rules* for the oscillator, rotator and the hydrogen atom

(easy to show from the results of p. 30 *et seq.*, p. 35 *et seq.*, and of pp. 1-12; cf. p. 101). Further, they also furnish satisfactory intensity relations for the fine structure of the Balmer lines in an electric field (p. 82 to p. 92). If only *one* proper vibration or only proper vibrations of *one* proper frequency are excited, then the electrical distribution becomes *static*, yet *stationary currents* may possibly be superimposed (magnetic atoms, p. 123). In this manner the *stability* of the normal state and its *lack of radiation* are explained.

The amplitudes of the partial moments are closely connected with those quantities ("matrix elements"), which determine the radiation, according to the formal theory of Heisenberg, Born, and Jordan. There can be demonstrated a far-reaching formal identity of the two theories (pp. 45-61), according to which not only do the calculated emission frequencies and selection and polarisation rules agree, but also the above-mentioned successful results of the intensity calculations are to be credited as much to the matrix theory as to the present one.

Everything up till now has referred in the first instance only to *conservative* systems, although some parts have reached their final formulation only in the sixth paper in connection with the treatment of non-conservative systems. For the latter, the wave equation used hitherto must be generalised into a *true* wave equation, which contains the *time* explicitly, and is valid not merely for vibrations purely sinusoidal with respect to time (with a frequency which appears in the equation as a proper value parameter), but for any arbitrary dependence on the time (pp. 102-104). From the wave equation generalised in this way, the interaction of the system with an incident light wave can be deduced, and hence a rational dispersion formula (pp. 104-117); in all this the electrodynamic hypothesis about ψ is retained. The generalisation for an arbitrary disturbance is indicated (p. 117 *et seq.*). Further, from the generalised wave equation an interesting conservation theorem for the "weight function" $\psi\bar{\psi}$ can be obtained (p. 121), which demonstrates the complete justification of the electrodynamic hypothesis frequently mentioned above, and which makes possible the deduction of the expressions for the components of the electric current density, in terms of the ψ -distribution (p. 122 *et seq.*).

Even the systems treated in the first five papers cannot be conservative in the literal sense of the word, inasmuch as they radiate energy; this must be accompanied by a change in the system. Thus there still seems to be something lacking in the wave law for the ψ -function,—corresponding to the "reaction of radiation" of the classical electron theory, which may result in a dying away of the higher vibrations in favour of the lower ones (p. 116). This necessary complement is still *missing*.

The form of the theory discussed so far corresponds to classical (*i.e.* non-relativistic) mechanics, and does not take magnetic fields into consideration. Therefore, neither the wave equation nor the

components of the four-current are invariant for the Lorentz transformation. For the *one-electron* problem an immediate relativistic-magnetic generalisation is readily suggested (pp. 118-120; the Lorentz-invariant expressions for the components of the four-current are not given in the text, but they can be got¹ from the "equation of continuity", which is to be formed in a way quite analogous to that in the non-relativistic case; cf. p. 122). Though this generalisation yields *formally* reasonable expressions for the wave lengths, polarisations, intensities, and selection in the natural fine structure and in the Zeeman pattern of the hydrogen atom, yet the actual diagram turns out quite wrong, for the reason that "half integers" appear as azimuthal quantum numbers in the Sommerfeld fine structure formula (p. 9 and p. 119; here the results only are given; V. Fock carried out the calculations quite independently in Leningrad, before my last paper was sent in, and also succeeded in deriving the relativistic equation from a variation principle. *Zeitschrift für Physik*, 38, p. 242, 1926). A correction is therefore necessary; all that can be said about it at present is that it must have the same significance for wave mechanics as the "spinning electron" of Uhlenbeck and Goudsmit has for the older quantum theory dealing with electronic orbits (p. 63); with this difference, however, that in the latter, together with the introduction of the "spinning electron", the half-integral form of the azimuthal quantum number must be postulated *ad hoc*, in order to avoid serious conflict with experiment even in the case of hydrogen; while wave mechanics (and also Heisenberg's quantum mechanics) *necessarily* yields halves of odd integers (German: *Halbzahligkeit*), and thus gives a hint, from the very beginning, of that further extension, which under the regime of the older theory was only shown to be necessary by more complicated phenomena, such as the Paschen-Back effect in hydrogen, anomalous Zeeman effects, structures of multiplets, the laws of Röntgen doublets and the analogy between them and the alkali doublets.

Addition in the second (German) edition: the first and second of the three new papers now added, namely, "The Compton Effect" and "The Energy-Momentum Theorem for Material Waves", are contributions to the four-dimensional relativistic form of wave mechanics discussed in the above paragraph. In connection with the first of these papers I should like above all to remark that, as Herr Ehrenfest has pointed out to me, the figure (p. 128) is incorrect: the pair of wave trains represented in the right half of the figure should coincide *completely* with the pair on the left, in respect of wave length and the orientation of their planes as well as in breadth of interference fringes (the broken lines).—The second paper, that on "The Energy-Momentum Theorem", throws a strong light on the difficulties which a merely *four-dimensional* theory of ψ -waves comes up against, despite the formally beautiful possibilities of development which present themselves here.—In the last paper, on "The Exchange of

¹ Cf. also a paper by W. Gordon on the Compton Effect, *Ztschr. f. Physik*, 40, p. 117, 1926.

Energy according to Wave Mechanics", the *many*-dimensional, non-relativistic form is again used. This paper is a first attempt to find out whether, with reference to Heisenberg's important discovery of the "quantum mechanics resonance phenomenon", it should not be possible to regard those very phenomena which seem to be decisive evidence for the existence of discrete energy levels, without this hypothesis, *merely* as resonance phenomena.

WAVE MECHANICS

Quantisation as a Problem of Proper Values (Part I)

(*Annalen der Physik* (4), vol. 79, 1926)

§ 1. IN this paper I wish to consider, first, the simple case of the hydrogen atom (non-relativistic and unperturbed), and show that the customary quantum conditions can be replaced by another postulate, in which the notion of "whole numbers", merely as such, is not introduced. Rather when integralness does appear, it arises in the same natural way as it does in the case of the *node-numbers* of a vibrating string. The new conception is capable of generalisation, and strikes, I believe, very deeply at the true nature of the quantum rules.

The usual form of the latter is connected with the Hamilton-Jacobi differential equation,

$$(1) \quad H\left(q, \frac{\partial S}{\partial q}\right) = E.$$

A solution of this equation is sought such as can be represented as the *sum* of functions, each being a function of one only of the independent variables q .

Here we now put for S a new unknown ψ such that it will appear as a *product* of related functions of the single co-ordinates, *i.e.* we put

$$(2) \quad S = K \log \psi.$$

The constant K must be introduced from considerations of dimensions; it has those of *action*. Hence we get

$$(1') \quad H\left(q, \frac{K}{\psi} \frac{\partial \psi}{\partial q}\right) = E.$$

Now we do *not* look for a solution of equation (1'), but proceed as follows. If we neglect the relativistic variation of mass, equation (1') can always be transformed so as to become a quadratic form (of ψ and its first derivatives) equated to zero. (For the *one-electron* problem

this holds even when mass-variation is not neglected.) We now seek a function ψ , such that for any arbitrary variation of it the integral of the said quadratic form, taken over the whole co-ordinate space,¹ is stationary, ψ being everywhere real, single-valued, finite, and continuously differentiable up to the second order. *The quantum conditions are replaced by this variation problem.*

First, we will take for H the Hamilton function for Keplerian motion, and show that ψ can be so chosen for *all positive*, but only for a *discrete set of negative values of E* . That is, the above variation problem has a discrete and a continuous spectrum of proper values.

The discrete spectrum corresponds to the Balmer terms and the continuous to the energies of the hyperbolic orbits. For numerical agreement K must have the value $h/2\pi$.

The choice of co-ordinates in the formation of the variational equations being arbitrary, let us take rectangular Cartesians. Then (1') becomes in our case

$$(1'') \quad \left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2 + \left(\frac{\partial\psi}{\partial z}\right)^2 - \frac{2m}{K^2}\left(E + \frac{e^2}{r}\right)\psi^2 = 0;$$

$$e = \text{charge, } m = \text{mass of an electron, } r^2 = x^2 + y^2 + z^2.$$

Our variation problem then reads

$$(3) \quad \delta J = \delta \iiint dx dy dz \left[\left(\frac{\partial\psi}{\partial x}\right)^2 + \left(\frac{\partial\psi}{\partial y}\right)^2 + \left(\frac{\partial\psi}{\partial z}\right)^2 - \frac{2m}{K^2}\left(E + \frac{e^2}{r}\right)\psi^2 \right] = 0,$$

the integral being taken over all space. From this we find in the usual way

$$(4) \quad \frac{1}{2}\delta J = \int df \delta\psi \frac{\partial\psi}{\partial n} - \iiint dx dy dz \delta\psi \left[\nabla^2\psi + \frac{2m}{K^2}\left(E + \frac{e^2}{r}\right)\psi \right] = 0.$$

Therefore we must have, firstly,

$$(5) \quad \nabla^2\psi + \frac{2m}{K^2}\left(E + \frac{e^2}{r}\right)\psi = 0,$$

and secondly,

$$(6) \quad \int df \delta\psi \frac{\partial\psi}{\partial n} = 0.$$

df is an element of the infinite closed surface over which the integral is taken.

(It will turn out later that this last condition requires us to supplement our problem by a postulate as to the behaviour of $\delta\psi$ at infinity, in order to ensure the existence of the above-mentioned *continuous* spectrum of proper values. See later.)

The solution of (5) can be effected, *for example*, in polar co-ordinates, r, θ, ϕ , if ψ be written as the *product* of three functions, each only of r , of θ , or of ϕ . The method is sufficiently well known. The function of the angles turns out to be a *surface harmonic*, and if that of r be called χ , we get easily the differential equation,

¹ I am aware this formulation is not entirely unambiguous.

$$(7) \quad \frac{d^2\chi}{dr^2} + \frac{2}{r} \frac{d\chi}{dr} + \left(\frac{2mE}{K^2} + \frac{2me^2}{K^2r} - \frac{n(n+1)}{r^2} \right) \chi = 0.$$

$n = 0, 1, 2, 3 \dots$

The limitation of n to integral values is *necessary* so that the surface harmonic may be *single-valued*. We require solutions of (7) that will remain finite for all non-negative real values of r . Now¹ equation (7) has *two* singularities in the complex r -plane, at $r=0$ and $r = \infty$, of which the second is an "indefinite point" (essential singularity) of *all* integrals, but the first on the contrary is not (for any integral). These two singularities form exactly the *bounding points of our real interval*. In such a case it is known now that the postulation of the *finiteness* of χ at the bounding points is equivalent to a *boundary condition*. The equation has *in general* no integral which remains finite at *both* end points; such an integral exists only for certain special values of the constants in the equation. It is now a question of defining these special values. This is the *jumping-off* point of the whole investigation.²

Let us examine first the singularity at $r=0$. The so-called *indicial* equation which defines the behaviour of the integral at this point, is

$$(8) \quad \rho(\rho - 1) + 2\rho - n(n + 1) = 0,$$

with roots

$$(8') \quad \rho_1 = n, \quad \rho_2 = -(n + 1).$$

The two canonical integrals at this point have therefore the exponents n and $-(n + 1)$. Since n is not negative, only the first of these is of use to us. Since it belongs to the greater exponent, it can be represented by an ordinary power series, which begins with r^n . (The other integral, which does not interest us, can contain a logarithm, since the difference between the indices is an integer.) The next singularity is at infinity, so the above power series is always convergent and represents a *transcendental integral function*. We therefore have established that :

The required solution is (except for a constant factor) a single-valued definite transcendental integral function, which at $r=0$ belongs to the exponent n .

We must now investigate the behaviour of this function at infinity on the positive real axis. To that end we simplify equation (7) by the substitution

$$(9) \quad \chi = r^\alpha U,$$

where α is so chosen that the term with $1/r^2$ drops out. It is easy to verify that then α must have one of the two values $n, -(n + 1)$. Equation (7) then takes the form,

¹ For guidance in the treatment of (7) I owe thanks to Hermann Weyl.

² For unproved propositions in what follows, see L. Schlesinger's *Differential Equations* (Collection Schubert, No. 13, Göschen, 1900, especially chapters 3 and 5).

$$(7') \quad \frac{d^2U}{dr^2} + \frac{2(\alpha+1)}{r} \frac{dU}{dr} + \frac{2m}{K^2} \left(E + \frac{e^2}{r} \right) U = 0.$$

Its integrals belong at $r=0$ to the exponents 0 and $-2\alpha-1$. For the α -value, $\alpha=n$, the *first* of these integrals, and for the second α -value, $\alpha=-(n+1)$, the *second* of these integrals is an integral function and leads, according to (9), to the desired solution, which is single-valued. We therefore lose nothing if we confine ourselves to *one* of the two α -values. Take, then,

$$(10) \quad \alpha = n.$$

Our solution U then, at $r=0$, belongs to the exponent 0. Equation (7') is called Laplace's equation. The general type is

$$(7'') \quad U'' + \left(\delta_0 + \frac{\delta_1}{r} \right) U' + \left(\epsilon_0 + \frac{\epsilon_1}{r} \right) U = 0.$$

Here the constants have the values

$$(11) \quad \delta_0 = 0, \quad \delta_1 = 2(\alpha+1), \quad \epsilon_0 = \frac{2mE}{K^2}, \quad \epsilon_1 = \frac{2me^2}{K^2}.$$

This type of equation is comparatively simple to handle for this reason: The so-called Laplace's transformation, which in general leads *again* to an equation of the *second* order, *here* gives one of the *first*. This allows the solutions of (7'') to be represented by complex integrals. The result ¹ only is given here. The integral

$$(12) \quad U = \int_L e^{zr} (z-c_1)^{\alpha_1-1} (z-c_2)^{\alpha_1-1} dz$$

is a solution of (7'') for a path of integration L , for which

$$(13) \quad \int_L \frac{d}{dz} [e^{zr} (z-c_1)^{\alpha_1} (z-c_2)^{\alpha_1}] dz = 0.$$

The constants $c_1, c_2, \alpha_1, \alpha_2$ have the following values. c_1 and c_2 are the roots of the quadratic equation

$$(14) \quad z^2 + \delta_0 z + \epsilon_0 = 0,$$

and

$$(14') \quad \alpha_1 = \frac{\epsilon_1 + \delta_1 c_1}{c_1 - c_2}, \quad \alpha_2 = -\frac{\epsilon_1 + \delta_1 c_2}{c_1 - c_2}.$$

In the case of equation (7') these become, using (11) and (10),

$$(14'') \quad c_1 = + \sqrt{\frac{-2mE}{K^2}}, \quad c_2 = - \sqrt{\frac{-2mE}{K^2}};$$

$$\alpha_1 = \frac{me^2}{K\sqrt{-2mE}} + n + 1, \quad \alpha_2 = - \frac{me^2}{K\sqrt{-2mE}} + n + 1.$$

The representation by the integral (12) allows us, not only to survey the asymptotic behaviour of the totality of solutions when r

¹ Cf. Schlesinger. The theory is due to H. Poincaré and J. Horn.

tends to infinity in a definite way, but also to give an account of this behaviour for one *definite* solution, which is always a much more difficult task.

We shall at first *exclude* the case where α_1 and α_2 are real integers. When this occurs, it occurs for both quantities simultaneously, and when, and only when,

$$(15) \quad \frac{me^2}{K\sqrt{-2mE}} = \text{a real integer.}$$

Therefore we assume that (15) is not fulfilled.

The behaviour of the totality of solutions when r tends to infinity in a definite manner—we think always of r becoming infinite through real positive values—is characterised¹ by the behaviour of the two linearly independent solutions, which we will call U_1 and U_2 , and which are obtained by the following *specialisations* of the path of integration L . In *each* case let z come from infinity and return there along the same path, in such a direction that

$$(16) \quad \lim_{z \rightarrow \infty} e^{zr} = 0,$$

i.e. the real part of zr is to become negative and infinite. In this way condition (13) is satisfied. In the *one* case let z make a circuit once round the point c_1 (solution U_1), and in the *other*, round c_2 (solution U_2).

Now for very large real positive values of r , these two solutions are represented *asymptotically* (in the sense used by Poincaré) by

$$(17) \quad \begin{cases} U_1 \sim e^{c_1 r} r^{-\alpha_1} (-1)^{\alpha_1} (e^{2\pi i \alpha_1} - 1) \Gamma(\alpha_1) (c_1 - c_2)^{\alpha_1 - 1}, \\ U_2 \sim e^{c_2 r} r^{-\alpha_2} (-1)^{\alpha_2} (e^{2\pi i \alpha_2} - 1) \Gamma(\alpha_2) (c_2 - c_1)^{\alpha_2 - 1}, \end{cases}$$

in which we are content to take the first term of the asymptotic series of integral negative powers of r .

We have now to distinguish between the two cases.

1. $E > 0$. This guarantees the non-fulfilment of (15), as it makes the left hand a pure imaginary. Further, by (14''), c_1 and c_2 also become pure imaginaries. The exponential functions in (17), since r is real, are therefore periodic functions which remain finite. The values of α_1 and α_2 from (14'') show that *both* U_1 and U_2 tend to zero like r^{-n-1} . *This must therefore be valid for our transcendental integralsolution U , whose behaviour we are investigating, however it may be linearly compounded from U_1 and U_2 .* Further, (9) and (10) show that the function χ , *i.e.* the transcendental integral solution of the *original* equation (7), always tends to zero like $1/r$, as it arises from U through multiplication by r^n . We can thus state:

The Eulerian differential equation (5) of our variation problem has, for every positive E , solutions, which are everywhere single-valued, finite, and continuous; and which tend to zero with $1/r$ at infinity, under continual oscillations. The surface condition (6) has yet to be discussed.

¹ If (15) is satisfied, at least one of the two paths of integration described in the text cannot be used, as it yields a vanishing result.

2. $E < 0$. In this case the possibility (15) is not *eo ipso* excluded, yet we will maintain that exclusion provisionally. Then by (14'') and (17), for $r \rightarrow \infty$, U_1 grows beyond all limits, but U_2 vanishes exponentially. Our integral function U (and the same is true for χ) will then remain finite if, and only if, U is identical with U_2 , save perhaps for a numerical factor. *This, however, can never be*, as is proved thus: If a closed circuit round both points c_1 and c_2 be chosen for the path L , thereby satisfying condition (13) since the circuit is *really closed* on the Riemann surface of the integrand, on account of $\alpha_1 + \alpha_2$ being an integer, then it is easy to show that the integral (12) represents *our integral function* U . (12) can be developed in a series of positive powers of r , which converges, at all events, for r sufficiently small, and since it satisfies equation (7'), it must coincide with the series for U . Therefore U is represented by (12) if L be a closed circuit round both points c_1 and c_2 . This closed circuit can be so distorted, however, as to make it appear additively combined from the two paths, considered above, which belonged to U_1 and U_2 ; and the factors are non-vanishing, 1 and $e^{2\pi i \alpha_1}$. Therefore U cannot coincide with U_2 , but must contain also U_1 . Q.E.D.

Our integral function U , which alone of the solutions of (7') is considered for our problem, is therefore not finite for r large, on the above hypothesis. Reserving meanwhile the question of *completeness*, i.e. the proving that our treatment allows us to find all the linearly independent solutions of the problem, then we may state:

For negative values of E which do not satisfy condition (15) our variation problem has no solution.

We have now only to investigate that discrete set of negative E -values which satisfy condition (15). α_1 and α_2 are then both integers. The first of the integration paths, which previously gave us the fundamental values U_1 and U_2 , must now undoubtedly be modified so as to give a non-vanishing result. For, since $\alpha_1 - 1$ is certainly positive, the point c_1 is neither a branch point nor a pole of the integrand, but an ordinary zero. The point c_2 can also become regular if $\alpha_2 - 1$ is also not negative. In *every* case, however, two suitable paths are readily found and the integration effected completely in terms of known functions, so that the behaviour of the solutions can be fully investigated.

Let

$$(15') \quad \frac{me^2}{K\sqrt{-2mE}} = l; \quad l = 1, 2, 3, 4 \dots$$

Then from (14'') we have

$$(14''') \quad \alpha_1 - 1 = l + n, \quad \alpha_2 - 1 = -l + n.$$

Two cases have to be distinguished: $l \leq n$ and $l > n$.

(a) $l \leq n$. Then c_2 and c_1 lose every singular character, but instead become starting-points or end-points of the path of integration, in order to fulfil condition (13). A third characteristic point here is at infinity (negative and real). Every path between two of these three points yields a solution, and of these three solutions there are two linearly in-

dependent, as is easily confirmed if the integrals are calculated out. In particular, the *transcendental integral solution* is given by the path from c_1 to c_2 . That *this* integral remains regular at $r=0$ can be seen at once without calculating it. I emphasize this point, as the actual calculation is apt to obscure it. However, the calculation does show that the integral becomes indefinitely great for positive, infinitely great values of r . One of the *other* two integrals remains *finite* for r large, but it becomes infinite for $r=0$.

Therefore when $l \leq n$ we get *no* solution of the problem.

(b) $l > n$. Then from (14'''), c_1 is a zero and c_2 a pole of the first order at least of the integrand. Two independent integrals are then obtained: one from the path which leads from $z = -\infty$ to the zero, intentionally avoiding the pole; and the other from the *residue* at the pole. The *latter* is the integral function. We will give its calculated value, but multiplied by r^n , so that we obtain, according to (9) and (10), the solution χ of the original equation (7). (The multiplying constant is arbitrary.) We find

$$(18) \quad \chi = f\left(r \frac{\sqrt{-2mE}}{K}\right); \quad f(x) = x^n e^{-x} \sum_{k=0}^{l-n-1} \frac{(-2x)^k (l+n-k)!}{k! (l-n-1-k)!}.$$

It is seen that this is a solution that can be utilised, since it remains finite for all real non-negative values of r . In addition, it satisfies the surface condition (6) because of its vanishing exponentially at infinity. Collecting then the results for E negative:

For E negative, our variation problem has solutions if, and only if, E satisfies condition (15). Only values smaller than l (and there is always at least one such at our disposal) can be given to the integer n , which denotes the order of the surface harmonic appearing in the equation. The part of the solution depending on r is given by (18).

Taking into account the constants in the surface harmonic (known to be $2n+1$ in number), it is further found that:

The discovered solution has exactly $2n+1$ arbitrary constants for any permissible (n, l) combination; and therefore for a prescribed value of l has l^2 arbitrary constants.

We have thus confirmed the main points of the statements originally made about the proper-value spectrum of our variation problem, but there are still deficiencies.

Firstly, we require information as to the completeness of the *collected* system of proper functions indicated above, but I will not concern myself with that in this paper. From experience of similar cases, it may be supposed that no proper value has escaped us.

Secondly, it must be remembered that the proper functions, ascertained for E positive, do not solve the variation problem as originally postulated, because they only tend to zero at infinity as $1/r$, and therefore $\partial\psi/\partial r$ only tends to zero on an infinite sphere as $1/r^2$. Hence the surface integral (6) is still of the same order as $\delta\psi$ at infinity. If it is desired therefore to obtain the continuous spectrum, another condition must be added to the *problem*, viz. that $\delta\psi$ is to vanish at

infinity, or at least, that it tends to a constant value independent of the direction of proceeding to infinity; in the latter case the surface harmonics cause the surface integral to vanish.

§ 2. Condition (15) yields

$$(19) \quad -E_l = \frac{me^4}{2K^2l^2}.$$

Therefore the well-known Bohr energy-levels, corresponding to the Balmer terms, are obtained, if to the constant K , introduced into (2) for reasons of dimensions, we give the value

$$(20) \quad K = \frac{h}{2\pi},$$

from which comes

$$(19') \quad -E_l = \frac{2\pi^2me^4}{h^2l^2}.$$

Our l is the principal quantum number. $n+1$ is analogous to the azimuthal quantum number. The splitting up of this number through a closer definition of the surface harmonic can be compared with the resolution of the azimuthal quantum into an "equatorial" and a "polar" quantum. These numbers *here* define the system of node-lines on the sphere. Also the "radial quantum number" $l-n-1$ gives exactly the number of the "node-spheres", for it is easily established that the function $f(x)$ in (18) has exactly $l-n-1$ positive real roots. The positive E -values correspond to the continuum of the hyperbolic orbits, to which one may ascribe, in a certain sense, the radial quantum number ∞ . The fact corresponding to this is the proceeding to infinity, under *continual* oscillations, of the functions in question.

It is interesting to note that the range, inside which the functions of (18) differ sensibly from zero, and outside which their oscillations die away, is of the *general order of magnitude* of the major axis of the ellipse in each case. The factor, multiplied by which the radius vector enters as the argument of the constant-free function f , is—naturally—the reciprocal of a length, and this length is

$$(21) \quad \frac{K}{\sqrt{-2mE}} = \frac{K^2l}{me^2} = \frac{h^2l}{4\pi^2me^2} = \frac{a_l}{l},$$

where a_l = the semi-axis of the l th elliptic orbit. (The equations follow from (19) plus the known relation $E_l = \frac{-e^2}{2a_l}$).

The quantity (21) gives the order of magnitude of the range of the roots when l and n are small; for then it may be assumed that the roots of $f(x)$ are of the order of unity. That is naturally no longer the case if the coefficients of the polynomial are large numbers. At present I will not enter into a more exact evaluation of the roots, though I believe it would confirm the above assertion pretty thoroughly.

§ 3. It is, of course, strongly suggested that we should try to connect the function ψ with some *vibration process* in the atom, which would more nearly approach reality than the electronic orbits, the real existence of which is being very much questioned to-day. I originally intended to found the new quantum conditions in this more intuitive manner, but finally gave them the above neutral mathematical form, because it brings more clearly to light what is really essential. The essential thing seems to me to be, that the postulation of "whole numbers" no longer enters into the quantum rules mysteriously, but that we have traced the matter a step further back, and found the "integralness" to have its origin in the finiteness and single-valuedness of a certain space function.

I do not wish to discuss further the possible representations of the vibration process, before more complicated cases have been calculated successfully from the new stand-point. It is not decided that the results will merely re-echo those of the usual quantum theory. For example, if the relativistic Kepler problem be worked out, it is found to lead in a remarkable manner to *half-integral partial quanta* (radial and azimuthal).

Still, a few remarks on the representation of the vibration may be permitted. Above all, I wish to mention that I was led to these deliberations in the first place by the suggestive papers of M. Louis de Broglie,¹ and by reflecting over the space distribution of those "phase waves", of which he has shown that there is always a *whole number*, measured along the path, present on each period or quasi-period of the electron. The main difference is that de Broglie thinks of progressive waves, while we are led to stationary proper vibrations if we interpret our formulae as representing vibrations. I have lately shown² that the Einstein gas theory can be based on the consideration of such stationary proper vibrations, to which the dispersion law of de Broglie's phase waves has been applied. The above reflections on the atom could have been represented as a generalisation from those on the gas model.

If we take the separate functions (18), multiplied by a surface harmonic of order n , as the description of proper vibration processes, then the quantity E must have something to do with the related *frequency*. Now in vibration problems we are accustomed to the "parameter" (usually called λ) being proportional to the *square* of the frequency. However, in the first place, such a statement in our case would lead to *imaginary* frequencies for the *negative E-values*, and, secondly, instinct leads us to believe that the energy must be proportional to the frequency itself and not to its square.

The contradiction is explained thus. There has been *no natural zero level* laid down for the "parameter" E of the variation equation (5), especially as the unknown function ψ appears multiplied by a function of r , which can be changed by a constant to meet a corresponding

¹ L. de Broglie, *Ann. de Physique* (10) 3, p. 22, 1925. (Thèses, Paris, 1924.)

² *Physik. Ztschr.* 27, p. 95, 1926.

change in the zero level of E . Consequently, we have to correct our anticipations, in that not E itself—continuing to use the same terminology—but E increased by a certain constant is to be expected to be proportional to the square of the frequency. Let this constant be now *very great* compared with all the admissible negative E -values (which are already limited by (15)). Then firstly, the frequencies will become *real*, and secondly, since our E -values correspond to only relatively small frequency *differences*, they will actually be very approximately proportional to these frequency differences. This, again, is all that our “quantum-instinct” can require, as long as the zero level of *energy* is not fixed.

The view that the frequency of the vibration process is given by

$$(22) \quad \nu = C' \sqrt{C + E} = C' \sqrt{C} + \frac{C'}{2\sqrt{C}} E + \dots,$$

where C is a constant very great compared with all the E 's, has still another very appreciable advantage. *It permits an understanding of the Bohr frequency condition.* According to the latter the *emission frequencies* are proportional to the E -*differences*, and therefore from (22) also to the differences of the proper frequencies ν of those hypothetical vibration processes. But these proper frequencies are all very great compared with the emission frequencies, and they agree very closely among themselves. The emission frequencies appear therefore as deep “difference tones” of the proper vibrations themselves. It is quite conceivable that on the transition of energy from one to another of the normal vibrations, *something*—I mean the light wave—with a *frequency* allied to each frequency *difference*, should make its appearance. One only needs to imagine that the light wave is causally related to the *beats*, which necessarily arise at each point of space during the transition; and that the frequency of the light is defined by the number of times per second the intensity maximum of the beat-process repeats itself.

It may be objected that these conclusions are based on the relation (22), in its *approximate* form (after expansion of the square root), from which the Bohr frequency condition itself seems to obtain the nature of an approximation. This, however, is merely apparently so, and it is wholly avoided when the *relativistic* theory is developed and makes a profounder insight possible. The large constant C is naturally very intimately connected with the rest-energy of the electron (mc^2). Also the seemingly *new* and *independent* introduction of the constant \hbar (already brought in by (20)), into the frequency condition, is cleared up, or rather avoided, by the relativistic theory. But unfortunately the correct establishment of the latter meets right away with certain difficulties, which have been already alluded to.

It is hardly necessary to emphasize how much more congenial it would be to imagine that at a quantum transition the energy changes over from one form of vibration to another, than to think

of a jumping electron. The changing of the vibration form can take place continuously in space and time, and it can readily last as long as the emission process lasts empirically (experiments on canal rays by W. Wien); nevertheless, if during this transition the atom is placed for a comparatively short time in an electric field which alters the proper frequencies, then the beat frequencies are immediately changed sympathetically, and for just as long as the field operates. It is known that this experimentally established fact has hitherto presented the greatest difficulties. See the well-known attempt at a solution by Bohr, Kramers, and Slater.

Let us not forget, however, in our gratification over our progress in these matters, that the idea of only *one* proper vibration being excited whenever the atom does not radiate—if we must hold fast to this idea—is very far removed from the *natural* picture of a vibrating system. We know that a macroscopic system does not behave like that, but yields in general a *pot-pourri* of its proper vibrations. But we should not make up our minds too quickly on this point. A *pot-pourri* of proper vibrations would also be permissible for a single atom, since thereby no beat frequencies could arise other than those which, according to experience, the atom is capable of emitting *occasionally*. The actual sending out of many of these spectral lines simultaneously by the same atom does not contradict experience. It is thus conceivable that only in the normal state (and approximately in certain “meta-stable” states) the atom vibrates with *one* proper frequency and just for this reason does *not* radiate, namely, because no beats arise. The *stimulation* may consist of a simultaneous excitation of one or of several other proper frequencies, whereby beats originate and evoke emission of light.

Under all circumstances, I believe, the proper functions, which belong to the *same* frequency, are in general all simultaneously stimulated. Multipleness of the proper values corresponds, namely, in the language of the previous theory to *degeneration*. To the reduction of the quantisation of degenerate systems probably corresponds the arbitrary partition of the energy among the functions belonging to *one* proper value.

Addition at the proof correction on 28.2.1926.

In the case of conservative systems in classical mechanics, the variation problem can be formulated in a neater way than was previously shown, and without express reference to the Hamilton-Jacobi differential equation. Thus, let $T(q, p)$ be the kinetic energy, expressed as a function of the co-ordinates and momenta, V the potential energy, and $d\tau$ the volume element of the space, “measured rationally”, i.e. it is not simply the product $dq_1 dq_2 dq_3 \dots dq_n$, but this divided by the square root of the discriminant of the quadratic form $T(q, p)$. (Cf. Gibbs’ *Statistical Mechanics*.) Then let ψ be such as to make the “Hamilton integral”

$$(23) \quad \int d\tau \left\{ K^2 T \left(q, \frac{\partial \psi}{\partial q} \right) + \psi^2 V \right\}$$

stationary, while fulfilling the *normalising, accessory condition*

$$(24) \quad \int \psi^2 d\tau = 1.$$

The proper values of this variation problem are then *the stationary values* of integral (23) and yield, according to our thesis, *the quantum-levels of the energy.*

It is to be remarked that in the quantity α_2 of (14'') we have essentially the well-known Sommerfeld expression $-\frac{B}{\sqrt{A}} + \sqrt{C}$. (Cf. *Atombau*, 4th (German) ed., p. 775.)

Physical Institute of the University of Zürich.

(Received January 27, 1926.)

Quantisation as a Problem of Proper Values (Part II)

(*Annalen der Physik* (4), vol. 79, 1926)

§ 1. The Hamiltonian Analogy between Mechanics and Optics

BEFORE we go on to consider the problem of proper values for further special systems, let us throw more light on the *general* correspondence which exists between the Hamilton-Jacobi differential equation of a mechanical problem and the "allied" *wave equation*, i.e. equation (5) of Part I. in the case of the Kepler problem. So far we have only briefly described this correspondence on its external analytical side by the transformation (2), which is in itself unintelligible, and by the equally incomprehensible transition from the *equating to zero* of a certain expression to the postulation that the *space integral* of the said expression shall be *stationary*.¹

The *inner* connection between Hamilton's theory and the process of wave propagation is anything but a new idea. It was not only well known to Hamilton, but it also served him as the starting-point for his theory of mechanics, which grew² out of his *Optics of Non-homogeneous Media*. Hamilton's variation principle can be shown to correspond to Fermat's *Principle* for a wave propagation in configuration space (*q*-space), and the Hamilton-Jacobi equation expresses Huygens' *Principle* for this wave propagation. Unfortunately this powerful and momentous conception of Hamilton is deprived, in most modern reproductions, of its beautiful raiment as a superfluous accessory, in favour of a more colourless representation of the analytical correspondence.³

¹ This procedure will *not be pursued further* in the present paper. It was only intended to give a provisional, quick survey of the external connection between the wave equation and the Hamilton-Jacobi equation. ψ is not actually the action function of a definite motion in the relation stated in (2) of Part I. On the other hand the connection between the wave equation and the variation problem is of course very real; the integrand of the stationary integral is the Lagrange function for the wave process.

² Cf. e.g. E. T. Whittaker's *Anal. Dynamics*, chap. xi.

³ Felix Klein has since 1891 repeatedly developed the theory of Jacobi from quasi-optical considerations in non-Euclidean higher space in his lectures on mechanics. Cf. F. Klein, *Jahresber. d. Deutsch. Math. Ver.* 1, 1891, and *Zeits. f. Math. u. Phys.* 46,

Let us consider the general problem of conservative systems in classical mechanics. The Hamilton-Jacobi equation runs

$$(1) \quad \frac{\partial W}{\partial t} + T\left(q_k, \frac{\partial W}{\partial q_k}\right) + V(q_k) = 0.$$

W is the action function, *i.e.* the time integral of the Lagrange function $T - V$ along a path of the system as a function of the end points and the time. q_k is a representative position co-ordinate; T is the kinetic energy as function of the q 's and momenta, being a quadratic form of the latter, for which, as prescribed, the partial derivatives of W with respect to the q 's are written. V is the potential energy. To solve the equation put

$$(2) \quad W = -Et + S(q_k),$$

and obtain

$$(1') \quad 2T\left(q_k, \frac{\partial W}{\partial q_k}\right) = 2(E - V).$$

E is an arbitrary integration constant and signifies, as is known, the energy of the system. Contrary to the usual practice, we have let the function W remain itself in (1'), instead of introducing the time-free function of the co-ordinates, S . That is a mere superficiality.

Equation (1') can now be very simply expressed if we make use of the method of Heinrich Hertz. It becomes, like all geometrical assertions in configuration space (space of the variables q_k), especially simple and clear if we introduce into this space a non-Euclidean metric by means of the kinetic energy of the system.

Let \bar{T} be the kinetic energy as function of the velocities \dot{q}_k , not of the momenta as above, and let us put for the line element

$$(3) \quad ds^2 = 2\bar{T}(\dot{q}_k, \dot{q}_k)dt^2.$$

The right-hand side now contains dt only externally and represents (since $\dot{q}_k dt = dq_k$) a quadratic form of the dq_k 's.

After this stipulation, conceptions such as angle between two line elements, perpendicularity, divergence and curl of a vector, gradient of a scalar, Laplacian operation (=div grad) of a scalar, and others, may be used in the same simple way as in three-dimensional Euclidean space, and we may use in our thinking the Euclidean three-dimensional representation with impunity, except that the analytical expressions for these ideas become a very little more complicated, as the line element (3) must everywhere replace the Euclidean line element. *We stipulate, that in what follows, all geometrical statements in q -space are to be taken in this non-Euclidean sense.*

One of the most important modifications for the calculation is

1901 (*Ges.-Abh.* ii. pp. 601 and 603). In the second note, Klein remarks reproachfully that his discourse at Halle ten years previously, in which he had discussed this correspondence and emphasized the great significance of Hamilton's optical works, had "not obtained the general attention, which he had expected". For this allusion to F. Klein, I am indebted to a friendly communication from Prof. Sommerfeld. See also *Atombau*, 4th ed., p. 803.

that we must distinguish carefully between covariant and contravariant components of a vector or tensor. But this complication is not any greater than that which occurs in the case of an oblique set of Cartesian axes.

The dq_k 's are the prototype of a contravariant vector. The coefficients of the form $2\bar{T}$, which depend on the q_k 's, are therefore of a covariant character and form the covariant fundamental tensor. $2T$ is the contravariant form belonging to $2\bar{T}$, because the momenta are known to form the covariant vector belonging to the speed vector \dot{q}_k , the momentum being the velocity vector in covariant form. The left side of (1') is now simply the contravariant fundamental form, in which the $\frac{\partial W}{\partial q_k}$'s are brought in as variables. The latter form the components of the vector,—according to its nature covariant,

$$\text{grad } W.$$

(The expressing of the kinetic energy in terms of momenta instead of speeds has then *this* significance, that covariant vector components can only be introduced in a contravariant form if something intelligible, *i.e.* invariant, is to result.)

Equation (1') is equivalent thus to the simple statement

$$(1'') \quad (\text{grad } W)^2 = 2(E - V),$$

or

$$(1''') \quad |\text{grad } W| = \sqrt{2(E - V)}.$$

This requirement is easily analysed. Suppose that a function W , of the form (2), has been found, which satisfies it. Then this function can be clearly represented for every definite t , if the family of surfaces $W = \text{const.}$ be described in q -space and to each member a value of W be ascribed.

Now, on the one hand, as will be shown immediately, equation (1''') gives an exact rule for constructing all the other surfaces of the family and obtaining their W -values from any single member, *if the latter and its W -value is known.* On the other hand, if the sole necessary data for the construction, *viz. one surface and its W -value be given quite arbitrarily,* then from the rule, which presents just *two* alternatives, there may be completed one of the functions W fulfilling the given requirement. Provisionally, the time is regarded as constant.—The construction rule therefore *exhausts* the contents of the differential equation; *each* of its solutions can be obtained from a suitably chosen surface and W -value.

Let us consider the construction rule. Let the value W_0 be given in Fig. 1 to an arbitrary surface. In order to find the surface $W_0 + dW_0$, take *either* side of the given surface as the positive one, erect the normal at each point of it and cut off (with due regard to the sign of dW_0) the *step*

$$(4) \quad ds = \frac{dW_0}{\sqrt{2(E - V)}}.$$

The locus of the end points of the steps is the surface $W_0 + dW_0$. Similarly, the family of surfaces may be constructed successively on both sides.

The construction has a *double* interpretation, as the *other* side of the given surface might have been taken as positive for the first step. This ambiguity does not hold for later steps, *i.e.* at any later stage of the process we cannot change arbitrarily the sign of the sides of the surface, at which we have arrived, as this would involve in general a discontinuity in the

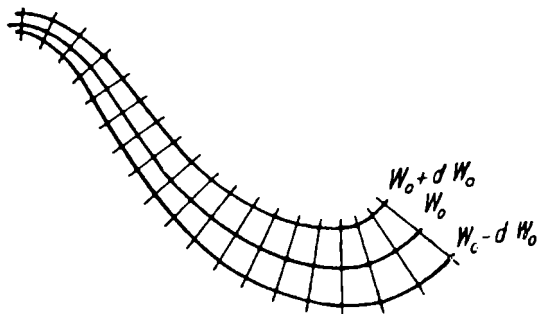


FIG. 1.

first differential coefficient of W . Moreover, the two families obtained in the two cases are clearly identical; the W -values merely run in the opposite direction.

Let us consider now the very simple dependence on the *time*. For this, (2) shows that at any later (or earlier) instant $t + t'$, the *same* group of surfaces illustrates the W -distribution, though different W -values are associated with the individual members, namely, from each W -value ascribed at time t there must be subtracted Et' . The W -values wander, as it were, from surface to surface according to a definite, simple law, and for positive E in the direction of W increasing. Instead of this, however, we may imagine that the *surfaces* wander in such a way that each of them continually takes the place and exact form of the following one, and always carries its W -value *with* it. The rule for this wandering is given by the fact that the surface W_0 at time $t + dt$ must have reached *that* place, which at t was occupied by the surface $W_0 + Edt$. This will be attained according to (4), if each point of the surface W_0 is allowed to move in the direction of the positive normal through a distance

$$(5) \quad ds = \frac{Edt}{\sqrt{2(E - V)}}$$

That is, the surfaces move with a *normal velocity*

$$(6) \quad u = \frac{ds}{dt} = \frac{E}{\sqrt{2(E - V)}}$$

which, when the constant E is given, is a pure function of position.

Now it is seen that our system of surfaces $W = \text{const.}$ can be conceived as the system of wave surfaces of a progressive but stationary wave motion in q -space, for which the value of the phase velocity at every point in the space is given by (6). For the normal construction

can clearly be replaced by the construction of elementary Huygens waves (with radius (5)), and then of their envelope. The "index of refraction" is proportional to the reciprocal of (6), and is dependent on the position but not on the direction. The q -space is thus optically non-homogeneous but is isotropic. The elementary waves are "spheres", though of course—let me repeat it expressly once more—in the sense of the line-element (3).

The function of action W plays the part of the *phase* of our wave system. The Hamilton-Jacobi equation is the expression of Huygens' principle. If, now, Fermat's principle be formulated thus,

$$(7) \quad 0 = \delta \int_{P_1}^{P_2} \frac{ds}{u} = \delta \int_{P_1}^{P_2} \frac{ds \sqrt{2(E-V)}}{E} = \delta \int_{t_1}^{t_2} \frac{2T}{E} dt = \frac{1}{E} \delta \int_{t_1}^{t_2} 2T dt,$$

we are led directly to Hamilton's principle in the form given by Maupertuis (where the time integral is to be taken with the usual grain of salt, *i.e.* $T + V = E = \text{constant}$, even during the variation). The "rays", *i.e.* the orthogonal trajectories of the wave surfaces, are therefore the *paths* of the system for the value E of the energy, in agreement with the well-known system of equations

$$(8) \quad p_k = \frac{\partial W}{\partial q_k}$$

which states, that a set of system paths can be derived from each special function of action, just like a fluid motion from its velocity potential.¹ (The momenta p_k form the covariant velocity vector, which equations (8) assert to be equal to the gradient of the function of action.)

Although in these deliberations on wave surfaces we speak of velocity of propagation and Huygens' principle, we must regard the analogy as one between mechanics and *geometrical* optics, and not physical or *undulatory* optics. For the idea of "rays", which is the essential feature in the mechanical analogy, belongs to *geometrical* optics; it is only clearly defined in the latter. Also Fermat's principle can be applied in geometrical optics without going beyond the idea of index of refraction. And the system of W -surfaces, regarded as wave surfaces, stands in a somewhat looser relationship to mechanical motion, inasmuch as the image point of the mechanical system in no wise moves along the ray with the wave velocity u , but, on the contrary, its velocity (for constant E) is proportional to $\frac{1}{u}$. It is given directly from (3) as

$$(9) \quad v = \frac{ds}{dt} = \sqrt{2T} = \sqrt{2(E-V)}.$$

¹ See especially A. Einstein, *Verh. d. D. Physik. Ges.* 19, pp. 77, 82, 1917. The framing of the quantum conditions here is the most akin, out of all the older attempts, to the present one. De Broglie has returned to it.

This non-agreement is obvious. Firstly, according to (8), the system's point velocity is *great* when $\text{grad } W$ is great, *i.e.* where the W -surfaces are closely crowded together, *i.e.* where u is small. Secondly, from the definition of W as the time integral of the Lagrange function, W *alters* during the motion (by $(T - V)dt$ in the time dt), and so the image point *cannot* remain continuously in contact with the same W -surface.

And important ideas in wave theory, such as amplitude, wave length, and frequency—or, speaking more generally, the wave *form*—do not enter into the analogy at all, as there exists no mechanical parallel; even of the wave function itself there is no mention beyond that W has the meaning of the *phase* of the waves (and this is somewhat hazy owing to the wave *form* being undefined).

If we find in the whole parallel merely a satisfactory means of contemplation, then this defect is not disturbing, and we would regard any attempt to supply it as idle trifling, believing the analogy to be precisely with *geometrical*, or at furthest, with a very primitive form of wave optics, and not with the fully developed undulatory optics. That geometrical optics is only a rough approximation for *Light* makes no difference. To *preserve* the analogy on the further development of the optics of q -space on the lines of wave theory, we must take good care not to depart markedly from the limiting case of geometrical optics, *i.e.* must choose¹ the *wave length* sufficiently small, *i.e.* small compared with all the path dimensions. Then the additions do not teach anything new; the picture is only draped with superfluous ornaments.

So we might think to begin with. But even the first attempt at the development of the analogy to the wave theory leads to such striking results, that a quite different suspicion arises: *we know to-day, in fact, that our classical mechanics fails for very small dimensions of the path and for very great curvatures.* Perhaps this failure is in strict analogy with the failure of geometrical optics, *i.e.* "the optics of infinitely small wave lengths", that becomes evident as soon as the obstacles or apertures are no longer great compared with the real, finite, wave length. Perhaps our classical mechanics is the *complete* analogy of geometrical optics and as such is wrong and not in agreement with reality; it fails whenever the radii of curvature and dimensions of the path are no longer great compared with a certain wave length, to which, in q -space, a real meaning is attached. Then it becomes a question of searching² for an undulatory mechanics, and the most obvious way is the working out of the Hamiltonian analogy on the lines of undulatory optics.

¹ Cf. for the optical case, A. Sommerfeld and Iris Runge, *Ann. d. Phys.* 35, p. 290, 1911. There (in the working out of an oral remark of P. Debye), it is shown, how the equation of *first* order and *second* degree for the *phase* ("Hamiltonian equation") may be accurately derived from the equation of the *second* order and *first* degree for the *wave function* ("wave equation"), in the limiting case of vanishing wave length.

² Cf. A. Einstein, *Berl. Ber.* p. 9 *et seq.*, 1925.

§ 2. "Geometrical" and "Undulatory" Mechanics

We will at first assume that it is fair, in extending the analogy, to imagine the above-mentioned wave system as consisting of *sine* waves. This is the simplest and most obvious case, yet the *arbitrariness*, which arises from the *fundamental significance* of this assumption, must be emphasized. The wave function has thus only to contain the time in the form of a factor, $\sin(\dots)$, where the argument is a linear function of W . The coefficient of W must have the dimensions of the reciprocal of action, since W has those of action and the phase of a sine has zero dimensions. We assume that it is quite universal, *i.e.* that it is not only independent of E , but also of the nature of the mechanical system. We may then at once denote it by $\frac{2\pi}{h}$. The time factor then is

$$(10) \quad \sin\left(\frac{2\pi W}{h} + \text{const.}\right) = \sin\left(-\frac{2\pi Et}{h} + \frac{2\pi S(q_k)}{h} + \text{const.}\right).$$

Hence the *frequency* ν of the waves is given by

$$(11) \quad \nu = \frac{E}{h}.$$

Thus we get the frequency of the q -space waves to be proportional to the energy of the system, in a manner which is not markedly artificial.¹ This is only true of course if E is absolute and not, as in classical mechanics, indefinite to the extent of an additive constant. By (6) and (11) the *wave length* is *independent* of this additive constant, being

$$(12) \quad \lambda = \frac{u}{\nu} = \frac{h}{\sqrt{2(E - V)}},$$

and we know the term under the root to be double the kinetic energy. Let us make a preliminary rough comparison of this wave length with the dimensions of the orbit of a hydrogen electron as given by classical mechanics, taking care to notice that a "step" in q -space has not the dimensions of length, but length multiplied by the square root of mass, in consequence of (3). λ has similar dimensions. We have therefore to divide λ by the dimension of the orbit, a cm., say, and by the square root of m , the mass of the electron. The quotient is of the order of magnitude of

$$\frac{h}{mva},$$

where v represents for the moment the electron's velocity (cm./sec.). The denominator mva is of the order of the mechanical moment of momentum, and this is at least of the order of 10^{-27} for Kepler orbits, as can be calculated from the values of electronic charge and mass

¹ In Part I. this appeared merely as an approximate equation, derived from a pure speculation.

independently of all quantum theories. We thus obtain the correct order for the *limit of the approximate region of validity of classical mechanics*, if we identify our constant h with Planck's quantum of action—and this is only a preliminary attempt.

If in (6), E is expressed by means of (11) in terms of ν , then we obtain

$$(6') \quad u = \frac{h\nu}{\sqrt{2(h\nu - V)}}.$$

The dependence of the wave velocity on the energy thus becomes a particular kind of dependence on the *frequency*, *i.e.* it becomes a *law of dispersion* for the waves. This law is of great interest. We have shown in § 1 that the wandering wave surfaces are only loosely connected with the motion of the system point, since their velocities are not equal and cannot be equal. According to (9), (11), and (6') the system's velocity v has thus also a concrete significance for the wave. We verify at once that

$$(13) \quad v = \frac{d\nu}{d\left(\frac{\nu}{u}\right)},$$

i.e. the velocity of the system point is that of a *group of waves*, included within a small range of frequencies (signal-velocity). We find here again a theorem for the "phase waves" of the electron, which M. de Broglie had derived, with essential reference to the relativity theory, in those fine researches,¹ to which I owe the inspiration for this work. We see that the theorem in question is of wide generality, and does not arise solely from relativity theory, but is valid for every conservative system of ordinary mechanics.

We can utilise this fact to institute a much more innate connection between wave propagation and the movement of the representative point than was possible before. We can attempt to build up a wave group which will have relatively small dimensions in every direction. Such a wave group will then presumably obey the same laws of motion as a single image point of the mechanical system. It will then give, so to speak, an *equivalent* of the image point, so long as we can look on it as being approximately confined to a point, *i.e.* so long as we can neglect any spreading out in comparison with the dimensions of the path of the system. This will only be the case when the path dimensions, and especially the radius of curvature of the path, are very great compared with the wave length. For, in analogy with ordinary optics, it is obvious from what has been said that not only must the dimensions of the wave group not be reduced below the order of magnitude of the wave length, but, on the contrary, the group must extend in all directions over a large number of wave lengths, if it is to be *approximately monochromatic*. This, however, must be postulated, since the wave group must move about as a whole with a definite

¹ L. de Broglie, *Ann. de Physique* (10) 3, p. 22, 1925. (Thèses, Paris, 1924.)

group velocity and correspond to a mechanical system of *definite energy* (cf. equation 11).

So far as I see, such groups of waves can be constructed on exactly the same principle as that used by Debye¹ and von Laue² to solve the problem in ordinary optics of giving an exact analytical representation of a cone of rays or of a sheaf of rays. From this there comes a very interesting relation to that part of the Hamilton-Jacobi theory not described in § 1, viz. the well-known derivation of the equations of motion in integrated form, by the differentiation of a complete integral of the Hamilton-Jacobi equation with respect to the constants of integration. As we will see immediately, the system of equations called after Jacobi is equivalent to the statement: the image point of the mechanical system continuously corresponds to *that* point, where a certain continuum of wave trains coalesces in *equal phase*.

In optics, the representation (strictly on the wave theory) of a "sheaf of rays" with a sharply defined finite cross-section, which proceeds to a focus and then diverges again, is thus carried out by Debye. A *continuum* of *plane* wave trains, each of which alone would fill the *whole* space, is superposed. The *continuum* is produced by letting the wave normal vary throughout the given solid angle. The waves then destroy one another almost completely by interference outside a certain double cone; they represent exactly, on the wave theory, the desired limited sheaf of rays and also the diffraction phenomena, necessarily occasioned by the limitation. We can represent in this manner an *infinitesimal* cone of rays just as well as a finite one, if we allow the wave normal of the group to vary only inside an infinitesimal solid angle. This has been utilised by von Laue in his famous paper on the degrees of freedom of a sheaf of rays.³ Finally, instead of working with waves, hitherto tacitly accepted as purely monochromatic, we can also allow the *frequency* to vary within an infinitesimal interval, and by a suitable distribution of the amplitudes and phases can confine the disturbance to a region which is relatively small in the longitudinal direction also. So we succeed in representing analytically a "parcel of energy" of relatively small dimensions, which travels with the speed of light, or when dispersion occurs, with the group velocity. Thereby is given the instantaneous *position* of the parcel of energy—if the detailed structure is not in question—in a very plausible way as that point of space where *all* the superposed plane waves meet in *exactly* agreeing phase.

We will now apply these considerations to the q -space waves. We select, at a definite time t , a definite point P of q -space, through which the parcel of waves passes in a given direction R , at that time. In addition let the mean frequency ν or the mean E -value for the packet be also given. These conditions correspond exactly to postulating that at a given time the mechanical system is starting from a given

¹ P. Debye, *Ann. d. Phys.* 30, p. 755, 1909.

² M. v. Laue, *idem* 44, p. 1197 (§ 2), 1914.

³ *Loc. cit.*

configuration with given velocity components. (Energy *plus* direction is equivalent to velocity components.)

In order to carry over the optical construction, we require firstly *one* set of wave surfaces with the desired frequency, *i.e.* *one* solution of the Hamilton-Jacobi equation (1') for the given E -value. This solution, W , say, is to have the following property: the surface of the set which passes through P at time t , which we may denote by

$$(14) \quad W = W_0,$$

must have its normal at P in the prescribed direction R . But this is still not enough. We must be able to vary to an infinitely small extent this set of waves W in an n -fold manner (n = number of degrees of freedom), so that the wave normal will sweep out an infinitely small $(n - 1)$ dimensional space angle at the point P , and so that the frequency $\frac{E}{\hbar}$ will vary in an infinitely small *one*-dimensional region, whereby care is taken that all members of the infinitely small n -dimensional continuum of sets of waves meet together at time t in the point P in exactly agreeing phase. Then it is a question of finding at any other time *where* that point lies at which this agreement of phases occurs.

To do this, it will be sufficient if we have at our disposal a solution W of the Hamilton-Jacobi equation, which is dependent not only on the constant E , here denoted by a_1 , but also on $(n - 1)$ additional constants $a_2, a_3 \dots a_n$, in such a way that it cannot be written as a function of less than n combinations of these n constants. For then we can, firstly, bestow on a_1 the value prescribed for E , and, secondly, define $a_2, a_3 \dots a_n$, so that the surface of the set passing through the point P has at P the prescribed normal direction. Henceforth we understand by $a_1, a_2 \dots a_n$, *these* values, and take (14) as the surface of *this* set, which passes through the point P at time t . Then we consider the *continuum of sets* which belongs to the a_k -values of an adjacent infinitesimal a_k -region. A member of this continuum, *i.e.* therefore *a set*, will be given by

$$(15) \quad W + \frac{\partial W}{\partial a_1} da_1 + \frac{\partial W}{\partial a_2} da_2 + \dots + \frac{\partial W}{\partial a_n} da_n = \text{const.}$$

for a *fixed* set of values of $da_1, da_2 \dots da_n$, and varying constant. That member of *this set*, *i.e.* therefore that single surface, which goes through P at time t will be defined by the following choice of the const.,

$$(15') \quad W + \frac{\partial W}{\partial a_1} da_1 + \dots + \frac{\partial W}{\partial a_n} da_n = W_0 + \left(\frac{\partial W}{\partial a_1}\right)_0 da_1 + \dots + \left(\frac{\partial W}{\partial a_n}\right)_0 da_n,$$

where $\left(\frac{\partial W}{\partial a_1}\right)_0$, etc., are the *constants* obtained by substituting in the differential coefficients the co-ordinates of the point P and the value t of the time (which latter really only occurs in $\frac{\partial W}{\partial a_1}$).

The surfaces (15') for all possible sets of values of $da_1, da_2 \dots da_n$,

form on their part *a set*. They all go through the point *P* at time *t*, their wave normals continuously sweep out a little $(n - 1)$ dimensional solid angle and, moreover, their *E*-parameter also varies within a small region. The set of surfaces (15') is so formed that each of the sets (15) supplies *one* representative to (15'), namely, that member which passes through *P* at time *t*.

We will now assume that the phase angles of the wave functions which belong to the sets (15) happen to agree precisely for those representatives which enter the set (15'). They agree therefore at time *t* at the point *P*.

We now ask: Is there, at *any arbitrary time*, a point where all surfaces of the set (15') cut one another, and in which, therefore, all the wave functions which belong to the sets (15) agree in phase? The answer is: *There exists* a point of agreeing phase, but it is *not* the common intersection of the surfaces of set (15'), for such does *not* exist at any subsequent arbitrary time. Moreover, the point of phase agreement arises in *such a way* that the sets (15) *continuously exchange* their representatives given to (15').

That is shown thus. There must hold

$$(16) \quad W = W_0, \quad \frac{\partial W}{\partial a_1} = \left(\frac{\partial W}{\partial a_1} \right)_0, \quad \frac{\partial W}{\partial a_2} = \left(\frac{\partial W}{\partial a_2} \right)_0 \dots \frac{\partial W}{\partial a_n} = \left(\frac{\partial W}{\partial a_n} \right)_0,$$

simultaneously for the common meeting point of all members of (15') at any time, because the da_1 's are arbitrary within a small region. In these $n + 1$ equations, the right-hand sides are constants, and the left are functions of the $n + 1$ quantities q_1, q_2, \dots, q_n, t . The equations are satisfied by the initial system of values, *i.e.* by the co-ordinates of *P* and the initial time *t*. For another arbitrary value of *t*, they will have *no* solutions in $q_1 \dots q_n$, but will *more than define* the system of these n quantities.

We may proceed, however, as follows. Let us leave the first equation, $W = W_0$, aside at first, and define the q_k 's as functions of the time and the constants according to the remaining n equations. Let this point be called *Q*. By it, naturally, the *first* equation will *not* be satisfied, but the left-hand side will differ from the right by a certain value. If we go back to the derivation of system (16) from (15'), what we have just said means that though *Q* is not a common point for the set of surfaces (15'), it is so, however, for a set which results from (15'), if we alter the right-hand side of equation (15') by an amount which is constant for all the surfaces. Let this new set be (15''). For it, therefore, *Q* is a common point. The new set results from (15'), as stated above, by an exchange of the representatives in (15'). This exchange is occasioned by the alteration of the constant in (15), *by the same amount*, for all representatives. Hence the *phase angle* is altered by the same amount for all representatives. The new representatives, *i.e.* the members of the set we have called (15''), which meet in the point *Q*, agree in phase angle just as the old ones did. This amounts therefore to saying :

The point Q which is defined as a function of the time by the n equations

$$(17) \quad \frac{\partial W}{\partial a_1} = \left(\frac{\partial W}{\partial a_1}\right)_0, \dots, \frac{\partial W}{\partial a_n} = \left(\frac{\partial W}{\partial a_n}\right)_0,$$

continues to be a point of agreeing phase for the whole aggregate of wave sets (15).

Of all the n -surfaces, of which Q is shown by (17) to be the common point, only the first is variable; the others remain fixed (only the first of equations (17) contains the time). The $n-1$ fixed surfaces determine the *path* of the point Q as their line of intersection. It is easily shown that this line is the orthogonal trajectory of the set $W = \text{const.}$ For, by hypothesis, W satisfies the Hamilton-Jacobi equation (1') identically in a_1, a_2, \dots, a_n . If we now differentiate the Hamilton-Jacobi equation with respect to a_k ($k=2, 3, \dots, n$), we get the statement that the normal to a surface, $\frac{\partial W}{\partial a_k} = \text{const.}$, is *perpendicular*, at every point on it, to the normal of the surface, $W = \text{const.}$, which passes through that point, *i.e.* that each of the two surfaces *contains* the normal to the other. If the line of intersection of the $n-1$ fixed surfaces (17) has no branches, as is generally the case, then must each line element of the intersection, as the *sole common* line element of the $n-1$ surfaces, coincide with the normal of the W -surface, passing through the same point, *i.e.* the line of intersection is the orthogonal trajectory of the W -surfaces. Q.E.D.

We may sum up the somewhat detailed discussion, which has led us to equations (17), in a much shorter or (so to speak) shorthand fashion, as follows: W denotes, apart from a universal constant $\left(\frac{1}{h}\right)$, the phase angle of the wave function. If we now deal not merely with *one*, but with a continuous manifold of wave systems, and if these are continuously arranged by means of any continuous parameters a_i , then the equations $\frac{\partial W}{\partial a_i} = \text{const.}$ express the fact that all infinitely adjacent individuals (wave systems) of this manifold agree in phase. These equations therefore define the geometrical locus of the points of agreeing phase. If the equations are sufficient, this locus shrinks to one point; the equations then define *the point* of phase agreement as a function of the time.

Since the system of equations (17) agrees with the known second system of equations of Jacobi, we have thus shown:

The point of phase agreement for certain infinitesimal manifolds of wave systems, containing n parameters, moves according to the same laws as the image point of the mechanical system.

I consider it a very difficult task to give an exact proof that the superposition of these wave systems really produces a noticeable disturbance in only a relatively small region surrounding the point of phase agreement, and that everywhere else they practically destroy

one another through interference, or that the above statement turns out to be true at least for a suitable choice of the amplitudes, and possibly for a special choice of the *form* of the wave surfaces. I will advance the physical hypothesis, which I wish to attach to what is to be proved, without attempting the proof. The latter will only be worth while if the hypothesis stands the test of trial *and* if its application should *require* the exact proof.

On the other hand, we may be sure that the region to which the disturbance may be confined still contains in all directions a great number of wave lengths. This is directly evident, firstly, because so long as we are only a *few* wave lengths distant from the point of phase agreement, then the agreement of phase is hardly disturbed, as the interference is still almost as favourable as it is at the point itself. Secondly, a glance at the three-dimensional Euclidean case of ordinary optics is sufficient to assure us of this general behaviour.

What I now categorically conjecture is the following :

The true mechanical process is realised or represented in a fitting way by the *wave processes* in q -space, and not by the motion of *image points* in this space. The study of the motion of image points, which is the object of classical mechanics, is only an approximate treatment, and has, as such, just as much justification as geometrical or "ray" optics has, compared with the true optical process. A macroscopic mechanical process will be portrayed as a wave signal of the kind described above, which can approximately enough be regarded as confined to a point compared with the geometrical structure of the path. We have seen that the same laws of motion hold exactly for such a signal or group of waves as are advanced by classical mechanics for the motion of the image point. This manner of treatment, however, loses all meaning where the structure of the path is no longer very large compared with the wave length or indeed is comparable with it. Then we *must* treat the matter strictly on the wave theory, *i.e.* we must proceed from the *wave equation* and not from the fundamental equations of mechanics, in order to form a picture of the manifold of the possible processes. These latter equations are just as useless for the elucidation of the micro-structure of mechanical processes as geometrical optics is for explaining the *phenomena of diffraction*.

Now that a certain interpretation of this micro-structure has been successfully obtained as an addition to classical mechanics, although admittedly under new and very artificial assumptions, an interpretation bringing with it practical successes of the highest importance, it seems to me very significant that these theories—I refer to the forms of quantum theory favoured by Sommerfeld, Schwarzschild, Epstein, and others—bear a very close relation to the Hamilton-Jacobi equation and the theory of its solution, *i.e.* to that form of classical mechanics which already points out most clearly the true undulatory character of mechanical processes. The Hamilton-Jacobi equation corresponds to Huygens' Principle (in its old simple form, not in the form due to Kirchhoff). And just as this, supplemented by

some rules which are not intelligible in geometrical optics (Fresnel's construction of zones), can explain to a great extent the phenomena of diffraction, so light can be thrown on the processes in the atom by the theory of the action-function. But we inevitably became involved in irremovable contradictions if we tried, as was very natural, to maintain also the idea of *paths of systems* in these processes; just as we find the tracing of the course of a *light ray* to be meaningless, in the neighbourhood of a diffraction phenomenon.

We can argue as follows. I will, however, not yet give a conclusive picture of the actual process, which positively cannot be arrived at from this starting-point but only from an investigation of the wave equation; I will merely illustrate the matter qualitatively. Let us think of a wave group of the nature described above, which in some way gets into a small closed "path", whose dimensions are of the order of the wave length, and therefore *small* compared with the dimensions of the wave group itself. It is clear that then the "system path" in the sense of classical mechanics, *i.e.* the path of the point of exact phase agreement, will completely lose its prerogative, because there exists a whole continuum of points before, behind, and near the particular point, in which there is almost as complete phase agreement, and which describe totally different "paths". In other words, the wave group not only fills the whole path domain all at once but also stretches far beyond it in all directions.

In *this* sense do I interpret the "phase waves" which, according to de Broglie, accompany the path of the electron; in the sense, therefore, that no special meaning is to be attached to the electronic path itself (at any rate, in the interior of the atom), and still less to the position of the electron on its path. And in this sense I explain the conviction, increasingly evident to-day, *firstly*, that real meaning has to be denied to the *phase* of electronic motions in the atom; *secondly*, that we can never assert that the electron at a definite instant is to be found on *any definite one* of the quantum paths, specialised by the quantum conditions; and *thirdly*, that the true laws of quantum mechanics do not consist of definite rules for the *single path*, but that in these laws the elements of the whole manifold of paths of a system are bound together by equations, so that apparently a certain reciprocal action exists between the different paths.¹

It is not incomprehensible that a careful analysis of the experimentally known quantities should lead to assertions of this kind, if the experimentally known facts are the outcome of such a structure of the real process as is here represented. All these assertions systematically contribute to the relinquishing of the ideas of "place of the electron" and "path of the electron". If these are not given up, contradictions remain. This contradiction has been so strongly felt that it has even been doubted whether what goes on in the atom could ever be described within the scheme of space and time. From the philo-

¹ Cf. especially the papers of Heisenberg, Born, Jordan, and Dirac quoted later, and further N. Bohr, *Die Naturwissenschaften*, January 1926.

sophical standpoint, I would consider a conclusive decision in this sense as equivalent to a complete surrender. For we cannot really alter our manner of thinking in space and time, and what we cannot comprehend within it we cannot understand at all. There *are* such things—but I do not believe that atomic structure is one of them. From our standpoint, however, there is no reason for such doubt, although or rather *because* its appearance is extraordinarily comprehensible. So might a person versed in geometrical optics, after many attempts to explain diffraction phenomena by means of the idea of the ray (trustworthy for his macroscopic optics), which always came to nothing, at last think that the *Laws of Geometry* are not applicable to diffraction, since he continually finds that light rays, which he imagines as *rectilinear* and *independent* of each other, now suddenly show, even in homogeneous media, the most remarkable *curvatures*, and obviously *mutually influence* one another. I consider this analogy as *very* strict. Even for the unexplained *curvatures*, the analogy in the atom is not lacking—think of the “non-mechanical force”, devised for the explanation of anomalous Zeeman effects.

In what way now shall we have to proceed to the undulatory representation of mechanics for those cases where it is necessary? We must start, not from the fundamental equations of mechanics, but from a wave equation for q -space and consider the manifold of processes possible *according to it*. The wave equation has not been explicitly used or even put forward in this communication. The only datum for its construction is the *wave velocity*, which is given by (6) or (6') as a function of the mechanical energy parameter or frequency respectively, and by this datum the wave equation is evidently not uniquely defined. It is not even decided that it must be definitely of the second order. Only the striving for simplicity leads us to try this to begin with. We will then say that for the wave function ψ we have

$$(18) \quad \text{div grad } \psi - \frac{1}{u^2} \ddot{\psi} = 0,$$

valid for all processes which only depend on the time through a factor $e^{2\pi i \nu t}$. Therefore, considering (6), (6'), and (11), we get, respectively,

$$(18') \quad \text{div grad } \psi + \frac{8\pi^2}{h^2} (h\nu - V)\psi = 0,$$

and

$$(18'') \quad \text{div grad } \psi + \frac{8\pi^2}{h^2} (E - V)\psi = 0.$$

The differential operations are to be understood with regard to the line element (3). But even under the postulation of second order, the above is not the only equation consistent with (6). For it is possible to generalize by replacing $\text{div grad } \psi$ by

$$(19) \quad f(q_k) \text{ div } \left(\frac{1}{f(q_k)} \text{ grad } \psi \right),$$

where f may be an arbitrary function of the q 's, which must depend in

some plausible way on E , $V(q_k)$, and the coefficients of the line element (3). (Think, *e.g.*, of $f=u$.) Our postulation is again dictated by the striving for simplicity, yet I consider in this case that a wrong deduction is not out of the question.¹

The substitution of a *partial* differential equation for the equations of dynamics in atomic problems appears at first sight a very doubtful procedure, on account of the multitude of solutions that such an equation possesses. Already classical dynamics had led not just to one solution but to a much too extensive manifold of solutions, viz. to a continuous set, while all experience seems to show that only a discrete number of these solutions is realised. The problem of the quantum theory, according to prevailing conceptions, is to select by means of the "quantum conditions" that discrete set of actual paths out of the continuous set of paths possible according to classical mechanics. It seems to be a bad beginning for a new attempt in this direction if the number of possible solutions has been *increased* rather than diminished.

It is true that the problem of classical dynamics also allows itself to be presented in the form of a *partial* equation, namely, the Hamilton-Jacobi equation. But the manifold of solutions of the problem does not correspond to the manifold of solutions of that equation. An arbitrary "complete" solution of the equation solves the mechanical problem *completely*; any *other* complete solution yields the same paths—they are only contained in another way in the manifold of paths.

Whatever the fear expressed about taking equation (18) as the foundation of atomic dynamics comes to, I will not positively assert that no further additional definitions will be required with it. But these will probably no longer be of such a completely strange and incomprehensible nature as the previous "quantum conditions", but will be of the type that we are accustomed to find in physics with a partial differential equation as initial or boundary conditions. They will be, in no way, *analogous* to the quantum conditions—because in all cases of classical dynamics, which I have investigated up till now, it turns out that equation (18) *carries within itself the quantum conditions*. It distinguishes in certain cases, and indeed in those where experience demands it, *of itself*, certain frequencies or energy levels as those which alone are possible for stationary processes, without any further assumption, other than the almost obvious demand that, as a physical quantity, the function ψ must be single-valued, finite, and continuous throughout configuration space.

Thus the fear expressed is transformed into its contrary, in any case in what concerns the energy levels, or let us say more prudently, the frequencies. (For the question of the "vibrational energy" stands by itself; we must not forget that it is only in the one electron problem that the interpretation as a vibration in real three-dimensional space is immediately suggested.) The definition of the quantum levels *no*

¹ The introduction of $f(q_k)$ means that not only the "density" but also the "elasticity" varies with the position.

longer takes place in two separated stages: (1) Definition of all paths dynamically possible. (2) *Discarding* of the greater part of those solutions and the selection of a few by special postulations; on the contrary, the quantum levels are *at once* defined as the *proper values* of equation (18), which carries in itself its natural boundary conditions.

As to how far an analytical simplification will be effected in this way in more complicated cases, I have not yet been able to decide. I should, however, expect so. Most of the analytical investigators have the feeling that in the two-stage process, described above, there must be yielded in (1) the solution of a more complicated problem than is really necessary for the final result: energy as a (usually) very simple rational function of the quantum numbers. Already, as is known, the application of the Hamilton-Jacobi method creates a great simplification, as the actual calculation of the mechanical solution is avoided. It is sufficient to evaluate the integrals, which represent the momenta, merely for a closed complex path of integration instead of for a variable upper limit, and this gives much less trouble. Still the complete solution of the Hamilton-Jacobi equation must really be known, *i.e.* given by quadratures, so that the integration of the mechanical problem must in principle be effected for arbitrary initial values. In seeking for the proper values of a differential equation, we must usually, in practice, proceed thus. We seek the solution, firstly, without regard to boundary or continuity conditions, and from the form of the solution then pick out those values of the parameters, for which the solution satisfies the given conditions. Part I. supplies an example of this. We see by this example also, however—what is typical of proper value problems—that the solution was only given *generally* in an extremely inaccessible analytical form [equation (12) *loc. cit.*], but that it is extraordinarily simplified for those proper values belonging to the “natural boundary condition”. I am not well enough informed to say whether *direct* methods have now been worked out for the calculation of the proper values. This is known to be so for the distribution of proper values of *high order*. But this limiting case is *not* of interest here; it corresponds to the classical, macroscopic mechanics. For spectroscopy and atomic physics, in general just the *first* 5 or 10 proper values will be of interest; even the *first alone* would be a great result—it defines the *ionisation potential*. From the idea, definitely outlined, that every problem of proper values allows itself to be treated as one of maxima and minima without direct reference to the differential equation, it appears to me very probable that direct methods will be found for the calculation, at least approximately, of the proper values, as soon as *urgent* need arises. At least it should be possible to test in individual cases whether the proper values, *known* numerically to all desired accuracy through spectroscopy, *satisfy* the problem or not.

I would not like to proceed without mentioning here that at the present time a research is being prosecuted by Heisenberg, Born,

Jordan, and other distinguished workers,¹ to remove the quantum difficulties, which has already yielded such noteworthy success that it cannot be doubted that it contains at least a part of the truth. In its *tendency*, Heisenberg's attempt stands very near the present one, as we have already mentioned. In its method, it is so totally different that I have not yet succeeded in finding the connecting link. I am distinctly hopeful that these two advances will not fight against one another, but on the contrary, just because of the extraordinary difference between the starting-points and between the methods, that they will supplement one another and that the one will make progress where the other fails. The strength of Heisenberg's programme lies in the fact that it promises to give the *line-intensities*, a question that we have not approached as yet. The strength of the present attempt—if I may be permitted to pronounce thereon—lies in the guiding, physical point of view, which creates a bridge between the macroscopic and microscopic mechanical processes, and which makes intelligible the outwardly different modes of treatment which they demand. For me, personally, there is a special charm in the conception, mentioned at the end of the previous part, of the emitted frequencies as "beats", which I believe will lead to an intuitive understanding of the intensity formulae.

§ 3. Application to Examples

We will now add a few more examples to the Kepler problem treated in Part I., but they will only be of the very simplest nature, since we have provisionally confined ourselves to *classical* mechanics, with no magnetic field.²

1. The Planck Oscillator. The Question of Degeneracy

Firstly we will consider the one-dimensional oscillator. Let the co-ordinate q be the displacement multiplied by the square root of the mass. The two forms of the kinetic energy then are

$$(20) \quad \bar{T} = \frac{1}{2}\dot{q}^2, \quad T = \frac{1}{2}p^2.$$

The potential energy will be

$$(21) \quad V(q) = 2\pi^2\nu_0^2q^2,$$

where ν_0 is the proper frequency in the mechanical sense. Then equation (18) reads in this case

$$(22) \quad \frac{d^2\psi}{dq^2} + \frac{8\pi^2}{h^2}(E - 2\pi^2\nu_0^2q^2)\psi = 0.$$

¹ W. Heisenberg, *Ztschr. f. Phys.* 33, p. 879, 1925; M. Born and P. Jordan, *ibid.* 34, p. 858, 1925; M. Born, W. Heisenberg, and P. Jordan, *ibid.* 35, p. 557, 1926; P. Dirac, *Proc. Roy. Soc.*, London, 109, p. 642, 1925.

² In relativity mechanics and taking a magnetic field into account the statement of the Hamilton-Jacobi equation becomes more complicated. In the case of a single electron, it asserts that the *four-dimensional* gradient of the action function, *diminished* by a given vector (the four-potential), has a constant value. The translation of this statement into the language of the wave theory presents a good many difficulties.

For brevity write

$$(23) \quad a = \frac{8\pi^2 E}{h^2}, \quad b = \frac{16\pi^4 \nu_0^2}{h^2}.$$

Therefore

$$(22') \quad \frac{d^2\psi}{dq^2} + (a - bq^2)\psi = 0.$$

Introduce as independent variable

$$(24) \quad x = q\sqrt[4]{b},$$

and obtain

$$(22'') \quad \frac{d^2\psi}{dx^2} + \left(\frac{a}{\sqrt{b}} - x^2\right)\psi = 0.$$

The proper values and functions of this equation are *known*.¹ The proper values are, with the notation used here,

$$(25) \quad \frac{a}{\sqrt{b}} = 1, 3, 5 \dots (2n + 1) \dots$$

The functions are the *orthogonal functions of Hermite*,

$$(26) \quad e^{-\frac{x^2}{2}} H_n(x).$$

$H_n(x)$ means the n th Hermite polynomial, which can be defined as

$$(27) \quad H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n},$$

or explicitly by

$$(27') \quad H_n(x) = (2x)^n - \frac{n(n-1)}{1!} (2x)^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2!} (2x)^{n-4} - + \dots$$

The first of these polynomials are

$$(27'') \quad \begin{array}{ll} H_0(x) = 1 & H_1(x) = 2x \\ H_2(x) = 4x^2 - 2 & H_3(x) = 8x^3 - 12x \\ H_4(x) = 16x^4 - 48x^2 + 12 & \dots \end{array}$$

Considering next the proper values, we get from (25) and (23)

$$(25') \quad E_n = \frac{2n+1}{2} h\nu_0; \quad n = 0, 1, 2, 3, \dots$$

Thus as quantum levels appear so-called "half-integral" multiples of the "quantum of energy" peculiar to the oscillator, *i.e.* the *odd* multiples of $\frac{h\nu_0}{2}$. The intervals between the levels, which alone are important for the radiation, are the same as in the former theory. It is remarkable that our quantum levels are *exactly* those of Heisenberg's theory. In the theory of *specific heat* this deviation from the previous

¹ Cf. Courant-Hilbert, *Methods of Mathematical Physics*, i. (Berlin, Springer, 1924), v. § 9, p. 261, eqn. 43, and further ii. § 10, 4, p. 76.

theory is not without significance. It becomes important first when the proper frequency ν_0 varies owing to the dissipation of heat. Formally it has to do with the old question of the "zero-point energy", which was raised in connection with the choice between the first and second forms of Planck's Theory. By the way, the additional term $\frac{h\nu_0}{2}$ also influences the law of the *band-edges*.

The *proper functions* (26) become, if we reintroduce the original q from (24) and (23),

$$(26') \quad \psi_n(q) = e^{-\frac{2\pi^2\nu_0 q^2}{h}} H_n\left(2\pi q \sqrt{\frac{\nu_0}{h}}\right).$$

Consideration of (27'') shows that the first function is a *Gaussian Error-curve*; the second vanishes at the origin and for x positive corresponds to a "Maxwell distribution of velocities" in two dimensions, and is continued in the manner of an odd function for x negative. The third function is even, is negative at the origin, and has two symmetrical zeros at $\pm \frac{1}{\sqrt{2}}$, etc. The curves can easily be sketched

roughly and it is seen that the roots of consecutive polynomials *separate* one another. From (26') it is also seen that the characteristic points of the proper functions, such as half-breadth (for $n=0$), zeros, and maxima, are, as regards order of magnitude, within the range of the classical vibration of the oscillator. For the classical *amplitude* of the n th vibration is readily found to be given by

$$(28) \quad q_n = \frac{\sqrt{E_n}}{2\pi\nu_0} = \frac{1}{2\pi} \sqrt{\frac{h}{\nu_0}} \sqrt{\frac{2n+1}{2}}.$$

Yet there is in general, as far as I see, no definite meaning that can be attached to the *exact* abscissa of the classical *turning points* in the graph of the proper function. It may, however, be conjectured, because the turning points have *this* significance for the phase space wave, that, at them, the square of the velocity of propagation becomes *infinite* and at greater distances becomes *negative*. In the differential equation (22), however, this only means the *vanishing* of the coefficient of ψ and gives rise to no singularities.

I would not like to suppress the remark here (and it is valid quite generally, not merely for the oscillator), that nevertheless this vanishing and becoming imaginary of the velocity of propagation is something which is very characteristic. It is the analytical reason for the selection of definite proper values, merely through the condition that the function should remain finite. I would like to illustrate this further. A wave equation with a *real* velocity of propagation means just this: there is an *accelerated* increase in the value of the function at all those points where its value is *lower* than the average of the values at neighbouring points, and vice versa. Such an equation, if not immediately and lastingly as in case of the *equation for the conduction of heat*, yet in the course of time, causes a *levelling*

of extreme values and does not permit at any point an excessive growth of the function. A wave equation with an *imaginary* velocity of propagation means the exact opposite: values of the function above the average of surrounding values experience an *accelerated increase* (or retarded decrease), and vice versa. We see, therefore, that a function represented by such an equation is in the greatest danger of growing beyond all bounds, and we must order matters skilfully to preserve it from this danger. The sharply defined proper values are just what makes this possible. Indeed, we can see in the example treated in Part I. that the demand for sharply defined proper values immediately ceases as soon as we choose the quantity E to be *positive*, as this makes the wave velocity real throughout all space.

After this digression, let us return to the oscillator and ask ourselves if anything is altered when we allow it two or more degrees of freedom (space oscillator, rigid body). If *different* mechanical proper frequencies (ν_0 -values) belong to the separate co-ordinates, then nothing is changed. ψ is taken as the *product* of functions, each of a single co-ordinate, and the problem splits up into just as many separate problems of the type treated above as there are co-ordinates present. The proper functions are products of Hermite orthogonal functions, and the proper values of the whole problem appear as sums of those of the separate problems, taken in every possible combination. No proper value (for the whole system) is multiple, if we presume that there is no rational relation between the ν_0 -values.

If, however, there is such a relation, then the same manner of treatment is still *possible*, but it is certainly not *unique*. Multiple proper values appear and the "separation" can certainly be effected in other co-ordinates, *e.g.* in the case of the isotropic space oscillator in spherical polars.¹

The proper values that we get, however, are certainly in each case exactly the same, at least in so far as we are able to prove the "completeness" of a system of proper functions, obtained in *one* way. We recognise here a complete parallel to the well-known relations which the method of the previous quantisation meets with in the case of *degeneracy*. Only in one point there is a not unwelcome formal difference. If we applied the Sommerfeld-Epstein quantum conditions *without* regard to a possible degeneracy then we always got the same energy levels, but reached different conclusions as to the paths permitted, according to the choice of co-ordinates.

Now that is *not* the case here. Indeed we come to a completely different system of proper functions, if we, for example, treat the vibration problem corresponding to unperturbed Kepler motion in

¹ We are led thus to an equation in r , which may be treated by the method shown in the Kepler problem of Part I. Moreover, the *one*-dimensional oscillator leads to the same equation if q^2 be taken as variable. I originally solved the problem directly in *that* way. For the hint that it was a question of Hermite polynomials, I have to thank Herr E. Fues. The polynomial appearing in the Kepler problem (eqn. 18 of Part I.) is the $(2n+1)$ th differential coefficient of the $(n+l)$ th polynomial of Laguerre, as I subsequently found.

parabolic co-ordinates instead of the polars used in Part I. However, it is not just the *single proper vibration* that furnishes a *possible state of vibration*, but an arbitrary, finite or infinite, *linear aggregate* of such vibrations. And as such the proper functions found in any second way may always be represented; namely, they may be represented as linear aggregates of the proper functions found in an arbitrary way, provided the latter form a *complete system*.

The question of how the energy is really distributed among the proper vibrations, which has not been taken into account here up till now, will, of course, have to be faced some time. Relying on the former quantum theory, we will be disposed to assume that in the degenerate case only the energy of the set of vibrations belonging to one definite proper value must have a certain prescribed value, which in the non-degenerate case belongs to one single proper vibration. I would like to leave this question still *quite open*—and also the question whether the discovered “energy levels” are really energy steps of the *vibration process* or whether they *merely* have the significance of its frequency. If we accept the beat theory, then the meaning of energy levels is no longer necessary for the explanation of sharp emission frequencies.

2. Rotator with Fixed Axis

On account of the lack of potential energy and because of the *Euclidean* line element, this is the simplest conceivable example of vibration theory. Let A be the moment of inertia and ϕ the angle of rotation, then we clearly obtain as the vibration equation

$$(29) \quad \frac{1}{A} \frac{d^2\psi}{d\phi^2} + \frac{8\pi^2 E}{h^2} \psi = 0,$$

which has the solution

$$(30) \quad \psi = \frac{\sin \left[\sqrt{\frac{8\pi^2 EA}{h^2}} \cdot \phi \right]}{\cos \left[\sqrt{\frac{8\pi^2 EA}{h^2}} \cdot \phi \right]}.$$

Here the argument must be an *integral* multiple of ϕ , simply because otherwise ψ would neither be single-valued nor continuous throughout the range of the co-ordinate ϕ , as we know $\phi + 2\pi$ has the same significance as ϕ . This condition gives the well-known result

$$(31) \quad E_n = \frac{n^2 h^2}{8\pi^2 A}$$

in *complete* agreement with the former quantisation.

No meaning, however, can be attached to the result of the application to band spectra. For, as we shall learn in a moment, it is a peculiar fact that our theory gives *another* result for the rotator with *free* axis. *And this is true in general.* It is not allowable in the applications of wave mechanics, to think of the freedom of movement of the system as being more strictly limited, in order to simplify calculation, than it *actually is*, even when we know from the integrals of the mechanical

equations that in a single movement certain definite freedoms are not made use of. For micro-mechanics, the fundamental system of mechanical equations is absolutely incompetent; the single paths with which it deals have now no separate existence. A wave process fills the *whole* of the phase space. It is well known that even the *number* of the dimensions in which a wave process takes place is very significant.

3. Rigid Rotator with Free Axis

If we introduce as co-ordinates the polar angles θ , ϕ of the radius from the nucleus, then for the kinetic energy as a function of the momenta we get

$$(32) \quad T = \frac{1}{2A} \left(p_\theta^2 + \frac{p_\phi^2}{\sin^2 \theta} \right).$$

According to its form this is the kinetic energy of a particle constrained to move on a spherical surface. The Laplacian operator is thus simply that part of the spatial Laplacian operator which depends on the polar angles, and the vibration equation (18'') takes the following form,

$$(33) \quad \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} + \frac{8\pi^2 A E}{\hbar^2} \psi = 0.$$

The postulation that ψ should be single-valued and continuous on the spherical surface leads to the proper value condition

$$(34) \quad \frac{8\pi^2 A}{\hbar^2} E = n(n+1); \quad n = 0, 1, 2, 3, \dots$$

The proper functions are known to be spherical surface harmonics. The energy levels are, therefore,

$$(34') \quad E_n = \frac{n(n+1)\hbar^2}{8\pi^2 A}; \quad n = 0, 1, 2, 3, \dots$$

This definition is different from all previous statements (except perhaps that of Heisenberg?). Yet, from various arguments from experiment we were led to put "half-integral" values for n in formula (31). It is easily seen that (34') gives practically the same as (31) with half-integral values of n . For

$$n(n+1) = \left(n + \frac{1}{2}\right)^2 - \frac{1}{4}.$$

The discrepancy consists only of a small additive constant; the level *differences* in (34') are the same as are got from "half-integral quantisation". This is true also for the application to short-wave bands, where the moment of inertia is not the same in the initial and final states, on account of the "electronic jump". For at most a small constant additional part comes in for *all* lines of a band, which is swamped in the large "electronic term" or in the "nuclear vibration term". Moreover, our previous analysis does not permit us to speak of this small part in any more definite way than as, say,

$$\frac{1}{4} \frac{\hbar^2}{8\pi^2} \left(\frac{1}{A} - \frac{1}{A'} \right).$$

The notion of the moment of inertia being fixed by "quantum conditions" for electronic motions and nuclear vibrations follows naturally from the whole line of thought developed here. We will show in the next section how we can treat, approximately at least, the nuclear vibrations and the rotations of the diatomic molecule simultaneously by a synthesis of the cases ¹ considered in 1 and 3.

I should like to mention also that the value $n=0$ corresponds not to the *vanishing* of the wave function ψ but to a *constant* value for it, and accordingly to a vibration with amplitude constant over the whole sphere.

4. Non-rigid Rotator (Diatomic Molecule)

According to the observation at the end of section 2, we must state the problem initially with all the six degrees of freedom that the rotator really possesses. Choose Cartesian co-ordinates for the two molecules, viz. x_1, y_1, z_1 ; x_2, y_2, z_2 , and let the masses be m_1 and m_2 , and r be their distance apart. The potential energy is

$$(35) \quad V = 2\pi^2\nu_0^2\mu(r - r_0)^2,$$

$$r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2.$$

where
Here

$$(36) \quad \mu = \frac{m_1 m_2}{m_1 + m_2}$$

may be called the "resultant mass". Then ν_0 is the mechanical proper frequency of the nuclear vibration, regarding the line joining the nuclei as fixed, and r_0 is the distance apart for which the potential energy is a minimum. These definitions are all in the sense of the usual mechanics.

For the vibration equation (18'') we get the following :

$$(37) \quad \left\{ \frac{1}{m_1} \left(\frac{\partial^2 \psi}{\partial x_1^2} + \frac{\partial^2 \psi}{\partial y_1^2} + \frac{\partial^2 \psi}{\partial z_1^2} \right) + \frac{1}{m_2} \left(\frac{\partial^2 \psi}{\partial x_2^2} + \frac{\partial^2 \psi}{\partial y_2^2} + \frac{\partial^2 \psi}{\partial z_2^2} \right) + \frac{8\pi^2}{h^2} [E - 2\pi^2\nu_0^2\mu(r - r_0)^2] \psi = 0. \right.$$

Introduce new independent variables $x, y, z, \xi, \eta, \zeta$, where

$$(38) \quad \begin{aligned} x &= x_1 - x_2; & (m_1 + m_2)\xi &= m_1 x_1 + m_2 x_2 \\ y &= y_1 - y_2; & (m_1 + m_2)\eta &= m_1 y_1 + m_2 y_2 \\ z &= z_1 - z_2; & (m_1 + m_2)\zeta &= m_1 z_1 + m_2 z_2. \end{aligned}$$

The substitution gives

$$(37') \quad \left\{ \frac{1}{\mu} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) + \frac{1}{m_1 + m_2} \left(\frac{\partial^2 \psi}{\partial \xi^2} + \frac{\partial^2 \psi}{\partial \eta^2} + \frac{\partial^2 \psi}{\partial \zeta^2} \right) + [a'' - b'(r - r_0)^2] \psi = 0, \right.$$

where for brevity

¹ Cf. A. Sommerfeld, *Atombau und Spektrallinien*, 4th edit., p. 833. We do not consider here the additional non-harmonic terms in the potential energy.

$$(39) \quad a'' = \frac{8\pi^2 E}{h^2}, \quad b' = \frac{16\pi^4 \nu_0^2 \mu}{h^2}.$$

Now we can put for ψ the product of a function of the relative co-ordinates x, y, z , and a function of the co-ordinates of the centre of mass ξ, η, ζ :

$$(40) \quad \psi = f(x, y, z) g(\xi, \eta, \zeta).$$

For g we get the defining equation

$$(41) \quad \frac{1}{m_1 + m_2} \left(\frac{\partial^2 g}{\partial \xi^2} + \frac{\partial^2 g}{\partial \eta^2} + \frac{\partial^2 g}{\partial \zeta^2} \right) + \text{const. } g = 0.$$

This is of the same form as the equation for the motion, under no forces, of a particle of mass $m_1 + m_2$. The constant would in this case have the meaning

$$(42) \quad \text{const.} = \frac{8\pi^2 E_t}{h^2},$$

where E_t is the energy of translation of the said particle. Imagine this value inserted in (41). The question as to the values of E_t admissible as proper values depends now on this, whether the whole infinite space is available for the original co-ordinates and hence for those of the centre of gravity without new potential energies coming in, or not. In the first case every non-negative value is permissible and every negative value not permissible. For when E_t is not negative and *only* then, (41) possesses solutions which do not vanish identically and yet remain finite in all space. If, however, the molecule is situated in a "vessel", then the latter must supply boundary conditions for the function g , or in other words, equation (41), on account of the introduction of further potential energies, will alter its form very abruptly at the walls of the vessel, and thus a discrete set of E_t -values will be selected as proper values. It is a question of the "Quantisation of the motion of translation", the main points of which I have lately discussed, showing that it leads to Einstein's Gas Theory.¹

For the factor f of the vibration function ψ , depending on the relative co-ordinates x, y, z , we get the defining equation

$$(43) \quad \frac{1}{\mu} \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right) + [a' - b'(r - r_0)^2] f = 0,$$

where for brevity we put

$$(39') \quad a' = \frac{8\pi^2(E - E_t)}{h^2}.$$

We now introduce instead of x, y, z , the spherical polars r, θ, ϕ (which is in agreement with the previous use of r). After multiplying by μ we get

$$(43') \quad \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2} \right\} + [\mu a' - \mu b'(r - r_0)^2] f = 0.$$

Now break up f . The factor depending on the angles is a surface harmonic. Let the order be n . The curled bracket is $-n(n+1)f$. Imagine this inserted and for simplicity let f now stand for the factor depending on r . Then introduce as new *dependent* variable

$$(44) \quad \chi = rf,$$

and as new *independent* variable

$$(45) \quad \rho = r - r_0.$$

The substitution gives

$$(46) \quad \frac{\partial^2 \chi}{\partial \rho^2} + \left[\mu a' - \mu b' \rho^2 - \frac{n(n+1)}{(r_0 + \rho)^2} \right] \chi = 0.$$

To this point the analysis has been exact. Now we will make an approximation, which I well know requires a stricter justification than I will give here. Compare (46) with equation (22') treated earlier. They agree in form and only differ in the coefficient of the unknown function by terms of the relative order of magnitude of $\frac{\rho}{r_0}$. This is seen, if we develop thus :

$$(47) \quad \frac{n(n+1)}{(r_0 + \rho)^2} = \frac{n(n+1)}{r_0^2} \left(1 - \frac{2\rho}{r_0} + \frac{3\rho^2}{r_0^2} - + \dots \right),$$

substitute in (46), and arrange in powers of ρ/r_0 . If we introduce for ρ a new variable differing only by a small constant, viz.

$$(48) \quad \rho' = \rho - \frac{n(n+1)}{r_0^3 \left(\mu b' + \frac{3n(n+1)}{r_0^4} \right)},$$

then equation (46) takes the form

$$(46') \quad \frac{\partial^2 \chi}{\partial \rho'^2} + \left(a - b \rho'^2 + \left[\frac{\rho'}{r_0} \right] \right) \chi = 0,$$

where we have put

$$(49) \quad \begin{cases} a = \mu a' - \frac{n(n+1)}{r_0^2} \left(1 - \frac{n(n+1)}{r_0^4 \mu b' + 3n(n+1)} \right) \\ b = \mu b' + \frac{3n(n+1)}{r_0^4}. \end{cases}$$

The symbol $\left[\frac{\rho'}{r_0} \right]$ in (46') represents terms which are small compared with the retained term of the order of $\frac{\rho'}{r_0}$.

Now we know that the *first* proper functions of equation (22'), to which we now compare (46'), only differ markedly from zero in a small range on both sides of the origin. Only those of higher order stretch gradually further out. For moderate orders, the domain for equation (46'), if we *neglect* the term $\left[\frac{\rho'}{r_0} \right]$ and bear in mind the

order of magnitude of molecular constants, is indeed small compared with r_0 . We thus conclude (without rigorous proof, I repeat), that we can in this way obtain a useful approximation for the first proper functions, within the region where they differ at all markedly from zero, and also for the first *proper values*. From the proper value condition (25) and omitting the abbreviations (49), (39'), and (39), though introducing the small quantity

$$(50) \quad \epsilon = \frac{n(n+1)\hbar^2}{16\pi^4\nu_0^2\mu^2r_0^4} = \frac{n(n+1)\hbar^2}{16\pi^4\nu_0^2A^2}$$

instead, we can easily derive the following *energy steps*,

$$(51) \quad \begin{cases} E = E_t + \frac{n(n+1)\hbar^2}{8\pi^2A} \left(1 - \frac{\epsilon}{1+3\epsilon}\right) + \frac{2l+1}{2} h\nu_0 \sqrt{1+3\epsilon} \\ (n=0, 1, 2 \dots; \quad l=0, 1, 2 \dots), \end{cases}$$

where

$$(52) \quad A = \mu r_0^2$$

is still written for the *moment of inertia*.

In the language of classical mechanics, ϵ is the square of the ratio of the frequency of rotation to the vibration frequency ν_0 ; it is therefore really a small quantity in the application to the molecule, and formula (51) has the usual structure, apart from this small correction and the other differences already mentioned. It is the synthesis of (25') and (34') to which E_t is added as representing the energy of translation. It must be emphasized that the value of the approximation is to be judged not only by the smallness of ϵ but also by l not being too large. *Practically*, however, only small numbers have to be considered for l .

The ϵ -corrections in (51) do *not yet* take account of deviations of the nuclear vibrations from the pure harmonic type. Thus a comparison with Kratzer's formula (*vide* Sommerfeld, *loc. cit.*) and with experience is impossible. I only desired to mention the case provisionally, as an example showing that the intuitive idea of the *equilibrium configuration* of the nuclear system retains its meaning in undulatory mechanics also, and showing the manner in which it does so, provided that the wave amplitude ψ is different from zero practically only in a small neighbourhood of the equilibrium configuration. The direct interpretation of this wave function of *six* variables in *three-dimensional* space meets, at any rate initially, with difficulties of an abstract nature.

The rotation-vibration-problem of the diatomic molecule will have to be re-attacked presently, the non-harmonic terms in the energy of binding *being taken into account*. The method, selected skilfully by Kratzer for the classical mechanical treatment, is also suitable for undulatory mechanics. If, however, we are going to push the calculation as far as is necessary for the fineness of band structure, then we must make use of the theory of the *perturbation of proper values and functions*, that is, of the alteration experienced by a definite proper value and the appertaining proper functions of a

differential equation, when there is added to the coefficient of the unknown function in the equation a small "disturbing term". This "perturbation theory" is the complete counterpart of that of classical mechanics, except that it is simpler because in undulatory mechanics we are always in the domain of *linear* relations. As a first approximation we have the statement that the perturbation of the proper value is equal to the perturbing term averaged "over the undisturbed motion".

The perturbation theory broadens the analytical range of the new theory extraordinarily. As an important practical success, let me say here that the *Stark effect* of the first order will be found to be really completely in accord with Epstein's formula, which has become unimpeachable through the confirmation of experience.

Zürich, Physical Institute of the University.

(Received February 23, 1926.)

The Continuous Transition from Micro- to Macro-Mechanics

(*Die Naturwissenschaften*, 28, pp. 664-666, 1926)

BUILDING on ideas of de Broglie¹ and Einstein,² I have tried to show³ that the usual differential equations of mechanics, which attempt to define the co-ordinates of a mechanical system as functions of the time, are no longer applicable for "small" systems; instead there must be introduced a certain *partial* differential equation, which defines a variable ψ ("wave function") as a function of the co-ordinates and the time. As in the differential equation of a vibrating string or of any other vibrating system, ψ is given as a superposition of pure time harmonic (*i.e.* "sinusoidal") vibrations, the frequencies of which agree exactly with the spectroscopic "term frequencies" of the micro-mechanical system. For example, in the case of the linear Planck oscillator⁴ where the energy function is

$$(1) \quad \frac{m}{2} \left(\frac{dq}{dt} \right)^2 + 2\pi^2 \nu_0^2 m q^2,$$

when we put, instead of the displacement q , the dimensionless variable

$$(2) \quad x = q \cdot 2\pi \sqrt{\frac{m\nu_0}{h}},$$

we get ψ as the superposition of the following proper vibrations:⁵

$$(3) \quad \left\{ \begin{array}{l} \psi_n = e^{-\frac{x^2}{2}} H_n(x) e^{2\pi i \nu_n t} \\ (\nu_n = \frac{2n+1}{2} \nu_0; \quad n = 0, 1, 2, 3 \dots) \end{array} \right.$$

The H_n 's are the polynomials⁶ named after Hermite. If they are

¹ L. de Broglie, *Ann. de Physique* (10), 3, p. 22, 1925 (Thèses, Paris, 1924).

² A. Einstein, *Berlin Ber.* 1925, p. 9 *et seq.*

³ *Ann. d. Physik*; the essays here collected.

⁴ *i.e.* a particle of mass m which, moving in a straight line, is attracted towards a fixed point in it, with a force proportional to its displacement q from this point; according to the usual mechanics, such a particle executes sine vibrations of frequency ν_0 .

⁵ i means $\sqrt{-1}$. On the right-hand side the real part is to be taken, as usual.

⁶ Cf. Courant-Hilbert, *Methoden der mathematischen Physik*, I. chap. ii. § 10, 4, p. 76 (Berlin, Springer, 1924).

multiplied by $e^{-\frac{x^2}{2}}$ and the "normalising factor" $(2^n n!)^{-\frac{1}{2}}$ they are called Hermite's orthogonal functions. They represent therefore the amplitudes of the proper vibrations.

The first five are represented in Fig. 1. The similarity between this and the well-known picture of the vibrations of a string is very great.

At first sight it appears very strange to try to describe a process, which we previously regarded as belonging to particle mechanics, by a system of such proper vibrations. For this chosen simple case, I would like to demonstrate here *in concreto* the transition to macroscopic mechanics, by showing that a *group* of proper vibrations of *high* order-number n ("quantum number") and of relatively small order-number differences ("quantum number differences") may represent a "particle", which is executing the "motion", expected

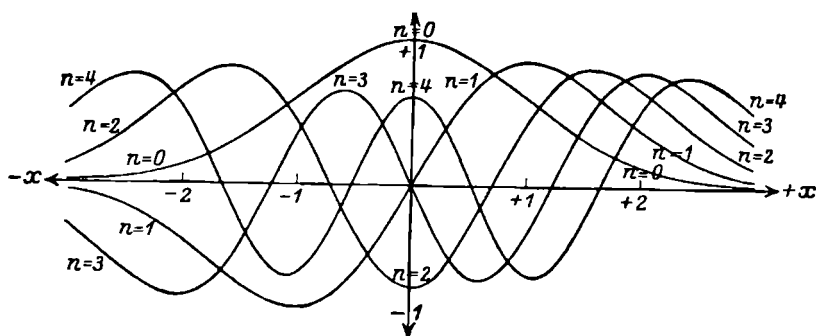


FIG. 1.—The first five proper vibrations of the Planck oscillator according to undulatory mechanics. Outside of the region $-3 \leq x \leq +3$ represented here, all five functions approach the x -axis in monotonic fashion.

from the usual mechanics, *i.e.* oscillating with the frequency ν_0 . I choose a number $A \gg 1$ (*i.e.* great compared with 1) and form the following aggregate of proper vibrations:

$$(4) \quad \psi = \sum_{n=0}^{\infty} \left(\frac{A}{2}\right)^n \frac{\psi_n}{n!} = e^{\pi i \nu_0 t} \sum_{n=0}^{\infty} \left(\frac{A}{2} e^{2\pi i \nu_0 t}\right)^n \frac{1}{n!} e^{-\frac{x^2}{2}} H_n(x).$$

Thus the *normalised* proper vibrations (see above) are taken with the coefficients

$$(5) \quad \frac{A^n}{\sqrt{2^n n!}},$$

and this, as is easily seen,¹ results in the singling out of a relatively small group in the neighbourhood of the n -value given by

$$(6) \quad n = \frac{A^2}{2}.$$

¹ $z^n/n!$ has, as function of n , for large values of z , a single extremely high and relatively very sharp maximum at $n=z$. By taking square roots and with $z=A^2/2$, we get the series of numbers (5).

The summation of the series (4) is made possible by the following identity¹ in x and s :

$$(7) \quad \sum_{n=0}^{\infty} \frac{s^n}{n!} e^{-\frac{x^2}{2}} H_n(x) = e^{-s^2 + 2sx - \frac{x^2}{2}}.$$

Thus

$$(8) \quad \psi = e^{\pi i \nu_0 t - \frac{A^2}{4} e^{4\pi i \nu_0 t} + A x e^{2\pi i \nu_0 t} - \frac{x^2}{2}}.$$

Now we take, as is provided for, the real part of the right-hand side and after a short calculation obtain

$$(9) \quad \psi = e^{\frac{A^2}{4} - \frac{1}{2}(x - A \cos 2\pi \nu_0 t)^2} \cos \left[\pi \nu_0 t + (A \sin 2\pi \nu_0 t) \cdot \left(x - \frac{A}{2} \cos 2\pi \nu_0 t \right) \right].$$

This is the *final result*, in which the first factor is our first interest. It represents a relatively tall and narrow "hump", of the form of a "Gaussian error-curve" which at a given moment lies in the neighbourhood of the position

$$(10) \quad x = A \cos 2\pi \nu_0 t.$$

The breadth of the hump is of the order of magnitude unity and therefore very small compared with A , by hypothesis. According to (10), the hump oscillates under exactly the same law as would operate in the usual mechanics for a particle having (1) as its energy function. The amplitude in terms of x is A , and thus in terms of q is

$$(11) \quad a = \frac{A}{2\pi} \sqrt{\frac{h}{m\nu_0}}.$$

Ordinary mechanics gives for the *energy* of a particle of mass m , which oscillates with this amplitude and with frequency ν_0 ,

$$(12) \quad 2\pi^2 a^2 \nu_0^2 m = \frac{A^2}{2} h \nu_0,$$

i.e. from (6) exactly $nh\nu_0$, where n is the average quantum number of the selected group. The "correspondence" is thus complete in this respect also.

The *second* factor in (9) is in general a function whose absolute value is small compared with unity, and which varies very rapidly with x and also t . It ploughs many deep and narrow furrows in the profile of the first factor, and makes a *wave group* out of it, which is represented—schematically only—in Fig. 2. The x -scale of Fig. 2 is naturally much smaller than that of Fig. 1; Fig. 2 requires to be magnified five times before being directly compared with Fig. 1. A more exact consideration of the second factor of (9) discloses the following interesting details, which cannot be seen in Fig. 2, which only represents *one* stage. The *number* and *breadth* of the "furrows" or "wavelets" within the particle vary with the time. The wavelets are most *numerous* and *narrowest* when passing through the centre $x=0$; they become *completely smoothed out* at the turning points $x = \pm A$, because

¹ Courant-Hilbert, *loc. cit.* eqn. (58).

there, by (10), $\cos 2\pi\nu_0 t = \pm 1$ and thus $\sin 2\pi\nu_0 t$ becomes equal to zero, so that the second factor of (9) is absolutely independent of x . The entire extension of the wave group ("density of the particle") remains, however, always the same. The variability of the "corrugation" is to be conceived as depending on the *velocity*, and, as such, is completely intelligible from all general aspects of undulatory mechanics—but I do not wish to discuss this further at present.

Our wave group always remains compact, and does not spread out into larger regions as time goes on, as we were accustomed to make it do, for example, in optics. It is admitted that this does not mean much in one dimension, and that a hump on a string will behave quite similarly. But it is easily seen that, by multiplying together two or three expressions like (4), written in x , in y , and in z respectively, we can represent also the *plane* and the *spatial* oscillator respectively, *i.e.* a plane or spatial wave group which moves round a harmonic ellipse.¹ Also such a wave group will remain compact, in contrast, *e.g.*, to a

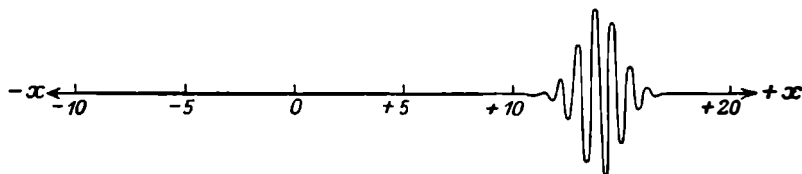


FIG. 2.—Oscillating wave group as the representation of a particle in wave mechanics.

wave packet in classical optics, which is dissipated in the course of time. The distinction may originate in the fact that our group is built up out of separate *discrete* harmonic components, and not out of a *continuum* of such.

I wish to mention, in conclusion, that a general additive constant, C , let us say, which should be added to all the ν_n 's in (3), (and corresponds to the "rest-energy" of the particle) does not alter the essentials. It only affects the square bracket in (9), adding $2\pi Ct$ thereto. Hence the oscillations *within* the wave group become very much quicker with respect to the *time*, while the oscillation of the group as a whole, given by (10), and its "corrugation", remain quite unaffected.

We can definitely foresee that, in a similar way, wave groups can be constructed which move round highly quantised Kepler ellipses and are the representation by wave mechanics of the hydrogen electron. But the technical difficulties in the calculation are greater than in the especially simple case which we have treated here.

¹ We may point out, in passing, the interesting fact that the quantum levels of the *plane* oscillator are *integral*, but for the *spatial* oscillator they again become "half-integral". Similarly for the rotator. This half-integralness, which is spectroscopically so significant, is thus connected with the "oddness" of the number of the dimensions of space.

On the Relation between the Quantum Mechanics of Heisenberg, Born, and Jordan, and that of Schrödinger

(*Annalen der Physik* (4), vol. 79, 1926)

§ 1. Introduction and Abstract

CONSIDERING the extraordinary differences between the starting-points and the concepts of Heisenberg's quantum mechanics¹ and of the theory which has been designated "undulatory" or "physical" mechanics,² and has lately been described here, it is very strange that these two new theories agree *with one another* with regard to the known facts, where they differ from the old quantum theory. I refer, in particular, to the peculiar "half-integralness" which arises in connection with the oscillator and the rotator. That is really very remarkable, because starting-points, presentations, methods, and in fact the whole mathematical apparatus, seem fundamentally different. Above all, however, the departure from classical mechanics in the two theories seems to occur in diametrically opposed directions. In Heisenberg's work the classical continuous variables are replaced by systems of discrete numerical quantities (matrices), which depend on a pair of integral indices, and are defined by *algebraic* equations. The authors themselves describe the theory as a "true theory of a discontinuum".³ On the other hand, wave mechanics shows just the reverse tendency; it is a step from classical point-mechanics towards a *continuum-theory*. In place of a process described in terms of a finite number of dependent variables occurring in a finite number of total differential equations, we have a continuous *field-like* process in

¹ W. Heisenberg, *Ztschr. f. Phys.* 33, p. 879, 1925; M. Born and P. Jordan, *idem* 34, p. 858, 1925, and 35, p. 557, 1926 (the latter in collaboration with Heisenberg). I may be allowed, for brevity's sake, to replace the three names simply by Heisenberg, and to quote the last two essays as "Quantum Mechanics I. and II." Interesting contributions to the theory have also been made by P. Dirac, *Proc. Roy. Soc.*, London, 109, p. 642, 1925, and *idem* 110, p. 561, 1926.

² E. Schrödinger. Parts I. and II. in this collection. These parts will be continued quite independently of the present paper, which is only intended to serve as a connecting link.

³ "Quantum Mechanics I." p. 879.

configuration space, which is governed by a single *partial* differential equation, derived from a principle of action. This principle and this differential equation replace the equations of motion *and* the quantum conditions of the older "classical quantum theory".¹

In what follows the very intimate *inner connection* between Heisenberg's quantum mechanics and my wave mechanics will be disclosed. From the formal mathematical standpoint, one might well speak of the *identity* of the two theories. The train of thought in the proof is as follows.

Heisenberg's theory connects the solution of a problem in quantum mechanics with the solution of a system of an infinite number of algebraic equations, in which the unknowns—infinite matrices—are allied to the classical position- and momentum-co-ordinates of the mechanical system, and functions of these, and obey peculiar *calculating rules*. (The relation is this: to *one* position-, *one* momentum-co-ordinate, or to *one* function of these corresponds always *one* infinite matrix.)

I will first show (§§ 2 and 3) how to each function of the position- and momentum-co-ordinates there may be related a matrix in such a manner, that these matrices, *in every case, satisfy* the formal calculating rules of Born and Heisenberg (among which I also reckon the so-called "quantum condition" or "interchange rule"; see below). This relation of matrices to functions is *general*; it takes no account of the *special* mechanical system considered, but is the same for all mechanical systems. (In other words: the particular Hamilton function does not enter into the connecting law.) However, the relation is still *indefinite* to a great extent. It arises, namely, from the *auxiliary introduction* of an *arbitrary* complete orthogonal system of functions having for domain *entire configuration space* (*N.B.*—not "*pq*-space", but "*q*-space"). The provisional *indefiniteness* of the relation lies in the fact that we can assign the *auxiliary rôle* to an *arbitrary* orthogonal system.

After matrices are thus constructed in a very general way, so as to satisfy the general rules, I will show the following in § 4. The *special* system of algebraic equations, which, in a *special* case, connects the *matrices* of the position and impulse co-ordinates with the *matrix* of the Hamilton function, and which the authors call "equations of motion", will be completely solved by assigning the auxiliary rôle to a *definite* orthogonal system, namely, to the system of *proper functions* of that partial differential equation which forms the basis of my wave mechanics. The solution of the natural *boundary-value problem* of this differential equation is *completely equivalent* to the solution of Heisenberg's algebraic problem. *All* Heisenberg's matrix elements, which

¹ My theory was inspired by L. de Broglie, *Ann. de Physique* (10) 3, p. 22, 1925 (*Theses*, Paris, 1924), and by brief, yet infinitely far-seeing remarks of A. Einstein, *Berl. Ber.*, 1925, p. 9 *et seq.* I did not at all suspect any relation to Heisenberg's theory at the beginning. I naturally knew about his theory, but was discouraged, if not repelled, by what appeared to me as very difficult methods of transcendental algebra, and by the want of perspicuity (*Anschaulichkeit*).

may interest us from the surmise that they define "transition probabilities" or "line intensities", can be actually evaluated by *differentiation and quadrature*, as soon as the *boundary-value problem* is solved. Moreover, in wave mechanics, these matrix elements, or quantities that are closely related to them, have the perfectly clear significance of amplitudes of the partial oscillations of the atom's electric moment. The intensity and polarisation of the emitted light is thus intelligible *on the basis of the Maxwell-Lorentz theory*. A short preliminary sketch of this relationship is given in § 5.

§ 2. The Co-ordination of an Operator and of a Matrix with a Well-arranged Function-symbol and the Establishment of the Product Rule

The starting-point in the construction of matrices is given by the simple observation that Heisenberg's peculiar calculating laws for functions of the *double* set of n quantities, $q_1, q_2, \dots, q_n; p_1, p_2, \dots, p_n$ (position- and canonically conjugate momentum-co-ordinates) agree exactly with the rules, which *ordinary analysis* makes *linear differential operators* obey in the domain of the *single* set of n variables, q_1, q_2, \dots, q_n . So the *co-ordination* has to occur in such a manner that each p_i in the *function* is to be replaced by the operator $\frac{\partial}{\partial q_i}$. Actually the operator $\frac{\partial}{\partial q_i}$ is *exchangeable* with $\frac{\partial}{\partial q_m}$, where m is arbitrary, but with q_m only, if $m \neq i$. The operator, obtained by interchange and subtraction when $m = i$, viz.

$$(1) \quad \frac{\partial}{\partial q_i} q_i - q_i \frac{\partial}{\partial q_i},$$

when applied to any arbitrary function of the q 's, *reproduces* the function, i.e. this operator gives *identity*. This simple fact will be reflected in the domain of matrices as Heisenberg's interchange rule.

After this preliminary survey, we turn to systematic construction. Since, as noticed above, the interchangeability does not *always* hold good, then a definite operator does not correspond uniquely to a definite "function in the usual sense" of the q 's and p 's, but to a "function-symbol written in a definite way". Moreover, since we can perform only the operations of addition and multiplication with the operators $\frac{\partial}{\partial q_k}$, the function of the q 's and p 's must be written as a regular power series in p at least, before we substitute $\frac{\partial}{\partial q_i}$ for p_i . It is sufficient to carry out the process for a single term of such a power series, and thus for a function of the following construction :

$$(2) \quad F(q_k, p_k) = f(q_1 \dots q_n) p_r p_s p_t g(q_1 \dots q_n) p_r h(q_1 \dots q_n) p_r p_s \dots$$

We wish to express this as a "well-arranged¹ function-symbol" and relate it to the following operator,

¹ Or "well-ordered."

$$(3) \quad [F, \cdot] = f(q_1 \dots q_n) K^3 \frac{\partial^3}{\partial q_r \partial q_s \partial q_t} g(q_1 \dots q_n) K \frac{\partial}{\partial q_r} \\ h(q_1 \dots q_n) K^2 \frac{\partial^2}{\partial q_r \partial q_s} \dots$$

wherein, somewhat more generally than in the preliminary survey, p_r is not replaced by $\frac{\partial}{\partial q_r}$ simply, but by $K \frac{\partial}{\partial q_r}$, and K stands for a universal constant. As an abbreviation for the operator arising out of the well-arranged function F , I have introduced the symbol $[F, \cdot]$ in passing (*i.e.* only for the purpose of the present proof). The function (in the usual sense) of $q_1 \dots q_n$, which is obtained by using the operator on another function (in the usual sense), $u(q_1 \dots q_n)$, will be denoted by $[F, u]$. If G is another well-arranged function, then $[GF, u]$ will denote the function u after the operator of F has first been used on it, and then the operator of G ; or, what is defined to be the same, when the operator of GF has been used. Of course this is not generally the same as $[FG, u]$.

Now we connect a *matrix* with a well-arranged function, like F , by means of its operator (3) and of an arbitrary complete orthogonal system having for its domain the whole of q -space. It is done as follows. For brevity we will simply write x for the group of variables q_1, q_2, \dots, q_n , as is usual in the theory of Integral Equations, and write $\int dx$ for an integral extending over the whole of q -space.

The functions

$$(4) \quad u_1(x) \sqrt{\rho(x)}, \quad u_2(x) \sqrt{\rho(x)}, \quad u_3(x) \sqrt{\rho(x)} \dots \text{ad inf.}$$

are now to form a complete orthogonal system, normalised to 1.

Let, therefore, in every case

$$(5) \quad \begin{cases} \int \rho(x) u_i(x) u_k(x) dx = 0 & \text{for } i \neq k \\ = 1 & \text{for } i = k. \end{cases}$$

Further, it is postulated that these functions vanish at the natural *boundary* of q -space (in general, infinity) in a way sufficient to cause the vanishing of certain boundary integrals which come in later on as secondary products after certain integrations by parts.

By the operator (3) we now relate the following *matrix*,

$$(6) \quad F^{kl} = \int \rho(x) u_k(x) [F, u_l(x)] dx,$$

to the function F represented by (2). (The way of writing the indices on the left-hand side must not suggest the idea of "contravariance"; from this point of view, here discarded, *one* index was formerly written above, and the other below; we write the matrix indices *above*, because later we will also have to write matrix elements, corresponding to the q 's and p 's, where the lower place is already occupied.) In words: a matrix element is computed by *multiplying* the function of

the orthogonal system denoted by the *row*-index (whereby we understand always u_i , not $u_i\sqrt{\rho}$), by the "density function" ρ , and by the result arising from using our operator on the orthogonal function corresponding to the *column*-index, and then by *integrating* the whole over the domain.¹

It is not very difficult to show that additive and multiplicative combination of well-arranged functions or of the appertaining operators works out as matrix addition and matrix multiplication of the allied matrices. For addition the proof is trivial. For multiplication the proof runs as follows. Let G be any other well-arranged function, like F , and

$$(7) \quad G^{lm} = \int \rho(x) u_l(x) [G, u_m(x)] dx,$$

the matrix corresponding. We wish to form the product matrix

$$(FG)^{km} = \sum_l F^{kl} G^{lm}.$$

Before writing it, let us transform the expression (6) for F^{kl} as follows. By a series of integrations by parts, the operator $[F, \cdot]$ is "revolved" from the function $u_l(x)$ to the function $\rho(x)u_k(x)$. By the expression "revolve" (instead of, say, "push") I wish to convey that the *sequence* of the operations reverses itself exactly thereby. The boundary integrals, which come in as "by-products", are to disappear (see above). The "revolved" operator, including the change of sign that accompanies an odd number of differentiations, will be denoted by $[\bar{F}, \cdot]$. For example, from (3) comes

$$(3') \quad [\bar{F}, \cdot] = (-1)^\tau \dots K^2 \frac{\partial^2}{\partial q_s \partial q_r} h(q_1 \dots q_n) K \frac{\partial}{\partial q_r} \\ g(q_1 \dots q_n) K^3 \frac{\partial^3}{\partial q_i \partial q_s \partial q_r} f(q_1 \dots q_n),$$

where τ = number of differentiations. By applying this symbol, we have

$$(6') \quad F^{kl} = \int u_l(x) [\bar{F}, \rho(x)u_k(x)] dx.$$

If we now calculate the product matrix, we get

$$(8) \quad \sum_l F^{kl} G^{lm} \\ = \sum_l \left\{ \int u_l(x) [\bar{F}, \rho(x)u_k(x)] dx \cdot \int \rho(x)u_l(x) [G, u_m(x)] dx \right\} \\ = \int [\bar{F}, \rho(x)u_k(x)] [G, u_m(x)] dx.$$

The last equation is simply the so-called "relation of completeness"

¹ More briefly: F^{kl} is the k th "development coefficient" of the operator used on the function u_l .

of our orthogonal system,¹ applied to the "development coefficients" of the functions

$$[G, u_m(x)] \quad \text{and} \quad \frac{1}{\rho(x)}[F, \rho(x)u_k(x)].$$

Now in (8), let us revolve, by further integrations by parts, the operator $[\bar{F}, \cdot]$ from the function $\rho(x)u_k(x)$ back again to the function $[G, u_m(x)]$, so that the operator regains its original form. We clearly get

$$(9) \quad (FG)^{km} = \sum_l F^{kl}G^{lm} = \int \rho(x)u_k(x)[FG, u_m(x)]dx.$$

On the left is the (km) th element of the product matrix, and on the right, by the law of connection (6), stands the (km) th element of the matrix, corresponding to the well-arranged product FG . Q.E.D.

§ 3. Heisenberg's Quantum Condition and the Rules for Partial Differentiation

Since operation (1) gave identity, then corresponding to the well-arranged function

$$(10) \quad p_i q_i - q_i p_i$$

we have the *operator*, multiplication by K , in accordance with our law of connection, in which we incorporated a universal constant K . Hence to function (10) corresponds the *matrix*

$$(11) \quad (p_i q_i - q_i p_i)^{ik} = K \int \rho(x)u_i(x)u_k(x)dx = 0 \quad \text{for } i \neq k \\ = K \quad \text{for } i = k.$$

That is Heisenberg's "quantum relation" if we put

$$(12) \quad K = \frac{h}{2\pi\sqrt{-1}},$$

and this may be assumed to hold from now on. It is understood that we could have also found relation (11) by taking the two matrices allied to q_i and p_i , viz.

$$(13) \quad q_i^{ik} = \int q_i \rho(x)u_i(x)u_k(x)dx, \\ p_i^{ik} = K \int \rho(x)u_i(x) \frac{\partial u_k(x)}{\partial q_i} dx,$$

multiplying them together in different sequence and subtracting the two results.

Let us now turn to the "rules for partial differentiation". A well-arranged function, like (2), is said to be differentiated partially with respect to q_i , when it is differentiated with respect to q_i without

¹ See, e.g., Courant-Hilbert, *Methods of Mathematical Physics*, I., p. 36. It is important to remember that the "relation of completeness" for the "development coefficients" is valid in every case, even when the developments themselves do not converge. If these do converge, then the equivalence (8) is directly evident.

altering the succession of the factors *at each place* where q_i appears in it, and all these results are added.¹ Then it is easy to show that the following equation between the operators is valid :

$$(14) \quad \left[\frac{\partial F}{\partial q_i}, \cdot \right] = \frac{1}{K} [p_i F - F p_i, \cdot].$$

The line of thought is this. Instead of really differentiating with respect to q_i , it is very convenient simply to prefix p_i to the function ; as it is, p_i must finally be replaced by $K \frac{\partial}{\partial q_i}$. Obviously I have to divide by K . Furthermore, when we apply the entire operator to any function u , the operator $\frac{\partial}{\partial q_i}$ will act not only on that part of F which contains q_i (as it *ought*), but also *wrongly* on the function u , affected by the entire operator. *This mistake is exactly corrected* by subtracting again the operation $[F p_i, \cdot]!$

Consider now partial differentiation with respect to a p_i . Its meaning for a well-arranged function, like (2), is a little simpler than in the case of $\frac{\partial}{\partial q_i}$, because the p 's only appear as power products.

We imagine every power of p_i to be resolved into single factors, *e.g.* think of $p_i p_i p_i$ instead of p_i^3 , and we can then say : in partial differentiation with respect to p_i , every *separate* p_i that appears in F is to be *dropped* once, all the other p_i 's remaining ; all the results obtained are to be added. What will be the effect on the operator (3) ?

“Every separate $K \frac{\partial}{\partial q_i}$ is to be dropped once, and all the results so obtained are to be added.”

I maintain that on this reasoning the operational equation

$$(15) \quad \left[\frac{\partial F}{\partial p_i}, \cdot \right] = \frac{1}{K} [F q_i - q_i F, \cdot]$$

is valid. Actually, I picture the operator $[F q_i, \cdot]$ as formed and now attempt to “push q_i through F from right to left”, that means, attempt to arrive at the operator $[q_i F, \cdot]$ through successive exchanges. This pushing through meets an obstacle only as often as I come against a $\frac{\partial}{\partial q_i}$. With the latter I may not interchange q_i simply, but have to replace

$$(16) \quad \frac{\partial}{\partial q_i} \text{ by } 1 + q_i \frac{\partial}{\partial q_i}$$

in the interior of the operator. The secondary products of the interchange, which are yielded by this “uniformising”, form just the

¹ We are naturally following Heisenberg faithfully in all these definitions. From a strictly logical standpoint the following proof is evidently superfluous, and we could have written down rules (14) and (15) right away, as they are proved in Heisenberg, and only depend upon the sum and product rules and the exchange rule (11) which we have proved.

desired "partial differential coefficients", as is easily seen. After the pushing-through process is finished, the operator $[q_l F, \cdot]$ still remains left over. It would be superfluous and therefore is explicitly subtracted in (15). Hence (15) is proved. The equations (14) and (15), which have been proved for *operators*, naturally hold good unchanged for the matrices belonging to the right-hand and left-hand sides, because by (6) one matrix, and one only, belongs to one linear operator (after the system $u_i(x)$ has been chosen once for all).¹

§ 4. The Solution of Heisenberg's Equations of Motion

We have now shown that matrices, constructed according to definitions (3) and (6) from well-arranged functions by the agency of an arbitrary, complete orthogonal system (4), satisfy all Heisenberg's calculating rules, including the interchange rule (11). Now let us consider a special mechanical problem, characterised by a definite Hamilton function

$$(17) \quad H(q_k, p_k).$$

The authors of quantum mechanics take this function over from *ordinary* mechanics, which naturally does *not* give it in a "well-arranged" form; for in ordinary analysis no stress is laid on the sequence of the factors. They therefore "normalise" or "symmetralise" the function in a definite manner for their purposes. For example, the usual mechanical function $q_k p_k^2$ is replaced by

$$\frac{1}{2}(p_k^2 q_k + q_k p_k^2)$$

¹ In passing it may be noted that the converse of this theorem is also true, at least in the sense that certainly *not more than one* linear differential operator can belong to a given *matrix*, according to our connecting law (6), when the orthogonal system and the density function are prescribed. For in (6), let the F^{kl} 's be given, let $[F, \cdot]$ be the linear operator we are *seeking* and which we *presume to exist*, and let $\phi(x)$ be a function of q_1, q_2, \dots, q_n , which is sectionally continuous and differentiable as often as necessary, but otherwise arbitrary. Then the *relation of completeness* applied to the functions $\phi(x)$ and $[F, u_k(x)]$ yields the following:

$$\int \rho(x) \phi(x) [F, u_k(x)] dx = \sum_i \left\{ \int \rho(x) \phi(x) u_i(x) dx \cdot \int \rho(x) u_i(x) [F, u_k(x)] dx \right\}.$$

The right-hand side can be regarded as definitely known, for in it occur only development coefficients of $\phi(x)$ and the prescribed matrix elements F^{ik} . By "revolving" (see above), we can change the left-hand side into the k th development coefficient of the function

$$\frac{[\overline{F}, \rho(x) \phi(x)]}{\rho(x)}.$$

Thus all the development coefficients of this function are uniquely fixed, and thus so is the function itself (Courant-Hilbert, p. 37). Since, however, $\rho(x)$ was fixed beforehand and $\phi(x)$ is a quite arbitrary function, we can say: the result of the action of the *revolved* operator on an *arbitrary* function, provided, of course, it can be submitted to the operator at all, is fixed *uniquely* by the matrix F^{ki} . This can only mean that the *revolved operator* is uniquely fixed, for the notion of "operator" is logically identical with the whole of the results of its action. By revolving the revolved operator, we obtain uniquely the operator we have sought, itself.

It is to be noted that the *developability* of the functions which appear is *not* necessarily postulated—we have not proved that a linear operator, corresponding to an arbitrary matrix, *always exists*.

or by

$$p_k q_k p_k$$

or by

$$\frac{1}{3}(p_k^2 q_k + p_k q_k p_k + q_k p_k^2),$$

which are all the same, according to (11). This function is then "well-arranged", *i.e.* the sequence of the factors is inviolable. I will not enter into the general rule for symmetricalising here; ¹ the idea, if I understand it aright, is that H^{ki} is to be a *diagonal matrix*, and in other respects the normalised function, regarded as one of ordinary analysis, is to be identical with the one originally given. ² We will satisfy these demands in a direct manner.

Then the authors postulate that the *matrices* q_i^{ik} , p_i^{ik} shall satisfy an infinite system of equations, as "equations of motion", and to begin with they write this system as follows:

$$(18) \quad \left. \begin{aligned} \left(\frac{dq_l}{dt}\right)^{ik} &= \left(\frac{\partial H}{\partial p_l}\right)^{ik} \\ \left(\frac{dp_l}{dt}\right)^{ik} &= \left(-\frac{\partial H}{\partial q_l}\right)^{ik} \end{aligned} \right\} \begin{array}{l} l = 1, 2, 3, \dots n \\ i, k = 1, 2, 3, \dots \text{ad. inf.} \end{array}$$

The upper pair of indices signifies, as before in F^{ki} , the respective element of the matrix belonging to the well-arranged function in question. The meaning of the partial differential coefficient on the right-hand side has just been explained, but *not* that of the $\frac{d}{dt}$ appearing on the left. By it the authors signify the following. It is to give a series of *numbers*

$$(19) \quad \nu_1, \nu_2, \nu_3, \nu_4, \dots \text{ad. inf.},$$

such that the above equations are fulfilled, when to the $\frac{d}{dt}$ is ascribed the meaning: multiplication of the (ik) th matrix element by $2\pi\sqrt{-1}$ ($\nu_i - \nu_k$). Thus, in particular,

$$(20) \quad \left\{ \begin{aligned} \left(\frac{dq_l}{dt}\right)^{ik} &= 2\pi\sqrt{-1}(\nu_i - \nu_k)q_l^{ik}; \\ \left(\frac{dp_l}{dt}\right)^{ik} &= 2\pi\sqrt{-1}(\nu_i - \nu_k)p_l^{ik}. \end{aligned} \right.$$

The series of numbers (19) is not defined in any way beforehand, but together with the matrix elements q_i^{ik} , p_i^{ik} , they form the numerical unknowns of the system of equations (18). The latter assumes the form

$$(18') \quad \left\{ \begin{aligned} (\nu_i - \nu_k)q_l^{ik} &= \frac{1}{h}(Hq_l - q_l H) \\ (\nu_i - \nu_k)p_l^{ik} &= \frac{1}{h}(Hp_l - p_l H) \end{aligned} \right.$$

¹ "Quantum Mechanics I." p. 873 *et seq.*

² The *stricter* postulation—"shall yield the same quantum-mechanical equations of motion"—I consider too narrow. It arises, in my opinion, from the fact that the authors confine themselves to *power products with regard also to the q_k 's*—which is unnecessary.

when we utilise the explanation of the symbols (20), and the calculating rules (14) and (15), and take account of (12).

We must thus satisfy *this* system of equations, and we have no means at our disposal, other than the suitable choice of the orthogonal system (4), which intervenes in the formation of the matrices. I now assert the following :

1. The equations (18') will in general be satisfied if we choose as the orthogonal system the *proper functions* of the natural boundary value problem of the following partial differential equation,

$$(21) \quad -[H, \psi] + E\psi = 0.$$

ψ is the unknown function of q_1, q_2, \dots, q_n ; E is the proper value parameter. Of course, as density function, $\rho(x)$ appears that function of q_1, \dots, q_n , by which equation (21) must be multiplied in order to make it self-adjoint. The quantities ν_i are found to be equal to the proper values E_i divided by h . H^{kl} becomes a diagonal matrix, with $H^{kk} = E_k$.

2. If the symmetricalising of the function H has been effected in a *suitable way*—the process of symmetricalising, in my opinion, has not hitherto been defined uniquely—then (21) is *identical with the wave equation which is the basis of my wave mechanics*.¹

Assertion 1 is almost directly evident, if we provisionally lay aside the questions whether equation (21) gives rise at all to an intelligible boundary value problem with the domain of entire q -space, and whether it can always be made self-adjoint through multiplication by a suitable function, etc. These questions are largely settled under heading 2. For now we have, according to (21) and the definitions of proper values and functions,

$$(22) \quad [H, u_i] = E_i u_i,$$

and thus from (6) we get

$$(23) \quad \begin{cases} H^{kl} = \int \rho(x) u_k(x) [H, u_l(x)] dx = E_i \int \rho(x) u_k(x) u_l(x) dx \\ = 0 \text{ for } l \neq k \\ = E_l \text{ for } l = k, \end{cases}$$

and, for example,

$$(24) \quad \begin{cases} (Hq_l)^{ik} = \sum_m H^{im} q_l^{mk} = E_i q_l^{ik} \\ (q_l H)^{ik} = \sum_m q_l^{im} H^{mk} = E_k q_l^{ik}, \end{cases}$$

so that the right-hand side of the first equation of (18') takes the value

$$(25) \quad \frac{E_i - E_k}{h} q_l^{ik}.$$

Similarly for the second equation. Thus everything asserted under 1 is proved.

¹ Equation (18''), Part II.

Let us turn now to assertion 2, which is, that there is agreement between the negatively taken operator of the Hamilton function (suitably symmetricalised) and the wave operator of wave mechanics. I will first illustrate by a simple example why the process of symmetricalisation seems to me to be, *in the first instance*, not unique. Let, for *one degree of freedom*, the *ordinary* Hamilton function be

$$(26) \quad H = \frac{1}{2}(p^2 + q^2).$$

Then it is admitted that we can take this function, just as it stands, unchanged, over to "quantum mechanics" as a "well-arranged" function. But we can also, and seemingly indeed with as much right *to begin with*, apply the well-arranged function

$$(27) \quad H = \frac{1}{2} \left(\frac{1}{f(q)} p f(q) p + q^2 \right),$$

where $f(q)$ is a function arbitrary within wide limits. $f(q)$ would appear in this case as a "density function" $\rho(x)$. (26) is quite evidently just a special case of (27), and the question arises, whether (and how) it is at all possible to distinguish the special case we are concerned with, *i.e.* for more complicated H -functions. Confining ourselves to power products only of the q_k 's (where we could then simply prohibit the "production of denominators") would be most inconvenient just in the most important applications. Besides, I believe that does *not* lead to correct symmetricalisation.

For the convenience of the reader, I will now give again a short derivation of the wave equation in a form suited to the present purpose, confining myself to the case of classical mechanics (without relativity and magnetic fields). Let, therefore,

$$(28) \quad H = T(q_k, p_k) + V(q_k),$$

T being a quadratic form in the p_k 's. Then the wave equation can be deduced¹ from the following variation problem,

$$(29) \quad \left\{ \begin{array}{l} \delta J_1 = \delta \int \left\{ \frac{\hbar^2}{4\pi^2} T \left(q_k, \frac{\partial \psi}{\partial q_k} \right) + \psi^2 V(q_k) \right\} \Delta_p^{-\frac{1}{2}} dx = 0, \\ \text{with the subsidiary condition} \\ J_2 = \int \psi^2 \Delta_p^{-\frac{1}{2}} dx = 1. \end{array} \right.$$

As above, $\int dx$ stands for $\int \dots \int dq_1 \dots dq_n$; $\Delta_p^{-\frac{1}{2}}$ is the reciprocal of the square root of the discriminant of the quadratic form T . This factor must not be omitted, because otherwise the whole process would not be invariant for point transformations of the q 's! By all means another *explicit* function of the q 's might appear as a factor, *i.e.* a function which would be invariant for a point transformation of the q 's. (For Δ_p , as is known, this is not the case. Otherwise we could omit $\Delta_p^{-\frac{1}{2}}$, if this extra function was given the value $\Delta_p^{\frac{1}{2}}$.)

If we indicate the derivative of T with respect to *that* argument,

¹ Equations (23) and (24) of Part I.

which originally was p_k , by the suffix p_k , we obtain, as the *result of the variation*,

$$(30) \quad \left\{ \begin{aligned} 0 &= \frac{1}{2}(\delta J_1 - E\delta J_2) \\ &= \int \left\{ -\frac{\hbar^2}{8\pi^2} \sum_k \frac{\partial}{\partial q_k} \left[\Delta_p^{-\frac{1}{2}} T_{p_k} \left(q_k, \frac{\partial \psi}{\partial q_k} \right) \right] \right. \\ &\quad \left. + (V(q_k) - E) \Delta_p^{-\frac{1}{2}} \psi \right\} \delta \psi dx ; \end{aligned} \right.$$

the Eulerian variation equation thus runs :

$$(31) \quad \frac{\hbar^2}{8\pi^2} \Delta_p^{\frac{1}{2}} \sum_k \frac{\partial}{\partial q_k} \left\{ \Delta_p^{-\frac{1}{2}} T_{p_k} \left(q_k, \frac{\partial \psi}{\partial q_k} \right) \right\} - V(q_k) \psi + E \psi = 0.$$

It is not difficult to see that this equation has the form of (21) if we remember our law connecting the operators, and consider

$$(32) \quad T(q_k, p_k) = \frac{1}{2} \sum_k p_k T_{p_k}(q_k, p_k)$$

the Eulerian equation for homogeneous functions, applied to the quadratic form T . In actual fact, if we detach the operator from the left side of (31), with the proper value term $E\psi$ removed, and replace in it $\frac{\hbar}{2\pi\sqrt{-1}} \frac{\partial}{\partial q_k}$ by p_k , then according to (32) we obtain the negatively taken Hamilton function (28). Thus the process of variation has given quite automatically a uniquely defined "symmetricalisation" of the operator, which makes it self-adjoint (except possibly for a common factor) and makes it invariant for point transformations, and which I would like to maintain, as long as there are no definite reasons for the appearance under the integrals (29) of the additional factor, already¹ mentioned as possible, and for a definite form of the latter.

Hence the solution of the whole system of matrix equations of Heisenberg, Born, and Jordan is reduced to the natural boundary value problem of a linear partial differential equation. If we have solved the boundary value problem, then by the use of (6) we can calculate by differentiations and quadratures every matrix element we are interested in.

As an illustration of what is to be understood by the *natural* boundary value problem, *i.e.* by the natural boundary conditions at the natural boundary of configuration space, we may refer to the worked examples.² It invariably turns out that the natural infinitely distant boundary forms a singularity of the differential equation and only allows of the one boundary condition—"remaining finite". This seems to be a general characteristic of those micro-mechanical problems with which the theory in the first place is meant to deal. If the domain of the position co-ordinates is artificially limited (example: a molecule in a "vessel"), then an essential allowance must be made for this limitation by the introduction of suitable potential energies in

¹ Cf. also *Ann. d. Phys.* 79, p. 362 and p. 510 (*i.e.* Parts I. and II.).

² In Parts I. and II. of this collection.

the well-known manner. Also the *vanishing* of the proper functions at the boundary generally occurs to an adequate degree, even if relations among *certain* of the integrals (6) are present, which necessitate a special investigation, and into which I will not enter at present. (It has to do with those matrix elements in the Kepler problem which, according to Heisenberg, correspond to the transition from one hyperbolic orbit to another.)

I have confined myself here to the case of classical mechanics, without magnetic fields, because the relativistic magnetic generalisation does not seem to me to be sufficiently clear yet. But we can scarcely doubt that the complete parallel between the two new quantum theories will still stand when this generalisation is obtained.

We conclude with a general observation on the whole formal apparatus of §§ 2, 3, and 4. The basic orthogonal system was regarded as an absolutely *discrete* system of functions. Now, in the most important applications this is *not* the case. Not only in the hydrogen atom but also in heavier atoms the wave equation (31) must possess a continuous proper value spectrum as well as a line spectrum. The former manifests itself, for example, in the continuous *optical* spectra which adjoin the limit of the series. It appeared better, provisionally, not to burden the formulae and the line of thought with this generalisation, though it is indeed indispensable. The chief aim of this paper is to work out, in the clearest manner possible, the formal connection between the two theories, and this is certainly not changed, in any essential point, by the appearance of a continuous spectrum. An important precaution that we have always observed is not to postulate, without further investigation, the convergence of the development in a series of proper functions. This precaution is especially demanded by the *accumulation of the proper values at a finite point* (viz. the limit of the series). This accumulation is most intimately connected with the appearance of the continuous spectrum.

§ 5. Comparison of the Two Theories. Prospect of a Classical Understanding of the Intensity and Polarisation of the Emitted Radiation

If the two theories—I might reasonably have used the singular—should¹ be tenable in the form just given, *i.e.* for more complicated systems as well, then every discussion of the superiority of the one over the other has only an illusory object, in a certain sense. For they are completely equivalent from the mathematical point of view, and it can only be a question of the subordinate point of convenience of calculation.

¹ There is a special reason for leaving this question open. The two theories initially take the energy function over from ordinary mechanics. Now in the cases treated the *potential* energy arises from the interaction of particles, of which perhaps *one*, at least, may be regarded in wave mechanics also as forming a point, on account of its great mass (*cf.* A. Einstein, *Berl. Ber.*, 1925, p. 10). We must take into account the possibility that it is no longer permissible to take over from ordinary mechanics the statement for the potential energy, if *both* "point charges" are really extended states of vibration, which penetrate each other.

To-day there are not a few physicists who, like Kirchhoff and Mach, regard the task of physical theory as being merely a mathematical description (*as economical as possible*) of the empirical connections between observable quantities, *i.e.* a description which reproduces the connection, as far as possible, without the intervention of unobservable elements. On this view, mathematical equivalence has almost the same meaning as physical equivalence. In the present case there might perhaps appear to be a certain superiority in the matrix representation because, through its stifling of intuition, it does not tempt us to form space-time pictures of atomic processes, which must perhaps remain uncontrollable. In this connection, however, the following *supplement* to the proof of equivalence given above is interesting. The equivalence *actually* exists, and it also exists *conversely*. Not only can the matrices be constructed from the proper functions as shown above, but also, conversely, the functions can be constructed from the numerically given matrices. Thus the functions do not form, as it were, an *arbitrary* and *special* "fleshy clothing" for the bare matrix skeleton, provided to pander to the need for intuitiveness. This really would establish the superiority of the matrices, from the epistemological point of view. We suppose that in the equations

$$(33) \quad q_l^{ik} = \int u_i(x) u_k(x) dx$$

the *left-hand* sides are given numerically and the functions $u_i(x)$ are to be found. (*N.B.*—The "density function" is omitted for simplicity; the $u_i(x)$'s *themselves* are to be orthogonal functions for the present.) We may then calculate by matrix multiplication (without, by the way, any "revolving", *i.e.* integration by parts) the following integrals,

$$(34) \quad \int P(x) u_i(x) u_k(x) dx,$$

where $P(x)$ signifies *any* power product of the q_i 's. The totality of these integrals, when i and k are fixed, forms what is called the totality of the "moments" of the function $u_i(x) u_k(x)$. And it is known that, under very general assumptions, a function is determined uniquely by the totality of its moments. So all the products $u_i(x) u_k(x)$ are uniquely fixed, and thus also the squares $u_i(x)^2$, and therefore also $u_i(x)$ itself. The only arbitrariness lies in the supplementary detachment of the density function $\rho(x)$, *e.g.* $r^2 \sin \theta$ in polar co-ordinates. No false step is to be feared there, certainly not so far as *epistemology* is concerned.

Moreover, the validity of the thesis that mathematical and physical equivalence mean the same thing, must itself be qualified. Let us think, for example, of the two expressions for the electrostatic energy of a system of charged conductors, the space integral $\frac{1}{2} \int E^2 d\tau$ and the sum $\frac{1}{2} \sum e_i V_i$ taken over the conductors. The two expressions

are completely equivalent in electrostatics; the one may be derived from the other by integration by parts. Nevertheless we intentionally prefer the first and say that *it* correctly localises the energy in space. In the domain of electrostatics this preference has admittedly no justification. On the contrary, it is due simply to the fact that the first expression remains useful in electrodynamics also, while the second does not.

We cannot yet say with certainty to which of the two new quantum theories preference should be given, from *this* point of view. As the natural advocate of one of them, I will not be blamed if I frankly—and perhaps not wholly impartially—bring forward the arguments in its favour.

Leaving aside the special optical questions, the problems which the course of development of atomic dynamics brings up for consideration are presented to us by experimental physics in an eminently intuitive form; as, for example, how two colliding atoms or molecules rebound from one another, or how an electron or α -particle is diverted, when it is shot through an atom with a given velocity and with the initial path at a given perpendicular distance from the nucleus. In order to treat such problems more particularly, it is necessary to survey clearly the transition between macroscopic, perceptual mechanics and the micro-mechanics of the atom. I have lately¹ explained how I picture this transition. Micro-mechanics appears as a refinement of macro-mechanics, which is necessitated by the geometrical and mechanical smallness of the objects, and the transition is of the same nature as that from geometrical to physical optics. The latter is demanded as soon as the wave length is no longer very great compared with the dimensions of the objects investigated or with the dimensions of the space inside which we wish to obtain more accurate information about the light distribution. To me it seems extraordinarily difficult to tackle problems of the above kind, as long as we feel obliged on epistemological grounds to repress intuition in atomic dynamics, and to operate only with such abstract ideas as transition probabilities, energy levels, etc.

An especially important question—perhaps the cardinal question of all atomic dynamics—is, as we know, that of the *coupling* between the dynamic process in the atom and the electromagnetic field, or whatever has to appear in the place of the latter. Not only is there connected with this the whole complex of questions of dispersion, of resonance and secondary-radiation, and of the natural breadth of lines, but, in addition, the specification of certain quantities in atomic dynamics, such as emission frequencies, line intensities, etc., has only a mere dogmatic meaning until this coupling is described mathematically in some form or other. Here, now, the matrix representation of atomic dynamics has led to the conjecture that in fact the electromagnetic field also *must* be represented otherwise, namely, by matrices, so that the coupling may be mathematically formulated. Wave mechanics

shows we are not compelled to do this in any case, for the mechanical field scalar (which I denote by ψ) is perfectly capable of entering into the unchanged Maxwell-Lorentz equations between the electromagnetic field vectors, as the "source" of the latter; just as, conversely, the electrodynamic potentials enter into the coefficients of the wave equation, which defines the field scalar.¹ In any case, it is worth while *attempting* the representation of the coupling in such a way that we bring into the unchanged Maxwell-Lorentz equations as *four-current* a four-dimensional vector, which has been suitably derived from the mechanical field scalar of the electronic motion (perhaps through the medium of the field vectors themselves, or the potentials). There even exists a hope that we can represent the wave equation for ψ equally well as a consequence of the Maxwell-Lorentz equations, namely, as an equation of continuity for electricity. The difficulty in regard to the problem of *several* electrons, which mainly lies in the fact that ψ is a function in *configuration* space, not in real space, must be mentioned. Nevertheless I would like to discuss the one-electron problem a little further, showing that it may be possible to give an extraordinarily clear interpretation of intensity and polarisation of radiation in this manner.

Let us consider the picture, on the wave theory, of the hydrogen atom, when it is in such a state that the field scalar ψ is given by a series of discrete proper functions, thus :

$$(35) \quad \psi = \sum_k c_k u_k(x) e^{\frac{2\pi\sqrt{-1}}{h} E_k t}$$

(x stands here for *three* variables, e.g. r, θ, ϕ ; the c_k 's are taken as real and it is correct to take the real part). We now make the *assumption* that the space density of electricity is given by the real part of

$$(36) \quad \psi \frac{\partial \bar{\psi}}{\partial t}.$$

The bar is to denote the conjugate complex function. We then calculate for the space density,

$$(37) \quad \text{space density} = 2\pi \sum_{(k, m)} c_k c_m \frac{E_k - E_m}{h} u_k(x) u_m(x) \sin \frac{2\pi t}{h} (E_m - E_k),$$

where the sum is to be taken once only over every combination (k, m). Only term *differences* enter (37) as frequencies. The former are so low that the length of the corresponding ether wave is large compared

¹ Similar ideas are expressed by K. Lanczos in an interesting note that has just appeared (*Ztschr. f. Phys.* 35, p. 812, 1926). This note is also valuable as showing that Heisenberg's atomic dynamics is capable of a continuous interpretation as well. However, Lanczos' work has fewer points of contact with the present work than at first it was thought to have. The determination of his formal system, which was provisionally left quite indefinite, is *not* to be sought by following the idea that in some way the symmetrical nucleus $K(s, \sigma)$ of Lanczos can be identified with the *Green's function* of our wave equation (21) or (31). For this Green's function, if it exists, has the quantum levels themselves as proper values. On the other hand, it is required that Lanczos' function should have the *reciprocals* of the quantum levels as proper values.

with atomic dimensions, that is, compared with the region within which (37) is markedly different from zero.¹ The radiation can therefore be estimated simply by the *dipole moment* which according to (37) the whole atom possesses. We multiply (37) by a Cartesian co-ordinate q_i , and by the "density function" $\rho(x)$, ($r^2 \sin \theta$ in the present case) and integrate over the whole space. According to (13), we get for the component of the dipole moment in the direction q_i ,

$$(38) \quad M_{q_i} = 2\pi \sum_{(k, m)} c_k c_m q_i^{km} \frac{E_k - E_m}{h} \sin \frac{2\pi t}{h} (E_m - E_k).$$

Thus we really get a "Fourier development" of the atom's electric moment, in which only term *differences* appear as frequencies. The Heisenberg matrix elements q_i^{km} come into the coefficients in such a manner that their co-operating influence on the intensity and polarisation of the part of the radiation concerned is completely intelligible on the grounds of classical electrodynamics.

The present sketch of the mechanism of radiation is far from completely satisfactory and is in no way final. Assumption (36) makes use, somewhat freely, of complex calculation, in order to put to one side undesired components of vibration whose radiation cannot be investigated at all in the simple way used for the dipole moment of the entire atom, because the corresponding ether wave lengths (about 0.01 Å) lie far below atomic dimensions. Moreover, if we integrate over all space, then by (5) the space density (37) gives zero and not, as is required, a finite value, independent of the time, which requires to be normalised to the electronic charge. In conclusion, for completeness, account should be taken of magnetic radiation, since if there is a spatial distribution of electric currents, radiation is possible without the appearance of an electric moment, *e.g.* with a frame aerial.

Nevertheless it appears to be a well-founded hope that a real understanding of the nature of emitted radiation will be obtained on the basis of one of the two very similar analytical mechanisms which have been sketched here.

(Received March 18, 1926).

¹ *Ann. d. Phys.* 79, p. 371, 1926, *i.e.* beginning of § 2, Part I. here.

Quantisation as a Problem of Proper Values (Part III)

PERTURBATION THEORY, WITH APPLICATION TO THE STARK EFFECT
OF THE BALMER LINES

(*Annalen der Physik* (4), vol. 80, 1926)

Introduction. Abstract

As has already been mentioned at the end of the preceding paper,¹ the available range of application of the proper value theory can by comparatively elementary methods be considerably increased beyond the "directly soluble problems"; for proper values and functions can readily be approximately determined for *such* boundary value problems as are sufficiently closely related to a directly soluble problem. In analogy with ordinary mechanics, let us call the method in question the *perturbation* method. It is based upon the important *property of continuity* possessed by proper values and functions,² principally, for our purpose, upon their *continuous* dependence on the *coefficients* of the differential equation, and less upon the extent of the domain and on the boundary conditions, since in our case the domain ("entire q -space") and the boundary conditions ("remaining finite") are generally the same for the unperturbed and perturbed problems.

The method is essentially the same as that used by Lord Rayleigh in investigating³ the vibrations of a string with *small inhomogeneities* in his *Theory of Sound* (2nd edit., vol. i., pp. 115-118, London, 1894). This was a particularly simple case, as the differential equation of the unperturbed problem had *constant* coefficients, and only the perturbing terms were arbitrary functions along the string. A complete generalisation is possible not merely with regard to these points, but also for the specially important case of *several* independent variables, *i.e.* for *partial* differential equations, in which *multiple proper values* appear in the unperturbed problem, and where the addition of a

¹ Last two paragraphs of Part II.

² Courant-Hilbert, chap. vi. §§ 2, 4, p. 337.

³ Courant-Hilbert, chap. v. § 5, 2, p. 241.

perturbing term causes the *splitting up* of such values and is of the greatest interest in well-known spectroscopic questions (Zeeman effect, Stark effect, Multiplicities). In the development of the perturbation theory in the following Section I., which really yields nothing new to the mathematician, I put less value on generalising to the *widest possible extent* than on bringing forward the very simple rudiments in the clearest possible manner. From the latter, any desired generalisation arises almost automatically when needed. In Section II., as an example, the Stark effect is discussed and, indeed, by *two* methods, of which the *first* is analogous to Epstein's method, by which he first solved¹ the problem on the basis of classical mechanics, supplemented by quantum conditions, while the *second*, which is much more general, is analogous to the method of secular perturbations.² The *first* method will be utilised to show that in wave mechanics also the perturbed problem can be "separated" in *parabolic* co-ordinates, and the perturbation theory will first be applied to the ordinary differential equations into which the original vibration equation is split up. The theory thus merely takes over the task which on the old theory devolved on Sommerfeld's elegant complex integration for the calculation of the quantum integrals.³ In the *second* method, it is found that in the case of the Stark effect an exact separation co-ordinate system exists, quite by accident, for the perturbed problem also, and the perturbation theory is applied directly to the *partial* differential equation. This latter proceeding proves to be more troublesome in wave mechanics, although it is theoretically superior, being more capable of generalisation.

Also the problem of the intensity of the components in the Stark effect will be shortly discussed in Section II. Tables will be calculated, which, as a whole, agree even better with experiment than the well-known ones calculated by Kramers with the help of the correspondence principle.⁴

The application (not yet completed) to the *Zeeman effect* will naturally be of much greater interest. It seems to be indissolubly linked with a correct formulation in the language of wave mechanics of the *relativistic* problem, because in the four-dimensional formulation the vector-potential automatically ranks equally with the scalar. It was already mentioned in Part I. that the relativistic hydrogen atom may indeed be treated without further discussion, but that it leads to "half-integral" azimuthal quanta, and thus contradicts experience. Therefore "something must still be missing". Since then I have learnt *what* is lacking from the most important publications of G. E. Uhlenbeck and S. Goudsmit,⁵ and then from oral and written communications from Paris (P. Langevin) and Copenhagen (W. Pauli),

¹ P. S. Epstein, *Ann. d. Phys.* 50, p. 489, 1916.

² N. Bohr, *Kopenhagener Akademie* (8), IV., 1, 2, p. 69 *et seq.*, 1918.

³ A. Sommerfeld, *Atombau*, 4th ed., p. 772.

⁴ H. A. Kramers, *Kopenhagener Akademie* (8), III., 3, p. 287, 1919.

⁵ G. E. Uhlenbeck and S. Goudsmit, *Physica*, 1925; *Die Naturwissenschaften*, 1926; *Nature*, 20th Feb., 1926; cf. also L. H. Thomas, *Nature*, 10th April, 1926.

viz., in the language of the theory of electronic orbits, the *angular momentum* of the electron round its axis, which gives it a *magnetic moment*. The utterances of these investigators, together with two highly significant papers by Slater¹ and by Sommerfeld and Unsöld² dealing with the Balmer spectrum, leave no doubt that, by the introduction of the paradoxical yet happy conception of the spinning electron, the orbital theory will be able to master the disquieting difficulties which have latterly begun to accumulate (anomalous Zeeman effect; Paschen-Back effect of the Balmer lines; irregular and regular Röntgen doublets; analogy of the latter with the alkali doublets, etc.). We shall be obliged to attempt to take over the idea of Uhlenbeck and Goudsmit into wave mechanics. I believe that the latter is a very fertile soil for this idea, since in it the electron is not considered as a point charge, but as continuously flowing through space,³ and so the unpleasing conception of a "rotating point-charge" is avoided. In the present paper, however, the taking over of the idea is not yet attempted.

To the *third section*, as "mathematical appendix", have been relegated numerous uninteresting calculations—mainly quadratures of products of proper functions, required in the second section. *The formulae of the appendix are numbered (101), (102), etc.*

I. PERTURBATION THEORY

§ 1. A Single Independent Variable

Let us consider a linear, homogeneous, differential expression of the second order, which we may assume to be in self-adjoint form without loss of generality, viz.

$$(1) \quad L[y] = py'' + p'y' - qy.$$

y is the dependent function; p , p' and q are continuous functions of the independent variable x and $p \geq 0$. A dash denotes differentiation with respect to x (p' is therefore the derivative of p , which is the condition for self-adjointness).

Now let $\rho(x)$ be another continuous function of x , which never becomes negative, and also in general does not vanish. We consider the proper value problem of Sturm and Liouville,⁴

$$(2) \quad L[y] + E\rho y = 0.$$

It is a question, first, of finding all *those* values of the constant E ("proper values") for which the equation (2) possesses solutions $y(x)$, which are continuous and not identically vanishing within a certain domain, and which satisfy certain "boundary conditions" at the bounding points; and secondly of finding these solutions ("proper

¹ J. C. Slater, *Proc. Amer. Nat. Acad.* 11, p. 732, 1925.

² A. Sommerfeld and A. Unsöld, *Ztschr. f. Phys.* 36, p. 259, 1926.

³ Cf. last two pages of previous paper.

⁴ Cf. Courant-Hilbert, chap. v. § 5, 1, p. 238 *et seq.*

functions”) themselves. In the cases treated in atomic mechanics, domain and boundary conditions are always “natural”. The domain, for example, reaches from 0 to ∞ , when x signifies the value of the radius vector or of an intrinsically positive parabolic co-ordinate, and the boundary conditions are in these cases : *remaining finite*. Or, when x signifies an azimuth, then the domain is the interval from 0 to 2π and the condition is : Repetition of the initial values of y and y' at the end of the interval (“periodicity”).

It is only in the case of the periodic condition that *multiple*, viz. *double-valued*, proper values appear for *one* independent variable. By this we understand that to the same proper value belong *several* (in the particular case, two) linearly independent proper functions. We will now exclude this case for the sake of simplicity, as it attaches itself easily to the developments of the following paragraph. Moreover, to lighten the formulae, we will not expressly take into account in the notation the possibility that a “band spectrum” (*i.e.* a *continuum* of proper values) may be present when the domain extends to infinity.

Let now $y = u_i(x)$, $i = 1, 2, 3, \dots$, be the series of Sturm-Liouville proper functions ; then the series of functions $u_i(x)\sqrt{\rho(x)}$, $i = 1, 2, 3, \dots$, forms a *complete orthogonal system* for the domain ; *i.e.* in the first place, if $u_i(x)$ and $u_k(x)$ are the proper functions belonging to the values E_i and E_k , then

$$(3) \quad \int \rho(x)u_i(x)u_k(x)dx = 0 \text{ for } i \neq k.$$

(Integrals without limits are to be taken over the domain, throughout this paper.) The expression “complete” signifies that an originally arbitrary continuous function is condemned to vanish identically, by the mere postulation that it must be orthogonal with respect to *all* the functions $u_i(x)\sqrt{\rho(x)}$. (More shortly : “There exists no further orthogonal function for the system.”) We can and will always regard the proper functions $u_i(x)$ in all general discussions as “normalised”, *i.e.* we imagine the constant factor, which is still arbitrary in each of them on account of the homogeneity of (2), to be defined *in such a way* that the integral (3) takes the value unity for $i = k$. Finally we again remind the reader that the proper values of (2) are certainly all *real*.

Let now the proper values E_i and functions $u_i(x)$ be *known*. Let us, from now on, direct our attention specially to a *definite* proper value, E_k say, and the corresponding function $u_k(x)$, and ask how these alter, when we do not alter the problem in any way other than by adding to the left-hand side of (2) a small “perturbing term”, which we will initially write in the form

$$(4) \quad -\lambda r(x)y.$$

In this λ is a small quantity (the perturbation parameter), and $r(x)$ is an arbitrary continuous function of x . It is therefore simply a matter of a slight alteration of the coefficient q in the differential expression (1). From the continuity properties of the proper quantities,

mentioned in the introduction, we now know that the altered Sturm-Liouville problem

$$(2') \quad L[y] - \lambda ry + E\rho y = 0$$

must have, in any case for a sufficiently small λ , proper quantities in the near neighbourhood of E_k and u_k , which we may write, by way of trial, as

$$(5) \quad E_k^* = E_k + \lambda\epsilon_k; \quad u_k^* = u_k(x) + \lambda v_k(x).$$

On substituting in equation (2'), remembering that u_k satisfies (2), neglecting λ^2 and cutting away a factor λ we get

$$(6) \quad L[v_k] + E_k\rho v_k = (r - \epsilon_k\rho)u_k.$$

For the defining of the perturbation v_k of the proper *function*, we thus obtain, as a comparison of (2) and (6) shows, a *non-homogeneous* equation, which belongs precisely to *that* homogeneous equation which is satisfied by our unperturbed proper function u_k (for in (6) the special proper value E_k stands in place of E). On the right-hand side of this non-homogeneous equation occurs, in addition to known quantities, the still unknown perturbation ϵ_k of the proper *value*.

This occurrence of ϵ_k serves for the calculation of this quantity *before* the calculation of v_k . It is known that the non-homogeneous equation—and this is *the starting-point of the whole perturbation theory*—for a proper *value* of the homogeneous equation possesses a solution *when, and only when*, its right-hand side is *orthogonal*¹ to the allied proper function (to all the allied functions, in the case of multiple proper values). (The physical interpretation of this mathematical theorem, for the vibrations of a string, is that if the force is in resonance with a proper vibration it must be distributed in a very special way over the string, namely, so that it does no work in the vibration in question; otherwise the amplitude grows beyond all limits and a stationary condition is impossible.)

The right-hand side of (6) must therefore be orthogonal to u_k , *i.e.*

$$(7) \quad \int (r - \epsilon_k\rho)u_k^2 dx = 0,$$

or

$$(7') \quad \epsilon_k = \frac{\int ru_k^2 dx}{\int \rho u_k^2 dx},$$

or, if we imagine u_i already normalised, then, more simply,

$$(7'') \quad \epsilon_k = \int ru_k^2 dx.$$

This simple formula expresses the perturbation of the proper value (of first order) in terms of the perturbing function $r(x)$ and the unperturbed proper function $u_k(x)$. If we consider that the proper

¹ Cf. Courant-Hilbert, chap. v. § 10, 2, p. 277.

value of our problem signifies mechanical energy or is analogous to it, and that the proper function u_k is comparable to "motion with energy E_k ", then we see in (7") the complete parallel to the well-known theorem in the perturbation theory of classical mechanics, viz. the perturbation of the energy, to a first approximation, is equal to the perturbing function, averaged over the unperturbed motion. (It may be remarked in passing that it is as a rule sensible, or at least aesthetic, to throw into bold relief the factor $\rho(x)$ in the integrands of *all* integrals taken over the entire domain. If we do this, then, in integral (7"), we must speak of $\frac{r(x)}{\rho(x)}$ and not $r(x)$ as the perturbing function, and make a corresponding change in the expression (4). Since the point is quite unimportant, however, we will stick to the notation already chosen.)

We have yet to define $v_k(x)$, the perturbation of the proper *function*, from (6). We solve¹ the non-homogeneous equation by putting for v_k a series of proper functions, viz.

$$(8) \quad v_k(x) = \sum_{i=1}^{\infty} \gamma_{ki} u_i(x),$$

and by developing the right-hand side, divided by $\rho(x)$, likewise in a series of proper functions, thus

$$(9) \quad \left(\frac{r(x)}{\rho(x)} - \epsilon_k \right) u_k(x) = \sum_{i=1}^{\infty} c_{ki} u_i(x),$$

where

$$(10) \quad \begin{cases} c_{ki} = \int (r - \epsilon_k \rho) u_k u_i dx \\ \quad = \int r u_k u_i dx \text{ for } i \neq k \\ \quad = 0 \quad \quad \quad \text{for } i = k. \end{cases}$$

The last equality follows from (7). If we substitute from (8) and (9) in (6) we get

$$(11) \quad \sum_{i=1}^{\infty} \gamma_{ki} (L[u_i] + E_k \rho u_i) = \sum_{i=1}^{\infty} c_{ki} \rho u_i.$$

Since now u_i satisfies equation (2) with $E = E_i$, it follows that

$$(12) \quad \sum_{i=1}^{\infty} \gamma_{ki} \rho (E_k - E_i) u_i = \sum_{i=1}^{\infty} c_{ki} \rho u_i.$$

By equating coefficients on left and right, all the γ_{ki} 's, except γ_{kk} , are defined. Thus

$$(13) \quad \gamma_{ki} = \frac{c_{ki}}{E_k - E_i} = \frac{\int r u_k u_i dx}{E_k - E_i} \text{ for } i \neq k,$$

while γ_{kk} , as may be understood, remains completely undefined. This indefiniteness corresponds to the fact that the postulation of

¹ Cf. Courant-Hilbert, chap. v. § 5, 1, p. 240, and § 10, p. 279.

normalisation is still available for us for the perturbed proper function. If we make use of (8) in (5) and claim for $u_k^*(x)$ the same normalisation as for $u_k(x)$ (quantities of the order of λ^2 being neglected), then it is evident that $\gamma_{kk}=0$. Using (13) we now obtain for the *perturbed proper function*

$$(14) \quad u_k^*(x) = u_k(x) + \lambda \sum'_{i=1}^{\infty} \frac{u_i(x) \int r u_i u_i dx}{E_k - E_i}.$$

(The dash on the sigma denotes that the term $i=k$ has not to be taken.) And the allied perturbed proper value is, from the above,

$$(15) \quad E_k^* = E_k + \lambda \int r u_k^2 dx.$$

By substituting in (2') we may convince ourselves that (14) and (15) do really satisfy the proper value problem to the proposed degree of approximation. This verification is necessary since the development, assumed in (5), in *integral* powers of the perturbation parameter is no necessary consequence of continuity.

The procedure, here explained in fair detail for the simplest case, is capable of generalisation in many ways. In the first place, we can of course consider the perturbation in a quite similar manner for the second, and then the third order in λ , etc., in each case obtaining first the next approximation to the proper value, and then the corresponding approximation for the proper function. In certain circumstances it may be advisable—just as in the perturbation theory of mechanics—to regard the perturbation function itself as a power series in λ , whose terms come into play one by one in the separate stages. These questions are discussed exhaustively by Herr E. Fues in work which is now appearing in connection with the application to the theory of *band spectra*.

In the second place, in quite similar fashion, we can consider also a perturbation of the term in y' of the differential operator (1) just as we have considered above the term $-qy$. The case is important, for the Zeeman effect leads without doubt to a perturbation of this kind—though admittedly in an equation with several independent variables. Thus the equation loses its self-adjoint form by the perturbation—not an essential matter in the case of a single variable. In a partial differential equation, however, this loss may result in the perturbed proper values no longer being real, though the perturbing term is real; and naturally also conversely, an imaginary perturbing term may have a real, physically intelligible perturbation as its consequence.

We may also go further and consider a perturbation of the term in y'' . Indeed it is quite possible, in general, to add an arbitrary "infinitely small" linear¹ and homogeneous differential operator, even of higher order than the second, as the perturbing term and to calculate the perturbations in the same manner as above. In these cases,

¹ Even the limitation "linear" is not absolutely necessary.

however, we would use with advantage the fact that the second and higher derivatives of the proper functions may be expressed by means of the differential equation itself, in terms of the zero and first derivatives, so that this general case may be reduced, in a certain sense, to the two special cases, first considered—perturbation of the terms in y and y' .

Finally, it is obvious that the extension to equations of order higher than the second is possible.

Undoubtedly, however, the most important generalisation is that to several independent variables, *i.e.* to partial differential equations. For *this* really is the problem in the general case, and only in exceptional cases will it be possible to split up the disturbed partial differential equation, by the introduction of suitable variables, into separate differential equations, each only with one variable.

§ 2. Several Independent Variables (Partial Differential Equation)

We will represent the several independent variables in the formulae symbolically by the *one* sign x , and briefly write $\int dx$ (instead of $\int \dots \int dx_1 dx_2 \dots$) for an integral extending over the multiply-dimensioned *domain*. A notation of this type is already in use in the theory of integral equations, and has the advantage, here as there, that the structure of the formulae is not altered by the increased number of variables as such, but only by *essentially* new occurrences, which *may* be related to it.

Let therefore $L[y]$ now signify a self-adjoint *partial* linear differential expression of the second order, whose explicit form we do not require to specify; and further let $\rho(x)$ again be a positive function of the independent variables, which does not vanish in general. The postulation “self-adjoint” is *now* no longer unimportant, as the property cannot now be generally gained by multiplication by a suitably chosen $f(x)$, as was the case with *one* variable. In the particular differential expression of wave mechanics, however, this is still the case, as it arises from a variation principle.

According to these definitions or conventions, we can regard equation (2) of § 1,

$$(2) \quad L[y] + E\rho y = 0,$$

as the formulation of the Sturm-Liouville proper value problem in the case of several variables also. Everything said there about the proper values and functions, their orthogonality, normalisation, etc., as also *the whole perturbation theory there developed*—in short, the whole of § 1—remains *valid without change, when* all the proper values are *simple*, if we use the abbreviated symbolism just agreed upon above. And only *one* thing does *not* remain valid, namely, that they *must* be simple.

Nevertheless, from the pure mathematical standpoint, the case when the roots are all distinct is to be regarded as the *general* case for several variables also, and multiplicity regarded as a special occurrence, which, it is admitted, *is the rule in applications*, on account of the specially simple and symmetrical structure of the differential expressions $L[y]$ (and the "boundary conditions") which appear. Multiplicity of the proper values corresponds to *degeneracy* in the theory of conditioned periodic systems and is therefore especially interesting for quantum theory.

A proper value E_k is called α -fold, when equation (2), for $E = E_k$, possesses not *one* but exactly α linearly independent solutions which satisfy the boundary conditions. We will denote these by

$$(16) \quad u_{k1}, u_{k2}, \dots, u_{k\alpha}.$$

Then it is true that each of these α proper functions is *orthogonal* to each of the *other* proper functions belonging to *another* proper value (the factor $\rho(x)$ being included; cf. (3)). On the contrary, these α functions are *not* in general orthogonal to *one another*, if we merely postulate that they are α linearly independent proper functions for the proper value E_k , and nothing more. For then we can equally well replace them by α arbitrary, linearly independent, linear aggregates (with constant coefficients) of themselves. We may express this otherwise, thus. The series of functions (16) is initially *indefinite* to the extent of a linear transformation (with constant coefficients), involving a non-vanishing determinant, and such a transformation *destroys*, in general, the mutual orthogonality.

But through such a transformation this mutual orthogonality can always be *brought about*, and indeed in an infinite number of ways; the latter property arising because *orthogonal* transformation does *not* destroy the mutual orthogonality. We are now accustomed to include this simply in *normalisation*, that orthogonality is secured for *all* proper functions, even for those which belong to the *same* proper value. We will assume that our u_{ki} 's are already *normalised* in this way, and of course for *each* proper value. Then we must have

$$(17) \quad \begin{cases} \int \rho(x) u_{ki}(x) u_{k'i'}(x) dx = 0 & \text{when } (k, i) \neq (k', i') \\ = 1 & \text{when } k' = k, \text{ as well as } i' = i. \end{cases}$$

Each of the finite series of proper functions u_{ki} , obtained for *constant* k and *varying* i , is then only still indefinite to this extent, that it is subject to an *orthogonal* transformation.

We will now discuss, first in words, without using formulae, the consequences which follow when a perturbing term is added to the differential equation (2). The addition of the perturbing term will, in general, remove the above-mentioned symmetry of the differential equation, to which the multiplicity of the proper values (or of certain of them) is due. Since, however, the proper values and functions are *continuously* dependent on the coefficients of the differential equation, a small perturbation causes a group of α proper values, which lie close

to one another and to E_k , to enter in place of the a -fold proper value E_k . The latter is *split up*. Of course, if the symmetry is not wholly destroyed by the perturbation, it may happen that the splitting up is not complete and that several proper values (still partly multiple) of, *in summa*, equal multiplicity merely appear in the place of E_k (“*partial removal of degeneracy*”).

As for the perturbed proper *functions*, those a members which belong to the a values arising from E_k must evidently also on account of continuity lie infinitely near the unperturbed functions belonging to E_k , viz. u_{ki} ; $i=1, 2, 3 \dots a$. Yet we must remember that the last-named series of functions, as we have established above, is indefinite to the extent of an *arbitrary orthogonal transformation*. One of the infinitely numerous definitions, which may be applied to the series of functions, u_{ki} ; $i=1, 2, 3 \dots a$, will lie infinitely near the series of perturbed functions; and if the value E_k is completely split up, it will be a *quite definite one*! For to the separate simple proper values, into which the value is split up, there belong proper functions which are quite uniquely defined.

This unique particular specification of the *unperturbed* proper functions (which may fittingly be designated as the “*approximations of zero order*” for the *perturbed* functions), which is defined by the nature of the perturbation, will naturally *not* generally coincide with that definition of the unperturbed functions which we chanced to adopt to begin with. Each group of the latter, belonging to a definite a -fold proper value E_k , will have first to be submitted to an orthogonal substitution, defined by the kind of perturbation, before it can serve as the starting-point, the “*zero approximation*”, for a more exact definition of the perturbed proper functions. *The defining of these orthogonal substitutions*—one for each multiple proper value—is *the only essentially new point* that arises because of the increased number of variables, or from the appearance of multiple proper values. The defining of these substitutions forms the exact counterpart to the finding of an approximate separation system for the perturbed motion in the theory of conditioned periodic systems. As we will see immediately, the definition of the substitutions can always be given in a theoretically simple way. It requires, for each a -fold proper value, merely the principal axes transformation of a quadratic form of a (and thus of a finite number of) variables.

When the substitution has once been accomplished, the calculation of the approximations of the *first order* runs almost word for word as in § 1. The sole difference is that the dash on the sigma in equation (14) must mean that in the summation *all* the proper functions belonging to the value E_k , *i.e.* *all* the terms whose denominators would vanish, must be left out. It may be remarked in passing that it is not at all necessary, in the calculation of *first* approximations, to have completed the orthogonal substitutions referred to for *all* multiple proper values, but it is sufficient to have done so for the value E_k , in whose splitting up we are interested. For the approximations

of higher order, we admittedly require them all. In all other respects, however, these higher approximations are from the beginning carried out exactly as for simple proper values.

Of course it may happen, as was mentioned above, that the value E_k , either generally or at the initial stages of the approximation, is not completely split up, and that multiplicities ("degeneracies") still remain. This is expressed by the fact that to the substitutions already frequently mentioned there still clings a certain indefiniteness, which either always remains, or is removed step by step in the later approximations.

Let us now represent these ideas by formulae, and consider as before the perturbation caused by (4), § 1,

$$(4) \quad -\lambda r(x)y,$$

i.e. we imagine the proper value problem belonging to (2) solved, and now consider the exactly corresponding problem (2'),

$$(2') \quad L[y] - \lambda ry + E\rho y = 0.$$

We again fix our attention on a definite proper value E_k . Let (16) be a system of proper functions belonging to it, which we assume to be normalised and orthogonal to one another in the sense described above, but *not yet* fitted to the particular perturbation in the sense explained, because to find the substitution that leads to *this fitting* is precisely our chief task! In place of (5), § 1, we must now put for the perturbed quantities the following,

$$(18) \quad E^*_{kl} = E_k + \lambda \epsilon_l; \quad u^*_{kl}(x) = \sum_{i=1}^a \kappa_{li} u_{ki}(x) + \lambda v_l(x) \\ (l=1, 2, 3 \dots a),$$

wherein the $v_l(x)$'s are functions, and the ϵ_l 's and the κ_{li} 's are systems of constants, which are still to be defined, but which we initially do not limit in any way, although we know that the system of coefficients κ_{li} must¹ form an orthogonal substitution. The index k should still be attached to the three types of quantity named, in order to indicate that the whole discussion refers to the k th proper value of the unperturbed problem. We refrain from carrying this out, in order to avoid the confusing accumulation of indices. The index k is to be assumed *fixed* in the whole of the following discussion, until the contrary is stated.

Let us select *one* of the perturbed proper functions and values by giving a definite value to the index l in (18), and let us substitute from (18) in the differential equation (2') and arrange in powers of λ . Then the terms independent of λ disappear exactly as in § 1, because the unperturbed proper quantities satisfy equation (2),

¹ It follows from the general theory that the perturbed system of functions $u^*_{kl}(x)$ must be orthogonal if the perturbation completely removes the degeneracy, and may be assumed orthogonal although that is not the case.

by hypothesis. Only terms containing the *first* power of λ remain, as we can strike out the others. Omitting a factor λ , we get

$$(19) \quad L[v_l] + E_k \rho v_l = \sum_{i=1}^{\alpha} \kappa_{li} (r - \epsilon_l \rho) u_{ki},$$

and thus obtain again for the definition of the perturbation v_l of the *functions* a *non-homogeneous* equation, to which corresponds as homogeneous equation the equation (2), with the particular value $E = E_k$, i.e. the equation satisfied by the set of functions u_{ki} ; $i = 1, 2, \dots \alpha$. The form of the left side of equation (19) is independent of the index l .

On the right side occur ϵ_l and κ_{li} , the constants to be defined, and we are thus enabled to evaluate them, even *before* calculating v_l . For, in order that (19) should have a solution at all, it is necessary and sufficient that its right-hand side should be orthogonal to *all* the proper functions of the homogeneous equation (2) belonging to E_k . Therefore, we must have

$$(20) \quad \begin{cases} \sum_{i=1}^{\alpha} \kappa_{li} \int (r - \epsilon_l \rho) u_{ki} u_{km} dx = 0 \\ (m = 1, 2, 3 \dots \alpha), \end{cases}$$

i.e. on account of the normalisation (17),

$$(21) \quad \begin{cases} \kappa_{lm} \epsilon_l = \sum_{i=1}^{\alpha} \kappa_{li} \int r u_{ki} u_{km} dx \\ (m = 1, 2, 3 \dots \alpha). \end{cases}$$

If we write, briefly, for the *symmetrical* matrix of constants, which can be evaluated by quadrature,

$$(22) \quad \begin{cases} \int r u_{ki} u_{km} dx = \epsilon_{im} \\ (i, m = 1, 2, 3 \dots \alpha), \end{cases}$$

then we recognise in

$$(21') \quad \begin{cases} \kappa_{lm} \epsilon_l = \sum_{i=1}^{\alpha} \kappa_{li} \epsilon_{mi} \\ (m = 1, 2, 3 \dots \alpha) \end{cases}$$

a system of α linear homogeneous equations for the calculation of the α constants κ_{lm} ; $m = 1, 2 \dots \alpha$, where the perturbation ϵ_l of the proper value still occurs in the coefficients, and is itself unknown. However, this serves for the calculation of ϵ_l before that of the κ_{lm} 's. For it is known that the linear homogeneous system (21') of equations has solutions if, and only if, its determinant vanishes. This yields the following algebraic equation of degree α for ϵ_l :

$$(23) \quad \begin{vmatrix} \epsilon_{11} - \epsilon_l & \epsilon_{12} & \dots & \epsilon_{1\alpha} \\ \epsilon_{21} & \epsilon_{22} - \epsilon_l & \dots & \epsilon_{2\alpha} \\ \dots & \dots & \dots & \dots \\ \epsilon_{\alpha 1} & \epsilon_{\alpha 2} & \dots & \epsilon_{\alpha\alpha} - \epsilon_l \end{vmatrix} = 0.$$

We see that the problem is completely identical with the transformation of the quadratic form in a variables, with coefficients ϵ_{mi} , to its principal axes. The "secular equation" (23) yields a roots for ϵ_l , the "reciprocal of the squares of the principal axes", which in general are different, and on account of the symmetry of the ϵ_{mi} 's *always real*. We thus get all the a perturbations of the proper values ($l=1, 2 \dots a$) at the same time, and would have *inferred* the splitting up of an a -fold proper value into exactly a simple values, generally different, even had we not assumed it already, as fairly obvious. For *each* of these ϵ_l -values, equations (21') give a system of quantities $\kappa_{li}; i=1, 2, \dots a$, and, as is known, *only one* (apart from a general constant factor), provided all the ϵ_l 's are really different. Further, it is known that the whole system of a^2 quantities κ_{li} forms an *orthogonal* system of coefficients, defining as usual, in the principal axes problem, the *directions* of the new co-ordinate axes with reference to the old ones. We may, and will, employ the undefined factors just mentioned to normalise the κ_{li} 's completely as "direction cosines", and this, as is easily seen, makes the perturbed proper functions $u_{ki}^*(x)$ turn out *normalised* again, according to (18), at least in the "zero approximation" (*i.e.* apart from the λ -terms).

If the equation (23) has multiple roots, then we have the case previously mentioned, when the perturbation does not completely remove the degeneration. The perturbed equation has then multiple proper values also and the definition of the constants κ_{li} becomes partially arbitrary. This has no consequence other than that (as is *always* the case with multiple proper values) we *must* and *may* acquiesce, even *after* the perturbation is applied, in a system of proper functions which in many respects is still arbitrary.

The main task is accomplished with this transformation to principal axes, and we will often find it sufficient in the applications in quantum theory to define the proper values to a first and the functions to zero approximation. The evaluation of the constants κ_{li} and ϵ_{li} cannot be carried out always, since it depends on the solution of an algebraic equation of degree a . At the worst there are methods¹ which give the evaluation to any desired approximation by a rational process. We may thus regard these constants as known, and will now give the calculation of the functions to the *first* approximation, for the sake of completeness. The procedure is exactly as in § 1.

We have to solve equation (19) and to that end we write v_l as a series of the *whole set* of proper functions of (2),

$$(24) \quad v_l(x) = \sum_{(k' i')} \gamma_{l, k' i'} u_{k' i'}(x).$$

The summation is to extend with respect to k' from 0 to ∞ , and, for each fixed value of k' , for i' varying over the finite number of proper functions which belong to $E_{k'}$. (Now, for the first time, we take account of proper functions which do *not* belong to the a -fold value

¹ Courant-Hilbert, chap. i. § 3. 3, p. 14.

E_k we are fixing our attention on.) Secondly, we develop the right-hand side of (19), divided by $\rho(x)$, in a series of the entire set of proper functions,

$$(25) \quad \sum_{i=1}^{\alpha} \kappa_{li} \left(\frac{r}{\rho} - \epsilon_l \right) u_{ki} = \sum_{(k'i')} c_{l, k'i'} u_{k'i'},$$

wherein

$$(26) \quad \left\{ \begin{aligned} c_{l, k'i'} &= \sum_{i=1}^{\alpha} \kappa_{li} \int (r - \epsilon_l \rho) u_{ki} u_{k'i'} dx \\ &= \sum_{i=1}^{\alpha} \kappa_{li} \int r u_{ki} u_{k'i'} dx \text{ for } k' \neq k \\ &= 0 \qquad \qquad \qquad \text{for } k' = k \end{aligned} \right.$$

(the last two equalities follow from (17) and (20) respectively). On substituting from (24) and (25) in (19), we get

$$(27) \quad \sum_{(k'i')} \gamma_{l, k'i'} (L[u_{k'i'}] + E_k \rho u_{k'i'}) = \sum_{(k'i')} c_{l, k'i'} \rho u_{k'i'}.$$

Since $u_{k'i'}$ satisfies equation (2) with $E = E_k$, this gives

$$(28) \quad \sum_{(k'i')} \gamma_{l, k'i'} \rho (E_k - E_{k'}) u_{k'i'} = \sum_{(k'i')} c_{l, k'i'} \rho u_{k'i'}.$$

By equating coefficients on right and left, all the $\gamma_{l, k'i'}$'s are defined, with the exception of those in which $k' = k$. Thus

$$(29) \quad \gamma_{l, k'i'} = \frac{c_{l, k'i'}}{E_k - E_{k'}} = \frac{1}{E_k - E_{k'}} \sum_{i=1}^{\alpha} \kappa_{li} \int r u_{ki} u_{k'i'} dx \text{ (for } k' \neq k),$$

while those γ 's for which $k' = k$ are of course not fixed by equation (19). This again corresponds to the fact that we have provisionally normalised the perturbed functions u^*_{kl} , of (18), only in the zero approximation (through the normalisation of the κ_{li} 's), and it is easily recognised again that we have to put the whole of the γ -quantities in question equal to zero, in order to bring about the normalisation of the u^*_{kl} 's even in the first approximation. By substituting from (29) in (24), and then from (24) in (18), we finally obtain for the *perturbed proper functions to a first approximation*

$$(30) \quad u^*_{kl}(x) = \sum_{i=1}^{\alpha} \kappa_{li} \left(u_{ki}(x) + \lambda \sum'_{(k'i')} \frac{u_{k'i'}(x)}{E_k - E_{k'}} \int r u_{ki} u_{k'i'} dx \right) \\ (l = 1, 2, \dots, \alpha).$$

The dash on the second sigma indicates that *all* the terms with $k' = k$ are to be omitted. In the application of the formula for an arbitrary k , it is to be observed that the κ_{li} 's, as obviously also the multiplicity α of the proper value E_k , to which we have specially directed our attention, still depend on the index k , though this is not expressed in the symbols. Let us repeat here that the κ_{li} 's are to be calculated as a system of solutions of equations (21'), normalised so that the sum of the squares is unity, where the coefficients of the equations are given by (22), while for the quantity ϵ_l in (21'), *one* of the roots of

(23) is to be taken. *This* root then gives the allied perturbed proper value, from

$$(31) \quad E^*_{kl} = E_k + \lambda \epsilon_l.$$

Formulae (30) and (31) are the generalisations of (14) and (15) of § 1.

It need scarcely be said that the extensions and generalisations mentioned at the end of § 1 can of course take effect here also. It is hardly worth the trouble to carry out these developments generally. We succeed best in any special case if we do not use ready-made formulae, but go directly by the simple fundamental principles, which have been explained, perhaps too minutely, in the present paper. I would only like to consider briefly the possibility, already mentioned at the end of § 1, that the equation (2) perhaps may lose (and indeed in the case of several variables irreparably lose), its self-adjoint character if the perturbing terms also contain derivatives of the unknown function. From general theorems we know that then the proper values of the perturbed equation no longer need to be real. We can illustrate this further. We can easily see, by carrying out the developments of this paragraph, that the elements of determinant (23) are *no longer symmetrical*, when the perturbing term contains derivatives. It is known that in this case the roots of equation (23) no longer require to be real.

The necessity for the expansion of certain functions in a series of proper functions, in order to arrive at the first or zero approximation of the proper values or functions, can become very inconvenient, and can at least complicate the calculation considerably in cases where an extended spectrum co-exists with the point spectrum and where the point spectrum has a limiting point (point of accumulation) at a finite distance. This is just the case in the problems appearing in the quantum theory. Fortunately it is often—perhaps always—possible, for the purpose of the perturbation theory, to free oneself from the generally very troublesome extended spectrum, and to develop the perturbation theory from an equation which does *not* possess such a spectrum, and whose proper values do *not* accumulate near a finite value, but grow beyond all limits with increasing index. We will become acquainted with an example in the next paragraph. Of course, this simplification is only possible when we are not interested in a proper value of the extended spectrum.

II. APPLICATION TO THE STARK EFFECT

§ 3. Calculation of Frequencies by the Method which corresponds to that of Epstein

If we add a potential energy $+eFz$ to the wave equation (5), Part I. (p. 2), of the Kepler problem, corresponding to the influence of an electric field of strength F in the positive z -direction, on a negative

electron of charge e , then we obtain the following wave equation for the Stark effect of the hydrogen atom,

$$(32) \quad \nabla^2\psi + \frac{8\pi^2m}{h^2}\left(E + \frac{e^2}{r} - eFz\right)\psi = 0,$$

which forms the basis of the remainder of this paper. In § 5 we will apply the general perturbation theory of § 2 directly to this partial differential equation. Now, however, we will lighten our task by introducing space parabolic co-ordinates $\lambda_1, \lambda_2, \phi$, by the following equations,

$$(33) \quad \begin{cases} x = \sqrt{\lambda_1\lambda_2} \cos \phi \\ \quad + \\ y = \sqrt{\lambda_1\lambda_2} \sin \phi \\ \quad + \\ z = \frac{1}{2}(\lambda_1 - \lambda_2). \end{cases}$$

λ_1 and λ_2 run from 0 to infinity; the corresponding co-ordinate surfaces are the two sets of confocal paraboloids of revolution, which have the origin as focus and the positive (λ_2) or negative (λ_1) z -axis respectively as axes. ϕ runs from 0 to 2π , and the co-ordinate surfaces belonging to it are the set of half planes limited by the z -axis. The relation of the co-ordinates is *unique*. For the functional determinant we get

$$(34) \quad \frac{\partial(x, y, z)}{\partial(\lambda_1, \lambda_2, \phi)} = \frac{1}{4}(\lambda_1 + \lambda_2).$$

The *space element* is thus

$$(35) \quad dx dy dz = \frac{1}{4}(\lambda_1 + \lambda_2) d\lambda_1 d\lambda_2 d\phi.$$

We notice, as consequences of (33),

$$(36) \quad x^2 + y^2 = \lambda_1\lambda_2; \quad r^2 = x^2 + y^2 + z^2 = \left\{\frac{1}{2}(\lambda_1 + \lambda_2)\right\}^2.$$

The expression of (32) in the chosen co-ordinates gives, if we multiply by (34)¹ (to restore the self-adjoint form).

$$(32') \quad \begin{cases} \frac{\partial}{\partial\lambda_1}\left(\lambda_1 \frac{\partial\psi}{\partial\lambda_1}\right) + \frac{\partial}{\partial\lambda_2}\left(\lambda_2 \frac{\partial\psi}{\partial\lambda_2}\right) + \frac{1}{4}\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right)\frac{\partial^2\psi}{\partial\phi^2} \\ \quad + \frac{2\pi^2m}{h^2}\left[E(\lambda_1 + \lambda_2) + 2e^2 - \frac{1}{2}eF(\lambda_1^2 - \lambda_2^2)\right]\psi = 0. \end{cases}$$

Here we can again take—and this is the why and wherefore of all “methods” of solving linear partial differential equations—the function ψ as the product of three functions, thus,

$$(37) \quad \psi = \Lambda_1\Lambda_2\Phi,$$

¹ So far as the actual details of the analysis are concerned, the simplest way to get (32'), or, in general, to get the wave equation for any special co-ordinates, is to transform not the wave equation itself, but the corresponding variation problem (cf. Part I. p. 12), and thus to obtain the wave equation afresh as an Eulerian variation problem. We are thus spared the troublesome evaluation of the *second* derivatives. Cf. Courant-Hilbert, chap. iv. § 7, p. 193.

each of which depends on only *one* co-ordinate. For these functions we get the ordinary differential equations

$$(38) \quad \begin{cases} \frac{\partial^2 \Phi}{\partial \phi^2} = -n^2 \Phi \\ \frac{\partial}{\partial \lambda_1} \left(\lambda_1 \frac{\partial \Lambda_1}{\partial \lambda_1} \right) + \frac{2\pi^2 m}{\hbar^2} \left(-\frac{1}{2} eF \lambda_1^2 + E \lambda_1 + e^2 - \beta - \frac{n^2 \hbar^2}{8\pi^2 m} \frac{1}{\lambda_1} \right) \Lambda_1 = 0, \\ \frac{\partial}{\partial \lambda_2} \left(\lambda_2 \frac{\partial \Lambda_2}{\partial \lambda_2} \right) + \frac{2\pi^2 m}{\hbar^2} \left(\frac{1}{2} eF \lambda_2^2 + E \lambda_2 + e^2 + \beta - \frac{n^2 \hbar^2}{8\pi^2 m} \frac{1}{\lambda_2} \right) \Lambda_2 = 0, \end{cases}$$

wherein n and β are two further "proper value-like" constants of integration (in addition to E), still to be defined. By the choice of symbol for the first of these, we have taken into account the fact that the first of equations (38) makes it take integral values, if Φ and $\frac{\partial \Phi}{\partial \phi}$ are to be continuous and single-valued functions of the azimuth ϕ .

We then have

$$(39) \quad \Phi = \frac{\sin}{\cos} n\phi$$

and it is evidently sufficient if we do not consider negative values of n . Thus

$$(40) \quad n = 0, 1, 2, 3 \dots$$

In the symbol used for the second constant β , we follow Sommerfeld (*Atombau*, 4th edit., p. 821) in order to make comparison easier. (Similarly, below, with A, B, C, D .) We treat the last two equations of (38) together, in the form

$$(41) \quad \frac{\partial}{\partial \xi} \left(\xi \frac{\partial \Lambda}{\partial \xi} \right) + \left(D\xi^2 + A\xi + 2B + \frac{C}{\xi} \right) \Lambda = 0,$$

where

$$(42) \quad \left. \begin{matrix} D_1 \\ D_2 \end{matrix} \right\} = \mp \frac{\pi^2 m e F}{\hbar^2}, \quad A = \frac{2\pi^2 m E}{\hbar^2}, \quad \left. \begin{matrix} B_1 \\ B_2 \end{matrix} \right\} = \frac{\pi^2 m}{\hbar^2} (e^2 \mp \beta), \quad C = -\frac{n^2}{4},$$

and the upper sign is valid for $\Lambda = \Lambda_1, \xi = \lambda_1$, and the lower one for $\Lambda = \Lambda_2, \xi = \lambda_2$. (Unfortunately, we have to write ξ instead of the more appropriate λ , to avoid confusion with the perturbation parameter λ of the general theory, §§ 1 and 2.)

If we omit initially in (41) the Stark effect term $D\xi^2$, which we conceive as a perturbing term (limiting case for vanishing field), then this equation has the same general structure as equation (7) of Part I., and the domain is also the same, from 0 to ∞ . The discussion is almost the same, word for word, and shows that non-vanishing solutions, which, with their derivatives, are continuous and remain finite within the domain, only exist if *either* $A > 0$ (extended spectrum, corresponding to hyperbolic orbits) *or*

$$(43) \quad \frac{B}{\sqrt{-A}} - \sqrt{-C} = k + \frac{1}{2}; \quad k = 0, 1, 2, \dots$$

If we apply this to the last two equations of (38) and distinguish the two k -values by suffixes 1 and 2, we obtain

$$(44) \quad \begin{cases} \sqrt{-A}(k_1 + \frac{1}{2} + \sqrt{-C}) = B_1 \\ \sqrt{-A}(k_2 + \frac{1}{2} + \sqrt{-C}) = B_2. \end{cases}$$

By addition, squaring and use of (42) we find

$$(45) \quad A = -\frac{4\pi^4 m^2 e^4}{h^4 l^2} \quad \text{and} \quad E = -\frac{2\pi^2 m e^4}{h^2 l^2}.$$

These are the well-known Balmer-Bohr elliptic levels, where as *principal quantum number* enters

$$(46) \quad l = k_1 + k_2 + n + 1.$$

We get the *discrete* term spectrum and the allied proper functions in a way *simpler* than that indicated, if we apply results already known in mathematical literature as follows. We transform first the dependent variable Λ in (41) by putting

$$(47) \quad \Lambda = \xi^{\frac{n}{2}} u$$

and then the independent ξ by putting

$$(48) \quad 2\xi\sqrt{-A} = \eta.$$

We find for u as a function of η the equation

$$(41') \quad \frac{d^2 u}{d\eta^2} + \frac{n+1}{\eta} \frac{du}{d\eta} + \left(\frac{D}{(2\sqrt{-A})^3 \eta} - \frac{1}{4} + \frac{B}{\sqrt{-A}} \frac{1}{\eta} \right) u = 0.$$

This equation is very intimately connected with the polynomials named after Laguerre. In the mathematical appendix, it will be shown that the product of $e^{-\frac{x}{2}}$ and the n th derivative of the $(n+k)$ th Laguerre polynomial satisfies the differential equation

$$(103) \quad y'' + \frac{n+1}{x} y' + \left(-\frac{1}{4} + \left(k + \frac{n+1}{2} \right) \frac{1}{x} \right) y = 0,$$

and that, for a fixed n , the functions named form the complete system of proper functions of the equation just written, when k runs through all non-negative integral values. Thus it follows that, for vanishing D , equation (41') possesses the proper functions

$$(49) \quad u_k(\eta) = e^{-\frac{\eta}{2}} L_{n+k}^n(\eta)$$

and the proper values

$$(50) \quad \frac{B}{\sqrt{-A}} = \frac{n+1}{2} + k \quad (k = 0, 1, 2 \dots)$$

—and no others! (See the mathematical appendix concerning the remarkable loss of the extended spectrum caused by the apparently inoffensive transformation (48); by this loss the development of the perturbation theory is made much easier.)

We have now to calculate the perturbation of the proper values (50) from the general theory of § 1, caused by including the D -term in (41'). The equation becomes self-adjoint if we multiply by η^{n+1} . The density function $\rho(x)$ of the general theory thus becomes η^n . As perturbation function $r(x)$ appears

$$(51) \quad -\frac{D}{(2\sqrt{-A})^3} \eta^{n+2}.$$

(We formally put the perturbation parameter $\lambda=1$; if we desired, we could identify D or F with it.) Now formula (7') gives, for the perturbation of the k th proper value,

$$(52) \quad \epsilon_k = -\frac{D}{(2\sqrt{-A})^3} \frac{\int_0^\infty \eta^{n+2} e^{-\eta[L_{n+k}^n(\eta)]^2} d\eta}{\int_0^\infty \eta^n e^{-\eta[L_{n+k}^n(\eta)]^2} d\eta}.$$

For the integral in the denominator, which merely provides for the normalisation, formula (115) of the appendix gives the value

$$(53) \quad \frac{[(n+k)!]^3}{k!},$$

while the integral in the numerator is evaluated in the same place, as

$$(54) \quad \frac{[(n+k)!]^3}{k!} (n^2 + 6nk + 6k^2 + 6k + 3n + 2).$$

Consequently

$$(55) \quad \epsilon_k = -\frac{D}{(2\sqrt{-A})^3} (n^2 + 6nk + 6k^2 + 6k + 3n + 2).$$

The condition for the k th perturbed proper value of equation (41') and therefore, naturally, also for the k th discrete proper value of the original equation (41) runs therefore

$$(56) \quad \frac{B}{\sqrt{-A}} = \frac{n+1}{2} + k + \epsilon_k$$

(ϵ_k is retained meantime for brevity).

This result is applied twice, namely, to the last two equations of (38) by substituting the two systems (42) of values of the constants A, B, C, D ; and it is to be observed that n is the *same* number in the two cases, while the two k -values are to be distinguished by the suffixes 1 and 2, as above. First we have

$$(57) \quad \begin{cases} \frac{B_1}{\sqrt{-A}} = \frac{n+1}{2} + k_1 + \epsilon_{k_1} \\ \frac{B_2}{\sqrt{-A}} = \frac{n+1}{2} + k_2 + \epsilon_{k_2} \end{cases}$$

whence comes

$$(58) \quad A = -\frac{(B_1 + B_2)^2}{(l + \epsilon_{k_1} + \epsilon_{k_2})^2}$$

(applying abbreviation (46) for the principal quantum number). In the approximation we are aiming at we may expand with respect to the small quantities ϵ_k and get

$$(59) \quad A = -\frac{(B_1 + B_2)^2}{l^2} \left[1 - \frac{2}{l}(\epsilon_{k_1} + \epsilon_{k_2}) \right].$$

Further, in the calculation of these small quantities, we may use the approximate value (45) for A in (55). We thus obtain, noticing the two D values, by (42),

$$(60) \quad \begin{cases} \epsilon_{k_1} = +\frac{Fh^4l^3}{64\pi^4m^2e^5}(n^2 + 6nk_1 + 6k_1^2 + 6k_1 + 3n + 2) \\ \epsilon_{k_2} = -\frac{Fh^4l^3}{64\pi^4m^2e^5}(n^2 + 6nk_2 + 6k_2^2 + 6k_2 + 3n + 2). \end{cases}$$

Addition gives, after an easy reduction,

$$(61) \quad \epsilon_{k_1} + \epsilon_{k_2} = \frac{3Fh^4l^4(k_1 - k_2)}{32\pi^4m^2e^5}.$$

If we substitute this, and the values of A , B_1 , and B_2 from (42) in (59), we get, after reduction,

$$(62) \quad E = -\frac{2\pi^2mc^4}{h^2l^2} - \frac{3}{8} \frac{h^2Fl(k_2 - k_1)}{\pi^2me}$$

This is our provisional *conclusion*; it is the well-known formula of Epstein for the term values in the Stark effect of the hydrogen spectrum.

k_1 and k_2 correspond fully to the parabolic quantum numbers; they are capable of taking the value zero. Also the integer n , which has evidently to do with the *equatorial* quantum number, may from (40) take the value zero. However, from (46) the sum of these three numbers must still be increased by unity in order to yield the principal quantum number. Thus $(n + 1)$ and not n corresponds to the equatorial quantum number. The value zero for the *latter* is thus *automatically* excluded by wave mechanics, just as by Heisenberg's mechanics.¹ *There is simply no proper function, i.e.* no state of vibration, which corresponds to such a meridional orbit. This important and gratifying circumstance was already brought to light in Part I. in counting the constants, and also afterwards in § 2 of Part I. in connection with the azimuthal quantum number, through the non-existence of states of vibration corresponding to *pendulum orbits*; its full meaning, however, only fully dawned on me through the remarks of the two authors just quoted.

¹ W. Pauli, jun., *Ztschr. f. Phys.* 36, p. 336, 1926; N. Bohr, *Die Naturw.* 1, 1926.

For later application, let us note the system of proper functions of equation (32) or (32') in "zero approximation", which belongs to the proper values (62). It is obtained from statement (37), from conclusions (39) and (49), and from consideration of transformations (47) and (48) and of the approximate value (45) of A . For brevity, let us call a_0 the "radius of the first hydrogen orbit". Then we get

$$(63) \quad \frac{1}{2l\sqrt{-A}} = \frac{h^2}{4\pi^2 m e^2} = a_0.$$

The proper functions (not yet normalised!) then read

$$(64) \quad \psi_{nk_1k_2} = \lambda_1^{\frac{n}{2}} \lambda_2^{\frac{n}{2}} e^{-\frac{\lambda_1 + \lambda_2}{2la_0}} L_{n+k_1}^n \left(\frac{\lambda_1}{la_0}\right) L_{n+k_2}^n \left(\frac{\lambda_2}{la_0}\right) \frac{\sin n\phi}{\cos n\phi}.$$

They belong to the proper values (62), where l has the meaning (46). To each non-negative integral trio of values n, k_1, k_2 belong (on account of the double symbol $\frac{\sin}{\cos}$) two proper functions or one, according as $n > 0$ or $n = 0$.

§ 4. Attempt to calculate the Intensities and Polarisation of the Stark Effect Patterns

I have lately shown¹ that from the proper functions we can calculate by differentiation and quadrature the elements of the *matrices*, which are allied in Heisenberg's mechanics to functions of the generalised position- and momentum-co-ordinates. For example, for the (rr')th element of the matrix, which according to Heisenberg belongs to the generalised co-ordinate q itself, we find

$$(65) \quad \left\{ \begin{aligned} q^{rr'} &= \int q \rho(x) \psi_r(x) \psi_{r'}(x) dx \\ &\cdot \left\{ \int \rho(x) [\psi_r(x)]^2 dx \cdot \int \rho(x) [\psi_{r'}(x)]^2 dx \right\}^{-\frac{1}{2}}. \end{aligned} \right.$$

Here, for our case, the separate indices *each* deputise for a *trio* of indices n, k_1, k_2 , and further, x represents the three co-ordinates r, θ, ϕ . $\rho(x)$ is the density function; in our case the quantity (34). (We may compare the self-adjoint equation (32') with the general form (2)). The "denominator" (. . .)^{-1/2} in (65) must be put in because our system (64) of functions is not yet normalised.

According to Heisenberg,² now, if q means a rectangular Cartesian co-ordinate, then the square of the matrix element (65) is to be a measure of the "probability of transition" from the r th state to the r' th, or, more accurately, a measure of the intensity of that part of the radiation, bound up with this transition, which is polarised in the q -direction. Starting from this, I have shown in the above paper that if we make

¹ Preceding paper of this collection.

² W. Heisenberg, *Ztschr. f. Phys.* 33, p. 879, 1925; M. Born and P. Jordan, *Ztschr. f. Phys.* 34, pp. 867, 886, 1925.

certain simple assumptions as to the electrodynamical meaning of ψ , the "mechanical field scalar", then the matrix element in question is susceptible of a very simple physical interpretation in wave mechanics, namely, *actually*: component of the amplitude of the periodically oscillating electric moment of the atom. The word *component* is to be taken in a double sense: (1) component in the q -direction, *i.e.* in the spatial direction in question, and (2) only the part of this spatial component which changes in a time-sinusoidal manner with exactly the frequency of the emitted light, $|E_r - E_r'|/h$. (It is a question then of a kind of Fourier analysis: not in harmonic frequencies, but in the actual frequencies of emission.) However, the idea of wave mechanics is not that of a sudden transition from one state of vibration to another, but according to it, the partial moment concerned—as I will briefly name it—arises from the *simultaneous existence* of the two proper vibrations, and lasts just as long as both are excited together.

Moreover, the above assertion that the q^{rr} 's are proportional to the partial moments is more accurately phrased thus. The ratio of, *e.g.*, q^{rr} to $q^{rr'}$ is equal to the ratio of the partial moments which arise when the proper function ψ_r and the proper functions $\psi_{r'}$ and $\psi_{r''}$ are stimulated, the first *with any strength whatever* and the last two with strengths *equal to one another—i.e.* corresponding to normalisation. To calculate the ratio of the *intensities*, the q -quotient must first be squared and then multiplied by the ratio of the fourth powers of the emission frequencies. The latter, however, has no part in the intensity ratio of the Stark effect components, for there we only compare intensities of lines which have practically the same frequency.

The known *selection and polarisation rules* for Stark effect components can be obtained, almost without calculation, from the integrals in the numerator of (65) and from the form of the proper functions in (64). They follow from the vanishing or non-vanishing of the integral with respect to ϕ . We obtain the components whose electric vector vibrates *parallel* to the field, *i.e.* to the z -direction, by replacing the q in (65) by z from (33). The expression for z , *i.e.* $\frac{1}{2}(\lambda_1 - \lambda_2)$, does *not* contain the azimuth ϕ . Thus we see at once from (64) that a non-vanishing result after integration with respect to ϕ can only arise if we combine proper functions whose n 's are *equal*, and thus whose equatorial quantum numbers are equal, being in fact equal to $n + 1$. For the components which vibrate *perpendicular* to the field, we must put q equal to x or equal to y (cf. equation (33)). Here $\cos \phi$ or $\sin \phi$ enters, and we see almost as easily as before, that the n -values of the two combined proper functions must differ exactly by unity, if the integration with respect to ϕ is to yield a non-vanishing result. Hence the known selection and polarisation rules are proved. Further, it should be recalled again that we do not require to exclude any n -value after additional reflection, as was necessary in the older theory in order to agree with experience. Our n is smaller by 1 than the equatorial quantum number, and right from the beginning cannot take negative values

(quite the same state of affairs exists, we know, in Heisenberg's theory).¹

The numerical evaluation of the integrals with respect to λ_1 and λ_2 which appear in (65) is exceptionally tedious, especially for those of the numerator. The same apparatus for calculating comes into play as served already in the evaluation of (52), only the matter is somewhat more detailed because the two (generalised) Laguerre polynomials, whose product is to be integrated, have not the same argument. By good luck, in the *Balmer lines*, which interest us principally, one of the two polynomials L_{n+k}^n , namely that relating to the doubly quantised state, is either a constant or is a linear function of its argument. The method of calculation is described more fully in the mathematical appendix. The following tables and diagrams give the results for the first four Balmer lines, in comparison with the known measurements and estimates of intensity, made by Stark² for a field strength of about 100,000 volts per centimetre. The first column indicates the state of polarisation, the second gives the combination of the terms in the usual manner of description, *i.e.* in *our* symbols: of the two trios of numbers ($k_1, k_2, n+1$) the *first* trio refers to the higher quantised state and the *second* to the doubly quantised state. The third column, with the heading Δ , gives the term decomposition in multiples of $3h^2 F/8\pi^2 me$, (see equation (62)). The next column gives the intensities observed by Stark, and 0 there signifies not observed. The question mark was put by Stark at such lines as clash either with irrelevant lines or with possible "ghosts" and thus cannot be guaranteed. On account of the unequal weakening of the two states of polarisation in the spectrograph, according to Stark his results for the \parallel and for the \perp components of vibration are not directly comparable with one another. Finally, the last column gives the results of our calculation in *relative numbers*, which are comparable for the collective components (\parallel and \perp) of *one* line, *e.g.* of H_α , but not for those of H_α with H_β , etc. These relative numbers are reduced to their *smallest integral* values, *i.e.* the numbers in each of the four tables are *prime* to each other.

¹ W. Pauli, jun., *Ztschr. f. Physik*, 36, p. 336, 1926.

² J. Stark, *Ann. d. Phys.* 48, p. 193, 1915.

INTENSITIES IN THE STARK EFFECT OF THE BALMER LINES

TABLE 1

 H_{α}

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
	(111) (011)	2	1	729
	(102) (002)	3	1.1	2304
	(201) (101)	4	1.2	1681
	(201) (011)	8	0	1
				Sum : 4715
⊥	(003) (002)	0	} 2.6 {	4608
	(111) (002)	0		882
	(102) (101)	1	1	1936
	(102) (011)	5	0	16
	(201) (002)	6	0	18
				Sum * : 4715

* Undisplaced components halved.

TABLE 2

 H_{β}

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
	(112) (002)	0	1.4	0
	(211) (101)	2	1.2	9
	—	(4)	1	0
	(211) (011)	6	4.8	81
	(202) (002)	8	9.1	384
	(301) (101)	10	11.5	361
	—	(12)	1	0
	(301) (011)	14	0	1
				Sum : 836
⊥	—	(0)	1.4	0
	(112) (011)	2	3.3	72
	(103) (002)	4	} 12.6 {	384
	(211) (002)	4		72
	(202) (101)	6	9.7	294
	—	(8)	1.3	0
	(202) (011)	10	1.1 ?	6
(301) (002)	12	1 ?	8	
				Sum : 836

INTENSITIES IN THE STARK EFFECT OF THE BALMER LINES

TABLE 3

 H_γ

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
	(221) (011)	2	1.6	15 625
	(212) (002)	5	1.5	19 200
	(311) (101)	8	1	1 521
	(311) (011)	12	2.0	16 641
	(302) (002)	15	7.2	115 200
	(401) (101)	18	10.8	131 769
	(401) (011)	22	1 ?	729
Sum : 300 685				
⊥	(113) (002)	0	} 7.2 {	115 200
	(221) (002)	0		26 450
	(212) (101)	3		46 128
	(212) (011)	7	1.2	5 808
	(203) (002)	10	} 4.3 {	76 800
	(311) (002)	10		11 250
	(302) (101)	13		83 232
	(302) (011)	17	1.1	2 592
	(401) (002)	20	1	4 050
Sum : * 300 685				

* Undisplaced components halved.

TABLE 4

 H_δ

Polarisation.	Combination.	Δ	Observed Intensity.	Calculated Intensity.
	(222) (002)	0	0	0
	(321) (101)	4	1	8
	(321) (011)	8	1.2	32
	(312) (002)	12	1.5	72
	(411) (101)	16	1.2	18
	(411) (011)	20	1.1	18
	(402) (002)	24	2.8	180
	(501) (101)	28	7.2	242
	(501) (011)	32	1 ?	2
	Sum: 572			
⊥	(222) (011)	2	1.3	36
	(213) (002)	6	} 3.2 {	162
	(321) (002)	6		36
	(312) (101)	10		98
	(312) (011)	14	1	2
	(303) (002)	18	} 2.0 {	90
	(411) (002)	18		9
	(402) (101)	22		125
	(402) (011)	26	1.3	5
(501) (002)	30	1 ?	9	
Sum : 572				

In the *diagrams* it is to be noticed that, on account of the huge differences in the theoretical intensities, some theoretical intensities



FIG. 1.— H_{α} \parallel -components.

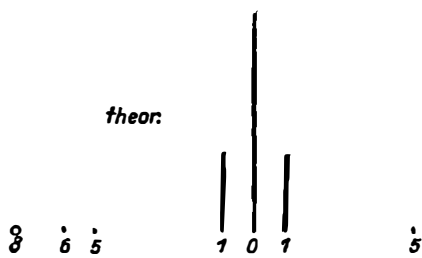
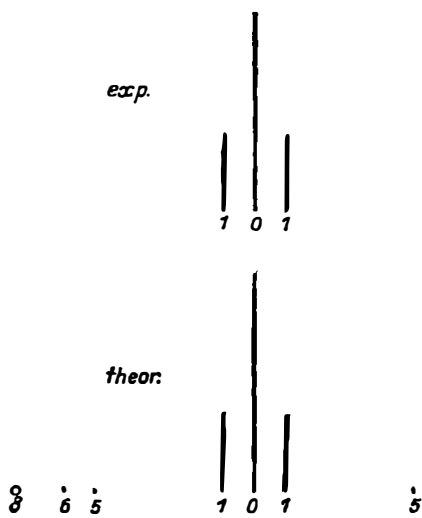


FIG. 2.— H_{α} \perp -components

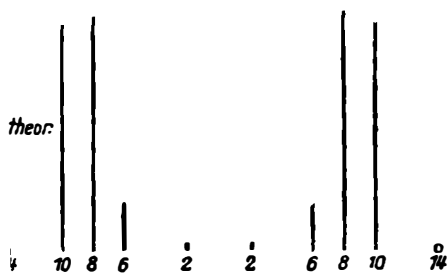
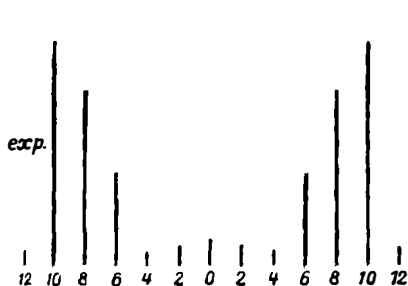


FIG. 3.— H_{β} \parallel -components.

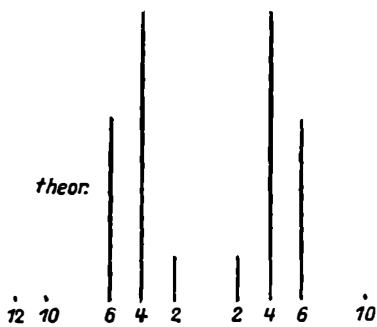
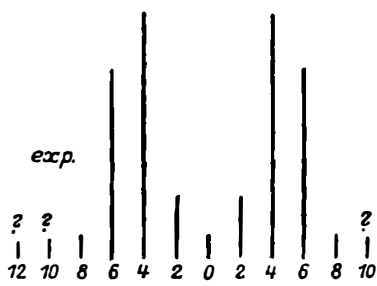


FIG. 4.— H_{β} \perp -components.

cannot be truly represented to scale, as they are much too small. These are indicated by *small circles*.

A consideration of the diagrams shows that the agreement is tolerably good for almost all the strong components, and taken all over it is somewhat better than for the values deduced from correspondence considerations.¹ Thus, for example, is removed one of the most serious contradictions which arose, in that the correspondence principle gave the ratio of the intensities of the two strong \perp -components of H_β , for $\Delta=4$ and 6, inversely and indeed very much out, in fact

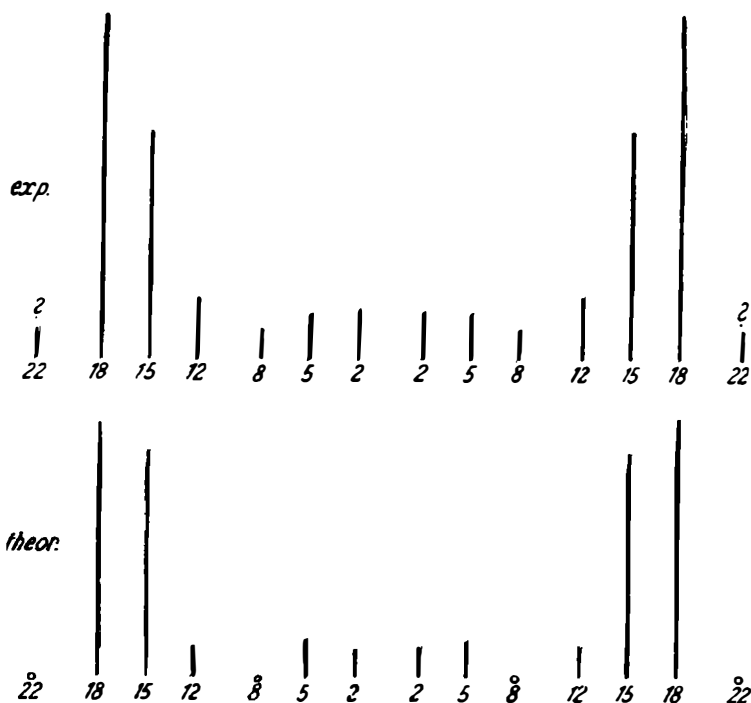


FIG. 5.— H_γ \parallel -components.

as almost 1:2, while experiment requires about 5:4. A similar thing occurs with the mean ($\Delta=0$) \perp -components of H_γ , which decidedly preponderate experimentally, but are given as *far* too weak by the correspondence principle. In *our* diagrams also, it is admitted that such "reciprocities" between the intensity ratios of intense components demanded by theory and by experiment are not entirely wanting. The theoretically most intense \parallel -component ($\Delta=3$) of H_α is furthest out; by experiment, it should lie *between* its neighbours in intensity. And the two strongest \parallel -components of H_β and two \perp -components ($\Delta=10, 13$) of H_γ are given "reciprocally" by the

¹ H. A. Kramers, *Dänische Akademie* (8), iii. 3, p. 333 *et seq.*, 1919.

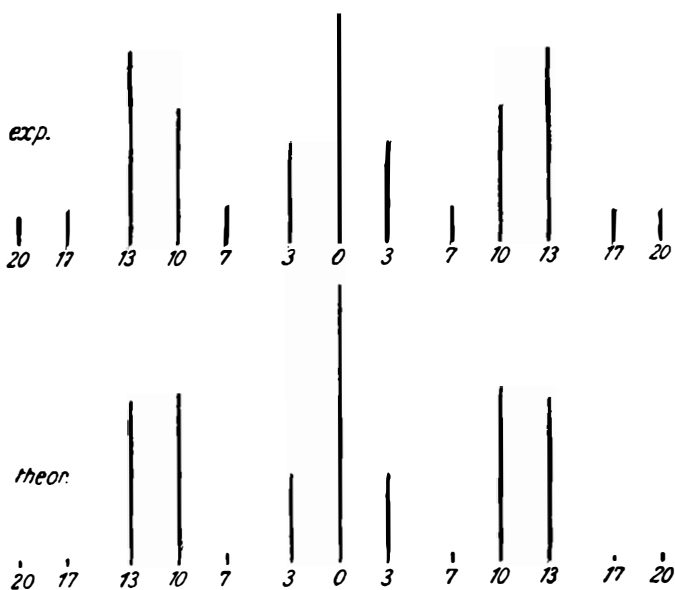


FIG. 6.— $H_\gamma \perp$ -components.

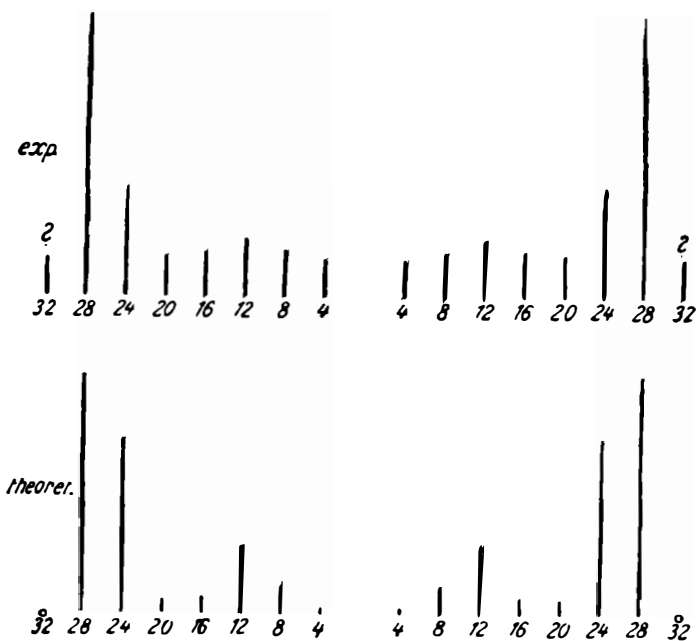


FIG. 7.— $H_\beta \parallel$ -components.

theory. Of course, in both cases the intensity ratios, both experimentally and theoretically, are pretty near unity.

Passing now to the weaker components, we notice first that the contradiction which exists for some weak observed components of H_β to the selection and polarisation rules, of course still remains in the new theory, since the latter gives these rules in conformity with the older theory. However, components which are extremely weak theoretically are for the most part *unobserved*, or the observations are *questionable*. The strength ratios of weaker components to one another or to stronger ones are *almost never* given even approximately correctly; cf. especially H_γ and H_δ . Such serious mistakes in the experimental determination of the blackening are of course out of the question.

Considering all this, we might feel inclined to be very sceptical of

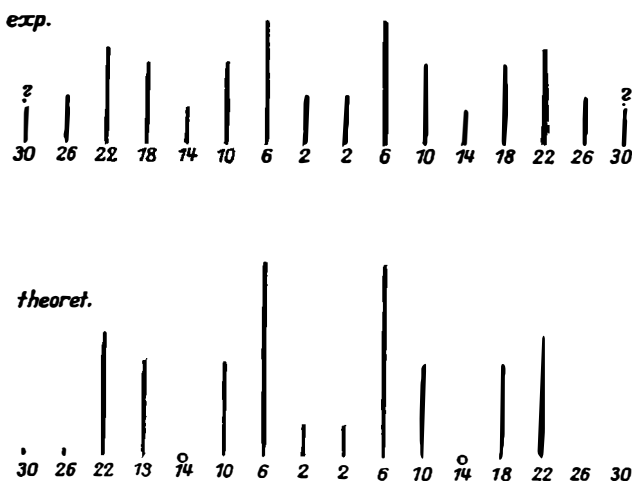


FIG. 8.— $H_\delta \perp$ -components.

the thesis that the integrals (65) or their squares are measures of intensity. I am far from wishing to represent this thesis as irrefutable. There are still many alterations conceivable, and these may, perhaps, be necessitated by internal reasons when the theory is further extended. Yet the following should be remembered. The whole calculation has been performed with the *unperturbed* proper functions, or more precisely, with the *zero* approximation to the perturbed ones (cf. above § 2). It, therefore, represents an approximation for a *vanishing* field strength! However, just for the weak or almost vanishing components we should expect theoretically a fairly powerful growth with increasing field strength, for the following reason. According to the view of wave mechanics, as explained at the beginning of this section, the integrals (65) represent the amplitudes of the electrical partial moments, which are produced by the distribution of charges which flow round about the nucleus within the atom's domain. When for

a line component we get as a zero approximation very weak or even vanishing intensity, this is not caused in any way by the fact that to the simultaneous existence of the two proper vibrations corresponds only an insignificant motion of electricity, or even none at all. The vibrating mass of electricity—if this vague expression is allowed—may be represented as the same in all components, on the ground of *normalisation*. Rather is the reason for the low line intensity to be found in a high degree of *symmetry* in the motion of the electricity, through which only a small, or even no, dipole moment arises (on the contrary, *e.g.*, only a four-pole moment). Therefore it is to be expected that the *vanishing* of a line component in presence of perturbations of any kind is a relatively *unstable* condition, since the symmetry is probably destroyed by the perturbation. And thus it may be expected that weak or vanishing components gain quickly in intensity with increasing field strength.

This has now actually been observed, and the intensity ratios, indeed, alter quite considerably with field strength, for strengths of about 10,000 gauss and upwards; and, if I understand aright, in the way¹ shown by the present general discussion. Certain information on the question whether this really explains these discrepancies could of course only be got from a continuation of the calculation to the next approximation, but this is *very* troublesome and complicated.

The present considerations are of course nothing but the “translation” into the language of the new theory of very well-known considerations which Bohr² has brought forward in connection with calculation of line intensities by means of the principle of correspondence.

The theoretical intensities given in the tables satisfy a fundamental requirement, which is set up not only by intuition but also by experiment,³ *viz.*, the sum of the intensities of the ||-components is equal to that of the ⊥-components. (Before adding, *undisplaced* components must be *halved*—as a compensation for the *duplication* of all the others, which occur on both sides.) This makes a very welcome “control” for the arithmetic.

It is also of interest to compare the *total intensities* of the four lines by using the four “sums” given in the tables. For this purpose I take back from my numerical calculations the four factors, which were omitted in order to represent the intensity ratios within each of the four line groups by the smallest integers possible, and multiply by them. Further, I multiply each of these four products by the *fourth* power of the appropriate emission frequency. Thus I obtain the following four numbers:

$$\begin{aligned} \text{for } H_{\alpha} \dots & \frac{2^6 \cdot 23 \cdot 41}{3^2 \cdot 5^9} = 0.003433 \dots \\ \text{for } H_{\beta} \dots & \frac{4 \cdot 11 \cdot 19}{3^{12}} = 0.001573 \dots \end{aligned}$$

¹ J. Stark, *Ann. d. Phys.* 43, p. 1001 *et seq.*, 1914.

² N. Bohr, *Dänische Akademie* (8), iv. 1. 1, p. 35, 1918.

³ J. Stark, *Ann. d. Phys.* 43, p. 1004, 1914.

$$\text{for } H_\gamma \dots \frac{2^6 \cdot 3^6 \cdot 11^2 \cdot 71}{5 \cdot 7^{13}} = 0.0008312 \dots$$

$$\text{for } H_\delta \dots \frac{11 \cdot 13}{2^{15} \cdot 3^2} = 0.0004849 \dots$$

I give these numbers *with still greater reserve* than the former ones because I am not sure, theoretically, about the *fourth* power of the frequency. Investigations¹ which I have lately published seem to call, perhaps, for the *sixth*. The above method of calculation corresponds exactly to the assumptions of Born, Jordan, and Heisenberg.² Fig. 9 represents the results diagrammatically.

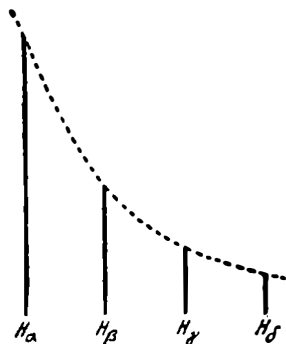


FIG. 9.—Total Intensities.

Actual measured intensities of emission lines, which are known to depend greatly on the conditions of excitation, naturally cannot here be used in a comparison with experience. From his researches³ on dispersion and magneto-rotation in the neighbourhood of H_α and H_β , R. Ladenburg has, with F. Reiche,⁴ calculated the value 4.5 (limits 3 and 6) for the ratio of the so-called "electronic numbers" of these two lines. If I assume that the above numbers may be taken as proportional to Ladenburg's⁵ expression,

$$\sum_i \frac{g_k}{g_i} a_{ki} \nu_0,$$

then they may be reduced to (relative) "electronic numbers" by division by ν_0^3 , i.e. by

$$\left(\frac{5}{36}\right)^3, \left(\frac{3}{16}\right)^3, \left(\frac{21}{100}\right)^3, \text{ and } \left(\frac{2}{9}\right)^3 \text{ respectively.}$$

Hence we obtain the four numbers,

$$1.281, 0.2386, 0.08975, 0.04418.$$

The ratio of the first to the second is 5.37, which agrees sufficiently with Ladenburg's value.

¹ Equation (38) at end of previous paper of this collection. The *fourth* allows for the fact that for the radiation it is a question of the square of the *acceleration* and not of the electric moment itself. In this equation (38) occurs explicitly another factor $(E_k - E_m)/h$. This is occasioned by the appearance of $\partial/\partial t$ in statement (36). *Addition at proof correction*: Now I recognise this $\partial/\partial t$ to be incorrect, though I hoped it would make the later relativistic generalisation easier. Statement (36), *loc. cit.*, is to be replaced by $\psi\dot{\psi}$. The above doubts about the *fourth* power are therefore dissolved.

² Cf. M. Born and P. Jordan, *Ztschr. f. Phys.* 34, p. 887, 1925.

³ R. Ladenburg, *Ann. d. Phys.* (4), 38, p. 249, 1912.

⁴ R. Ladenburg and F. Reiche, *Die Naturwissenschaften*, 1923, p. 584.

⁵ Cf. Ladenburg-Reiche, *loc. cit.*, the first formula in the second column, p. 584. The factor ν_0 in the above expression comes from the fact that the "transition probability" a_{ki} is still to be multiplied by the "energy quantum" to give the intensity of the radiation.

§ 5. Treatment of the Stark Effect by the Method which corresponds to that of Bohr

Mainly to give an *example* of the general theory of § 2, I wish to outline *that* treatment of the proper value problem of equation (32), which must have been adopted, if we had *not* noticed that the perturbed equation is also exactly “separable” in parabolic co-ordinates. We therefore now keep to the polar co-ordinates r, θ, ϕ , and thus replace z by $r \cos \theta$. We also introduce a new variable η for r by the transformation

$$(66) \quad 2r \sqrt{-\frac{8\pi^2 m E}{h^2}} = \eta,$$

(which is closely akin to transformation (48) for the parabolic co-ordinate ξ). For one of the unperturbed proper values (45), we get from (66)

$$(66') \quad \eta = \frac{2r}{a_0},$$

where a_0 is the same constant as in (63). (“Radius of the innermost hydrogen orbit.”) If we introduce this and the unperturbed value (45) into the equation (32), which is to be treated, then we obtain

$$(67) \quad \nabla'^2 \psi + \left(-\frac{1}{4} - g\eta \cos \theta + \frac{l}{\eta} \right) \psi = 0,$$

where for brevity

$$(68) \quad g = \frac{a_0^2 F l^3}{4e}.$$

The dash on the Laplacian operator is merely to signify that in it the letter η is to be written for the radius vector.

In equation (67) we *conceive* l to be the proper value, and the term in g to be the perturbing term. The fact that the perturbing term *contains* the proper value need not trouble us in the first approximation. If we neglect the perturbing term, the equation has as proper values the natural numbers

$$(69) \quad l = 1, 2, 3, 4 \dots$$

and no others. (The extended spectrum is again cut out by the artifice (66), which would be valuable for closer approximations.) The allied *proper functions* (not yet normalised) are

$$(70) \quad \psi_{lnm} = P_n^m(\cos \theta) \frac{\cos}{\sin} (m\phi) \cdot \eta^n e^{-\frac{\eta}{2}} L_{n+l}^{2n+1}(\eta).$$

Here P_n^m signifies the m th “associated” Legendre function of the n th order, and L_{n+l}^{2n+1} is the $(2n+1)$ th derivative of the $(n+l)$ th Laguerre polynomial.¹ So we must have

$$n < l,$$

¹ I lately gave the proper functions (70) (see Part I.), but without noticing their connection with the Laguerre polynomials. For the proof of the above representation, see the Mathematical Appendix, section 1.

otherwise L_{n+l}^{2n+1} would vanish, because the number of differentiations would be greater than the degree. With reference to this, the numbering of the spherical surface harmonics shows that l is an l^2 -fold proper value of the unperturbed equation. We now investigate the *splitting up* of a *definite* value of l , supposed fixed in what follows, due to the addition of the perturbing term.

To do this we have, *in the first place*, to normalise our proper functions (70), according to § 2. From an uninteresting calculation, which is easily performed with the aid of the formulae in the appendix,¹ we get as the normalising factor

$$(71) \quad \frac{1}{\sqrt{\pi}} \sqrt{\frac{2n+1}{2}} \sqrt{\frac{(n-m)!}{(n+m)!}} \sqrt{\frac{(l-n-1)!}{[(n+l)!]^3}},$$

if $m \neq 0$, but, for $m=0$, $\frac{1}{\sqrt{2}}$ times this value. *Secondly*, we have to calculate the symmetrical matrix of constants ϵ_{im} , according to (22). The r there is to be identified² with our perturbing function $-g\eta^3 \cos \theta \sin \theta$, and the proper functions, there called u_{ki} , are to be identified with our functions (70). The fixed suffix k , which characterises the proper value, corresponds to the *first* suffix l of ψ_{lnm} , and the *other* suffix i of u_{ki} corresponds now to the *pair* of suffixes n, m in ψ_{lnm} . The matrix (22) of constants forms in our case a square of l^2 rows and l^2 columns. The quadratures are easily carried out by the formulae of the appendix and yield the following results. Only those elements of the matrix are different from zero, for which the two proper functions $\psi_{lnm}, \psi_{ln'm'}$, to be combined, satisfy the following conditions simultaneously:

1. The *upper indices* of the "associated Legendre functions" must agree, *i.e.* $m = m'$.

2. The *orders* of the two Legendre functions must differ exactly by unity, *i.e.* $|n - n'| = 1$.

3. To each trio of indices lnm , if $m \neq 0$, there belong, according to (70), *two* Legendre functions, and thus also *two* proper functions ψ_{lnm} , which only differ from each other in that one contains a factor $\cos m\phi$ and the other $\sin m\phi$. The third condition reads: we may only combine sine with sine, or cosine with cosine, and not sine with cosine.

The remaining non-vanishing elements of the desired matrix would have to be characterised from the beginning by *two* index-pairs (n, m) and $(n+1, m)$. (We renounce any idea of showing the fixed index l explicitly.) Since the matrix is symmetrical, *one* index pair (n, m) is sufficient, if we stipulate that the first index, *i.e.* n , shall mean the *greater* of the two orders n, n' , in every case.

¹ It is to be noticed that the *density function*, generally denoted by $\rho(x)$, reads as $\eta \sin \theta$ in equation (67), because the equation must be multiplied by $\eta^2 \sin \theta$, in order to acquire self-adjoint form.

Then the calculation gives

$$(72) \quad \epsilon_{nm} = -6lg \sqrt{\frac{(l^2 - n^2)(n^2 - m^2)}{4n^2 - 1}}$$

We have now to form the determinant (22) out of these elements. It is advantageous to *arrange* its rows *as well as* its columns on the following principle. (To fix our ideas, let us speak of the columns, and therefore of the index-pair characterising the *first* of the two Legendre functions.) Thus: first come all terms with $m=0$, then all with $m=1$, then all with $m=2$, etc., and finally, all terms with $m=l-1$, which last is the greatest value that m (like n) can take. *Inside* each of these groups, let us arrange the terms thus: first, all terms with $\cos m\phi$, and then all with $\sin m\phi$. Within these "half groups" let us arrange them in order of increasing n , which runs through the values $m, m+1, m+2 \dots l-1$, i.e. $(l-m)$ values in all.

If we carry this out, we find that the non-vanishing elements (72) are exclusively confined to the two secondary diagonals, which lie immediately alongside the principal diagonal. On the latter are the proper value perturbations which are to be found, but taken negatively, while everywhere else are zeros. Further, the two secondary diagonals are interrupted by zeros at *those* places, where they break through the *boundaries* between the so-called "half-groups", in very convenient fashion. Hence the whole determinant *breaks up* into a product of just so many smaller determinants as there are "half-groups" present, viz. $(2l-1)$. It will be sufficient if we consider one of them. We write it here, denoting the desired perturbation of the proper value by ϵ (without suffix):

$$(73) \quad \begin{vmatrix} -\epsilon & \epsilon_{m+1,m} & 0 & 0 & \dots & 0 \\ \epsilon_{m+1,m} & -\epsilon & \epsilon_{m+2,m} & 0 & \dots & 0 \\ 0 & \epsilon_{m+2,m} & -\epsilon & \epsilon_{m+3,m} & \dots & 0 \\ 0 & 0 & \epsilon_{m+3,m} & -\epsilon & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \epsilon_{l-1,m} & -\epsilon \end{vmatrix}$$

If we divide each term here by the common factor $6lg$ of the ϵ_{nm} 's (cf. (72)), and for the moment regard as the unknown

$$(74) \quad k^* = -\frac{\epsilon}{6lg},$$

the above equation of the $(l-m)$ th degree has the roots

$$(75) \quad k^* = \pm(l-m-1), \quad \pm(l-m-3), \quad \pm(l-m-5) \dots$$

where the series stops with ± 1 or 0 (inclusive) according as the degree $l-m$ is *even* or *odd*. The proof of this is unfortunately *not* to be found in the appendix, as I have not been successful in obtaining it.

If we form the series (75) for each of the values $m=0, 1, 2 \dots (l-1)$, then we have in the numbers

$$(76) \quad \epsilon = -6lgk^*$$

the complete set of *perturbations of the principal quantum number* l . In order to find the perturbed *proper values* E (term-levels) of the equation (32), we have only to substitute (76) in

$$(77) \quad E = -\frac{2\pi^2 m e^4}{\hbar^2(l + \epsilon)^2}$$

taking into account the signification of the abbreviations g (see (68)) and a_0 (see (63)).

After reducing this gives

$$(78) \quad E = -\frac{2\pi^2 m e^4}{\hbar^2 l^2} - \frac{3}{8} \frac{\hbar^2 F l k^*}{\pi^2 m e}$$

Comparison with (62) shows that k^* is the *difference* $k_2 - k_1$ of the parabolic quantum numbers. From (75), bearing in mind the range of values of m referred to above, we see that k^* may also take the same values as the difference just mentioned, viz. 0, 1, 2 . . . $(l-1)$. Also, if we take the trouble to work it out, we will find for the *multiplicity*, in which k^* and the difference $k_2 - k_1$ appear, the same value, viz. $l - |k^*|$.

We have thus obtained the proper value perturbations of the first order also from the general theory. The next step would be the solution of the system (21') of linear equations of the general theory for the κ -quantities. These would then yield, according to (18) (provisionally putting $\lambda=0$), the perturbed proper functions of zero order; this is nothing more than a representation of the proper functions (64) as linear forms of the proper functions (70). In our case the solution of (21') would naturally be anything but unique, on account of the considerable multiplicity of the roots ϵ . The solution is made much simpler if we notice that the equations break up into *just as many* groups, viz. $(2l-1)$, or, retaining the former expression, *half-groups*, with completely separated variables, as the determinant investigated above contains factors like (73); and if we further notice that it is allowable, after we have chosen a definite ϵ -value, to regard only the variables κ of a *single* half-group as different from zero, of that half-group, in fact, for which the determinant (73) vanishes for the chosen ϵ -value. The definition of this half-group of variables is then *unique*.

But our object, viz. to illustrate the general method of § 2 by an example, has been sufficiently attained. Since the continuation of the calculation is of no special physical interest, I have not troubled to bring the determinantal quotients, which we immediately obtain for the coefficients κ , into a clearer form, or to work out the transformation to principal axes in any other way.

On the whole, we must admit that in the present case the method of secular perturbations (§ 5) is considerably more troublesome than the direct application of a system of separation (§ 3). I believe that this may also be true in other cases. In ordinary mechanics it is, as we know, usually quite the reverse.

III.—MATHEMATICAL APPENDIX

Prefatory Note :—It is not intended to supply in uninterrupted detail all the calculations omitted from the text. Without that, the present paper has already become too long. In general, only those methods of calculation will be briefly described which another might utilise with advantage in similar work, if something better does not occur to him—as it may easily do.

§ 1. The Generalised Laguerre Polynomials and Orthogonal Functions

The k th Laguerre polynomial $L_k(x)$ satisfies the differential equation ¹

$$(101) \quad xy'' + (1-x)y' + ky = 0.$$

If we first replace k by $n+k$, and then differentiate n times, we find that the n th derivative of the $(n+k)$ th Laguerre polynomial, which we will always denote by L_{n+k}^n , satisfies the equation

$$(102) \quad xy'' + (n+1-x)y' + ky = 0.$$

Moreover, by an easy transformation, we find that for $e^{-\frac{x}{2}}L_{n+k}^n(x)$ the following equation holds,

$$(103) \quad y'' + \frac{n+1}{x}y' + \left(-\frac{1}{4} + \left(k + \frac{n+1}{2}\right)\frac{1}{x}\right)y = 0.$$

This found an application in equation (41') of § 3. The allied generalised Laguerre orthogonal functions are

$$(104) \quad x^{\frac{n}{2}}e^{-\frac{x}{2}}L_{n+k}^n(x).$$

Their equation, it may be remarked in passing, is

$$(105) \quad y'' + \frac{1}{x}y' + \left(-\frac{1}{4} + \left(k + \frac{n+1}{2}\right)\frac{1}{x} - \frac{n^2}{4x^2}\right)y = 0.$$

Let us turn to equation (103), and consider there that n is a fixed (real) integer, and k is the proper value parameter. Then, according to what has been said, in the domain $x \geq 0$, at any rate, the equation has the proper functions,

$$(106) \quad e^{-\frac{x}{2}}L_{n+k}^n(x),$$

belonging to the proper values,

$$(107) \quad k = 0, 1, 2, 3, \dots$$

In the text it is maintained that it has no further values, and, above all, that it possesses no continuous spectrum. This seems paradoxical, for the equation

$$(108) \quad \frac{d^2y}{d\xi^2} + \frac{n+1}{\xi} \frac{dy}{d\xi} + \left(-\frac{1}{(2k+n+1)^2} + \frac{1}{\xi}\right)y = 0$$

¹ Courant-Hilbert, chap. ii. § 11, 5, p. 78, equation (72).

into which (103) is transformed by the substitution

$$(109) \quad \xi = \left(k + \frac{n+1}{2} \right) x$$

does possess a continuous spectrum, if in it we regard

$$(110) \quad E = -\frac{1}{(2k+n+1)^2}$$

as proper value parameter, viz. all positive values of E are proper values (cf. Part I., analysis of equation (7)). The reason why *no* proper values k of (103) can correspond to these positive E -values is that by (110) the k -values in question would be complex, and this is impossible, according to general theorems.¹ Each *real* proper value of (103), by (110), gives rise to a *negative* proper value of (108). Moreover, we know (cf. Part I.) that (108) possesses absolutely no negative proper values other than those that arise, as in (110), from the series (107). There thus remains only the one possibility, that in the series (107) certain negative k -values are lacking, which appear on solving (110) for k , on account of the double-valuedness when extracting the root. But this also is impossible, because the k -values in question turn out to be algebraically less than $-\frac{n+1}{2}$ and thus, from general theorems,² cannot be proper values of equation (103). The series of values (107) is thus complete. Q.E.D.

The above supplements the proof that the functions (70) are the proper functions of (67) (with the perturbing term suppressed), allied to the proper values (69). We have only to write the solutions of (67) as a product of a function of θ , ϕ and a function of η . The equation in η can readily be brought to the form of (105), the only difference being that our present n is there always an odd number, namely, the $(2n+1)$ which is to be found there.

§ 2. Definite Integrals of Products of Two Laguerre Orthogonal Functions

The Laguerre polynomials can all be obtained, in the following manner, as coefficients of the powers of the auxiliary variable t , in the expansion in a series of a so-called "generating function"³

$$(111) \quad \sum_{k=0}^{\infty} L_k(x) \frac{t^k}{k!} = \frac{e^{-\frac{xt}{1-t}}}{1-t}.$$

If we replace k by $n+k$ and then differentiate n times with respect to x , we obtain the generating function of our generalised polynomials,

$$(112) \quad \sum_{k=0}^{\infty} L_{n+k}^n(x) \frac{t^k}{(n+k)!} = (-1)^n \frac{e^{-\frac{xt}{1-t}}}{(1-t)^{n+1}}.$$

¹ Courant-Hilbert, chap. iii. § 4, 2, p. 115.

² Courant-Hilbert, chap. v. § 5, 1, p. 240.

³ Courant-Hilbert, chap. ii. § 11, 5, p. 78, equation (68).

In order to evaluate with its help integrals such as appeared for the first time in the text in expression (52), or, more generally, such as were necessary in § 4 for the calculation of (65), and also in § 5, we proceed as follows. We write (112) over again, providing both the fixed index n and the varying index k with a dash, and replacing the undefined t by s . These two equations are then multiplied together, *i.e.* left side by left side, and right side by right. Then we multiply further by

$$(113) \quad x^p e^{-x}$$

and integrate with respect to x from 0 to ∞ . p is to be a positive integer—this being sufficient for our purpose. The integration is practicable by elementary methods on the right-hand side, and we get

$$(114) \quad \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \frac{t^k s^{k'}}{(n+k)! (n'+k')!} \int_0^{\infty} x^p e^{-x} L_{n+k}^n(x) L_{n'+k'}^{n'}(x) dx \\ = (-1)^{n+n'} p! \frac{(1-t)^{p-n} (1-s)^{p-n'}}{(1-ts)^{p+1}}.$$

We have now, on the left, the desired integrals like pearls on a string, and we merely detach the one we happen to need by searching on the right for the coefficient of $t^k s^{k'}$. This coefficient is always a simple sum, and, in fact, in the cases occurring in the text, always a finite sum with very few terms (up to three). In general, we have

$$(115) \quad \left\{ \int_0^{\infty} x^p e^{-x} L_{n+k}^n(x) L_{n'+k'}^{n'}(x) dx = p! (n+k)! (n'+k')! \right. \\ \left. \cdot \sum_{\tau=0}^{\leq k, k'} (-1)^{n+n'+k+k'+\tau} \binom{p-n}{k-\tau} \binom{p-n'}{k'-\tau} \binom{-p-1}{\tau} \right\}.$$

The sum stops after the smaller of the two numbers k, k' . It often, in actual fact, begins at a positive value of τ , as binomial coefficients, whose lower number is greater than the upper, vanish. For example, in the integral in the denominator of (52), we put $p=n=n'$, and $k'=k$. Then τ can take only the *one* value k , and we can establish statement (53) of the text. In the integral of the numerator in (52), only p has another value, namely $p=n+2$. τ now takes the values $k-2, k-1$, and k , and after an easy reduction we get formula (54) of the text. In the very same way the integrals appearing in § 5 are evaluated by Laguerre polynomials.

We can now, therefore, regard integrals of the type of (115) as known, and we have only to concern ourselves with those occurring in § 4 in the calculation of intensities (*cf.* expression (65) and functions (64) which have to be substituted there). In this type, the two Laguerre orthogonal functions, whose product is to be integrated, *have not the same argument*, but, for example, in our case, have the arguments $\lambda_1/l a_0$ and $\lambda_1/l' a_0$, where l and l' are the principal quantum numbers of the two levels that we have combined. Let us consider, as typical, the integral

$$(116) \quad J = \int_0^{\infty} x^p e^{-\frac{\alpha+\beta}{2}x} L_{n+k}^n(\alpha x) L_{n'+k'}^{n'}(\beta x) dx.$$

Now we can proceed in a superficially different way. At first, the former procedure still goes on smoothly; only on the right-hand side of (114) a somewhat more complicated expression appears. In the denominator occurs the power of a quadrinomial instead of that of a binomial, as before. And this makes the matter somewhat confusing, for the right-hand side of (114) becomes five-fold instead of three-fold, and thus the right side of (115) becomes a three-fold instead of a simple sum. I found that the following substitution made things clearer:

$$(117) \quad \frac{\alpha + \beta}{2} x = y.$$

Hence

$$(118) \quad \begin{cases} \alpha x = \left(1 + \frac{\alpha - \beta}{\alpha + \beta}\right) y \\ \beta x = \left(1 - \frac{\alpha - \beta}{\alpha + \beta}\right) y. \end{cases}$$

After expanding the two polynomials in their Taylor series, which are finite and have similar polynomials as coefficients, we get, using the abbreviations

$$(119) \quad \sigma = \frac{2}{\alpha + \beta}, \quad \gamma = \frac{\alpha - \beta}{\alpha + \beta},$$

the following,

$$(120) \quad J = \sigma^{p+1} \sum_{\lambda=0}^k \sum_{\mu=0}^k (-1)^\mu \frac{\gamma^{\lambda+\mu}}{\lambda! \mu!} \int_0^\infty y^{p+\lambda+\mu} L_n^{\lambda+\mu}(y) L_n^{\lambda+\mu}(y) dy.$$

Thus the calculation of J is reduced to the simpler type of integral (115). In the case of the Balmer lines, the double sum in (120) is comparatively tractable, for one of the two k -values, namely, the one referring to the two-quantum level, never exceeds unity, and thus λ may have two values at most, and, as it turns out, μ four values at most. The circumstance that out of the polynomials referring to the two-quantum level, none but

$$L_0 = 1, \quad L_1 = -x + 1, \quad L_1^1 = -1,$$

appear, permits further simplifications. Nevertheless we must calculate out a number of tables, and it is much to be regretted that the figures given in the tables of the text for the intensities do not allow their general construction to be seen. By good fortune the additive relations between the \parallel - and the \perp components hold good, so that we may, with some *probability*, feel ourselves safe from arithmetical blunders at least.

§ 3. Integrals with Legendre Functions

There are three simple integral relations between associated Legendre functions, which are necessary for the calculations in § 5. For the convenience of others, I will state them here, because I was

not able to discover them in any of the places I searched. We use the customary definition,

$$(121) \quad P_n^m(\cos \theta) = \sin^m \theta \frac{d^m P_n(\cos \theta)}{(d \cos \theta)^m}.$$

Then the following holds,

$$(122) \quad \int_0^\pi [P_n^m(\cos \theta)]^2 \sin \theta d\theta = \frac{2}{2n+1} \frac{(n+m)!}{(n-m)!}$$

(the normalising relation).

Moreover,

$$(123) \quad \left\{ \int_0^\pi P_n^m(\cos \theta) P_{n'}^m(\cos \theta) \cos \theta \sin \theta d\theta = 0 \right.$$

for $|n - n'| \neq 1$.

On the other hand,

$$(124) \quad \int_0^\pi P_n^m(\cos \theta) P_{n-1}^m(\cos \theta) \cos \theta \sin \theta d\theta$$

$$= \frac{n+m}{2n+1} \int_0^\pi [P_{n-1}^m(\cos \theta)]^2 \sin \theta d\theta = \frac{2(n+m)!}{(4n^2-1)(n-m-1)!}.$$

The last two relations decide the "selection" of the determinantal terms on page 95 of the text. They are, moreover, of fundamental importance for the theory of spectra, for it is obvious that the selection principle for the azimuthal quantum number depends on them (and on two others which have $\sin^2 \theta$ in place of $\cos \theta \sin \theta$).

Addition at Proof Correction

Hr. W. Pauli, jun., informs me that he has arrived at the following closed formulae for the total intensity of the lines in the Lyman and Balmer series, through a modification of the method given in section 2 of the Appendix. For the Lyman series these are

$$\nu_{l,1} = R \left(\frac{1}{1^2} - \frac{1}{l^2} \right); \quad J_{l,1} = \frac{2^7 \cdot (l-1)^{2l-1}}{l \cdot (l+1)^{2l+1}};$$

and for the Balmer series

$$\nu_{l,2} = R \left(\frac{1}{2^2} - \frac{1}{l^2} \right); \quad J_{l,2} = \frac{4^3 \cdot (l-2)^{2l-3}}{l \cdot (l+2)^{2l+3}} (3l^2 - 4)(5l^2 - 4).$$

The total emission intensities (square of amplitudes into fourth power of the frequency) are proportional to these expressions, within the series in question. The numbers obtained from the formula for the Balmer series are in complete agreement with those given on pp. 91, 92.

Zürich, Physical Institute of the University.
(Received May 10, 1926.)

Quantisation as a Problem of Proper Values (Part IV¹)

(*Annalen der Physik* (4), vol. 81, 1926)

ABSTRACT: § 1. Elimination of the energy-parameter from the vibration equation. The real wave equation. Non-conservative systems. § 2. Extension of the perturbation theory to perturbations which explicitly contain the time. Theory of dispersion. § 3. Supplementing § 2. Excited atoms, degenerate systems, continuous spectrum. § 4. Discussion of the resonance case. § 5. Generalisation for an arbitrary perturbation. § 6. Relativistic-magnetic generalisation of the fundamental equations. § 7. On the physical significance of the field scalar.

§ 1. Elimination of the Energy-parameter from the Vibration Equation. The Real Wave Equation. Non-conservative Systems

The wave equation (18) or (18'') of Part II., viz.

$$(1) \quad \nabla^2\psi - \frac{2(E-V)}{E^2} \frac{\partial^2\psi}{\partial t^2} = 0$$

or

$$(1') \quad \nabla^2\psi + \frac{8\pi^2}{h^2}(E-V)\psi = 0,$$

which forms the *basis* for the re-establishment of mechanics attempted in this series of papers, suffers from the disadvantage that it expresses the law of variation of the "mechanical field scalar" ψ , neither *uniformly* nor *generally*. Equation (1) contains the energy- or frequency-parameter E , and is valid, as is expressly emphasized in Part II., with a *definite* E -value inserted, for processes which depend on the time exclusively through a *definite* periodic factor:

$$(2) \quad \psi \sim \text{real part of } \left(e^{\pm \frac{2\pi i E t}{h}} \right).$$

Equation (1) is thus not really any more general than equation (1'),

¹ Cf. *Ann. d. Phys.* 79, pp. 361, 489; 80, p. 437, 1926 (Parts I., II., III.); further, on the connection with Heisenberg's theory, *ibid.* 79, p. 734 (p. 45).

which takes account of the circumstance just mentioned and does not contain the time at all.

Thus, when we designated equation (1) or (1'), on various occasions, as "the wave equation", we were really wrong and would have been more correct if we had called it a "vibration-" or an "amplitude-" equation. However, we found it sufficient, because to it is linked the Sturm-Liouville proper value problem—just as in the mathematically strictly analogous problem of the free vibrations of strings and membranes—and not to the *real* wave equation.

As to this, we have always postulated up till now that the potential energy V is a pure function of the co-ordinates and does *not* depend explicitly on the time. There arises, however, an urgent need for the extension of the theory to *non-conservative* systems, because it is only in that way that we can study the behaviour of a system under the influence of prescribed external forces, e.g. a light wave, or a strange atom flying past. Whenever V contains the time explicitly, it is manifestly *impossible* that equation (1) or (1') should be satisfied by a function ψ , the method of dependence of which on the time is as given by (2). We then find that the amplitude equation is no longer sufficient and that we must search for the real wave equation.

For conservative systems, the latter is easily obtained. (2) is equivalent to

$$(3) \quad \frac{\partial^2 \psi}{\partial t^2} = -\frac{4\pi^2 E^2}{h^2} \psi.$$

We can eliminate E from (1') and (3) by differentiation, and obtain the following equation, which is written in a symbolic manner, easy to understand:

$$(4) \quad \left(\nabla^2 - \frac{8\pi^2}{h^2} V\right)^2 \psi + \frac{16\pi^2}{h^2} \frac{\partial^2 \psi}{\partial t^2} = 0.$$

This equation must be satisfied by every ψ which depends on the time as in (2), though *with E arbitrary*, and consequently also by every ψ which can be expanded in a Fourier series with respect to the time (naturally with functions of the co-ordinates as coefficients). Equation (4) is thus evidently *the uniform and general wave equation for the field scalar ψ* .

It is evidently no longer of the simple type arising for vibrating membranes, but is of the *fourth* order, and of a type similar to that occurring in many problems in the theory of elasticity.¹ However, we need not fear any excessive complication of the theory, or any necessity to revise the previous methods, associated with equation (1'). If V does *not* contain the time, we can, proceeding from (4), apply (2), and then split up the operator as follows:

$$(4') \quad \left(\nabla^2 - \frac{8\pi^2}{h^2} V + \frac{8\pi^2}{h^2} E\right) \left(\nabla^2 - \frac{8\pi^2}{h^2} V - \frac{8\pi^2}{h^2} E\right) \psi = 0.$$

¹ E.g., for a vibrating plate, $\nabla^2 \nabla^2 u + \frac{\partial^2 u}{\partial t^2} = 0$. Cf. Courant-Hilbert, chap. v. § 8, p. 256.

By way of trial, we can resolve this equation into two "alternative" equations, namely, into equation (1') and into another, which only differs from (1') in that its proper value parameter will be called minus E , instead of plus E . According to (2) this does not lead to new solutions. The decomposition of (4') is not absolutely cogent, for the theorem that "a product can only vanish when at least one factor vanishes" is not valid for operators. This lack of cogency, however, is a feature common to all the methods of solution of partial differential equations. The procedure finds its subsequent justification in the fact that we can prove the *completeness* of the discovered proper functions, as functions of the co-ordinates. This completeness, coupled with the fact that the imaginary part as well as the real part of (2) satisfies equation (4), allows arbitrary initial conditions to be fulfilled by ψ and $\partial\psi/\partial t$.

Thus we see that the wave equation (4), which contains in itself the law of dispersion, can really stand as the basis of the theory previously developed for conservative systems. The generalisation for the case of a time-varying potential function nevertheless demands caution, because terms with time derivatives of V may then appear, about which no information can be given to us by equation (4), owing to the way we obtained it. In actual fact, if we attempt to apply equation (4) as it stands to non-conservative systems, we meet with complications, which seem to arise from the term in $\partial V/\partial t$. Therefore, in the following discussions, I have taken a somewhat different route, which is much easier for calculations, and which I consider is justified in principle.

We need not raise the order of the wave equation to four, in order to get rid of the energy-parameter. The dependence of ψ on the time, which must exist if (1') is to hold, can be expressed by

$$(3') \quad \frac{\partial\psi}{\partial t} = \pm \frac{2\pi i}{h} E\psi$$

as well as by (3). We thus arrive at one of the two equations

$$(4'') \quad \nabla^2\psi - \frac{8\pi^2}{h^2} V\psi \mp \frac{4\pi i}{h} \frac{\partial\psi}{\partial t} = 0.$$

We will require the complex wave function ψ to satisfy one of these two equations. Since the conjugate complex function $\bar{\psi}$ will then satisfy the other equation, we may take the real part of ψ as the real wave function (if we require it). In the case of a conservative system (4'') is essentially equivalent to (4), as the real operator may be split up into the product of the two conjugate complex operators if V does not contain the time.

§ 2. Extension of the Perturbation Theory to Perturbations containing the Time explicitly. Theory of Dispersion

Our main interest is not in systems for which the time and spatial variations of the potential energy V are of the same order of magnitude,

but in systems, conservative in themselves, which are *perturbed* by the addition of small given functions of the time (and of the co-ordinates) to the potential energy. Let us, therefore, write

$$(5) \quad V = V_0(x) + r(x, t),$$

where, as often before, x represents the whole of the configuration co-ordinates. We regard the unperturbed proper value problem ($r=0$) as *solved*. Then the perturbation problem can be solved by *quadratures*.

However, we will not treat the general problem immediately, but will select the problem of the *dispersion theory* out of the vast number of weighty applications which fall under this heading, on account of its striking importance, which really justifies a separate treatment in any case. Here the perturbing forces originate in an alternating electric field, homogeneous and vibrating synchronously in the domain of the atom; and thus, if we have to do with a linearly polarised monochromatic light of frequency ν , we write

$$(6) \quad r(x, t) = A(x) \cos 2\pi\nu t,$$

and hence

$$(5') \quad V = V_0(x) + A(x) \cos 2\pi\nu t.$$

Here $A(x)$ is the negative product of the light-amplitude and the co-ordinate function which, *according to ordinary mechanics*, signifies the component of the electric moment of the atom in the direction of the electric light-vector (say $-F\sum e_i z_i$, if F is the light-amplitude, e_i, z_i the charges and z -co-ordinates of the particles, and the light is polarised in the z -direction). We borrow the time-variable part of the potential function from ordinary mechanics with just as much or as little right as previously, *e.g.* in the Kepler problem, we borrowed the *constant* part.

Using (5'), equation (4'') becomes

$$(7) \quad \nabla^2 \psi - \frac{8\pi^2}{h^2} (V_0 + A \cos 2\pi\nu t) \psi \mp \frac{4\pi i}{h} \frac{\partial \psi}{\partial t} = 0.$$

For $A=0$, these equations are changed by the substitution

$$(8) \quad \psi = u(x) e^{\pm \frac{2\pi i E t}{h}}$$

(which is now to be taken in the literal sense, and does *not* imply *pars realis*) into the amplitude equation (1') of the unperturbed problem, and we know (cf. § 1) that the totality of the solutions of the unperturbed problem is found in this way. Let

$$E_k \text{ and } u_k(x); \quad k = 1, 2, 3, \dots$$

be the proper values and normalised proper functions of the unperturbed problem, which we regard as *known*, and which we will assume to be *discrete* and *different* from one another (non-degenerate system with no continuous spectrum), so that we may not become involved in secondary questions, requiring special consideration.

Just as in the case of a perturbing potential independent of the

time, we will have to seek solutions of the perturbed problem in the neighbourhood of *each* possible solution of the unperturbed problem, and thus in the neighbourhood of an arbitrary linear combination of the u_k 's, which has constant co-efficients [from (8), the u_k 's to be combined with the appropriate time factors $e^{\pm \frac{2\pi i E_k t}{\hbar}}$]. The solution of the perturbed problem, lying in the neighbourhood of a *definite* linear combination, will have the following physical meaning. It will be *this solution* which first appears, if, when the light wave arrived, precisely that definite linear combination of free proper vibrations was present (perhaps with trifling changes during the "excitation").

Since, however, the equation of the perturbed problem is also *homogeneous*—let this want of analogy with the "forced vibrations" of acoustics be expressly emphasized—it is evidently sufficient to seek the perturbed solution in the neighbourhood of each *separate*

$$(9) \quad u_k(x)e^{\pm \frac{2\pi i E_k t}{\hbar}},$$

as we may then linearly combine these *ad lib.*, just as for unperturbed solutions.

To solve the first of equations (7) we therefore now put

$$(10) \quad \psi = u_k(x)e^{\frac{2\pi i E_k t}{\hbar}} + w(x, t).$$

[The lower symbol, *i.e.* the second of equations (7), is henceforth left on one side, as it would not yield anything new.] The additional term $w(x, t)$ can be regarded as small, and its product with the perturbing potential neglected. Bearing this in mind while substituting from (10) in (7), and remembering that $u_k(x)$ and E_k are proper functions and values of the unperturbed problem, we get

$$(11) \quad \left\{ \begin{aligned} \nabla^2 w - \frac{8\pi^2}{\hbar^2} V_0 w - \frac{4\pi i}{\hbar} \frac{\partial w}{\partial t} &= \frac{8\pi^2}{\hbar^2} A \cos 2\pi \nu t \cdot u_k e^{\frac{2\pi i E_k t}{\hbar}}, \\ &= \frac{4\pi^2}{\hbar^2} A u_k \cdot \left(e^{\frac{2\pi i t}{\hbar}(E_k + h\nu)} + e^{\frac{2\pi i t}{\hbar}(E_k - h\nu)} \right). \end{aligned} \right.$$

This equation is readily, and really *only*, satisfied by the substitution

$$(12) \quad w = w_+(x)e^{\frac{2\pi i t}{\hbar}(E_k + h\nu)} + w_-(x)e^{\frac{2\pi i t}{\hbar}(E_k - h\nu)},$$

where the two functions w_{\pm} respectively obey the two equations

$$(13) \quad \nabla^2 w_{\pm} + \frac{8\pi^2}{\hbar^2}(E_k \pm h\nu - V_0)w_{\pm} = \frac{4\pi^2}{\hbar^2} A u_k.$$

This step is essentially *unique*. At first sight, we apparently can add to (12) an arbitrary aggregate of unperturbed proper vibrations. But this aggregate would necessarily be assumed small, of the first order (since this has been assumed for w), and thus does not interest us at present, as it could only produce perturbations of the second order at most.

In equations (13) we have at last those *non-homogeneous* equations we might have expected to encounter—in spite of the lack of analogy with real forced vibrations, as emphasized above. This lack of analogy is extraordinarily important and manifests itself in equations (13) in the two following particulars. *Firstly*, as the “second member” (“exciting force”), the perturbation function $A(x)$ does not appear *alone*, but *multiplied* by the amplitude of the free vibration already present. This is indispensable if the physical facts are to be properly taken into account, for the reaction of an atom to an incident light wave depends almost entirely on the *state* of the atom at that time, whereas the forced vibrations of a membrane, plate, etc., are known to be quite independent of the proper vibrations which may be superimposed on them, and thus would yield an obviously wrong representation of our case. *Secondly*, in place of the proper value on the left-hand side of (13), *i.e.* as “exciting frequency”, we do not find the frequency ν of the perturbing force *alone*, but rather in one case added to, and in the other subtracted from, that of the free vibration already present. This is equally indispensable. Otherwise the proper frequencies *themselves*, which correspond to the *term-frequencies*, would function as *resonance-points*, and not the *differences* of the proper frequencies, as is demanded, and is really given by equation (13). Moreover, we see with satisfaction that the latter gives *only* the differences between a proper frequency *which is actually excited* and all the others, and *not* the differences between pairs of proper frequencies, of which *no member* is excited.

In order to investigate this more closely, let us complete the solution. By well-known methods¹ we find, as *simple* solutions of (13),

$$(14) \quad w_{\pm}(x) = \frac{1}{2} \sum_{n=1}^{\infty} \frac{a'_{kn} u_n(x)}{E_k - E_n \pm h\nu},$$

where

$$(15) \quad a'_{kn} = \int A(x) u_k(x) u_n(x) \rho(x) dx.$$

$\rho(x)$ is the “density function”, *i.e.* that function of the position-co-ordinates with which equation (1') must be multiplied to make it self-adjoint. The $u_n(x)$'s are assumed to be normalised. It is further postulated that $h\nu$ does not agree exactly with any of the differences $E_k - E_n$ of the proper values. This “resonance case” will be dealt with later (cf. § 4).

If we now form from (14), using (12) and (10), the entire perturbed vibration, we get

$$(16) \quad \left\{ \psi = u_k(x) e^{\frac{2\pi i E_k t}{h}} + \frac{1}{2} \sum_{n=1}^{\infty} a'_{kn} u_n(x) \left(\frac{e^{\frac{2\pi i t}{h} (E_k + h\nu)}}{E_k - E_n + h\nu} + \frac{e^{\frac{2\pi i t}{h} (E_k - h\nu)}}{E_k - E_n - h\nu} \right) \right\}.$$

Thus in the perturbed case, along with each *free* vibration $u_k(x)$ occur in small amplitude all those vibrations $u_n(x)$, for which $a'_{kn} \neq 0$.

¹ Cf. Part III. §§ 1 and 2, text beside equations (8) and (24).

The latter are exactly those, which, if they exist as free vibrations along with u_k , give rise to a radiation, which is (wholly or partially) polarised in the direction of polarisation of the incident wave. For apart from a factor, a'_{kn} is just the component amplitude, in this direction of polarisation, of the atom's *electric moment*, which is oscillating with frequency $(E_k - E_n)/h$, according to wave mechanics, and which appears when u_k and u_n exist together.¹ The simultaneous oscillation, however, takes place with neither the proper frequency E_n/h , peculiar to these vibrations, nor the frequency ν of the light wave, but rather with the sum and difference of ν and E_k/h (i.e. the frequency of the *one existing free vibration*).

The real or the imaginary part of (16) can be considered as the *real* solution. In the following, however, we will operate with the complex solution itself.

To see the significance that our result has in the theory of dispersion, we must examine the radiation arising from the simultaneous existence of the excited forced vibrations and the free vibration, already present. For this purpose, we form, following the method we² have always adopted above—a criticism follows in § 7—the product of the complex wave function (16) and its conjugate, i.e. the norm of the complex wave function ψ . We notice that the perturbing terms are small, so that squares and products may be neglected. After a simple reduction³ we obtain

$$(17) \quad \psi\bar{\psi} = u_k(x)^2 + 2 \cos 2\pi\nu t \sum_{n=1}^{\infty} \frac{(E_k - E_n)a'_{kn}u_k(x)u_n(x)}{(E_k - E_n)^2 - \hbar^2\nu^2}.$$

According to the *heuristic hypothesis* on the electro-dynamical significance of the field scalar ψ , the present quantity—apart from a multiplicative constant—represents the electrical density as a function of the space co-ordinates and the time, if x stands for only three space co-ordinates, i.e. if we are dealing with the problem of *one* electron. We remember that the same hypothesis led us to correct selection and polarisation rules and to a very satisfactory representation of intensity relationships in our discussion of the hydrogen Stark effect. By a natural generalisation of this hypothesis—of which more in § 7—we regard the following as representing in the general case the density of the electricity, which is “associated” with *one* of the particles of classical mechanics, or which “originates in it”, or which “corresponds to it in wave mechanics”: the *integral* of $\psi\bar{\psi}$ taken over all those co-ordinates of the system, which in classical mechanics fix the

¹ Cf. what follows, and § 7.

² Cf. end of paper on Quantum Mechanics of Heisenberg, etc., and also the Calculation of Intensities in the Stark Effect in Part III. At the first quoted place, the real part of $\psi\bar{\psi}$ was proposed instead of $\psi\bar{\psi}$. This was a mistake, which was corrected in Part III.

³ We assume as previously, for the sake of simplicity, the proper functions $u_n(x)$ to be *real*, but notice that it may sometimes be much more convenient or even imperative to work with complex aggregates of the real proper functions, e.g. in the proper functions of the Kepler problem to work with $e^{\pm m\phi i}$ instead of $\frac{\cos m\phi}{\sin m\phi}$.

position of the *rest of the* particles, multiplied by a certain constant, the classical "charge" of the first particle. The resultant density of charge at any point of space is then represented by the *sum* of such integrals taken over all the particles.

Thus in order to find any space component whatever of the total wave-mechanical dipole moment as a function of the time, we must, on this hypothesis, multiply expression (17) by that function of the co-ordinates which gives that particular dipole-component in *classical mechanics* as a function of the configuration of the point system, e.g. by

$$(18) \quad M_y = \sum e_i y_i,$$

if we are dealing with the dipole moment in the y -direction. Then we have to integrate over *all* the configuration co-ordinates.

Let us work this out, using the abbreviation

$$(19) \quad b_{kn} = \int M_y(x) u_k(x) u_n(x) \rho(x) dx.$$

Let us elucidate further the definition (15) of the a'_{kn} 's by recalling that if the incident electric light-vector is given by

$$(20) \quad \mathcal{E}_z = F \cos 2\pi\nu t,$$

then

$$(21) \quad \begin{cases} A(x) = -F \cdot M_z(x), \\ \text{where } M_z(x) = \sum e_i z_i. \end{cases}$$

If we put, in analogy with (19),

$$(22) \quad a_{kn} = \int M_z(x) u_k(x) u_n(x) \rho(x) dx,$$

then $a'_{kn} = -F a_{kn}$, and by carrying out the proposed integration we find,

$$(23) \quad \int M_y \psi \bar{\psi} \rho dx = a_{kk} + 2F \cos 2\pi\nu t \sum_{n=1}^{\infty} \frac{(E_n - E_k) a_{kn} b_{kn}}{(E_k - E_n)^2 - \hbar^2 \nu^2}$$

for the resulting electric moment, to which the secondary radiation, caused by the incident wave (20), is to be attributed.

The radiation depends of course only upon the second (time-variable) part, while the first part represents the time-constant dipole moment, which is possibly connected with the originally existing free vibration. This variable part seems fairly promising and may meet all the demands we are accustomed to make on a "dispersion formula". Above all, let us note the appearance of those so-called "negative" terms, which—in the usual phraseology—correspond to the probability of transition to a lower level ($E_n < E_k$), and to which Kramers¹ was the first to direct attention, from a correspondence

¹ H. A. Kramers, *Nature*, May 10, 1924; *ibid.* August 30, 1924; Kramers and W. Heisenberg, *Ztschr. f. Phys.* 31, p. 681, 1925. The description given in the latter paper of the polarisation of the scattered light (equation 27) from correspondence principles, is almost identical *formally* with ours.

standpoint. Generally, our formula—despite very different ways of thought and expression—may be characterised as really identical in form with Kramer's formula for secondary radiation. The important connection between a_{kn} , b_{kn} , the coefficients of the secondary and of the spontaneous radiation, is brought out, and indeed the secondary radiation is also described accurately with respect to its condition of polarisation.¹

I would like to believe that the absolute value of the scattered radiation or of the induced dipole moment is also given correctly by formula (23), although it is obviously within the bounds of possibility that an error in the numerical factor may have occurred in applying the heuristic hypothesis introduced above. At any rate the physical dimensions are right, for from (18), (19), (21), and (22) a_{kn} and b_{kn} are electric moments, since the squared integrals of the proper functions were normalised to unity. If ν is far removed from the emission frequency in question, the ratio of the induced to the spontaneous dipole moment is of the same order of magnitude as the ratio of the additional potential energy $F a_{kn}$ to the "energy step" $E_k - E_n$.

§ 3. Supplements to § 2. Excited Atoms, Degenerate Systems, Continuous Spectrum

For the sake of clearness, we have made some special assumptions, and put many questions aside, in the preceding paragraph. These have now to be discussed by way of supplement.

First: what happens when the light wave meets the atom, when the latter is in a state in which not merely *one* free vibration, u_k , is excited as hitherto assumed, but several, say two, u_k and u_l ? As remarked above, we have in the perturbed case simply to combine additively the two perturbed solutions (16) corresponding to the suffix k and the suffix l , after we have provided them with constant (possibly complex) coefficients, which correspond to the *strength* presumed for the free vibrations, and to the phase relationship of their stimulation. Without actually performing the calculation, we see that in the expression for $\psi\bar{\psi}$ and also in the expression (23) for the resulting electric moment, there then occurs *not merely* the corresponding linear aggregate of the terms previously obtained, *i.e.* of the expressions (17) or (23) written with k , and then with l . We have in addition "combination terms", namely, considering *first* the greatest order of magnitude, a term in

$$(24) \quad u_k(x)u_l(x)e^{\frac{2\pi i}{h}(E_k - E_l)t},$$

which gives again the *spontaneous* radiation, bound up with the co-

¹ It is hardly necessary to say that the two directions which, for simplicity, we have designated as "z-direction" and "y-direction" do not require to be exactly perpendicular to one another. The one is the direction of polarisation of the incident wave; the other is that polarisation component of the secondary wave, in which we are specially interested.

existence of the two *free* vibrations ; and *secondly* perturbing terms of the first order, which are proportional to the perturbing field amplitude, and which correspond to the interaction of the forced vibrations belonging to u_k with the free vibration u_l —and of the forced vibrations belonging to u_l with u_k . The *frequency* of these new terms appearing in (17) or (23) is *not* ν but

$$(25) \quad |\nu \pm (E_k - E_l)/h|,$$

as can easily be seen, still without carrying out the calculation. (New “resonance denominators”, however, do *not* occur in these terms.) Thus we have to do here with a secondary radiation, whose frequency neither coincides with the exciting light-frequency nor with a spontaneous frequency of the system, but is a combination frequency of both.

The existence of this remarkable kind of secondary radiation was first postulated by Kramers and Heisenberg (*loc. cit.*), from correspondence considerations, and then by Born, Heisenberg, and Jordan from consideration of Heisenberg’s quantum mechanics.¹ As far as I know, it has not yet been demonstrated experimentally. The *present* theory also shows distinctly that the occurrence of this scattered radiation is dependent on special conditions, which demand researches expressly arranged for the purpose. Firstly, *two* proper vibrations u_k and u_l must be *strongly* excited, so that all experiments made on atoms in their normal state—as happens in the vast majority of cases—are to be rejected. Secondly, at least *one* third state of proper vibration must *exist* (*i.e.* must be possible—it need not be *excited*), which leads to powerful spontaneous emission, when combined with u_k as well as with u_l . For the extraordinary scattered radiation, which is to be discovered, is proportional to the product of the spontaneous emission coefficients in question ($a_{kn}b_{ln}$ and $a_{ln}b_{kn}$). The combination (u_k, u_l) need not, in itself, cause a strong emission. It would not matter if—to use the language of the older theory—this was a “forbidden transition”. Yet in practice we must also demand that the line (u_k, u_l) should actually be emitted strongly during the experiment, for this is the only means of assuring ourselves that *both* proper vibrations are strongly excited in the same individual atoms and in a sufficiently great number of them. If we reflect now that in the powerful term-series mostly examined, *i.e.* in the ordinary *s*-, *p*-, *d*-, *f*-series, the relations are generally such that two terms, which combine strongly with a third, do not do so with one another, then a special choice of the object and conditions of the research seems really necessary, if we are to expect the desired scattered radiation with any certainty, especially as its frequency is not that of the exciting light and *thus* it does not produce dispersion or rotation of the plane of polarisation, but can only be observed as light scattered on all sides.

As far as I see, the above-mentioned dispersion theory of Heisenberg, Born, and Jordan does *not* allow of such reflections as we

¹ Born, Heisenberg, and Jordan, *Ztschr. f. Phys.* 35, p. 572, 1926.

have just made, in spite of its great formal similarity to the present one. For it only considers *one* way in which the atom reacts to incident radiation. It conceives the atom as a timeless entity, and up till now is not able to express in its language the undoubted fact that the atom can be in *different* states at different times, and thus, as has been proved, reacts in different ways to incident radiation.¹

Let us turn now to another question. In § 2 the collective proper values were postulated to be *discrete* and *different* from one another. We now drop the second hypothesis and ask: what is altered when *multiple* proper values occur, *i.e.* when *degeneracy* is present? Perhaps we expect that complications then arise, similar to those we met in the case of a time-constant perturbation (Part III. § 2), *i.e.* that a system of proper functions of the unperturbed atom, suited to the particular perturbation, must be defined by the solution of a "secular equation", and applied to carry out the perturbation calculation. This is indeed so in the case of an *arbitrary* perturbation, represented by $r(x, t)$ as in equation (5), but *not* so in the case of a perturbation by a light wave (equation (6))—at any rate, for our usual first approximation, and as long as we suppose that the light frequency ν does not coincide with any of the spontaneous emission frequencies considered. Then the parameter value in the double equation (13), for the amplitudes of the perturbed vibrations, is *not* a proper value, and the pair of equations has always the unambiguous pair of solutions (14), in which no vanishing denominators occur even when E_k is a multiple value. Thus the terms in the sum for which $E_n = E_k$ are *not*, as might be thought, to be omitted, any more than the term for $n = k$ itself. It is worth noticing that through these terms—if one of them occurs really, *i.e.* with non-vanishing a_{kn} —the frequency $\nu = 0$ also appears among the resonance frequencies. These terms do *not*, of course, contribute to the "ordinary" scattered radiation, as we see from (23), since $E_k - E_n = 0$.

The simplification, that we do not require to consider specially any possible degeneracy present, at least in a first approximation, is always available² when the time-averaged value of the perturbation function vanishes, or what is the same thing, when the latter's Fourier expansion in terms of the time contains no constant, *i.e.* time-independent, term. This is the case for a light wave.

While our *first* postulation about the proper values—that they should be *simple*—has thus shown itself to be really a superfluous precaution, a dropping of the *second*—that they should be absolutely discrete—while leading to no alterations in *principle*, brings about, however, very considerable alterations in the external appearance of the calculation, inasmuch as *integrals* taken over the continuous spectrum of equation (1') are to be added to the discrete sums in (14), (16), (17), and (23). The theory of such representations by integrals has been

¹ Cf. especially the concluding words of Heisenberg's latest exposition of his theory, *Math. Ann.* 95, p. 683, 1926, in connection with this difficulty of comprehending the course of an event in time.

² Further discussed in § 5.

developed by H. Weyl,¹ and though only for ordinary differential equations, the extension to partials is permissible. In all brevity, the state of the case is this.² If the homogeneous equation belonging to the non-homogeneous equations (13), *i.e.* the vibration equation (1') of the unperturbed system, possesses in addition to a point-spectrum a continuous one, which stretches, say, from $E=a$ to $E=b$, then an arbitrary function $f(x)$ naturally cannot be developed thus,

$$(26) \quad f(x) = \sum_{n=1}^{\infty} \phi_n \cdot u_n(x), \quad \text{where } \phi_n = \int f(x)u_n(x)\rho(x)dx$$

in terms of the normalised discrete proper functions $u_n(x)$ alone, but there must be added an integral expansion in terms of the proper solutions $u(x, E)$, which belong to the proper values $a \leq E \leq b$, and so we have

$$(27) \quad f(x) = \sum_{n=1}^{\infty} \phi_n \cdot u_n(x) + \int_a^b u(x, E)\phi(E)dE,$$

where to emphasize the analogy we have intentionally chosen the same letter for the "coefficient function" $\phi(E)$ as for the discrete coefficients ϕ_n . If now we have *normalised*, once for all, the proper solution $u(x, E)$ by associating with it a suitable function of E , in such a way that

$$(28) \quad \int dx\rho(x) \int_E^{E'+\Delta} u(x, E)u(x, E')dE' = 1 \text{ or } = 0$$

according to whether E belongs to the interval $E', E' + \Delta$ or not, then in (27) under the integral sign we substitute from

$$(29) \quad \phi(E) = \lim_{\Delta=0} \frac{1}{\Delta} \int \rho(\xi)f(\xi) \cdot \int_E^{E+\Delta} u(\xi, E')dE' \cdot d\xi,$$

wherein the *first* integral sign refers as always to the domain of the group of variables x .³ Assuming (28) to be fulfilled and expansion (27) to exist—which statements are proved by Weyl for ordinary differential equations—the definition of the "coefficient functions" from (29) is almost as obvious as the well-known definition of the Fourier coefficients.

The most important and difficult task in any concrete case is the carrying out of the normalisation of $u(x, E)$, *i.e.* the finding of that function of E by which we have to multiply the (as yet not normalised) proper solution of the continuous spectrum, in order that condition (28) may be satisfied. The above-quoted works of Herr Weyl contain very valuable guidance for this practical task, and also

¹ H. Weyl, *Math. Ann.* 68, p. 220, 1910; *Gött. Nachr.* 1910. Cf. also E. Hilb, *Sitz.-Ber. d. Physik. Mediz. Soc. Erlangen*, 43, p. 68, 1911; *Math. Ann.* 71, p. 76, 1911. I have to thank Herr Weyl not only for these references but also for very valuable oral instruction in these not very simple matters.

² I have to thank Herr Fues for this exposition.

³ As Herr E. Fues informs me, we can very often omit the limiting process in practice and write $u(\xi, E)$ for the inner integral, *viz.* always, when $\int \rho(\xi)f(\xi)u(\xi, E)d\xi$ exists.

some worked-out examples. An example from atomic dynamics on the intensities of band spectra is worked out by Herr Fues in a paper appearing in the present issue of *Annalen der Physik*.

Let us apply this to our problem, *i.e.* to the solution of the pair of equations (13) for the amplitudes w_{\pm} of the perturbed vibrations, where we postulate as usual that the *one* excited *free* vibration, u_k , belongs to the discrete point-spectrum. We develop the right-hand side of (13) according to the scheme (27) thus,

$$(30) \quad \frac{4\pi^2}{\hbar^2} A(x) u_k(x) = \frac{4\pi^2}{\hbar^2} \sum_{n=1}^{\infty} a'_{kn} u_n(x) + \frac{4\pi^2}{\hbar^2} \int_a^b u(x, E) a'_k(E) dE,$$

in which a'_{kn} is given by (15), and $a'_k(E)$ from (29) by

$$(15') \quad a'_k(E) = \lim_{\Delta=0} \frac{1}{\Delta} \int \rho(\xi) A(\xi) u_k(\xi) \cdot \int_E^{E+\Delta} u(\xi, E') dE' \cdot d\xi.$$

If we imagine expansion (30) put into (13), and then expand also the desired solution $w_{\pm}(x)$ similarly in terms of the proper solutions $u_n(x)$ and $u(x, E)$, and notice that for the last-named functions the left side of (13) takes the value

$$\frac{8\pi^2}{\hbar^2} (E_k \pm h\nu - E_n) u_n(x)$$

or

$$\frac{8\pi^2}{\hbar^2} (E_k \pm h\nu - E) u(x, E),$$

then by "comparison of coefficients" we obtain as the generalisation of (14)

$$(14') \quad w_{\pm}(x) = \frac{1}{2} \sum_{n=1}^{\infty} \frac{a'_{kn} u_n(x)}{E_k - E_n \pm h\nu} + \frac{1}{2} \int_a^b \frac{a'_k(E) u(x, E)}{E_k - E \pm h\nu} dE.$$

The further procedure is completely analogous to that of § 2.

Finally, we get as *additional term* for (23)

$$(23') \quad + 2 \cos 2\pi\nu t \int d\xi \rho(\xi) M_{\nu}(\xi) u_k(\xi) \int_a^b \frac{(E_k - E) a'_k(E) u(\xi, E)}{(E_k - E)^2 - \hbar^2 \nu^2} dE.$$

Here, perhaps, we may not always change the order of integration without further examination, because the integral with respect to ξ may possibly not converge. However, we can—as an intuitive makeshift for a strict passage to the limit, which may be dispensed with here—decompose

the integral \int_a^b into many small parts, each having a range Δ , which is

sufficiently small to allow us to regard all the functions of E in question as constant in each part, with the exception of $u(x, E)$, for we know from the general theory that its integral cannot be obtained through such a fixed partition, which is independent of ξ . We can then take the remaining functions out of the partial integrals, and as *additional term for the dipole moment* (23) of the secondary radiation, obtain finally exactly the following,

$$(23'') \quad 2F \cos 2\pi\nu t \int_a^b \frac{(E - E_k)\alpha_k(E)\beta_k(E)}{(E_k - E)^2 - \hbar^2\nu^2} dE,$$

where

$$(22') \quad \alpha_k(E) = \lim_{\Delta=0} \frac{1}{\Delta} \int \rho(\xi) M_z(\xi) u_k(\xi) \cdot \int_E^{E+\Delta} u(\xi, E') dE' \cdot d\xi,$$

$$(19') \quad \beta_k(E) = \lim_{\Delta=0} \frac{1}{\Delta} \int \rho(\xi) M_\nu(\xi) u_k(\xi) \cdot \int_E^{E+\Delta} u(\xi, E') dE' \cdot d\xi$$

(please note the complete analogy with the formulae with the same numbers but without the dashes in § 2).

The preceding sketch of the calculation is of course only a general outline, given merely to show that the much-discussed influence of the continuous spectrum on dispersion, which experiment¹ appears to indicate as existing, is required by the present theory exactly in the form expected, and to outline the way in which the calculation of the problem is to be tackled.

§ 4. Discussion of the Resonance Case

Up till now we have always assumed that the frequency ν of the light wave does not agree with any of the emission frequencies that have to be considered. We now assume that, say,

$$(31) \quad h\nu = E_n - E_k > 0,$$

and we revert, moreover, to the limiting conditions of § 2 for the sake of simplicity (simple, discrete proper values, one single free vibration u_k excited). In the pair of equations (13), the proper value parameter then takes the values

$$(32) \quad E_k \pm E_n \mp E_k = \begin{cases} E_n \\ 2E_k - E_n, \end{cases}$$

i.e. for the upper sign there appears a *proper value*, namely, E_n . The two cases are possible. Firstly, the right side of equation (13) multiplied by $\rho(x)$, may be *orthogonal* to the proper function $u_n(x)$ corresponding to E_n , *i.e.* we have

$$(33) \quad \int A(x) u_k(x) u_n(x) \rho(x) dx = a'_{kn} = 0,$$

which means, physically, that if u_k and u_n exist together as free vibrations they will give rise to no spontaneous emission or to one which is polarised perpendicularly to the direction of polarisation of the incident light. In this case the critical equation (13) also again possesses a solution, which now, as before, is given by (14), in which the catastrophic term vanishes. This means physically—in the old phraseology—that a “forbidden transition” cannot be stimulated through resonance, or that a “transition”, even if not forbidden,

¹ K. F. Herzfeld and K. L. Wolf, *Ann. d. Phys.* 76, p. 71, 567, 1925; H. Kollmann and H. Mark, *Die Nw.* 14, p. 648, 1926.

cannot be caused by light which is vibrating perpendicularly to the direction of polarisation of that light which would be emitted by the "spontaneous transition".

Otherwise, secondly, (33) is *not* fulfilled. Then the critical equation possesses *no* solution. Statement (10), which assumes a vibration which differs very *little*—by quantities of the order of the light amplitude F —from the originally existing free vibration, and is the *most general possible under this assumption, thus does not then lead to the goal*. No solution, therefore, exists which only differs by quantities of the order of F from the original free vibration. The incident light has thus a *varying influence* on the state of the system, *which bears no relation to the magnitude of the light amplitude*. What influence? We can judge this, still without further calculation, if we start out from the case where the resonance condition (31) is not exactly but only approximately fulfilled. Then we see from (16) that $u_n(x)$ is excited in unusually strong forced vibrations, on account of the small denominator, and that—not less important—the frequency of these forced vibrations approaches the natural proper frequency E_n/h of the proper vibration u_n . (All this is, indeed, very *similar* to, yet in a way of its own *different* from, the resonance phenomena encountered elsewhere; otherwise I would not discuss it so minutely.)

In a gradual approach to the critical frequency, the proper vibration u_n , formerly not excited, whose possible existence is responsible for the crisis, is stimulated to a stronger and stronger degree, and with a frequency more and more closely approaching its own proper frequency. In contradistinction to ordinary resonance phenomena there comes a point, and that even before the critical frequency is reached, where our solution does not represent the circumstances correctly any longer, even under the assumption that our obviously "undamped" wave postulation is strictly correct. For we have in fact regarded the forced vibration w as small compared with the existing free vibration and neglected a squared term (in equation (11)).

I believe that the present discussion has already shown, with sufficient clearness, that in the resonance case the theory will actually give the result it ought to give, in order to agree with Wood's resonance phenomenon: an increase of the proper vibration u_n , which causes the crisis, to a finite magnitude comparable with that of the originally existing u_k , from which, of course, "spontaneous emission" of the spectral line (u_k, u_n) results. I do not wish, however, to attempt to work out the calculation of the resonance case fully here, because the result would be of little value, so long as the *reaction* of the emitted radiation on the emitting system is not taken into account. Such a reaction must exist, not only because there is, no ground at all for differentiating on principle between the light wave which is incident from outside, and that which is emitted by the system itself, but also because otherwise, if several proper

vibrations were simultaneously excited in a system left to itself, the spontaneous emission would continue indefinitely. This required back-coupling must act so that in this case, along with the light emission, the higher proper vibrations gradually die down, and, finally, the fundamental vibration, corresponding to the normal state of the system, alone remains. The back-coupling is evidently exactly analogous to the reaction of radiation $\left(\frac{2e^2}{3mc^3}\ddot{v}\right)$ in the classical electron theory.

This analogy also allays the increasing apprehension caused by the previous neglect of this back-coupling. The influence of the relevant term (probably no longer linear) in the wave equation will generally be small, just as in the electron the back pressure of radiation is generally very small compared with the force of inertia and the external field strength. In the resonance case, however—just as in the electron theory—the coupling with the proper light wave will be of the same order as that with the incident wave, and must be taken into account, if the “equilibrium” between the different proper vibrations, which sets in for the given irradiation, is to be correctly computed.

Let it be expressly remarked, however, that the back-coupling term is *not necessary for averting a resonance catastrophe!* Such can never occur in any circumstances, because according to the theorem of the *persistence of normalisation*, proved below in § 7, the configuration space integral of $\psi\bar{\psi}$ always remains normalised to the same value, even under the influence of arbitrary external forces—and indeed quite automatically, as a consequence of the wave equation (4"). The amplitudes of the ψ -vibrations, therefore, cannot grow indefinitely; they have, “on the average”, always the same value. If *one* proper vibration waxes, then another must, therefore, wane.

§ 5. Generalisation for an Arbitrary Perturbation

If an *arbitrary* perturbation is in question as was assumed in equation (5) at the beginning of § 2, then we shall expand the perturbation energy $r(x, t)$ as a Fourier series or Fourier integral in terms of the time. The terms of this expansion have, then, the form (6) of the perturbation potential of a light wave. We see immediately that on the right-hand side of equation (11) we then simply get two *series* (or, possibly, integrals) of imaginary powers of e , instead of merely two terms. If none of the exciting frequencies coincide with a critical frequency, we get the solution in exactly the same way as described in § 2, but, naturally, as Fourier series (or possibly Fourier integrals) of the time. It serves no purpose to write down the formal expansions here, and a more exact working out of separate problems lies outside the scope of the present paper. Yet an important point, already touched upon in § 3, must be mentioned.

Among the critical frequencies of equation (13), the frequency $\nu=0$, from $E_k - E_k=0$, also generally figures. For in this case also one proper value, namely, E_k , appears on the left side as proper value

parameter. Thus, if the frequency 0, *i.e.* a term independent of the time, occurs in the Fourier expansion of the perturbation function $r(x, t)$, we cannot reach our goal by exactly the earlier method. We easily see, however, how it must be modified, for the case of a time-constant perturbation is known from previous work (cf. Part III.). We have then to consider, at the same time, a small alteration and possibly a splitting up of the proper value or values of the excited free vibrations, *i.e.* in the indices of the powers of e in the first term on the right hand of equation (10) we have to replace E_k by E_k plus a small constant, the perturbation of the proper value. Exactly as described in Part III., § 1 and § 2, this perturbation is defined by the postulation that the right side of the critical Fourier component of our equation (13) is to be orthogonal to u_k (or possibly to *all* the proper functions belonging to E_k).

The number of special problems, which fall under the question formulated in the present paragraph, is extraordinarily great. By superposing the perturbations due to a constant electric or magnetic field and a light wave, we obtain magnetic and electric double refraction, and magnetic rotation of the plane of polarisation. Resonance radiation in a magnetic field also comes under this heading, but for this purpose we must first obtain an exact solution for the resonance case discussed in § 4. Further, we can treat the action of an α -particle or electron flying past the atom¹ in this way, if the encounter is not too close for the perturbation of each of the two systems to be calculable from the undisturbed motion of the other. All these questions are mere matters of calculation as soon as the proper values and functions of the unperturbed systems are known. It is, therefore, to be hoped that we will succeed in defining these functions, at least approximately, for heavier atoms also, in analogy with the approximate definition of the Bohr electronic orbits which belong to different types of terms.

§ 6. Relativistic-magnetic Generalisation of the Fundamental Equations

As an appendix to the physical problems just mentioned, in which the magnetic field, which has hitherto been completely ignored in this series of papers, plays an important part, I would like to give, briefly, the probable relativistic-magnetic generalisation of the basic equations (4''), although I can only do this meantime for the one electron problem, and only with the greatest possible reserve—the latter for two reasons. *Firstly*, the generalisation is provisionally based on a purely formal analogy. *Secondly*, as was mentioned in Part I., though it does formally lead in the Kepler problem to Sommerfeld's fine-structure formula with, in fact, the "half-integral" azimuthal and radial quantum, which is generally regarded as correct to-day,

¹ A very interesting and successful attempt to compare the action of flying charged particles with the action of light waves, through a Fourier decomposition of their field, is to be found in a paper by E. Fermi, *Ztschr. f. Phys.* 29, p. 313, 1924.

nevertheless there is *still lacking* the *supplement*, which is necessary to secure numerically correct diagrams of the splitting up of the hydrogen lines, and which is given in Bohr's theory by Goudsmit and Uhlenbeck's electronic spin.

The Hamilton - Jacobi partial differential equation for the Lorentzian electron can readily be written :

$$(34) \quad \left\{ \begin{aligned} & \left(\frac{1}{c} \frac{\partial W}{\partial t} + \frac{e}{c} V \right)^2 - \left(\frac{\partial W}{\partial x} - \frac{e}{c} \mathfrak{A}_x \right)^2 - \left(\frac{\partial W}{\partial y} - \frac{e}{c} \mathfrak{A}_y \right)^2 \\ & - \left(\frac{\partial W}{\partial z} - \frac{e}{c} \mathfrak{A}_z \right)^2 - m^2 c^2 = 0. \end{aligned} \right.$$

Here e , m , c are the charge and mass of the electron, and the velocity of light ; V , \mathfrak{A} are the electro-magnetic potentials of the external electro-magnetic field at the position of the electron, and W is the action function.

From the classical (relativistic) equation (34) I am now attempting to derive the *wave equation* for the electron, by the following *purely formal* procedure, which, we can verify easily, will lead to equations (4"), if it is applied to the Hamiltonian equation of a particle moving in an arbitrary field of force in ordinary (non-relativistic) mechanics. *After* the squaring, in equation (34), I replace the *quantities*

$$(35) \quad \left\{ \begin{aligned} & \frac{\partial W}{\partial t}, \quad \frac{\partial W}{x \partial x}, \quad \frac{\partial W}{\partial y}, \quad \frac{\partial W}{\partial z}, \\ & \text{by the respective operators} \\ & \pm \frac{h}{2\pi i} \frac{\partial}{\partial t}, \quad \pm \frac{h}{2\pi i} \frac{\partial}{\partial x}, \quad \pm \frac{h}{2\pi i} \frac{\partial}{\partial y}, \quad \pm \frac{h}{2\pi i} \frac{\partial}{\partial z}. \end{aligned} \right.$$

The double linear operator, so obtained, is applied to a wave function ψ and the result put equal to zero, thus :

$$(36) \quad \nabla^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} \mp \frac{4\pi i e}{hc} \left(\frac{V}{c} \frac{\partial \psi}{\partial t} + \mathfrak{A} \text{ grad } \psi \right) + \frac{4\pi^2 e^2}{h^2 c^2} \left(V^2 - \mathfrak{A}^2 - \frac{m^2 c^4}{e^2} \right) \psi = 0.$$

(The symbols ∇^2 and grad have here their elementary three-dimensional Euclidean meaning.) The pair of equations (36) would be the possible relativistic-magnetic generalisation of (4") for the case of a single electron, and should likewise be understood to mean that the complex wave function has to satisfy either the one or the other equation.

From (36) the fine structure formula of Sommerfeld for the hydrogen atom may be obtained by exactly the same method as is described in Part I., and also we may derive (neglecting the term in \mathfrak{A}^2) the normal Zeeman effect as well as the well-known selection and polarisation rules and intensity formulae. They follow from the integral relations between Legendre functions introduced at the end of Part III.

For the reasons given in the first section of this paragraph, I

withhold the detailed reproduction of these calculations meantime, and also in the following final paragraph refer to the "classical", and not to the still incomplete relativistic-magnetic version of the theory.

§ 7. On the Physical Significance of the Field Scalar

The heuristic hypothesis of the electro-dynamical meaning of the field scalar ψ , previously employed in the *one*-electron problem, was extended off-hand to an arbitrary system of charged particles in § 2, and there a more exhaustive description of the procedure was promised. We had calculated the density of electricity at an arbitrary point in space as follows. We selected *one* particle, kept the trio of co-ordinates that describes *its* position in ordinary mechanics fixed; integrated $\psi\bar{\psi}$ over all the rest of the co-ordinates of the system and multiplied the result by a certain constant, the "charge" of the selected particle; we did a similar thing for each particle (trio of co-ordinates), in each case giving the selected particle the same position, namely, the position of that point of *space* at which we desired to know the electric density. The latter is equal to the algebraic sum of the partial results.

This rule is now equivalent to the following conception, which allows the true meaning of ψ to stand out more clearly. $\psi\bar{\psi}$ is a kind of *weight-function* in the system's configuration space. The *wave-mechanical* configuration of the system is a *superposition* of many, strictly speaking of *all*, point-mechanical configurations kinematically possible. Thus, each point-mechanical configuration contributes to the true wave-mechanical configuration with a certain *weight*, which is given precisely by $\psi\bar{\psi}$. If we like paradoxes, we may say that the system exists, as it were, simultaneously in all the positions kinematically imaginable, but not "equally strongly" in all. In macroscopic motions, the weight-function is practically concentrated in a small region of positions, which are practically indistinguishable. The centre of gravity of this region in configuration space travels over distances which are macroscopically perceptible. In problems of microscopic motions, we are in any case interested *also*, and in certain cases even *mainly*, in the varying *distribution* over the region.

This new interpretation may shock us at first glance, since we have often previously spoken in such an intuitive concrete way of the " ψ -vibrations" as though of something quite real. But there is something tangibly real behind the present conception also, namely, the very real electro-dynamically effective fluctuations of the electric space-density. The ψ -function is to do no more and no less than permit of the totality of these fluctuations being mastered and surveyed mathematically by a single partial differential equation. We have repeatedly called attention¹ to the fact that the ψ -function itself cannot and may not be interpreted directly in terms of three-dimensional space—however much the one-electron problem tends to mislead us on this point—

¹ End of Part II. (p. 39); paper on Heisenberg's quantum mechanics (p. 60).

because it is in general a function in configuration space, not real space.

Concerning such a weight-function in the above sense, we would wish its integral over the whole configuration space to remain constantly normalised to the same unchanging value, preferably to unity. We can easily verify that this is necessary if the total charge of the system is to remain constant on the above definitions. Even for non-conservative systems, this condition must obviously be postulated. For, naturally, the charge of a system is not to be altered when, *e.g.*, a light wave falls on it, continues for a certain length of time, and then ceases. (*N.B.*—This is also valid for ionisation processes. A disrupted particle is still to be included in the system, until the separation is also *logically*—by decomposition of configuration space—completed.)

The question now arises as to whether the postulated *persistence of normalisation* is actually guaranteed by equations (4''), to which ψ is subject. If this were not the case, our whole conception would practically break down. Fortunately, it is the case. Let us form

$$(37) \quad \frac{d}{dt} \int \psi \bar{\psi} \rho dx = \int \left(\psi \frac{\partial \bar{\psi}}{\partial t} + \bar{\psi} \frac{\partial \psi}{\partial t} \right) \rho dx.$$

Now, ψ satisfies one of the two equations (4''), and $\bar{\psi}$ the other. Therefore, apart from a multiplicative constant, this integral becomes

$$(38) \quad \int (\psi \nabla^2 \bar{\psi} - \bar{\psi} \nabla^2 \psi) \rho dx = 2i \int (J \nabla^2 R - R \nabla^2 J) \rho dx,$$

where for the moment we put

$$\psi = R + iJ.$$

According to Green's theorem, integral (38) vanishes identically; the sole necessary condition that functions R and J must satisfy for this—vanishing in sufficient degree at infinity—means physically nothing more than that the system under consideration should practically be confined to a *finite* region.

We can put this in a somewhat different way, by not immediately integrating over the whole configuration space, but by merely changing the time-derivative of the weight-function into a divergence by Green's transformation. Through this we get an insight into the question of the flow of the weight-function, and thus of electricity. The two equations

$$(4'') \quad \begin{aligned} \frac{\partial \psi}{\partial t} &= \frac{h}{4\pi i} \left(\nabla^2 - \frac{8\pi^2}{h^2} V \right) \psi \\ \frac{\partial \bar{\psi}}{\partial t} &= -\frac{h}{4\pi i} \left(\nabla^2 - \frac{8\pi^2}{h^2} V \right) \bar{\psi} \end{aligned}$$

are multiplied by $\rho \bar{\psi}$ and $\rho \psi$ respectively, and added. Hence

$$(39) \quad \frac{\partial}{\partial t} (\rho \psi \bar{\psi}) = \frac{h}{4\pi i} \rho \cdot (\bar{\psi} \nabla^2 \psi - \psi \nabla^2 \bar{\psi}).$$

To carry out *in extenso* the transformation of the right-hand side, we must remember the explicit form of our many-dimensional, non-Euclidean, Laplacian operator:¹

$$(40) \quad \rho \nabla^2 = \sum_k \frac{\partial}{\partial q_k} \left[\rho T_{p_k} \left(q_l, \frac{\partial \psi}{\partial q_l} \right) \right].$$

By a small transformation we readily obtain

$$(41) \quad \frac{\partial}{\partial t} (\rho \psi \bar{\psi}) = \frac{\hbar}{4\pi i} \sum_k \frac{\partial}{\partial q_k} \left[\rho \bar{\psi} T_{p_k} \left(q_l, \frac{\partial \psi}{\partial q_l} \right) - \rho \psi T_{p_k} \left(q_l, \frac{\partial \bar{\psi}}{\partial q_l} \right) \right].$$

The right-hand side appears as the divergence of a many-dimensional real vector, which is evidently to be interpreted as the *current density of the weight-function* in configuration space. Equation (41) is the *continuity equation* of the weight-function.

From it we can obtain the *equation of continuity of electricity*, and, indeed, a separate equation of this sort is valid for the charge density "originating from each separate particle". Let us fix on the *ath* particle, say. Let its "charge" be e_a , its mass m_a , and let its co-ordinate space be described by Cartesians x_a, y_a, z_a , for the sake of simplicity. We denote the product of the differentials of the *remaining* co-ordinates shortly by dx' . Over the latter, we integrate equation (41), keeping x_a, y_a, z_a , *fixed*. As the result, all terms except three disappear from the right-hand side, and we obtain

$$(42) \quad \left\{ \begin{aligned} \frac{\partial}{\partial t} \left[e_a \int \psi \bar{\psi} dx' \right] &= \frac{\hbar e_a}{4\pi i m_a} \left\{ \frac{\partial}{\partial x_a} \left[\int \left(\bar{\psi} \frac{\partial \psi}{\partial x_a} - \psi \frac{\partial \bar{\psi}}{\partial x_a} \right) dx' \right] \right. \\ &+ \frac{\partial}{\partial y_a} \left[\int \left(\bar{\psi} \frac{\partial \psi}{\partial y_a} - \psi \frac{\partial \bar{\psi}}{\partial y_a} \right) dx' \right] + \dots \left. \right\} \\ &= \frac{\hbar e_a}{4\pi i m_a} \operatorname{div}_a \left[\int \left(\bar{\psi} \operatorname{grad}_a \psi - \psi \operatorname{grad}_a \bar{\psi} \right) dx' \right]. \end{aligned} \right.$$

In this equation, *div* and *grad* have the usual three-dimensional Euclidean meaning, and x_a, y_a, z_a are to be interpreted as Cartesian co-ordinates of real space. The equation is the continuity equation of *that* charge density which "originates from the *ath* particle". If we form all the others in an analogous fashion, and add them together, we obtain the total equation of continuity. Of course, we must emphasize that the interpretation of the integrals on the right-hand side as *components of the current density*, is, as in all such cases, not absolutely compulsory, because a divergence-free vector could be added thereto.

To give an example, in the conservative *one-electron* problem, if ψ is given by

¹ Cf. paper on Heisenberg's theory, equation (31). The quantity there denoted by $\Delta_p - \frac{1}{2}$ is our "density function" $\rho(x)$ (e.g. $r^2 \sin \theta$ in spherical polars). T is the kinetic energy as function of the position co-ordinates and *momenta*, the suffix at T denoting differentiation with respect to a momentum. In equations (31) and (32), *loc. cit.*, unfortunately by error the suffix k is used twice, once for the summation and then also as a representative suffix in the argument of the functions.

$$(43) \quad \psi = \sum_k c_k u_k e^{2\pi i \nu_k t + i \theta_k} \quad (c_k, \theta_k \text{ real constants}),$$

we get for the *current density* J

$$(44) \quad J = \frac{he_1}{2\pi m_1} \sum_{(k, l)} c_k c_l (u_l \text{ grad } u_k - u_k \text{ grad } u_l) \cdot \sin [2\pi(\nu_k - \nu_l)t + \theta_k - \theta_l].$$

We see, and this is valid for conservative systems generally, that, if only a single proper vibration is excited, the current components disappear and the distribution of electricity is constant in time. The latter is also immediately evident from the fact that $\psi\bar{\psi}$ becomes constant with respect to the time. This is still the case even when several proper vibrations are excited, if they all belong to the same proper value. On the other hand, the current density then no longer needs to vanish, but there may be present, and generally is, a *stationary* current distribution. Since the one or the other occurs in the unperturbed normal state at any rate, we may in a certain sense speak of a *return to electrostatic and magnetostatic atomic models*. In this way the lack of radiation in the normal state would, indeed, find a startlingly simple explanation.

I hope and believe that the present statements will prove useful in the elucidation of the magnetic properties of atoms and molecules, and further for explaining the flow of electricity in solid bodies.

Meantime, there is no doubt a certain crudeness in the use of a *complex* wave function. If it were unavoidable *in principle*, and not merely a facilitation of the calculation, this would mean that there are in principle *two* wave functions, which must be used *together* in order to obtain information on the state of the system. This somewhat unacceptable inference admits, I believe, of the very much more congenial interpretation that the state of the system is given by a real function and its time-derivative. Our inability to give more accurate information about this is intimately connected with the fact that, in the pair of equations (4''), we have before us only the *substitute*—extraordinarily convenient for the calculation, to be sure—for a real wave equation of probably the fourth order, which, however, I have not succeeded in forming for the non-conservative case.

Zürich, Physical Institute of the University.

(Received June 23, 1926.)

The Compton Effect

(*Annalen der Physik* (4), vol. 82, 1927)

It is well known that according to the wave theory of light all changes in the *frequency* and in the *wave-normal* can be predicted by means of very simple and general considerations with respect to *phase*, without going into any details of the process. I mean considerations of the following kind :

Let the xy -plane be the surface of separation of two media with refractive indices n (for $z > 0$) and n' (for $z < 0$), and let a wave of light with the phase

$$2\pi\nu\left[t - \frac{n}{c}(ax + \beta y + \gamma z)\right]$$

fall on it from the positive z -direction. If we assign to the refracted wave the phase

$$2\pi\nu'\left[t - \frac{n'}{c}(a'x + \beta'y + \gamma'z) + \delta\right],$$

and stipulate that for $z=0$ the phase difference between the waves is to be *constant*, *i.e.* independent of x , y , t , we obtain

$$\nu' = \nu, \quad n'a' = na, \quad n'\beta' = n\beta,$$

i.e. Snell's law of refraction. The reasoning is so general that it holds without alteration, *e.g.* for crystals, and is immediately transferable to the case of a moving surface of separation. A more detailed investigation of the electromagnetic process only becomes necessary if we are also concerned with the *intensities* (Fresnel's formulae of reflection).

If now we are right in supposing that in de Broglie's waves we have at our disposal a means (ranking on a level with wave optics) for mastering those processes which have previously been thought of exclusively as motions of corpuscles, it is to be expected, nay, even demanded, that we should be able, by means of quite simple phase considerations of the kind mentioned above, to explain the connection between the changes in direction and frequency of the ether wave which occur in the Compton effect and the change of velocity of the electron. For according to de Broglie's idea the latter also can be described as the change of direction and frequency of a wave, namely,

of a de Broglie wave. A more detailed investigation of the wave mechanics of the process, such as W. Gordon¹ has recently carried out with complete success, is unnecessary, except for the determination of the intensities. As this is rather lengthy and involved, the simple intuitive treatment given below, which gives everything *but* the intensity, may be welcome.

We start from a result of classical optics. "If in a transparent, homogeneous, and isotropic medium, the refractive index of which depends on the density, a ray of light of wave-length λ crosses a wave of compression (sound wave) of wave-length Λ , then (as L. Brillouin² has shown by purely classical reasoning) the ray of light is in part regularly reflected from the planes of the sound wave, provided that the two wave-lengths and the glancing angle θ are connected by Bragg's relation (well known in the theory of X-ray reflection)

$$(1) \quad 2\Lambda \sin \theta = \lambda$$

for *first-order* reflection ($=\lambda$, not $=k\lambda$). This approximation holds good, provided we can regard the velocity of light as very great compared to the velocity of sound. More accurately, the circumstances are the same as in the case of a *moving* mirror: the angle of reflection is not exactly equal to the angle of incidence, the ray of light undergoes a Doppler displacement, and equation (1) also must be corrected as in the case of a *moving* crystal."

These sentences are taken from an earlier paper,³ in which it is satisfactorily proved that Brillouin's result can also be obtained on the hypothesis that exchanges of energy and momentum between the two waves proceed by quanta. At that time we were all of the opinion that our whole interpretation of Nature must ultimately rest on such quantum balances, and we rejoiced whenever a trustworthy classical result could be transferred from the old foundation to the new without any trouble. Now we are going in the reverse direction, so to speak. We show that the wave mechanics can provide for the Compton relationships an interpretation which is closely related to the above-mentioned result of Brillouin's, and which is just as simple as the quantum considerations of momentum-energy.

A plane sine wave

$$(2) \quad \psi \sim e^{\frac{2\pi i}{h} \left[h\nu t - \frac{h\sqrt{v^2 - v_0^2}}{c} (\alpha x + \beta y + \gamma z) \right]},$$

where

$$\alpha^2 + \beta^2 + \gamma^2 = 1, \quad \nu_0 = m_0 c^2 / h,$$

(m_0 = rest-mass of the electron, h = Planck's constant, c = velocity of light)

¹ W. Gordon, *Ztschr. f. Phys.* 40, p. 117, 1926. Herr Gordon kindly let me see the manuscript of his paper, and hence I was led to the following simple way of looking at the matter, which is, in a nutshell, the basis of Herr Gordon's own treatment.

² L. Brillouin, *Ann. d. Phys.* 17, p. 88, 1923.

³ E. Schrödinger, *Physik. Ztschr.* 25, p. 89, 1924.

for space with no field, is a solution of the relativistic ψ -wave equation which has recently been proposed on many sides,¹

$$\nabla^2\psi - \frac{1}{c^2}\ddot{\psi} - \frac{4\pi^2\nu_0^2}{c^2}\psi = 0.$$

According to de Broglie, the above solution belongs to an electron moving with energy $h\nu$ in the direction α , β , γ . Hence we deduce in a well-known way that

$$\frac{h\nu}{c}, \quad \frac{h\sqrt{\nu^2 - \nu_0^2}}{c} \cdot \alpha, \quad \frac{h\sqrt{\nu^2 - \nu_0^2}}{c} \cdot \beta, \quad \frac{h\sqrt{\nu^2 - \nu_0^2}}{c} \cdot \gamma,$$

is the four-vector "energy-momentum" of the corresponding electron. From the wave standpoint we will call it the four-vector "propagation", and we shall apply this expression to the coefficients of ct , $-x$, $-y$, $-z$, in the phase (dropping the factor $2\pi/h$) of a *completely arbitrary plane sine wave*, be it a ψ -wave, an ether wave, or any other. "Propagation" is therefore a purely wave-kinematical idea; its components are

$$(3) \quad \frac{h}{c} \cdot \text{frequency}, \quad \frac{h\alpha}{\text{wave length}}, \quad \frac{h\beta}{\text{wave length}}, \quad \frac{h\gamma}{\text{wave length}},$$

where α , β , γ are the direction cosines of the normal to the wave. For an ether wave these quantities coincide with the quantum theory values of energy and momentum. These references to quantum magnitudes, however, are only for the purpose of making it subsequently easier to identify our results with Compton's: we shall work with the purely wave-kinematic idea of propagation (3). By the *three-vector propagation*, we mean, of course, the projection in space, *i.e.* the vector (3) with the first component omitted.

According to the hypothesis of wave mechanics, which up to now has always proved trustworthy, it is not the ψ -function itself but the square of its absolute value that is given a physical meaning, namely, density of electricity.² A single ψ -wave of type (2) therefore produces a density distribution which is constant both in space and time. If, however, we superpose two such waves, the constants of the second being ν' , α' , β' , γ' , we easily see that a "wave of electrical density" arises from their combination, with a propagation vector which is the vector difference of the propagation vectors of the two constituent ψ -waves. If we denote the latter vectors symbolically by A , A' , that of the density wave is³

$$(4) \quad D = A - A'.$$

Now it is this density wave that takes the place of the sound wave of

¹ O. Klein, *Ztschr. f. Phys.* 37, p. 895, 1926; E. Schrödinger, *Ann. d. Phys.* 81, p. 109, 1926; V. Fock, *Ztschr. f. Phys.* 38, p. 242, 1926; Th. de Donder and H. van den Dungen, *Compt. rend.*, July 5, 1926; L. de Broglie, *Compt. rend.*, July 26, 1926; J. Kudar, *Ann. d. Phys.* 81, p. 632, 1926; W. Gordon, *loc. cit.*

² The relativistic refinement of this statement (W. Gordon, *loc. cit.*) does not affect our case at all.

³ The sign is of small importance, as changing it merely causes the ψ -waves to change places.

Brillouin's paper. If we assume that a light wave is reflected from it as from a moving mirror, subject, however, to fulfilment of Bragg's law, then, as we shall show, our four waves, namely, the two ψ -waves A and A' and the incident and reflected light waves, stand exactly in the Compton relationship. The difference between this case and Brillouin's case of reflection from a sound wave is quantitative only, inasmuch as the velocity of our density wave D is not in general small compared with the velocity of light; on the contrary, arbitrary values up to the velocity of light may occur (but, as is easily verified, never greater than the velocity of light).

The proof of our assertion is easily produced. In fact we do not really require to investigate reflection at a *moving* mirror. As all the four waves, and of course their propagation vectors also, are invariant with respect to Lorentz transformation, we can bring the density wave to rest by means

of such a transformation. The first component (the time component) of its propagation vector then vanishes. Moreover, the frequency (and wave length) of the light wave are *not* changed during the reflection in this case, *i.e.* the time component of *its* propagation vector remains unaltered on reflection. Finally, Bragg's relationship holds exactly in the form (1), if λ denotes the wave length of the light wave, Λ that of the density wave, and θ the glancing angle. It can be put in the form

$$(5) \quad 2\frac{h}{\lambda} \sin \theta = \frac{h}{\Lambda},$$

which is illustrated in the accompanying figure (Fig. 1), in which the equality of the angles of incidence and reflection is also taken into account.

Equation (5) therefore expresses the fact that the three-vector of the incident light wave and the three-vector of D are together equal to the three-vector of the reflected light wave. Moreover, according to what has been said above, a similar relationship holds for the time components also: for D , the time component is *zero*, and for the light wave it is *unaltered* on reflection. If we call the propagation four-vectors for the incident and reflected light waves L and L' respectively, we can sum up all this in the single four-vector equation ¹

$$(6) \quad L + D = L',$$

¹ The sign of D in (6) is of small importance, as changing it merely causes the light waves to change places.

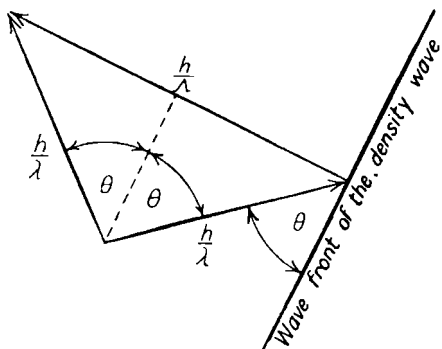


FIG. 1.

which must now hold for an arbitrary four-dimensional system of co-ordinates. Combined with (4), this equation gives

$$(7) \quad L + A = L' + A'.$$

If we bear in mind the significance of the components of L , L' from the light quantum point of view, and of those of A , A' according to de Broglie's correlation of ψ -waves and electrons, equation (7) agrees exactly with the statement of Compton's energy-momentum theory of the Compton effect.

It is quite interesting to notice the *complete reciprocity* between the ψ -waves on the one hand and the light waves on the other. The phenomenon may equally well be regarded as a Bragg's reflection of a ψ -wave at the system of *interference fringes* produced by two light

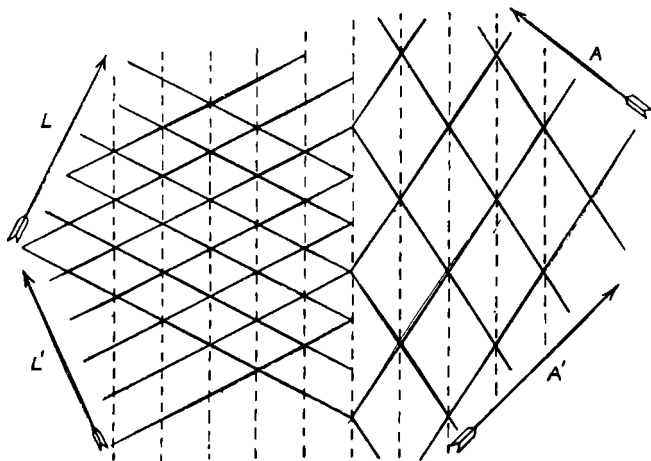


FIG. 2.

waves crossing one another. In the special system of co-ordinates used above, this system is *at rest* and is identical with O. Wiener's system of stationary light waves. The relationships (4) and (6) express the fact that the system of interference fringes and the density wave coincide, both having the propagation vector D . The special system of co-ordinates is just the one which W. Pauli¹ formerly found the most convenient for the investigation of the Compton effect.

Fig. 2 is an attempt to represent the relationships between the four wave fronts penetrating through each other and the common stationary wave (in dotted lines) in the special space-time system.² To avoid confusing the diagram, the two light wave fronts are drawn

¹ W. Pauli, jun., *Ztschr. f. Phys.* 18, p. 272, 1923.

² Note added to the English edition:

Professor Ehrenfest has drawn my attention to the fact that unfortunately Fig. 2 is *wrong*. The pair of waves (A , A') ought to be equal to (L' , L) in *every* respect, i.e. in wave length and direction also—the wave normal of A' being parallel to that of L , and the wave normal of A to that of L' .

in the left half only, and the two ψ -wave fronts in the right half only. The *arrows* are normal to the wave fronts and indicate their directions of propagation. Their length has no meaning. The reader should imagine them pushed together into the middle of the figure by parallel translation, so that the feathers of L' and A' and the heads of L and A all coincide in one point. It is easily deduced from the figure that Bragg's equation (1) holds for either pair of waves (L, L') and (A, A'), if we regard the stationary wave as the "crystal". We may therefore say :

The direction and frequency laws of the Compton effect are completely equivalent to the statement that the pair of light waves and the pair of ψ -waves concerned stand in the Bragg relationship for reflection of the first order (generalised for a moving crystal) to one and the same "space lattice". The space lattice considered can at the outset have an arbitrary position and an arbitrary distance between the plane layers, and a velocity of translation less than that of light but otherwise arbitrary.

I should now like to deal with one objection to the principle of the method. It might be said : yes, but in the Compton effect the original data are *one* light wave, and *one* electron moving in a specified way, i.e. *one* ψ -wave, as we say : now wherever does the *second* suitably chosen ψ -wave come from, which, together with the given one, is to form a suitable "Bragg's mirror" for the given light wave? The reply is that such simple considerations of phase as we have employed here are of course absolutely inadequate for the answering of such questions. By means of these simple considerations we investigate the Compton phenomenon *in a steady state*, so to speak, in which the primary wave of one kind is continually transformed into a secondary wave by reflection at the system of interference fringes of the other kind, and *vice versa*. This is just how we proceed in analogous discussions in optics, even when we study the subject much more accurately by means of a detailed theory. There also we do not in general study the first appearance, e.g. of a reflected and a refracted wave at the head of the primary wave train, but we make an *assumption* not merely for the incident wave alone but likewise for *all* the other waves whose appearance can be foreseen ; and by this assumption we seek to represent a stationary state which will satisfy all the conditions of the problem.

Zürich, Physical Institute of the University.
(Received November 30, 1926.)

The Energy-Momentum Theorem for Material Waves

(*Annalen der Physik* (4), vol. 82, 1927)

THE *Hamiltonian principle*, from which the exact relativistic differential equation for de Broglie's waves can be obtained,¹ appears to justify completely the hopes which I had set upon an intimate blending together of wave mechanics and classical electrodynamics.² If to the integrand (the "Lagrangian function") we add $\mathfrak{H}^2 - \mathfrak{E}^2$, the well-known Lagrangian function of the electromagnetic field when no charges are present, then by varying the potentials as well as the ψ -function we obtain simultaneously the four wave equations for the former with the components of the four-current on the right-hand sides, *i.e.* the complete electrodynamics. This is due to the fact, first noticed by Gordon, *loc. cit.*, that differentiation of the Lagrangian function for de Broglie's waves with respect to a component of potential gives the corresponding component of the four-current. A most important further result is the *energy-momentum theorem* for the whole field, from which the contribution of the charges, *i.e.* of the ψ -function, to the energy-momentum tensor may easily be deduced. It is quite clear to me that all this must somehow be included in the very general theories proposed by O. Klein³ and de Donder.⁴ It does not seem superfluous, however, to set forth these connections in as simple a form as possible without referring to the theory of gravitation or the interesting fifth dimension, especially if we bear in mind the very considerable gulf which still yawns between experiment and this beautiful and self-contained theory of the field (see the end of this paper).

We apply the wave equation and Hamilton's principle in the form given by Gordon (*loc. cit.*). The former (always summing from 1 to 4 for indices which occur twice) is as follows :

¹ O. Klein, *Ztschr. f. Phys.* 37, p. 895, 1926 ; V. Fock, *ibid.* 38, p. 242, 1926 ; J. Kudar, *Ann. d. Phys.* 81, p. 632, 1926 ; W. Gordon, *Ztschr. f. Phys.* 40, p. 117, 1926.

² *Ann. d. Phys.* 79, p. 754, 1926 (p. 60 above).

³ *Loc. cit.*

⁴ Th. de Donder and H. van den Dungen, *Compt. rend.*, July 5, 1926.

$$(1) \quad \left[\left(\frac{\partial}{\partial x_a} + i\phi_a \right) \left(\frac{\partial}{\partial x_a} + i\phi_a \right) - k^2 \right] \psi = 0,$$

where

$$x_1, x_2, x_3 = x, y, z; \quad x_4 = ict;$$

$$\phi_1, \phi_2, \phi_3 = \frac{2\pi e}{hc} \cdot \mathfrak{U}_x, \mathfrak{U}_y, \mathfrak{U}_z; \quad \phi_4 = \frac{2\pi e}{hc} \cdot iV.$$

$$(2) \quad k^2 = \frac{4\pi^2 m_0^2 c^2}{h^2}.$$

\mathfrak{U}, V , are the potentials; e, m_0, c, h , the familiar universal constants; $i = \sqrt{-1}$. It must specially be emphasized that the relationship to actual reality is *not* obliterated by the introduction of four-vectors with an imaginary fourth component. This procedure is merely a formal device of calculation, used in order that we may not be obliged to insert the fourth term specially in all four-sums on account of its different sign. The passage to the conjugate complex quantity therefore only affects i when it occurs explicitly, and the ψ -function.

According to Gordon (*loc. cit.*), (1) can be derived from a four-dimensional Hamiltonian integral with the (real) Lagrangian function

$$(3) \quad L_m = (\psi_a + i\phi_a\psi)(\bar{\psi}_a - i\phi_a\bar{\psi}) + k^2\psi\bar{\psi},$$

where the bar denotes the conjugate complex quantity, and we have put for brevity

$$\psi_a = \frac{\partial\psi}{\partial x_a}, \quad \bar{\psi}_a = \frac{\partial\bar{\psi}}{\partial x_a}.$$

The index a must therefore take effect *after* the bar (see above). In the formation of the variational derivatives, ψ and $\bar{\psi}$ are to be varied *independently*, as Gordon has observed. It is easy to see that this comes to the same thing as varying the real and imaginary parts of ψ independently (which would really be the rational procedure). Hence one of the variational derivatives is

$$(5) \quad \frac{\partial}{\partial x_a} \left(\frac{\partial L_m}{\partial \bar{\psi}_a} \right) - \frac{\partial L_m}{\partial \bar{\psi}} = 0,$$

which is identical with (1); the other gives nothing new. By multiplying (5) by $\bar{\psi}$, we easily obtain

$$(6) \quad \frac{\partial}{\partial x_a} \left(\bar{\psi} \frac{\partial L_m}{\partial \bar{\psi}_a} \right) = \bar{\psi}_a \frac{\partial L_m}{\partial \bar{\psi}_a} + \bar{\psi} \frac{\partial L_m}{\partial \bar{\psi}} = L_m;$$

the latter equality follows from the fact that L_m is homogeneous and of the first degree in the five quantities $\bar{\psi}, \bar{\psi}_a$. The right-hand side is unaltered by passage to the conjugate complex quantities; and hence by subtraction

$$(7) \quad \frac{\partial}{\partial x_a} \left(\bar{\psi} \frac{\partial L_m}{\partial \bar{\psi}_a} - \psi \frac{\partial L_m}{\partial \psi_a} \right) = 0.$$

This, according to Gordon, is the *equation of continuity for electricity*. We notice that

$$(8) \quad \bar{\psi} \frac{\partial L_m}{\partial \bar{\psi}_a} - \psi \frac{\partial L_m}{\partial \psi_a} = i \frac{\partial L_m}{\partial \phi_a}.$$

We define the *four-current* as follows,

$$(9) \quad s_a = -\lambda \frac{\partial L_m}{\partial \phi_a},$$

where λ is a universal real constant to be chosen later. By the quantities s_a we mean the four quantities which in Lorentz's theory were given by

$$(10) \quad s_1, s_2, s_3 = \frac{v}{c}, \quad s_4 = i\rho.$$

We now extend our Lagrangian function (3) in such a way that by varying the quantities ϕ_a in it we obtain (as is possible, in consequence of (9)) the laws of the electromagnetic field. We put

$$(11) \quad L_e = \frac{1}{4} f_{\alpha\beta} f_{\alpha\beta} = \frac{1}{4} \left(\frac{\partial \phi_\beta}{\partial x_\alpha} - \frac{\partial \phi_\alpha}{\partial x_\beta} \right) \left(\frac{\partial \phi_\beta}{\partial x_\alpha} - \frac{\partial \phi_\alpha}{\partial x_\beta} \right),$$

where for brevity

$$(12) \quad f_{\alpha\beta} = \frac{\partial \phi_\beta}{\partial x_\alpha} - \frac{\partial \phi_\alpha}{\partial x_\beta}.$$

From (2) and familiar formulae, we see that the quantities f have the following meanings :

$$(13) \quad \begin{aligned} f_{14} &= -\frac{2\pi e}{hc} i\mathfrak{E}_x, & f_{24} &= -\frac{2\pi e}{hc} i\mathfrak{E}_y, & f_{34} &= -\frac{2\pi e}{hc} i\mathfrak{E}_z; \\ f_{23} &= \frac{2\pi e}{hc} \mathfrak{H}_x, & f_{31} &= \frac{2\pi e}{hc} \mathfrak{H}_y, & f_{12} &= \frac{2\pi e}{hc} \mathfrak{H}_z; \end{aligned}$$

where \mathfrak{E} , \mathfrak{H} represent the field in the customary units. As Lagrangian function we now take

$$(14) \quad L = L_m + L_e,$$

and obtain in a familiar way

$$(15) \quad \frac{\partial f_{\alpha\beta}}{\partial x_\alpha} = \frac{\partial L}{\partial \phi_\beta} = -\frac{s_\beta}{\lambda}$$

by varying ϕ_β . If the constant λ is given the value

$$\lambda = \frac{hc}{8\pi^2 e},$$

the equations (15) represent the so-called *second quartet of the Maxwell-Lorentz* equations, while the *first quartet* is identically satisfied by (12).

Using (12) and Maxwell's subsidiary condition $\left(\frac{\partial \phi_a}{\partial x_a} = 0 \right)$, the equations (15) become the wave equations for the potentials,

$$(15') \quad \frac{\partial^2 \phi_\beta}{\partial x_a \partial x_a} = -\frac{s_\beta}{\lambda}.$$

From (15') (together with Maxwell's subsidiary condition), it is now easy to verify that

$$(16) \quad \frac{\partial T_{\rho\sigma}}{\partial x_\sigma} = -\frac{f_{\rho\sigma} s_\sigma}{\lambda},$$

where

$$(17) \quad T_{\rho\sigma} = f_{\rho\alpha} f_{\sigma\alpha} - \delta_{\rho\sigma} L_e$$

is Maxwell's well-known stress-energy-momentum-tensor (apart from a universal constant). In Lorentz's work the right-hand side of (16) means the energy or momentum abstracted from the electrons by the forces of the field. In virtue of (9) and the ψ -wave-equation (5), the right-hand side of (16) may likewise be represented as the divergence of a tensor, namely, of the energy-momentum-tensor of the charge (or of the "matter"). In the first place, we have

$$(18) \quad -\frac{f_{\rho\sigma} s_\sigma}{\lambda} = \left(\frac{\partial \phi_\sigma}{\partial x_\rho} - \frac{\partial \phi_\rho}{\partial x_\sigma} \right) \frac{\partial L_m}{\partial \phi_\sigma} = \frac{\partial L_m}{\partial \phi_\sigma} \frac{\partial \phi_\sigma}{\partial x_\rho} - \frac{\partial}{\partial x_\sigma} \left(\phi_\sigma \frac{\partial L_m}{\partial \phi_\sigma} \right),$$

where the latter equation follows from the fact that the four-current is free of sources (from equations (7) and (8)). Further, we note that

$$(19) \quad \frac{\partial L_m}{\partial x_\rho} = \frac{\partial L_m}{\partial \phi_\sigma} \frac{\partial \phi_\sigma}{\partial x_\rho} + \frac{\partial L_m}{\partial \bar{\psi}} \bar{\psi}_\rho + \frac{\partial L_m}{\partial \psi} \psi_\rho + \frac{\partial L_m}{\partial \bar{\psi}_\sigma} \frac{\partial \bar{\psi}_\sigma}{\partial x_\rho} + \frac{\partial L_m}{\partial \psi_\sigma} \frac{\partial \psi_\sigma}{\partial x_\rho}.$$

Since by (4), however,

$$(20) \quad \frac{\partial \psi_\sigma}{\partial x_\rho} = \frac{\partial \psi_\rho}{\partial x_\sigma}, \text{ etc.,}$$

it is possible to transform the last two terms in (19) as is done in integration by parts; after this transformation four terms cancel, on account of (5), and we obtain

$$(21) \quad \frac{\partial L_m}{\partial x_\rho} = \frac{\partial L_m}{\partial \phi_\sigma} \frac{\partial \phi_\sigma}{\partial x_\rho} + \frac{\partial}{\partial x_\sigma} \left(\bar{\psi}_\rho \frac{\partial L_m}{\partial \bar{\psi}_\sigma} + \psi_\rho \frac{\partial L_m}{\partial \psi_\sigma} \right).$$

We subtract this from equation (18), and we have

$$(22) \quad \left\{ \begin{aligned} -\frac{f_{\rho\sigma} s_\sigma}{\lambda} &= \frac{\partial L_m}{\partial x_\rho} - \frac{\partial}{\partial x_\sigma} \left(\bar{\psi}_\rho \frac{\partial L_m}{\partial \bar{\psi}_\sigma} + \psi_\rho \frac{\partial L_m}{\partial \psi_\sigma} + \phi_\rho \frac{\partial L_m}{\partial \phi_\sigma} \right) \\ &= \frac{\partial}{\partial x_\sigma} \left(\delta_{\rho\sigma} L_m - \bar{\psi}_\rho \frac{\partial L_m}{\partial \bar{\psi}_\sigma} - \psi_\rho \frac{\partial L_m}{\partial \psi_\sigma} - \phi_\rho \frac{\partial L_m}{\partial \phi_\sigma} \right) \\ &= -\frac{\partial S_{\rho\sigma}}{\partial x_\sigma}, \end{aligned} \right.$$

where we introduce the *energy-tensor of the charges or of the "matter"*

$$(23) \quad S_{\rho\sigma} = \bar{\psi}_\rho \frac{\partial L_m}{\partial \bar{\psi}_\sigma} + \psi_\rho \frac{\partial L_m}{\partial \psi_\sigma} + \phi_\rho \frac{\partial L_m}{\partial \phi_\sigma} - \delta_{\rho\sigma} L_m.$$

From (16) and (22) we obtain

$$(24) \quad \frac{\partial}{\partial x_\sigma} (T_{\rho\sigma} + S_{\rho\sigma}) = 0$$

for the combined laws of conservation of energy and momentum for the electromagnetic field and de Broglie's wave-field taken together.

Calculation shows that the tensor $S_{\rho\sigma}$ is *symmetrical*. We easily find the explicit expression

$$(25) \quad \begin{cases} S_{\rho\sigma} = \bar{\psi}_\rho \psi_\sigma + \bar{\psi}_\sigma \psi_\rho + i\phi_\sigma (\bar{\psi}_\rho \psi - \psi_\rho \bar{\psi}) \\ \quad + i\phi_\rho (\bar{\psi}_\sigma \psi - \psi_\sigma \bar{\psi}) + 2\psi \bar{\psi} \phi_\rho \phi_\sigma - \delta_{\rho\sigma} L_m, \end{cases}$$

or the following one, which is more closely related to the form of L_m given in (3),

$$(25') \quad \begin{cases} S_{\rho\sigma} = (\psi_\rho + i\phi_\rho \psi)(\bar{\psi}_\sigma - i\phi_\sigma \bar{\psi}) \\ \quad + (\psi_\sigma + i\phi_\sigma \psi)(\bar{\psi}_\rho - i\phi_\rho \bar{\psi}) - \delta_{\rho\sigma} L_m. \end{cases}$$

In contradistinction to $T_{\sigma\sigma}$, the Laue scalar (diagonal sum) $S_{\sigma\sigma}$ does *not* vanish. We easily obtain

$$(26) \quad S_{\sigma\sigma} = -2(L_m + k^2 \psi \bar{\psi}).$$

The complete tensor can also be represented by means of the complete Lagrangian function in the following way, well known from similar cases :

$$(27) \quad \begin{cases} T_{\rho\sigma} + S_{\rho\sigma} = \frac{\partial L}{\partial \left(\frac{\partial \phi_\rho}{\partial x_\alpha} \right)} \frac{\partial \phi_\sigma}{\partial x_\alpha} + \frac{\partial L}{\partial \left(\frac{\partial \phi_\alpha}{\partial x_\rho} \right)} \frac{\partial \phi_\alpha}{\partial x_\sigma} + \frac{\partial L}{\partial \bar{\psi}_\rho} \bar{\psi}_\sigma \\ \quad + \frac{\partial L}{\partial \psi_\rho} \psi_\sigma + \frac{\partial L}{\partial \phi_\rho} \phi_\sigma - \delta_{\rho\sigma} L. \end{cases}$$

This is analogous to the representation of the Hamiltonian function by means of the Lagrangian function in *point* mechanics.

It should be remembered that our tensor components $S_{\rho\sigma}$ and $T_{\rho\sigma}$ have the physical dimensions cm.^{-4} . Before being applied they must be multiplied by the constant

$$\frac{\hbar^2 c^2}{32\pi^3 e^2},$$

which is of the dimensions of the square of a charge, in order that they may physically represent energy, momentum, and stress. (*N.B.* Other defects in the dimensions are to be rectified, as is known, by means of powers of c .)

If we now ask ourselves whether this self-contained theory of the field—apart from the provisional neglect of the electronic spin—corresponds to reality in the way we had previously hoped for from such theories, the question must be answered *in the negative*. The examples worked out, particularly that of the H-atom, show in fact

that we have *not* to insert in the wave equation (1) those potentials which result from the potential equations (15') with the four-current (9). On the contrary, we know that in the case of the H-atom we have to substitute the *given* potentials of the nucleus and of possible "external" electromagnetic fields for the ϕ_a 's in (1), and solve the equation for ψ . The distribution of current produced by this ψ is then calculated from (9), and from the distribution the potentials produced by it are found by (15'). By adding the latter to the potentials given in advance, we obtain those potentials which define the external action of the atom as a whole. We thus obtain (with a suitable normalisation of ψ , for which, it must be admitted, a proof by the theory of the field is also wanting) the *neutralisation* of the nuclear charge at greater distances on the one hand, and on the other hand the *radiation*. With reference to the attempt, which it would now be natural to make, to substitute these newly found potentials in equation (1), and thus to calculate a "second approximation", it is to be remarked that we must not on any account proceed in this way with the *neutralisation potential*, as it would *completely* alter the values of the terms, and hence would make many more stages of approximation necessary. These, even if the process converges at all, certainly do *not* lead to the correct hydrogen terms, much less (in the case of nuclear charge 2) to the helium *atom* terms. On the contrary, we should very probably obtain the required *radiation correction*¹ by dealing with the *radiation potentials* in the way described, if we suppose that *one* proper vibration is *strongly* excited but all the others only very feebly.

Hence there is something which intrudes into the self-contained system of field equations in a peculiar way. This is not yet fully intelligible at present, but it must be considered in connection with the two following facts :

1. The exchange of energy and momentum between the electromagnetic field and "matter" does *not* in reality take place continuously as the expression (24) in terms of the field would lead us to believe.

2. In Lorentz's theory also we have to substitute in the first instance only the fields of the *other* electrons in the equations of motion of the single electron, and not its own individual field. The reaction of the latter has *already* been almost entirely taken account of as electromagnetic *mass*, *in setting up the equations of motion*. The corresponding term in equation (1) is the term with k^2 . The reaction of radiation results in a second approximation from the reaction of the electron's own field in Lorentz's theory also.

The question whether the solution of the difficulty is really to be found only in the purely *statistical* interpretation of the field theory which has been proposed in several quarters² must for the present be left unsettled. Personally I no longer regard this interpretation

¹ Cf. *Ann. d. Phys.* 81, p. 129 *et seq.*, 1926 (p. 116 *et seq.*).

² M. Born, *Ztschr. f. Phys.* 38, p. 803, 1926; 40, p. 167, 1926; P. A. M. Dirac, *Proc. Roy. Soc., A*, 112, p. 661, 1926; and W. Gordon, *loc. cit.*

as a finally satisfactory one,¹ even if it proves useful in practice. To me it seems to mean a renunciation, much too fundamental in principle, of all attempt to understand the individual process.

A *brighter* side of the difficulty in question deserves to be mentioned. By interrupting the completeness of the system of field equations in her actual behaviour, Nature accommodates herself to our mathematical powers to an astonishing extent. Even the theory of the hydrogen atom would become immeasurably complicated from the mathematical point of view, if the ϕ_a 's did not stand for *given* potential values in equation (1), but if instead we had to add to them those which are to be calculated by means of (9) and (15') from ψ , which is itself unknown.

Zürich, Physical Institute of the University.

(Received December 10, 1926.)

¹ Cf. *Die Naturwissenschaften*, 12, p. 720, 1924.

The Exchange of Energy according to Wave Mechanics

(*Annalen der Physik* (4), vol. 83, 1927)

A SERIES of papers which have appeared in this journal¹ forms the starting-point of the present note. Here, in fact, we shall apply the *many-dimensional* form of "wave mechanics" to which these papers are almost exclusively devoted, and which can be brought into coincidence with the Heisenberg-Dirac quantum mechanics, instead of that *four-* (or according to O. Klein *five-*) dimensional form,² which corresponds to de Broglie's original conception and possibly strikes more closely at the root of the matter, but which is meanwhile only prospective in character, because we do not yet understand how to formulate the problem for *more* than one electron by means of it. I must ask leave to develop afresh here some important matters which have previously been expounded by others (Heisenberg, Dirac, Jordan). For I should like to remain intelligible to those even who have not yet made themselves familiar with the use of the new number-systems (matrices, *q*-numbers) employed by those writers.³

§ 1. The Method of Variation of Constants⁴

More general methods,⁵ which are far superior for many purposes, have since been given for the treatment of the perturbation problem

¹ "Quantisation as a Problem of Proper Values", Parts I-IV. above; henceforth referred to as Q. I-IV.

² O. Klein, *Ztschr. f. Phys.* 37, p. 895, 1926; W. Gordon, *ibid.* 40, p. 117, 1926; Q. IV. p. 117 above; E. Schrödinger, *Ann. d. Phys.* 82, pp. 124 and 130 above, etc.

³ The difficulties which are very generally experienced may be compared with the following. If someone, e.g., in a lecture, began by expounding the old action-at-a-distance theory of electricity in Cartesian co-ordinates, and then introduced vector analysis for the first time while passing to Maxwell's theory, the listener would have great difficulty in distinguishing between the physically new matter and the new form. (Similarly, e.g., in P. A. M. Dirac's paper, *Proc. Roy. Soc., A*, 14, p. 250, § 3, one may easily overlook the fact that here a quite new physical hypothesis has just been introduced, namely, a "successive" or "double" application of the process which Heisenberg calls "passage to matrices" and Dirac "passage to *q*-numbers", and which I call "passage to wave-mechanics".)

⁴ P. A. M. Dirac, *Proc. Roy. Soc., A*, 112, p. 674, 1926.

⁵ Cf. especially M. Born, *Ztschr. f. Phys.* 40, p. 172, 1926.

solved in Q. III. (§§ 1 and 2). We consider a conservative system; let its wave equation (Q. IV., equation (4''))

$$(1) \quad \nabla^2 \psi - \frac{8\pi^2}{\hbar^2} V \psi - \frac{4\pi i}{\hbar} \dot{\psi} = 0$$

have the normalised proper solutions

$$(2) \quad \psi_k e^{\frac{2\pi i E_k t}{\hbar}},$$

where ψ_k depends only on the co-ordinates of the system.¹ ψ_k therefore satisfies the equation

$$(3) \quad \nabla^2 \psi_k + \frac{8\pi^2}{\hbar^2} (E_k - V) \psi_k = 0,$$

in which the time does not appear. The general solution of (1) is

$$(4) \quad \psi = \sum_k c_k \psi_k e^{\frac{2\pi i E_k t}{\hbar}},$$

where the c_k 's are arbitrary constants, complex in general, which we call the *amplitudes*. (The squares of their absolute values we call for short the squares of the amplitudes.)

We now bring into play a small perturbation, constant with respect to time, by replacing V in (1) by $V + r$, where r is a small function of the co-ordinates only. We again attempt to make (4) satisfy the equation perturbed in this way, by regarding the amplitudes as slowly varying functions of the time. By substituting (4) in the (perturbed) equation (1) and bearing (3) in mind, we obtain

$$(5) \quad -\frac{8\pi^2}{\hbar^2} r \sum_k c_k \psi_k e^{\frac{2\pi i E_k t}{\hbar}} - \frac{4\pi i}{\hbar} \sum_k \dot{c}_k \psi_k e^{\frac{2\pi i E_k t}{\hbar}} = 0$$

for this functional dependence on the time. As necessary and sufficient condition for the vanishing of the left-hand side we employ the condition that it should be orthogonal to every function of the complete orthogonal system ψ_l . We thus obtain the infinite set of equations

$$(6) \quad \dot{c}_l = \frac{2\pi i}{\hbar} \sum_k \epsilon_{kl} c_k e^{\frac{2\pi i (E_k - E_l) t}{\hbar}},$$

where

$$(7) \quad \epsilon_{kl} = \int r \psi_k \psi_l dx.$$

The equations (6) take everything into account.

If now all the proper value differences are large compared with the "elements of the matrix of perturbation" ϵ_{kl} , each c_k ($k \neq l$) can be taken as approximately constant throughout the period of the exponential factor standing beside it; thus all these terms produce only small oscillatory perturbations in c_l . It is only for the sum-

¹ The wave function ψ must be considered as essentially complex. We assume the functions of the co-ordinates ψ_k to be real, merely in order to simplify the formulae.

term $k=l$ that this does not hold, because then the exponential factor becomes unity. Apart from these small perturbations, we therefore have

$$(8) \quad \dot{c}_l = \frac{2\pi i}{h} \epsilon_{ll} c_l; \quad c_l = c_l^0 e^{\frac{2\pi i \epsilon_{ll} t}{h}}.$$

Hence the *absolute values* of the amplitudes remain altogether unchanged (to this approximation), but their *phases* undergo secular alterations (which can also be interpreted as *perturbations of proper values*; cf. Q. III.).

If, on the contrary, proper value differences which are comparable with, or even small compared to, the perturbation quantities ϵ_{kl} occur in the unperturbed problem, then the amplitudes of all those proper vibrations which belong to such a group of neighbouring proper values are coupled together by the equations (6), even in the approximation hitherto considered, in such a way that only the sum of the squares of the amplitudes remains constant, instead of, as before, the square of a single amplitude. We see this as follows. We shall consider in particular the case of an α -fold proper value. Let c_l be the amplitude of a corresponding proper vibration. On the right-hand side of (6) α exponentials then become equal to unity, and α secular terms remain, in the approximation considered, namely, just those amplitudes which belong to the proper value in question. The same happens in all the α equations (6) in which one of these amplitudes occurs on the left-hand side. We thus obtain for the determination of these amplitudes the finite and self-contained system of equations

$$(9) \quad \dot{c}_l = \frac{2\pi i}{h} \sum_{k=1}^{\alpha} \epsilon_{kl} c_k; \quad l=1, 2, \dots, \alpha,$$

where for simplicity we have numbered the α amplitudes in question from 1 to α . Accordingly, in general an exchange takes place between the amplitudes belonging to one and the same proper value, and—in the approximation considered—with such amplitudes only. If we multiply (9) by the conjugate complex quantity c_l^* , take the real parts, and sum for all values of l , the right-hand side vanishes (on account of the symmetry of the ϵ_{kl} 's), *i.e.*

$$(10) \quad \sum_{l=1}^{\alpha} c_l c_l^* = \text{const.}$$

is an integral of (9). Apart from this, the equations are of course very easy to integrate, since the ϵ_{kl} 's are constant. We are led precisely to the transformation to principal axes given in Q. III., pp. 73-4. The solution is identical with what were there called the "perturbed solution of zero approximation" and the "perturbed proper values of the first approximation".

§ 2. The Explanation according to Wave Mechanics of the Quantum Exchange of Energy

The very simple state of affairs which has just been described now provides, as Heisenberg¹ and Jordan² have remarked, the explanation according to wave mechanics of the fact which can quite well be characterised as the empirical basis of the quantum theory, namely, that physical systems apparently influence each other only when they agree in respect of a "difference of level", or nearly so; and that the influence is always exerted on the four critical levels only, and, moreover, always in such a way that one of the two systems is raised to its higher level at the expense of the other, which undergoes an "equivalent" opposite displacement.

Thus, if we have two systems with the wave equations

$$(11) \quad \nabla_1^2 \psi - \frac{8\pi^2}{h^2} V_1 \psi - \frac{4\pi i}{h} \dot{\psi} = 0$$

(proper functions ψ_k for E_k)

and

$$(12) \quad \nabla_2^2 \phi - \frac{8\pi^2}{h^2} V_2 \phi - \frac{4\pi i}{h} \dot{\phi} = 0$$

(proper functions ϕ_l for F_l)

and imagine them united into *one* system ("with vanishing coupling"), the wave equation of the latter will, as is easily seen, be

$$(13) \quad (\nabla_1^2 + \nabla_2^2) \Psi - \frac{8\pi^2}{h^2} (V_1 + V_2) \Psi - \frac{4\pi i}{h} \dot{\Psi} = 0,$$

with the proper functions $\psi_k \phi_l$ for the proper values $E_k + F_l$. Let us now add a small coupling term r to $V_1 + V_2$, as in § 1. It will then be a question of whether or not the imagined combination has caused new degenerations, or approximate degenerations (*i.e.* multiple proper values, or proper values close together), to appear. If this does not happen, *i.e.* if all the proper values $E_k + F_l$ differ sufficiently, the two systems do not influence each other in the first approximation, that considered in § 1. If, however, new degenerations occur in (13), a secular exchange of amplitudes takes place.

For example, let

$$(14) \quad E_k + F_{l'} = E_{k'} + F_l$$

for four special values k, k', l, l' (this just means that the two systems agree in respect of the proper value difference $E_k - E_{k'} = F_l - F_{l'}$). Then the *two* proper functions

$$(15) \quad \psi_k \phi_{l'} \text{ and } \psi_{k'} \phi_l$$

correspond to the proper value (14). If their amplitudes are c_1, c_2 .

¹ W. Heisenberg, *Ztschr. f. Phys.* 38, p. 411, 1926; 40, p. 501, 1926.

² P. Jordan, *ibid.* 40, p. 661, 1927.

an exchange will take place between them in accordance with the equations

$$(16) \quad \begin{cases} \dot{c}_1 = \frac{2\pi i}{h}(\epsilon_{11}c_1 + \epsilon_{12}c_2) \\ \dot{c}_2 = \frac{2\pi i}{h}(\epsilon_{12}c_1 + \epsilon_{22}c_2), \end{cases}$$

where the constants ϵ_{ik} are defined by appropriate application of equation 7, § 1.

Evidently we now have to interpret, *e.g.*, an increase in the amplitude corresponding to $\psi_k\phi_l$ at the expense of the other one, in the double sense that just as in the one system the amplitude of ψ_k increases at the expense of that of $\psi_{k'}$, so in the other system the amplitude of ϕ_l increases at the expense of that of $\phi_{l'}$. We can picture the state of affairs as follows: the wave function of the whole system describes at any moment the state of the first system (if we overlook the small coupling and the existence of the second system), and the reverse statement is equally true. To be sure, simple numbers no longer appear as amplitudes, but instead we have linear combinations of the proper functions of the *other* system, *i.e.* according to the present interpretation, of a completely external system. This, however, does not cause any serious difficulty. In the calculation of any physical quantity relating to the system under consideration, we have simply to integrate over the co-ordinates of the external system in a way similar to that described already in Q. IV., § 7. Thus we obtain, *e.g.*, the sum of the squares of the amplitudes of all those proper functions of the whole system which contain ϕ_l , for the *square of the amplitude* of ϕ_l .¹

We thus find that without assuming discrete energy levels and quantum exchange of energy, and even without having to consider any meaning for the proper values other than *frequencies*, we can give a simple explanation of the fact that physical interaction chiefly takes place between *those* systems in which, according to the older conception, "the same energy element occurs". As Heisenberg points out, it is a question of a simple resonance phenomenon with beats, similar to the phenomenon of the so-called sympathetic pendulum. Without quantum postulates we have arrived at an effect which is exactly the same *as if* the quantum postulates were in force. This "as-if" situation is not new to us. The spontaneously emitted frequencies are also obtained, as if the proper values were discrete energy levels and Bohr's frequency condition held good.

¹ The inconvenient fact that if we confine ourselves to the present simple method of calculation the external proper functions cannot be got rid of once for all, *i.e.* that the complex amplitude of ϕ_l in the isolated system cannot be obtained by itself, appears to be inherent in the nature of the case. For it is not possible actually to do away with the coupling, unless another system, namely, the radiation (or the "ether"), is taken into account as well. The Coulomb coupling terms cease to describe the state of affairs correctly long before they have become negligibly small; they would have to be altered by taking the radiation from one system to the other into account.

According to the fundamental principles of research which are commonly regarded as correct, does not the foregoing compel us to exercise the utmost caution with respect to the quantum postulates, even (I almost feel inclined to say) to distrust them—quite apart from their axiomatic unintelligibility? From the psychological point of view it is clear that as soon as the conception of the “terms” as discrete energy levels had been introduced, we were obliged to see a corroboration of that conception in every new exchange phenomenon discovered, even if there is really nothing present in nature beyond the resonance phenomenon just discussed. The reader is *not* to object: oh, but the conception of the terms as energy levels is raised above all doubts by *researches on electronic impact*, if by nothing else: surely you will not doubt that the potential difference fallen through is a measure of the kinetic energy of the single electron? My reply is this: Yes, I do question whether it is not very much more to the point to push the idea of the frequency of the de Broglie wave into the foreground instead of that of the “kinetic energy of the single electron”. It is known that in passing through a potential difference these waves undergo just that change of frequency which corresponds to the acquired kinetic energy. Further, the wave equation gives just those deviated ray-paths which are actually observed in the determination of e/m and v .

I cannot help feeling that to admit the quantum postulates *in addition* to the resonance phenomenon is to accept *two* explanations for the same thing. But two explanations are like two excuses: one is certainly untrue, and usually both. In the concluding section we shall add another “as-if” situation to the one described here.

§ 3. A Statistical Hypothesis

If in the case of prolonged interaction of two systems we try to obtain an expression for the average distribution of the amplitudes from the equations (9), then, just as in the analogous case in classical mechanics, the attempt will not succeed without a special supplementary hypothesis of a statistical nature. Like the fundamental equations of mechanics, the equations (9) are clearly unaffected by changing the sign of the time; this change can be compensated for by interchanging i and $-i$ (a change of sign of all the phases, corresponding to the change of sign of all the velocities in classical mechanics). This alone shows that there is no “equalising tendency” inherent in the resonance process itself. In fact, calculation shows that the time-averaged values of the squares of the amplitudes in general depend on their initial values. In order to obtain statistical expressions, a hypothesis as to the *a priori* probability of the initial values is therefore necessary. It appears that only *one* hypothesis is possible if it is to satisfy the following requirements:

(1) The hypothesis is to be independent of the *instant of time* for which it is stated; *i.e.* the probability that given values of the ampli-

tudes shall occur is not to alter with the passage of time, in consequence of the operation of equations (9).

(2) The hypothesis is to be independent of the particular one of the infinite number of completely equivalent orthogonal systems (arising from arbitrary orthogonal substitution among the proper functions corresponding to the same proper value) for which it is stated (cf. Q. III., p. 70 *et seq.*).

We may easily convince ourselves that in these circumstances the only possible assumption is the following: the probability density in a space in which the real and imaginary parts of the amplitudes are taken as rectangular co-ordinates is a function only of the *sums* of the squares of the amplitudes corresponding to the numerically different proper values.

It follows from this assumption that the average values of the squares of the amplitudes corresponding to the *same* proper value are equal, by symmetry; *i.e.* each partial sum of these quantities is now proportional to the number of terms in the sum. This is the only consequence that we shall employ in what follows, and that only for cases of extremely high degeneracy and only for partial sums with an extremely great number of terms.

We must refrain from the attempt, by means of anything analogous to the quasi-ergodic¹ hypothesis, to set up these average values as correct time-averages. The equations (9) are much too transparent to be satisfied by any such hypothesis (they possess at least a independent holomorphic integrals, namely, the squares of the amplitudes of the "normal vibrations"). The case is quite analogous to that of idealised solids, in which the constancy of the squares of the amplitudes of the normal vibrations really *ought* to prevent us from applying any statistical reasoning.

I should not like to leave unmentioned the fact that the same assumption regarding the squares of the amplitudes of the proper vibrations corresponding to the same proper value was necessary in the case of the *Stark effect*, in order that correct intensity ratios for the fine structure components should be obtained (cf. Q. III., p. 83).

§ 4. Arbitrary System in a Heat-bath

We return to the considerations of § 2. We will now assume that in the whole system the proper value (14) *only* is excited initially (and therefore permanently). Further, we will now assume that $E_k, E_{k'}, F_l, F_{l'}$, the four proper values of the separate systems which come into consideration, and which in § 2 we tacitly assumed to be *simple*, exhibit multiplicities of orders $a_k, a_{k'}, a_l, a_{l'}$. The proper value (14) then becomes $(a_k a_{l'} + a_{k'} a_l)$ -fold, for instead of the *two* degenerate proper functions (15) there appear *two groups*, in number $a_k a_{l'}$ and $a_{k'} a_l$ respectively. According to the statistical hypothesis of § 3, the

¹ Note to English edition. The "ergodic hypothesis" (Boltzmann) is what Maxwell called the "principle of continuity of path".

sum of the squares of the amplitudes of the first group bears the ratio

$$(17) \quad a_{k\alpha r} : a_{k\alpha l}$$

to the sum of the squares of the amplitudes of the second group. From what was said at the end of § 2, this is also the ratio of the sum of the squares of the amplitudes of all the proper vibrations corresponding to E_k to the sum of the squares of the amplitudes of all the proper vibrations corresponding to E_k in the first system, considered as isolated.

Thus, according to our statistical hypothesis, the interaction with the external system forces a quite definite value upon the initially *undetermined* ratio of the sums of squares of amplitudes corresponding to *different* proper values. This value is given by the "cross"-products of the degrees of degeneracy. (By "cross"-product the following is meant: the "upper" level of the system itself is to be combined with the lower level of the external system, and *vice versa*).—For brevity we will from this point call the sum of the squares of the amplitudes corresponding to a proper value the *excitation strength* of that proper value.

We now pass on to a somewhat more complicated case. We shall still, however, abide by the condition that in the whole system only *one* proper value, which we shall call E , is permanently excited. The second system (ϕ_l, F_l), which we will now call the *heat-bath*, is to be an extremely large system with an exceedingly dense proper value spectrum, such that for *every* E_k of the first system, which we will call the *thermometer*, there always exists a proper value of the heat-bath, F_r , satisfying the equation

$$(18) \quad F_r = E - E_k,$$

where F_r is to be multiple, even to a high order.

Hence a quite definite ratio is given perforce to the excitation strengths of all the proper values E_k of the thermometer; in fact, they are proportional to the products

$$(19) \quad a_{k\alpha r}.$$

The ratios of the a_r 's, however, can be determined in a very general way. For the question of the multiplicity a_r of the proper value F_r of the heat-bath, *i.e.* of the number of essentially different proper functions of the heat-bath which correspond to this proper value, is clearly identical with the question: in how many essentially different ways could we dispose of the *energy* F_r in the heat-bath *if* the latter were "energy-quantised"? This, however, is exactly the question which would be raised in connection with the calculation of the *entropy* of the heat-bath according to Planck's quantum statistics; in the latter, the entropy is taken to be equal to k times the logarithm of the number sought (k is Boltzmann's constant). The only difference¹ is that it is *sufficient* if we put the question in terms of a hypo-

¹ There are also of course the well-known small differences in the special specification of the "energy levels" by the new quantum mechanics as compared with that

thetical period—the result of the enumeration is of course independent of the form of the statement.

That is, we have

$$k \log a_r = S(E - E_k),$$

where the right-hand side denotes the entropy which the heat-bath with energy $E - E_k$ possesses, according to Planck's quantum statistics. By (19), the excitation strengths of the proper values E_k of the thermometer are therefore proportional to the quantities

$$(20) \quad a_k e^{\frac{1}{k} S(E - E_k)}$$

(excuse the occurrence of the letter k in a different sense). Now if the heat-bath is very large, we may set

$$(21) \quad \begin{cases} S(E - E_k) = S(E) - \left(\frac{\partial S}{\partial E}\right)_E \cdot E_k \\ \quad \quad \quad = S(E) - \frac{E_k}{T}, \end{cases}$$

where T denotes the temperature of the heat-bath for the energy E , calculated according to Planck. That is, instead of the ratios (20) we may use the following,

$$(22) \quad a_k e^{-\frac{E_k}{kT}}.$$

Thus we have obtained the important result: The average excitation strengths of the proper values of a system in the heat-bath are proportional to the relative numbers—according to the old quantum statistics—of the members of a canonical aggregate, which occur in the separate states considered as quantised. The multiplicities of the proper values of the system in question appear as “quantum weights”.

We can also get rid of the original assumption that a single proper value E is excited in the *whole* system. This procedure corresponds exactly to that made in classical statistics when we start from a micro-canonical aggregate and prove that a small partial system is distributed canonically in phase. If we wish, however, we can always make a canonical distribution for the whole system in addition; the result for the partial system remains unaltered. Of course the same is true in our case also.

The result (22) should in principle suffice to enable us to transfer all the important results of the old quantum statistics, in particular the statistical theory of gases, of solids, and of the “hohlraum” (Planck's radiation formula)—since all can be based on this formula—

by the old (“half-integral” quantisation, etc.). Further, we must remark that as regards what writers at present like to call the *kind of statistics* (Bose and Einstein, Fermi, etc.) absolutely nothing is prejudiced by the very general developments of the text. The distinction only appears when we give effect to a Pauli or a Heisenberg prohibition for the proper functions, or when we come to regard certain distributions of energy as essentially different, or not, in Planck's enumeration.

into the new theory without difficulty; of course the larger or smaller alterations alluded to in the footnote on pp. 144-5 must be made. I should like to lay special emphasis on the fact that this transference is *possible*, even *without* the support of the quantum postulates.

If the reader likes, however, he can understand everything that has been said in this paper in accordance with Born's theory,¹ in which the postulates are retained and the squares of the amplitudes are not interpreted as simultaneous excitation strengths in the single system, but merely as probabilities (relative frequencies of occurrence) of the discrete quantum states in a virtual aggregate. I have tried to think over the question whether we might from this point of view be able to do without the statistical hypothesis of § 3. This does not seem to be the case. According to Born, the alteration of the "probability field" as time goes on is compulsorily (causally) controlled by the wave equation, and consequently the alteration in time of the "probability amplitudes" is controlled by the equations (9). Hence the objection to reversal mentioned in § 3 now applies to the alteration in time of the probability amplitudes. So far as I can see, we can therefore never reach a one-way (irreversible) course without a supplementary hypothesis about the relative probability of the various possible distributions of the initial values of the probability amplitudes. I am averse to this conception, not so much on account of its complexity as on account of the fact that a theory which demands our assent to an absolute primary probability as a law of Nature should at least repay us by freeing us from the old "ergodic difficulties" and enabling us to understand the one-way course of natural processes without further supplementary assumptions.

Zürich, Physical Institute of the University.

(Received June 10, 1927.)

¹ M. Born, *Ztschr. f. Phys.* 37, p. 863; 38, p. 803; 40, p. 167, 1926.

Four Lectures on Wave Mechanics

Four Lectures on
Wave Mechanics

Delivered at the Royal Institution, London, on
5th, 7th, 12th, and 14th March, 1928

BY

DR. ERWIN SCHRÖDINGER

Professor of Theoretical Physics in the
University of Berlin

Dedicated
to the memory of
Fritz Hasenöhrl

Contents

FIRST LECTURE

1. Derivation of the fundamental idea of wave mechanics from Hamilton's analogy between ordinary mechanics and geometrical optics
2. Ordinary mechanics only an approximation, which no longer holds for very small systems
3. Bohr's stationary energy-levels derived as the frequencies of proper vibrations of the waves

SECOND LECTURE

4. Rough description of the wave-systems in the hydrogen atom. Degeneracy. Perturbation
5. The physical meaning of the wave function. Explanation of the selection rules and of the rules for the polarization of spectral lines
6. Derivation of the wave equation (properly speaking) which contains the time
7. An atom as perturbed by an alternating electric field

THIRD LECTURE

8. Theory of secondary radiation and dispersion
9. Theory of resonance radiation, and of changes of the state of the atom produced by incident radiation whose frequency coincides, or nearly coincides, with a natural emission frequency

10. Extension of wave mechanics to systems other than a single mass-point
11. Examples: the oscillator, the rotator

FOURTH LECTURE

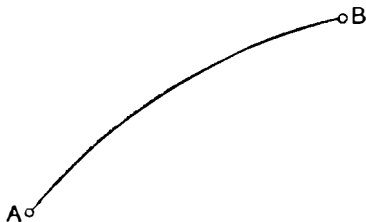
12. Correction for motion of the nucleus in the hydrogen atom
13. Perturbation of an arbitrary system
14. Interaction between two arbitrary systems
15. The physical meaning of the generalized ψ -function

Four Lectures on Wave Mechanics

FIRST LECTURE

1. Derivation of the fundamental idea of wave mechanics from Hamilton's analogy between ordinary mechanics and geometrical optics.

When a *mass-point* m moves in a conservative field of force, described by the potential energy $V(x, y, z)$, then, if you let it start from a given point A with a given velocity, i.e. with a given energy E , you will be able to get it into another arbitrarily chosen point B by suitably "aiming", i.e. by letting it start in a quite definitely chosen *direction*. There is in general *one* definite dynamical orbit which leads from A to B *with a given energy*. This orbit possesses the property that



$$\delta \int_A^B 2T dt = 0, \quad . . . \quad (1)$$

and is defined by this property (Hamilton's principle in the form given to it by Maupertuis). Here T means the

kinetic energy of the mass-point, and the equation means: consider the manifold of *all* orbits leading from *A* to *B* and subject to the law of conservation of energy ($T + V = E$); among them the actual dynamical orbit is distinguished by the fact that, *for it* and for all infinitely adjacent orbits of the manifold, the \int_A^B has the *same* value up to small quantities of the *second* order (the words “infinitely adjacent” being taken to define the *first* order of smallness). Calling $w = \frac{ds}{dt}$ the velocity of the mass-point, we have

$$2T = mw^2 = m \left(\frac{ds}{dt} \right)^2 = 2(E - V) = \frac{ds}{dt} \sqrt{2m(E - V)},$$

by means of which equation (1) can be transformed into

$$\delta \int_A^B \sqrt{2m(E - V)} ds = 0. \quad . \quad . \quad (2)$$

This form has the advantage that the variational principle is applied to a purely geometrical integral, which does not contain the time-variable, and further, that the condition of constant energy is automatically taken care of.

Hamilton found it useful to compare equation (2) with *Fermat's* principle, which tells us that in an optically non-homogeneous medium the actual light rays, i.e. the tracks along which energy is propagated, are determined by the “law of minimum time” (as it is usually called). Let fig. 1 *now* refer to an optical medium of arbitrary non-homogeneity, e.g. the earth's atmosphere; then, if you have a searchlight at *A*, furnishing a well-defined beam, it will in general be possible to illuminate an arbitrarily chosen point *B* by suitably *aiming* at it with the searchlight. There is one definite light-path leading

from A to B , which obeys, and is uniquely defined by, the law

$$\delta \int_A^B \frac{ds}{u} = 0. \quad (3)$$

Here ds , as before, means the element of the path, and u is the velocity of light, a function of the co-ordinates x, y, z .

The two laws contained in equations (2) and (3) respectively become *identical*, if we postulate that

$$u = \frac{C}{\sqrt{2m(E - V)}}, \quad . . . (4)$$

where C must be independent of x, y, z but may depend on E . Thus we have made a mental picture of an optical medium, in which the manifold of possible light-rays coincides with the manifold of dynamical orbits of a mass-point m moving *with given energy* E in a field of force $V(x, y, z)$. The fact that u , the velocity of light, depends not only on the co-ordinates but also on E , the total energy of the mass-point, is of the utmost importance.

This fact enables us to push the analogy a step farther by picturing the dependence on E as dispersion, i.e. as a dependence on *frequency*. For this purpose we must attribute to our light-rays a definite frequency ν , depending on E . We will (arbitrarily) put

$$E = h\nu \quad (5)$$

(h being Planck's constant), without dwelling much on this assumption, which is very suggestive to modern physicists. Then this non-homogeneous and dispersive medium provides in its *rays* a picture of *all* the dynamical orbits of our particle. Now we can proceed a stage farther, putting the question: can we make a small

“point-like” *light-signal* move *exactly* like our mass-point? (Hitherto we have only secured the geometrical identity of *orbits*, quite neglecting the question of time-rate.) At first sight this seems impossible, since the velocity of the mass-point,

$$w = \frac{1}{m} \sqrt{2m(E - V)}, \quad . . . \quad (6)$$

is (along the path, i.e. with constant E) *inversely* proportional to the light-velocity u (see equation (4); C depends on E only). But we must remember that u is of course the ordinary *phase-velocity*, whereas a small light-signal moves with the so-called *group-velocity*, say g , which is given by

$$\frac{1}{g} = \frac{d}{d\nu} \left(\frac{\nu}{u} \right),$$

or, in our case, following equation (5), by

$$\frac{1}{g} = \frac{d}{dE} \left(\frac{E}{u} \right). \quad \quad (7)$$

We will try to make $g = w$. The only means we have at our disposal for this purpose is a suitable choice of C , the arbitrary function of E that appeared in equation (4). From (4), (6), and (7), the postulate $g = w$ becomes

$$\begin{aligned} \frac{d}{dE} \left(\frac{E \sqrt{2m(E - V)}}{C} \right) \\ = \frac{m}{\sqrt{2m(E - V)}} \equiv \frac{d}{dE} \left(\sqrt{2m(E - V)} \right); \end{aligned}$$

hence

$$\left(\frac{E}{C} - 1 \right) \sqrt{2m(E - V)}$$

is constant with respect to E . Since V contains the co-ordinates and C must be a function of E only, this

relation can obviously be secured in a general way only by making the first factor vanish. Hence

$$\frac{E}{C} - 1 = 0, \quad \text{or} \quad C = E,$$

which gives equation (4) the special form

$$u = \frac{E}{\sqrt{2m(E - V)}} \cdot \cdot \cdot \cdot \quad (8)$$

This assumption about phase-velocity is the only one which will secure absolute coincidence between the dynamical laws of motion of the mass-point and the optical laws of motion of light-signals in our imagined light-propagation. It is worth while mentioning that, according to (8),

$$u = \frac{\text{energy}}{\text{momentum}} \cdot \cdot \cdot \cdot \quad (8')$$

There is still *one* arbitrariness in the definition of u , viz.: E may obviously be changed by an arbitrary additive constant, if the same constant is added to $V(x, y, z)$. This arbitrariness cannot be overcome in the non-relativistic treatment and we are not going to deal with the relativistic one in the present lectures.

Now the fundamental idea of wave-mechanics is the following. The phenomenon, of which we believed we had given an adequate description in the old mechanics by describing the motion of a mass-point, i.e. by giving its co-ordinates x, y, z as functions of the time variable t , is to be described correctly—according to the new ideas—by describing a definite wave-motion, which takes place among waves of the type considered, i.e. of the definite frequency and velocity (and hence of the definite wavelength) which we ascribed to what we called “light” in

the preceding. The mathematical description of a wave-motion will be furnished not by a limited number of functions of the one variable t , but by a continuous manifold, so to speak, of such functions, viz. by a function (or possibly by several functions) of x , y , z , and t . These functions will be subject to a *partial* differential equation, viz. to some sort of *wave equation*.

The statement that what *really* happens is correctly described by describing a wave-motion does not necessarily mean exactly the same thing as: what really *exists* is the wave-motion. We shall see later on that in generalizing to an *arbitrary* mechanical system we are led to describe what really happens in such a system by a wave-motion in the generalized space of its co-ordinates (q -space). Though the latter has quite a definite physical meaning, it cannot very well be said to “exist”; hence a wave-motion in this space cannot be said to “exist” in the ordinary sense of the word either. It is merely an adequate mathematical description of what happens. It may be that also in the case of one single mass-point, with which we are now dealing, the wave-motion must not be taken to “exist” in *too* literal a sense, although the configuration space happens to coincide with ordinary space in this particularly simple case.

2. Ordinary mechanics only an approximation, which no longer holds for very small systems.

In replacing the ordinary mechanical description by a wave-mechanical description our object is to obtain a theory which comprises both ordinary mechanical phenomena, in which quantum conditions play no appreciable part, and, on the other hand, typical quantum phenomena. The hope of reaching this object resides in the following

analogy. Hamilton's wave-picture, worked out in the way discussed above, contains *something* that corresponds to ordinary mechanics, viz. the *rays* correspond to the mechanical *paths*, and *signals* move like *mass-points*. But the description of a wave-motion in terms of *rays* is merely an approximation (called "geometrical optics" in the case of light-waves). It only holds if the structure of the wave phenomenon that we happen to be dealing with is coarse compared with the wave-length, and as long as we are only interested in its "coarse structure". The detailed fine structure of a wave phenomenon can never be revealed by a treatment in terms of rays ("geometrical optics"), and there always exist wave-phenomena which are altogether so minute that the ray-method is of no use and furnishes no information whatever. Hence in replacing ordinary mechanics by wave mechanics we may hope on the one hand to retain ordinary mechanics as an approximation which is valid for the coarse "macro-mechanical" phenomena, and on the other hand to get an explanation of those minute "micro-mechanical" phenomena (motion of the electrons in the atom), about which ordinary mechanics was quite unable to give any information. At least it was unable to do so without making very artificial accessory assumptions, which really formed a much more important part of the theory than the mechanical treatment itself.*

* To give an example: the actual application of the rules for quantization to the several-electron problem was, strange to say, not hindered by the fact that nobody in the world ever knew how to *enunciate* them for a non-conditionally periodic system! We simply *took* the problem of several bodies to be conditionally periodic, though it was perfectly well known that it was not. This shows, I think, that ordinary mechanics was not made use of in a very serious manner, otherwise the said application would have been as impossible as the application of penal law to the motion of the planets.

The step which leads from ordinary mechanics to wave mechanics is an advance similar in kind to Huygens' theory of light, which replaced Newton's theory. We might form the symbolic proportion:

$$\begin{aligned} \text{Ordinary mechanics} &: \text{Wave mechanics} \\ &= \text{Geometrical optics} : \text{Undulatory optics.} \end{aligned}$$

Typical quantum phenomena are analogous to typical wave phenomena like diffraction and interference.

For the conception of this analogy it is of considerable importance that the failure of ordinary mechanics does occur in dealing with very *tiny* systems. We can immediately control the order of magnitude at which a complete failure is to be expected, and we shall find that it is exactly the right one. The wave-length, say λ , of our waves is (see equations (5) and (8))

$$\lambda = \frac{u}{\nu} = \frac{h}{\sqrt{2m(E - V)}} = \frac{h}{mv}, \quad \dots \quad (9)$$

i.e. Planck's constant divided by the momentum of the mass-point. Now take, for the sake of simplicity, a circular orbit of the hydrogen-model, of radius a , but not necessarily a "quantized" one. Then we have by ordinary mechanics (without applying quantum rules):

$$mva = n \frac{h}{2\pi},$$

where n is any real positive number (which for Bohr's quantized circles would be 1, 2, 3 . . . ; the occurrence of h in the latter equation is for the moment only a convenient way of expressing the order of magnitude). Combining the last two equations, we get

$$\frac{\lambda}{a} = \frac{2\pi}{n}.$$

Now in order that we may be justified in the application of ordinary mechanics it is necessary that the dimensions of the path calculated in this way should turn out to be large compared with the wave-length. This is seen to be the case as long as the "quantum number" n is large compared with unity. As n becomes smaller and smaller, the ratio of λ to a becomes less and less favourable. A complete failure of ordinary mechanics is to be expected precisely in the region where we actually meet with it, viz. where n is of the order of unity, as it would be for orbits of the normal size of an atom (10^{-8} cm.).

3. Bohr's stationary energy-levels derived as the frequencies of proper vibrations of the waves.

Let us now consider the wave-mechanical treatment of a case which is inaccessible to ordinary mechanics; say, to fix our ideas, the wave-mechanical treatment of what in ordinary mechanics is called the motion of the electron in the hydrogen atom.

In what way are we to attack this problem?

Well, in very much the same way as we would attack the problem of finding the possible movements (vibrations) of an elastic body. Only, in the latter case the problem is complicated by the existence of *two* types of waves, longitudinal and transverse. To avoid this complication, let us consider an elastic fluid contained in a given enclosure. For the pressure, p , say, we should have a wave equation

$$\nabla^2 p - \frac{1}{u^2} \ddot{p} = 0, \quad (10)$$

u being the *constant* velocity of propagation of longitudinal waves, the only waves possible in the case of a

fluid. We should have to try to find the most general solution of this partial differential equation that satisfies certain boundary conditions at the surface of the vessel. The standard way of solving is to try

$$p(x, y, z, t) = \psi(x, y, z) e^{2\pi i \nu t},$$

which gives for ψ the equation

$$\nabla^2 \psi + \frac{4\pi^2 \nu^2}{u^2} \psi = 0, \quad . . . \quad (10')$$

ψ being subject to the same boundary conditions as p . We then meet with the well-known fact that a regular solution ψ satisfying the equation and the boundary conditions cannot be obtained for *all* values of the coefficient of ψ , i.e. for *all* frequencies ν , but only for an infinite set of discrete frequencies $\nu_1, \nu_2, \nu_3, \dots, \nu_k, \dots$, which are called the characteristic or proper frequencies (*Eigenfrequenzen*) of the problem or of the body. Call ψ_k the solution (ordinarily unique apart from a multiplying constant) that belongs to ν_k , then—since the equation and the boundary conditions are homogeneous—

$$p = \sum_k c_k \psi_k e^{2\pi i (\nu_k t + \theta_k)} \quad . . . \quad (11)$$

will, with arbitrary constants c_k, θ_k , be a more general solution and will indeed be *the* general solution, if the set of quantities (ψ_k, ν_k) is complete. (For physical applications we shall of course have to use the real part of the expression (11).)

In the case of the waves which are to replace in our thought the motion of the electron, there must also be some quantity p , subject to a wave equation like equation (10), though we cannot yet tell the physical meaning of p . Let us put this question aside for the moment.

In equation (10) we shall have to put (see above)

$$u = \frac{E}{\sqrt{2m(E - V)}} \cdot \cdot \cdot \cdot \quad (8)$$

This is not a constant; it depends (1) on E , that is, essentially on the frequency ν ($= E/h$); (2) on the co-ordinates x, y, z , which are contained in the potential energy V . These are the two complications as compared with the simple case of a vibrating fluid body considered above. Neither of them is serious. By the first, the dependence on E , we are restricted in that we can apply the wave equation only to a function p whose dependence on the time is given by

$$p \sim e^{\frac{2\pi i E t}{h}},$$

whence

$$\ddot{p} = -\frac{4\pi^2 E^2}{h^2} p \cdot \cdot \cdot \cdot \quad (12)$$

We need not mind that, since it is precisely the same assumption (*Ansatz*) as would be made in any case in the standard method of solution. Substituting from (12) and (8) in (10) and replacing the letter p by ψ (to remind us that now, just as before, we are investigating a function of the co-ordinates only), we obtain

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2} (E - V) \psi = 0. \quad \cdot \cdot \quad (13)$$

We now see that the *second* complication (the dependence of u on V , i.e. on the co-ordinates) merely results in a somewhat more interesting form of equation (13) as compared with (10'), the quantity multiplying ψ being no longer a *constant*, but depending on the co-ordinates. This was really to be expected, since an equation that is to embody the mechanical problem cannot very well help

containing the potential energy of the problem. A simplification in the problem of the "mechanical" waves (as compared with the fluid problem) consists in the absence of boundary conditions.

I thought the latter simplification fatal when I first attacked these questions. Being insufficiently versed in mathematics, I could not imagine how proper vibration frequencies could appear *without* boundary conditions. Later on I recognized that the more complicated form of the coefficients (i.e. the appearance of $V(x, y, z)$) takes charge, so to speak, of what is ordinarily brought about by boundary conditions, namely, the selection of definite values of E .

I cannot enter into this rather lengthy mathematical discussion here, nor into the detailed process of finding the solutions, though the method is practically the same as in ordinary vibration problems, namely: introducing an appropriate set of co-ordinates (e.g. spherical or elliptical, according to the form of the function V) and putting ψ equal to a *product* of functions, each of which contains one co-ordinate only. I will state the result straightforwardly for the case of the hydrogen atom. Here we have to put

$$V = -\frac{e^2}{r} + \text{const.}, \quad . . . \quad (14)$$

r being the distance from the nucleus. Then it is found that not for all, but only for the following values of E , is it possible to find regular, one-valued, and finite solutions ψ :

$$\left. \begin{array}{l} \text{(A) } E_n = \text{const.} - \frac{2\pi^2 me^4}{h^2 n^2}; \quad n = 1, 2, 3, 4 \dots \\ \text{(B) } E > \text{const.} \end{array} \right\} \quad (14')$$

The constant is the same as in (14) and is (in non-relativistic wave mechanics) meaningless, except that we cannot very well give it the value which is usually adopted for the sake of simplicity, viz. zero. For then all the values (A) would become negative. And a negative frequency, if it means anything at all, means the same as the positive frequency of the same absolute value. Then it would be mysterious why all positive frequencies should be allowed, but only a discrete set of negative ones. But the question of this constant is of no importance here.

You see that our differential equation automatically selects as the allowed E -values (A) the energy-levels of the elliptic orbits quantized according to Bohr's theory; (B) all energy-levels belonging to hyperbolic orbits. This is very remarkable. It shows that, whatever the waves may mean physically, the theory furnishes a method of quantization which is absolutely free from arbitrary postulates that this or that quantity must be an integer. Just to give an idea *how* the integers occur here: if e.g. ϕ is an azimuthal angle and the wave amplitude turns out to contain a factor $\cos m\phi$, m being an arbitrary constant, then m must *necessarily* be chosen integral, since otherwise the wave function would not be *single-valued*.

You will be interested in the form of the wave functions ψ which belong to the E -values mentioned above, and will inquire whether any observable facts can be explained by them. This *is* the case, but the matter is rather intricate.

SECOND LECTURE

4. Rough description of the wave-systems in the hydrogen atom. Degeneracy. Perturbation.

The chief property of the amplitude functions is that those which belong to the discrete set of values E_n ("elliptic orbits") fall off very rapidly with the distance from the nucleus, viz. like an exponential $e^{-\text{const. } r}$, which practically restricts them to a region of precisely the same order of magnitude as the corresponding Bohr orbit. The others, which belong to hyperbolic levels, fall off much less rapidly, viz. only like r^{-1} .

The detailed behaviour of the "elliptic" functions within the said region cannot very well be described in a unique way, for the following reason. To *one* value E_n there belongs in general not only one, but precisely n^2 independent solutions of the wave equation. From the mathematical point of view this is an exception due to the particular form of the potential energy V , especially to its spherical symmetry. This multiplicity of solutions belonging to one proper value corresponds to the well-known multiplicity of orbits belonging to the same energy-level in Bohr's theory. It is there called "degeneracy", and we will keep this expression in wave mechanics also. Now, since the equation is linear and homogeneous, any linear aggregate with quite arbitrary coefficients will also be a solution belonging to *the same*

proper value. It is well known that in such a case no set of solutions is in any way distinguished from any other set, derived from the first by forming a set of independent linear aggregates, equal in number to the first set. By this process of forming linear aggregates we can reach solutions which exhibit a very different behaviour. To give an example: from a set of solutions whose node-surfaces are (1) concentric spheres, (2) co-axial cones, (3) planes passing through the cone-axis, you can form other solutions, in which the concentric spheres and the co-axial cones are replaced by two sets of confocal paraboloids. This is only one of the simplest cases. In general, taking arbitrary coefficients, the system of node-surfaces will be *much* more complicated.

This multiplicity of solutions, or, as is often said, of the proper values (which, by the way, is well known from ordinary vibration problems), is of the utmost importance in the case of the atom. If there is no multiplicity (e.g. for the lowest frequency, $n = 1$), then a slight alteration of the potential energy V , corresponding e.g. to the application of a weak external electric field, will cause nothing but a slight displacement of the proper value and a slight alteration of the proper solution—just as a small piece of metal attached to a tuning fork would slightly alter its pitch and its form of vibration. But a multiple (say α -fold) proper value shows its actual multiplicity in this case in that it splits up into α slightly different proper values; every one of them has now a quite definite proper function, which differs very little from a quite definite linear aggregate of the proper functions that belonged to the multiple value. This splitting up may, theoretically, be caused by the very slightest disturbance, yet may differ widely for two

disturbances that are different in character. For instance, a homogeneous electric field produces the parabolic node-surfaces mentioned before, whereas a magnetic field produces the spheres and cones.

It need hardly be said that this splitting up corresponds in the two cases just mentioned to the splitting up of the hydrogen lines in the Zeeman and Stark effects. The displacement of the lines is quantitatively described by the new theory just as it was by the older one. But something more is described, which was inaccessible to the older theory, namely, the state of *polarization* of the lines, their *intensities*, and, in particular, the *absence* of a lot of lines which we should expect to appear if we took into account *all* the possible differences of the split energy levels. We shall see this presently.

5. The physical meaning of the wave function. Explanation of the selection rules and of the rules for the polarization of spectral lines.

The high importance of the perturbation effects consists in the fact that as soon as the degeneracy is removed we have to deal with uniquely defined proper functions ψ_k and can now more easily test any hypothesis about the physical meaning of the quantity called ψ .

Let us call

$$E_k = h\nu_k \quad \text{and} \quad \psi_k(x, y, z)$$

the proper values, proper frequencies, and proper functions of a problem, whose potential energy V we suppose sufficiently unsymmetrical to do away with all degeneracy. Then

$$\psi = \sum_k c_k \psi_k e^{2\pi i(\nu_k t + \theta_k)}, \quad . . . \quad (15)$$

with arbitrary constants c_k , θ_k , will describe the most

general "vibration" of the system.* In order to avoid ambiguity, since every ψ_k in itself is only defined apart from an arbitrary multiplying constant, we shall subject the ψ_k 's to the normalizing condition

$$\iiint \psi_k^2 dx dy dz = 1. \quad . \quad . \quad (16)$$

Perhaps this is the place to mention a very important property which the ψ_k 's possess automatically, viz. they are "orthogonal" to each other:

$$\iiint \psi_k \psi_l dx dy dz = 0 \quad \text{for } k \neq l, \quad . \quad (17)$$

and they form a *complete* orthogonal set; a function which is orthogonal to them *all* must necessarily vanish. (These properties are important for the development of an arbitrary function in a series in terms of the ψ_k 's, but we shall not enter upon that here, as we do not need it for the moment.)

Now return to the general vibration function (15). We put the question: is it possible to ascribe a definite physical meaning to the quantity ψ in such a way that the emission of light with frequencies

$$\nu_{kk'} = \nu_k - \nu_{k'}$$

becomes intelligible? Yes, it is, but—strange to say—only if we make use of the *complex* ψ -function as it stands, instead of its real part, as we are accustomed to do in ordinary vibration problems.

* Here we have not taken into account the "continuous spectrum", corresponding to the hyperbolic orbits. We may either suppose these modes of vibration to be absent or we may take the Σ to include, as a limiting case, the integral, which would have to be added in order to take proper account of the continuous region of proper values. At all events I wished to avoid encumbering the formulæ more than necessary.

The hypothesis which we have to admit is very simple, namely that the square of the absolute value of ψ is proportional to an electric density, which causes emission of light according to the laws of ordinary electrodynamics. Since the square of the absolute value of ψ is formed by multiplying ψ by the conjugate complex quantity (which we will call $\bar{\psi}$), a glance at the expression (15) shows that the terms which compose $\psi\bar{\psi}$ contain the time in the form of cosine factors of the desired frequencies $\nu_k - \nu_{k'}$. More precisely, let us put, for the charge-density ρ ,

$$\rho = -e\psi\bar{\psi} = -e \sum_k \sum_{k'} c_k c_{k'} \psi_k \bar{\psi}_{k'} e^{2\pi i[(\nu_k - \nu_{k'})t + \theta_k - \theta_{k'}]}, \quad (18)$$

where e means the absolute electronic charge. Integrating this over the whole space and making use of equations (16) and (17), we find for the *total* charge

$$-e \sum_k c_k^2,$$

which shows that we shall have to postulate

$$\sum_k c_k^2 = 1$$

in order to make the total charge equal to the electronic charge (which we feel inclined to do).

It was said before that ψ , and hence ρ , is practically confined to a very small region of a few Ångstrom units. Since the wave-lengths of the light-radiations $\nu_k - \nu_{k'}$ are very large compared with this region, it is well known that the radiation of the fluctuating density ρ will be very nearly the same as that of an electric dipole whose electric moment has (e.g.) the z -component

$$\mathbf{M}_z = \iiint z \rho \, dx \, dy \, dz$$

(and similarly formed x - and y -components). Calcu-

lating M_z from (18), we find after an easy reduction

$$M_z = - \sum_k c_k^2 a_{kk} - 2 \sum_{(k, k')} c_k c_{k'} a_{kk'} \cos [2\pi(\nu_k - \nu_{k'})t + \theta_k - \theta_{k'}]. \quad (19)$$

Here $a_{kk'}$ is an abbreviation for the following *constant*:

$$a_{kk'} = e \iiint z \psi_k \psi_{k'} dx dy dz, \quad . \quad . \quad (20)$$

and $\Sigma_{(k, k')}$ means a sum over all the *pairs* (k, k') . Hence the

squares of these integrals (and the corresponding integrals relating to the x - and y -directions) determine the intensity of emitted light of frequency $|\nu_k - \nu_{k'}|$. The intensity is not determined by them alone; the amplitude-constants c_k also come into play, of course. But this is quite satisfactory. For the integrals $a_{kk'}$ are determined by the *nature* of the system, i.e. by its proper functions, regardless of its *state*. $a_{kk'}$ is the amplitude of the corresponding dipole, which would be produced by the proper vibrations $\psi_k, \psi_{k'}$, if only these two were excited, and with equal strength $\left(c_k = c_{k'} = \frac{1}{\sqrt{2}}\right)$.

The first sum in (19) is of no interest in our investigation of the emitted radiation, since it means a component of electric moment that is constant in time.

The correctness of our $\psi\bar{\psi}$ -hypothesis has been checked by calculating the $a_{kk'}$'s in those cases where the ψ_k 's are sufficiently well defined, namely in the case of the Zeeman and Stark effects. The so-called rules of selection and polarization and the intensity-distribution in these patterns are described by the $a_{kk'}$'s in the following obvious way, and the description is in complete agreement with experiment:

The *absence* of a line which might be expected to occur ("selection-rule") is described by the *vanishing* of the corresponding $a_{kk'}$, and of the two other constants relating to the x - and y -directions.

The *linear polarization* of a line in a definite direction is described by the fact that *only* the constant $a_{kk'}$ relating to *this* direction differs from zero, whereas the two other constants vanish. In a similar way the circular polarization, say in the xy -plane, is indicated by (1) vanishing of the z -constant, (2) equality of the x - and y -constants, and (3) a phase-difference of $\pi/2$ between the corresponding cosine-functions in equation (19).

Finally, the *intensity relations* between the non-vanishing components in the Stark and Zeeman patterns of hydrogen are correctly indicated by the relations between the squares of the $a_{kk'}$'s in question; which is satisfactory, since the assumption that the c_k 's will be equal for the fine-structure components of one *level* is very suggestive, notwithstanding our lack of knowledge of the c_k 's in other respects.

Of course it is impossible to set forth in this lecture any of the calculations that led to the results just given; they would fill pages and pages, and are not at all difficult, but *very* tedious. In spite of their tediousness, it is rather fascinating to see all the well-known but not understood "rules" come out one after the other as the result of *very* familiar elementary and absolutely cogent analysis, like e.g. the fact that $\int_0^{2\pi} \cos m\phi \cos n\phi d\phi$ vanishes unless $n = m$. Once the hypothesis about $\psi\bar{\psi}$ has been made, no accessory hypothesis is needed or is possible; none could help us if the "rules" did not come out correctly. But fortunately they do.

I think I ought to draw attention to another fact which was only briefly mentioned at the beginning, namely, that the very fundamental "frequency-rule" of Bohr,

$$\nu_{kk'} = \nu_k - \nu_{k'} = \frac{1}{h}(E_k - E_{k'}),$$

may also be said to be explained by the $\psi\bar{\psi}$ -hypothesis. Something exists in the atom which actually vibrates with the *observed* frequency, viz. a certain part of the electric density-distribution or, if you prefer, of $\psi\bar{\psi}$.

This might lead us to suspect that only the square of its absolute value, and not the ψ -function itself, has a real meaning. And this suspicion again might arouse the desire to replace the wave equation by an equation which describes the behaviour of $\psi\bar{\psi}$ directly. To remove this desire, I will remind you of a case in which a similar desire might occur for exactly similar reasons; yet all of you will confess that it would be fatal to pursue it.

Maxwell's equations describe the behaviour of the electromagnetic vectors. But these are not really accessible to observation. The only things that are observable are the ponderomotive forces, or, if you please, the energy, since the forces are caused by virtual energy-differences. But all these quantities (energy, Maxwellian-stresses) are *quadratic* functions of the field-vectors. Therefore we might desire to replace Maxwell's equations by others, that determine the observable *quadratic* functions of the field-vectors directly. But everyone will agree that this would at all events mean an immense complication and that it would not really be possible to do without Maxwell's equations.

6. Derivation of the wave equation (properly speaking) which contains the time.

The equation

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2} (E - V) \psi = 0, \quad . \quad (13)$$

which we have used for the investigation of the hydrogen atom, only furnishes the distribution in space of the amplitude of the vibration, the dependence on time always being given by

$$\psi \sim e^{\frac{2\pi i E t}{h}} . \quad . \quad . \quad (21)$$

The value of the frequency, E , is present in the equation, so that we are really dealing with a family of equations, each of the members being valid for one particular frequency only. The state of things is exactly the same as in ordinary vibration problems; our equation corresponds to the so-called "amplitude equation" (see section 3, equation (10')),

$$\nabla^2 \psi + \frac{4\pi^2 \nu^2}{u^2} \psi = 0, \quad . \quad . \quad (10')$$

and not to

$$\nabla^2 p - \frac{1}{u^2} \ddot{p} = 0, \quad . \quad . \quad (10)$$

from which the former is derived in the manner described above (namely by supposing p to be a sine-function of the time). In our case the problem is to make the analogous step in the reverse direction, i.e. to remove the parameter E from the amplitude equation and introduce time-derivatives instead. This is easily done. Take *one* of the family (13) (with a particular value of E), then by (21) we have

$$\dot{\psi} = \frac{2\pi i E}{h} \psi \quad \text{or} \quad E\psi = \frac{h}{2\pi i} \dot{\psi}.$$

Using this, we get from (13)

$$\nabla^2\psi - \frac{4\pi mi}{h} \dot{\psi} - \frac{8\pi^2 mV}{h^2} \psi = 0. \quad (22)$$

The *same* equation is reached whatever the value of E may have been (for E has been eliminated). Hence equation (22) will be valid for an arbitrary linear aggregate of proper vibrations, i.e. for the most general wave-motion that is a solution of the problem.

We may tentatively go a step farther and try to use it also in the case where the potential energy V contains the time-variable explicitly. It is by no means obvious that this is a correct generalization, for terms with \dot{V} , &c., might be missing—they could not possibly enter into equation (22), in view of the way we have reached this equation. But success will justify our procedure. Of course it would have been nonsense to introduce the assumption that V contained the time explicitly in equation (13), since the condition (21), by which this equation is restricted, would make it impossible to satisfy (13) in the case of an arbitrarily varying V -function.

7. An atom as perturbed by an alternating electric field.

This generalization enables us to solve the important problem: how does an atom behave under the influence of an external alternating electric field, i.e. under the influence of an incident wave of light? This is a very important question: for it contains not only the mechanism of secondary radiation and, in particular, of resonance-radiation, but also the theory of the changes of state of the atom under the influence of an incident radiation of appropriate frequency, and in addition the theory of

refraction and dispersion; for it is well known that dispersion—I mean the phenomenon of a refractive index—is caused by the superposition on the primary radiation of all the secondary wavelets, which every single atom of the body emits under the action of, and in phase with, the primary radiation. If an incident electric vector \mathbf{E} causes every atom to emit a secondary wavelet *such* as would be emitted by a dipole of the electric moment

$$\mathbf{M} = \alpha \mathbf{E} \quad (23)$$

(α being a constant), and if Z atoms are present in unit volume, then they produce an increase in the refractive index of

$$2\pi Z\alpha. \quad (24)$$

Hence, studying the value of α (which usually depends on the frequency) means studying the phenomena of refraction and dispersion.

To investigate the behaviour of an atom in an external alternating electric field, let us take V in equation (22) to be composed of two parts, one describing the internal electrostatic field of the atom, V_0 , and one describing the light-field, $Ae^z \cos 2\pi\nu t$; A , ν mean the amplitude and the frequency of the light-field, which we suppose polarized in the direction of z . (The negative sign of the electronic charge has been taken account of; our e is a positive number.) Hence equation (22) becomes

$$\nabla^2 \psi - \frac{4\pi m i}{h} \dot{\psi} - \frac{8\pi^2 m}{h^2} (V_0 + Ae^z \cos 2\pi\nu t) \psi = 0. \quad (25)$$

We shall take A to be very small compared with the internal field (described by V_0) and solve the equation by approximation. If A were zero, by assuming (21) we

should get back to equation (13) (only with the notation V_0 instead of V). We shall assume the problem of the unperturbed atom to be completely solved, its normalized proper functions and proper values being

$$\psi_k \text{ and } E_k (= h\nu_k).$$

Hence the most general solution of (25), when $A = 0$, is

$$\psi = \sum_k c_k \psi_k e^{2\pi i \nu_k t}, \quad (26)$$

the c_k 's being arbitrary complex constants.

We shall try to satisfy equation (25), with A also present, by (26), but with the c_k 's varying slightly with time (method of variation of constants). Taking this into account, and also the facts that ψ_k , $h\nu_k$ are proper functions and proper values of the unperturbed equation, we easily obtain, by substituting (26) in (25):

$$\sum_k \dot{c}_k \psi_k e^{2\pi i \nu_k t} = \frac{2\pi i}{h} A e z \cos 2\pi \nu t \sum_k c_k \psi_k e^{2\pi i \nu_k t}. \quad (27)$$

This equation will be satisfied if (identically with respect to time) *all* the coefficients of the expansion of its left-hand side with respect to the *complete* system of orthogonal functions, ψ_k , are identical with the respective coefficients of the expansion of its right-hand side. Hence multiply by ψ_l and integrate over the whole space. Put for abbreviation (see section 5):

$$a_{kl} = e \iiint \psi_k \psi_l z \, dx \, dy \, dz. \quad (20)$$

Then, owing to the normalization and orthogonality of the ψ_k 's, we get

$$\begin{aligned} \dot{c}_l e^{2\pi i \nu_l t} &= \frac{2\pi i}{h} A \cos 2\pi \nu t \sum_k a_{kl} c_k e^{2\pi i \nu_k t} \\ & \quad (l = 1, 2, 3, 4, \dots). \quad . . . (28) \end{aligned}$$

This infinite set of ordinary differential equations is *equivalent* to (27). Isolating \dot{c}_l and splitting up the cosine into exponentials, we write it as follows:

$$\dot{c}_l = + \frac{\pi i A}{h} \sum_k a_{kl} c_k [e^{2\pi i(\nu_k - \nu_l + \nu)t} + e^{2\pi i(\nu_k - \nu_l - \nu)t}]. \quad (28')$$

Hitherto we have not made use of any approximation process. We will now do this in two different ways, one leading to the theory of secondary radiation (excluding the case of resonance) and of dispersion, the other furnishing the case of resonance and the changes of state of the atom.

THIRD LECTURE

8. Theory of secondary radiation and dispersion.

In equation (28') we shall at first assume that all the aggregates

$$\nu_k - \nu_l \pm \nu,$$

which appear in the exponents are *large* as compared with the order of magnitude of

$$\frac{Aa_{kl}c_k}{h}.$$

This means that the *difference* between the incident frequency and any one of the frequencies of spontaneous emission is large compared with the frequency that would correspond to the *potential energy* which the atom acquires in the external field. (Exclusion of exact or near *resonance*.) With this assumption the equations (28') show that *all* the time-derivatives of the c_l 's are small compared to the time-derivatives of the exponentials. After having stated this, let us take any one of the exponentials on the right-hand side of any one of the equations (28'). We may assume its coefficient c_k to be constant during a period of the exponential. Hence *this* term will only cause a *small* periodic oscillation of the c_l (on the left), which is restored (or nearly so) after

the exponential has gone through a period. But the same holds for *all* the exponentials. Hence *all* the c 's execute a vast number of small oscillations around their mean values, oscillations which would of course vanish with vanishing A . We may therefore replace the c 's on the *right-hand* side of equation (28') by constants, viz. by their mean values, since by neglecting the small oscillations here only terms in A^2 are dropped. We shall write c_k^0 for the said constants. The equations are now easily integrated. We get

$$c_l = c_l^0 + \frac{A}{2h} \sum_k a_{kl} c_k^0 \left[\frac{e^{2\pi i(\nu_k - \nu_l + \nu)t}}{\nu_k - \nu_l + \nu} + \frac{e^{2\pi i(\nu_k - \nu_l - \nu)t}}{\nu_k - \nu_l - \nu} \right].$$

Hence the l th term in our solution (26) will be:

$$c_l \psi_l e^{2\pi i \nu_l t} = c_l^0 \psi_l e^{2\pi i \nu_l t} + \frac{A \psi_l}{2h} \sum_k a_{kl} c_k^0 \left[\frac{e^{2\pi i(\nu_k + \nu)t}}{\nu_k - \nu_l + \nu} + \frac{e^{2\pi i(\nu_k - \nu)t}}{\nu_k - \nu_l - \nu} \right]. \quad (29)$$

Though we have not yet reached a point that can be compared with experiment, we will give in words the description of what happens, according to equation (29), under the influence of an incident light-wave. Every proper vibration ψ_l , whether it is itself excited from the beginning or not, is compelled to execute a multitude of small additional forced oscillations, namely two "in honour" of every proper vibration ψ_k that *is* excited appreciably ($c_k^0 \neq 0$). The frequencies of the two forced oscillations that ψ_l executes "in honour" of ψ_k , as we said, are $\nu_k \pm \nu$, i.e. the sum and difference of the incident frequency and the frequency of the "honoured" proper vibration. Their amplitudes are proportional to the amplitudes both of the external field and of the "honoured" vibration; they also contain as a factor a_{kl} , the constant

which governs the intensity of the spontaneous emission of frequency $|\nu_k - \nu_l|$. Further, in the two forced amplitudes two "resonance-denominators" appear, causing one of the two amplitudes to increase rapidly when the incident frequency approaches the frequency of spontaneous radiation $|\nu_k - \nu_l|$.

Before forming the complete solution from (26) and (29) we will restrict ourselves to the most important case, viz. that in which only *one* free vibration is excited, say ψ_k :

$$c_k^0 = 1 \quad c_l^0 = 0 \quad \text{for } l \neq k.$$

We may think of ψ_k as corresponding to the normal state. Then on the right-hand side of equation (29) the first term (except for $l = k$) and the summation sign are dropped, and we get for the complete solution (equation (26), in which k is to be replaced by l):

$$\psi = \psi_k e^{2\pi i \nu_k t} + \frac{A}{2\hbar} \sum_l a_{kl} \psi_l \left[\frac{e^{2\pi i (\nu_k + \nu) t}}{\nu_k - \nu_l + \nu} + \frac{e^{2\pi i (\nu_k - \nu) t}}{\nu_k - \nu_l - \nu} \right]. \quad (30)$$

(Note that *now* the exponentials are independent of the index of summation, l ; only *two* frequencies of forced vibration are present.)

To get information about the secondary radiation we form the component * \mathbf{M}_z of the resultant electric moment from (30). Neglecting small terms of the second order (proportional to A^2), we find after reduction:

$$\begin{aligned} \mathbf{M}_z = & -e \iiint \psi \bar{\psi} z \, dx \, dy \, dz = -a_{kk} \\ & + \frac{2}{\hbar} A \cos 2\pi \nu t \sum_l \frac{(\nu_l - \nu_k) a_{kl}^2}{(\nu_l - \nu_k)^2 - \nu^2}. \end{aligned} \quad (31)$$

* In general, for an anisotropic atom, there will be an \mathbf{M}_y and an \mathbf{M}_x (orthogonal to the polarization of the incident radiation) as well. We will not deal with them here.

The first term ($-a_{kk}$) is independent of the time; it is the constant electric moment due to the excitation of the free vibration ψ_k . It is of no interest here. The second term determines the secondary wavelet. It is seen to coincide in frequency with the incident electric force ($A \cos 2\pi\nu t$). Its phase is the same or opposite, depending on whether $\nu \lesseqgtr \nu_l - \nu_k$, just as in the classical theory. (This holds if ψ_k corresponds to the normal state, so that $\nu_l - \nu_k$ is always positive; if it is negative, the reverse is true; Kramers' terms of the dispersion formula.) The quantity α of equation (23), which by the expression (24) determines the contribution to the refractive index, is found from the second term on the right-hand side of (31) by dropping $A \cos 2\pi\nu t$. The denominators $(\nu_l - \nu_k)^2 - \nu^2$ furnish the phenomenon of anomalous dispersion in the neighbourhood of all *those emission* (or absorption) frequencies that involve the index k of ψ_k —remember that we supposed only this *one* free vibration to be excited. The quantity a_{kl}^2 in the numerator is the same as that which determines the intensity of spontaneous emission $|\nu_k - \nu_l|$. In all these respects the formula is a complete copy of the old Helmholtz formula (supplemented by Kramers' "negative" terms) and is thought to be in complete agreement with experiment.

Two additional points are worth mentioning. You know that Thomas and Kuhn formed a hypothesis concerning the *sum* of all the coefficients in the dispersion formula, in our case

$$\frac{2}{h} \sum_l (\nu_l - \nu_k) a_{kl}^2.$$

According to them it is to be equal to the value of the

coefficient for *one* elastically bound electron, i.e. it must be equal to

$$\frac{e^2}{4\pi^2 m}$$

(multiplied by *one*, in our case, for we are dealing with the *one*-electron atom; in general, multiplied by an integer). The equality of the two above-mentioned quantities can be *proved* for our dispersion formula—but the proof is a little lengthy, and I will therefore omit it.

The second remark is the following. Perhaps you remember the statement, first made by Smekal, that there should also exist secondary radiations, whose frequencies differ from the frequency ν of the incident radiation (*therefore* without phase relation, *therefore* without influence on the refraction phenomenon). The frequencies expected are

$$\nu \pm (\nu_k - \nu_{k'}).$$

Secondary radiations of precisely these frequencies are furnished by the present theory, if we give up our simplifying assumption that only *one free* vibration is excited, and suppose at least two of them, say ψ_k and $\psi_{k'}$, to be present.

9. Theory of resonance radiation, and of changes of the state of the atom produced by incident radiation whose frequency coincides, or nearly coincides, with a natural emission frequency.

At the beginning of the last section we had to make the assumption that all the aggregates like

$$\nu_k - \nu_l \pm \nu$$

are of appreciable size, which means that the frequency

of the incident light, ν , is excluded from the immediate neighbourhood of any natural frequency of the atom under consideration. We will now consider an incident frequency which is very close to one of the natural frequencies. To fix our ideas, let

$$\nu_k - \nu_l + \nu \text{ be very small and } \nu_l > \nu_k$$

(“very small” means: of the order of magnitude of Aa_{kl}/h or smaller, possibly vanishing). Returning to equation (28'), you will now find on the right-hand side of this system of equations altogether *two* exponentials which vary slowly, viz.

$$e^{2\pi i(\nu_k - \nu_l + \nu)t} \quad \text{and} \quad e^{2\pi i(\nu_l - \nu_k - \nu)t},$$

the former appearing in the l th equation, the latter in the k th equation. These terms (as we shall see presently) now cause very appreciable “secular” changes in the two quantities c_k and c_l , however small the amplitude A of the incident wave may be. All the other exponentials will only cause small periodic disturbances, as before. It is therefore reasonable to drop them altogether, since we are now dealing with a much coarser phenomenon (viz. appreciable secular variations of c_k and c_l). We might even suppose all the other c 's to be zero; this would have no effect, since they are certainly constant within the degree of accuracy we are aiming at. For determining c_k and c_l we get from (28') the two simple equations

$$\left. \begin{aligned} \dot{c}_l &= i\sigma c_k e^{i\epsilon t}, \\ \dot{c}_k &= i\sigma c_l e^{-i\epsilon t}, \end{aligned} \right\} \dots \dots \dots (32)$$

with the abbreviations

$$\sigma = \frac{\pi A a_{kl}}{h}, \quad \epsilon = \nu_k - \nu_l + \nu. \quad \dots \dots (33)$$

To solve them, we introduce new variables x, y by putting

$$c_l = x e^{\frac{i\epsilon t}{2}}, \quad c_k = y e^{-\frac{i\epsilon t}{2}}. \quad \dots \quad (34)$$

The result can be written

$$\left(\frac{d}{dt} + \frac{i\epsilon}{2}\right)x = i\sigma y,$$

$$\left(\frac{d}{dt} - \frac{i\epsilon}{2}\right)y = i\sigma x.$$

These equations have *constant* coefficients and are readily solved by familiar methods. The solution can be written in the following form:

$$\left. \begin{aligned} x &= \rho e^{i(\gamma t + \phi)} + \mu \rho' e^{-i(\gamma t + \phi')}, \\ y &= \mu \rho e^{i(\gamma t + \phi)} - \rho' e^{-i(\gamma t + \phi')}, \end{aligned} \right\} \dots \quad (35)$$

with the abbreviations

$$\gamma = \sqrt{\frac{\epsilon^2}{4} + \sigma^2}, \quad \mu = \frac{\gamma + \frac{\epsilon}{2}}{\sigma}, \quad \dots \quad (36)$$

whereas ρ, ρ', ϕ, ϕ' are *arbitrary* real constants, non-negative if you like. We can put (35) in the form:

$$\left. \begin{aligned} x &= e^{\frac{i(\phi - \phi')}{2}} [(\rho + \mu \rho') \cos \theta + i(\rho - \mu \rho') \sin \theta], \\ y &= e^{\frac{i(\phi - \phi')}{2}} [(\mu \rho - \rho') \cos \theta + i(\mu \rho + \rho') \sin \theta], \end{aligned} \right\} \quad (37)$$

with the abbreviation

$$\theta = \gamma t + \frac{\phi + \phi'}{2}. \quad \dots \quad (38)$$

From (37) we can easily form the squares of the absolute values of x and y , that is (by equation (34)) of c_l and c_k , and we can thus get information about the varying dis-

tribution of intensity between the two vibrations in question—which is the point of main interest. We obtain

$$\left. \begin{aligned} |c_l|^2 &= |x|^2 = (\rho - \mu\rho')^2 + 4\mu\rho\rho' \cos^2 \theta, \\ |c_k|^2 &= |y|^2 = (\mu\rho - \rho')^2 + 4\mu\rho\rho' \sin^2 \theta. \end{aligned} \right\} \quad (39)$$

The sum of the intensities is constant, as might have been anticipated. It may be taken to consist of three parts, two "portions" fixed invariably to the two vibration-levels, the third (viz. $4\mu\rho\rho'$) oscillating slowly between them. To fix our ideas, let us take the case where at a certain time all the intensity was stored up in *one* vibration, say the lower one, c_k . Choosing the corresponding value of t so as to make $\cos \theta = 0$, this requires

$$\rho' = \frac{\rho}{\mu}.$$

We then find for the ratio between the oscillating portion of the intensity and its total amount

$$\frac{4\mu\rho\rho'}{(\mu\rho + \rho')^2} = \frac{4}{\left(\mu + \frac{1}{\mu}\right)^2} = \frac{\sigma^2}{\gamma^2} = \frac{\sigma^2}{\sigma^2 + \frac{\epsilon^2}{4}} \quad (40)$$

(by using the fact, obvious from (36), that

$$\mu = \frac{\gamma + \frac{\epsilon}{2}}{\sigma} = \frac{\sigma}{\gamma - \frac{\epsilon}{2}}).$$

We see that when $\epsilon = 0$ the *total* intensity is oscillating. By (33), $\epsilon = 0$ means the case of *sharp* resonance. If the resonance is not complete, then (40) shows that only a certain fraction of the intensity oscillates, and that this fraction becomes inappreciable when the lack of resonance, ϵ , becomes large compared with the quantity

σ defined by (33). (The order of magnitude of σ is the *potential energy* (divided by h) which the atom acquires in the electric field of the light-wave, owing to the electric moment which is due to the co-operation of the k th and l th modes of vibration.) The quantity σ would, in a certain sense, give a measure of the natural sharpness of the resonance-line, if it were possible to form a universal idea of the amplitude A of the incident light. We shall not enter upon this question here.

The theory put forward here in its rough features describes both the change of state of the atom produced by radiation of appropriate frequency and the appearance of resonance-radiation. For of course the presence of the *two* vibrations ψ_k and ψ_l will give rise to their natural emission. It is worth while mentioning that on account of the exponentials appearing in equation (34) this emission should *not* have exactly the frequency $\nu_l - \nu_k$, but a frequency exactly equal to ν , the frequency of the incident light-wave.

10. Extension of wave mechanics to systems other than a single mass-point.

Hitherto we have applied the method of wave mechanics only to a very simple system, viz. a single mass-point moving in a field of force which was either constant or varying with the time. We will now proceed to a quite arbitrary mechanical system. We might have done this before; all that has been said about the influence of an alternating field would apply with very slight modification to an arbitrary system, e.g. to the many-electron atom. But I thought it better to have a clear and simple case before our mental eye.

The derivation of the fundamental wave equation

put forward in the first lecture is very easily generalized to a quite arbitrary system, the only difference being that the "space" in which the wave-propagation takes place is no longer ordinary three-dimensional space but the "configuration space".

Let us recall the Hamilton-Maupertuis principle from which we started, namely,

$$\delta \int_A^B 2T dt = 0 \quad (1)$$

and which we transformed into

$$\delta \int_A^B \sqrt{2m(E - V)} ds = 0, \quad . . . (2)$$

by putting

$$2T = mw^2 = m \left(\frac{ds}{dt} \right)^2 = 2(E - V) = \frac{ds}{dt} \sqrt{2m(E - V)}.$$

We then compared it with Fermat's principle for a wave-propagation:

$$\delta \int_A^B \frac{ds}{u} = 0, \quad (3)$$

which led us to

$$u = \frac{C}{\sqrt{2m(E - V)}} \quad (4)$$

Now, *in general* T is *not* of the simple form $\frac{m}{2} \left(\frac{ds}{dt} \right)^2$ but

$$2T = \sum_l \sum_k b_{lk} \dot{q}_l \dot{q}_k, \quad . . . (41)$$

where the b_{lk} 's are functions of the generalized coordinates q_l . We now *define* a line-element ds in the generalized q -space by

$$2T = \sum_l \sum_k b_{lk} \dot{q}_l \dot{q}_k = \left(\frac{ds}{dt} \right)^2,$$

or

$$ds^2 = \sum_l \sum_k b_{lk} dq_l dq_k. \quad . . . (42)$$

The generalized non-Euclidean geometry, which is defined by the latter formula, is exactly the one which Heinrich Hertz used in his famous mechanics and which allowed him to treat the motion of an arbitrary system formally as the motion of a single mass-point (in a non-Euclidean, many-dimensional space). Introducing this geometry here, we easily see that all the considerations of the first lecture which led us to the fundamental wave equation may be transferred, even with a slight formal simplification, viz. that we have to put $m = 1$. In exactly the same way as before we obtain

$$u = \frac{E}{\sqrt{2(E - V)}},$$

and finally for the wave (or rather amplitude) equation:

$$\nabla^2 \psi + \frac{8\pi^2}{h^2} (E - V) \psi = 0. \quad . \quad . \quad (43)$$

For the wave equation properly speaking we get, just as before (section 6),

$$\nabla^2 \psi - \frac{4\pi i}{h} \dot{\psi} - \frac{8\pi^2 V}{h^2} \psi = 0. \quad . \quad . \quad (44)$$

But, of course, ∇^2 is now to be understood *not* as the simple Laplacian in three dimensions nor as the simple Laplacian in a many-dimensional Euclidean space (i.e. the sum of the second derivatives with respect to the single co-ordinates), but it is to be understood as the well-known generalization of the Laplacian in the case of a general line-element like (42). In the treatment of general problems we can usually avoid writing down the explicit expression for this operation; we need only know that it is a self-adjoint differential operator of the second order. (Never mind whether you know what "self-

adjoint " means, it is of no importance for the moment.) Yet for the sake of completeness I will put down the general expression for ∇^2 . Let a_{lk} be the minor corresponding to b_{lk} , divided by the determinant $\Sigma \pm b_{lk}$. Let a be the determinant of the a_{lk} 's. Then

$$\nabla^2 \equiv a^{\frac{1}{2}} \sum_l \frac{\partial}{\partial q_l} \left(a^{-\frac{1}{2}} \sum_k a_{lk} \frac{\partial}{\partial q_k} \right). \quad \dots \quad (45)$$

In the case of a single mass-point of mass m , treated in Cartesian co-ordinates, this reduces to $\frac{1}{m}$ times the elementary ∇^2 -operator (viz. $\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$). Or, if you chose to describe the motion of a single mass-point by any other co-ordinates, e.g. polar or elliptic, you would get $\frac{1}{m}$ times the expression for the elementary ∇^2 transformed to those co-ordinates. If the system consists of n free mass-points, you get the sum of their elementary ∇^2 -operators each divided by the appropriate mass.

The theory in its present form is applicable to systems of any number of degrees of freedom more than, equal to, or less than, three. I shall give a rapid account of a few examples without going through the details of calculation unless they present some physical interest.

11. Examples: the oscillator, the rotator.

Take the *one*-dimensional harmonic oscillator. The expression for the energy in ordinary mechanics may be taken to be

$$T + V = \frac{m}{2} \dot{q}^2 + 2\pi^2 \nu_0^2 m q^2$$

(we have expressed the coefficient of the potential energy in terms of the classical proper frequency ν_0 which it

produces). This easily leads to the amplitude equation:

$$\frac{1}{m} \frac{d^2\psi}{dq^2} + \frac{8\pi^2}{h^2} (E - 2\pi^2\nu_0^2 q^2) \psi = 0.$$

It can be shown that this equation has solutions which are finite along the real q -axis, for the following values of E only:

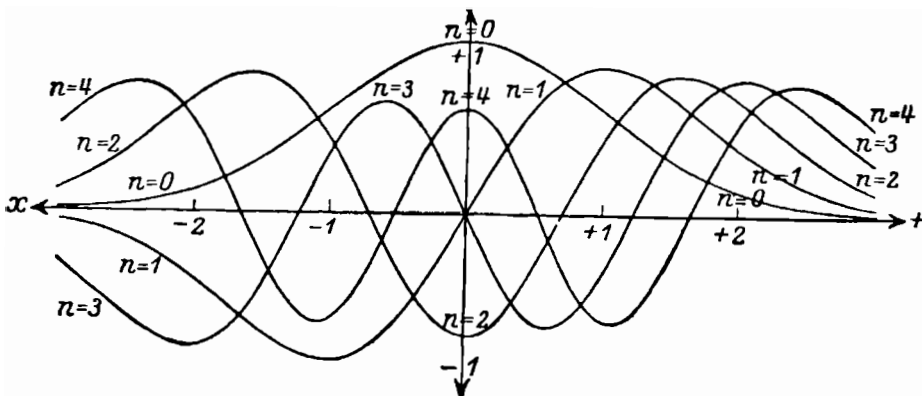
$$E_n = (n + \frac{1}{2})h\nu_0; \quad n = 0, 1, 2, 3. \quad . \quad (46)$$

The proper functions are the so-called Hermite orthogonal functions

$$\psi_n = (2^n n!)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} H_n(x) \quad . \quad . \quad (47)$$

with
$$x = q \cdot 2\pi \sqrt{\frac{m\nu_0}{h}}.$$

$H_n(x)$ is the so-called n th Hermite polynomial. A graph of the first five functions (47) is given in the figure.



The first five proper vibrations of the Planck oscillator according to undulatory mechanics. Outside of the region $-3 \leq x \leq +3$ represented here, all five functions approach the x -axis in monotonic fashion.

Though theoretically they extend to infinity, they are practically restricted by the exponential to a domain of the order of magnitude of the amplitude of the corresponding

classical mass-point. (This is very easy to prove.) We have not discussed the physical meaning of our generalized ψ -function. Yet the following statement is of interest. If the ψ_n 's were the proper functions of a one-electron problem and q one of the rectangular co-ordinates, we would (following our $\psi\bar{\psi}$ -hypothesis) estimate the intensity of emission of frequency $\frac{1}{h} |E_n - E_k|$, polarized in the direction of q , by the square of the integral

$$\int q \psi_k \psi_n dq.$$

If we try to do the same *here*, we get a most satisfactory result, viz. the integral *vanishes*, unless

$$|k - n| = 1.$$

This means that all the emission frequencies except $1 \cdot \nu_0$ are excluded. We shall return later to the question of the physical meaning of ψ in the general case.

Take as a second example another one-dimensional problem: the simple rotator with its axis fixed in space. Here all the energy is kinetic, viz.:

$$\frac{A}{2} \left(\frac{d\phi}{dt} \right)^2,$$

where A = moment of inertia, ϕ = angle of rotation. The amplitude equation becomes

$$\frac{1}{A} \frac{d^2\psi}{d\phi^2} + \frac{8\pi^2 E}{h^2} \psi = 0,$$

which has the solutions:

$$\psi = \frac{\sin}{\cos} \left[\sqrt{\frac{8\pi^2 EA}{h^2}} \phi \right].$$

Obviously ψ must be restricted to be periodic in ϕ with

period 2π . Hence the coefficient of ϕ must be an integer; this condition furnishes the proper values

$$E_n = \frac{n^2 h^2}{8\pi^2 A}; \quad n = 0, 1, 2, 3, \dots, \quad (48)$$

in *complete* agreement with the older form of quantum theory. Let us try to get an estimate for the intensity of radiation in the same formal way as before. If, in ordinary mechanics, an electrified particle were fixed to the rotator at a distance a from the centre of gravity, its rectangular co-ordinates would be

$$\left. \begin{matrix} x \\ y \end{matrix} \right\} = a \left\{ \begin{matrix} \cos \\ \sin \end{matrix} \right\} \phi.$$

Now form

$$\int_0^{2\pi} \psi_n \psi_k \left\{ \begin{matrix} x \\ y \end{matrix} \right\} d\phi = a \int_0^{2\pi} \left\{ \begin{matrix} \sin \\ \cos \end{matrix} \right\} n\phi \left\{ \begin{matrix} \sin \\ \cos \end{matrix} \right\} k\phi \left\{ \begin{matrix} \sin \\ \cos \end{matrix} \right\} \phi d\phi.$$

Since the product of the first two $\left\{ \begin{matrix} \sin \\ \cos \end{matrix} \right\}$ functions can always be expressed by the sum or difference of $\left\{ \begin{matrix} \sin \\ \cos \end{matrix} \right\}$ $(n \pm k)\phi$, it is easily recognized that none of the eight quantities comprised in the above formula differs from zero, unless either $|n + k|$ or $|n - k|$ is unity; or, what amounts essentially to the same, unless

$$|n - k| = 1.$$

This is the well-known selection-rule for the rotator. It is interesting to treat the rotator again without the assumption that its axis is rigidly fixed in direction. We find for the amplitude equation

$$\nabla_{\theta, \phi}^2 \psi + \frac{8\pi^2 AE}{h^2} \psi = 0.$$

Here $\nabla_{\theta, \phi}^2$ means *that* part of the elementary ∇^2 -operator

(when expressed in polar co-ordinates) that contains the differentiations with respect to the angles θ , ϕ only. It is known that the above equation only has finite single-valued solutions when the constant is the product of two successive integers:

$$\frac{8\pi^2 AE}{h^2} = n(n+1); \quad n = 0, 1, 2, \dots,$$

and that the solution is a spherical harmonic of order n . (The proper value E_n is $(2n+1)$ -fold degenerate, since there are $2n+1$ independent spherical harmonics of order n .) This furnishes the proper values

$$E_n = \frac{n(n+1)h^2}{8\pi^2 A}; \quad . . . \quad (49)$$

this means essentially that "half-integers" are to be inserted in place of n in the "classical" formula (48). (For $n(n+1) = (n + \frac{1}{2})^2 - \frac{1}{4}$, and a common constant in *all* the E_n 's cancels out in forming their differences.) It is known that the representation of band-spectra very often *compelled* the use of "half-integers", and it seems that all of them are *compatible* with the new formula. (Of course formula (49) is the correct one to use, and not (48), because the axis of a molecule is never rigidly fixed.) The selection rule comes out in exactly the same way as in the former case, only by a more troublesome calculation.

FOURTH LECTURE

12. Correction for motion of the nucleus in the hydrogen atom.

In the first lecture we treated the hydrogen atom as a one-body problem, as if the nucleus were fixed in space. In ordinary mechanics it is well known that if we start with the problem of two bodies (of masses m and M), we can split it in two, viz.:

(1) Uniform rectilinear motion of the centre of gravity (inertial motion).

(2) Keplerian motion around a fixed centre of a body with the "combined mass" μ , such that

$$\frac{1}{\mu} = \frac{1}{m} + \frac{1}{M}. \quad (50)$$

According to Bohr's theory, this refined treatment of the hydrogen atom is quantitatively supported by the slight difference in frequency between the Helium⁺-lines and *those* hydrogen-lines which would exactly coincide with them if the nucleus had infinite mass. (In other words, the slight difference between the Rydberg constant for He⁺ and for H is quantitatively accounted for by taking into account the slight movement of the nucleus; Sommerfeld.)

We meet with exactly the same state of affairs in wave mechanics. The six-dimensional amplitude equation for the two-body problem is:

$$\frac{1}{m} \nabla_1^2 \psi + \frac{1}{M} \nabla_2^2 \psi + \frac{8\pi^2}{h^2} (E - V) \psi = 0. \quad (51)$$

By ∇_1^2 and ∇_2^2 we mean the elementary Laplacians with respect to the co-ordinates of the electron (x_1, y_1, z_1) and of the nucleus (x_2, y_2, z_2) . About V we need only make the assumption that it depends on

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

only. Now, instead of x_1, \dots, z_2 , introduce the co-ordinates of the centre of gravity (ξ, η, ζ) and the relative co-ordinates of m with respect to M (say x, y, z). We can easily prove that

$$\frac{1}{m} \nabla_1^2 \psi + \frac{1}{M} \nabla_2^2 \psi = \frac{1}{m+M} \nabla_{\xi, \eta, \zeta}^2 \psi + \frac{1}{\mu} \nabla_{x, y, z}^2 \psi.$$

The meaning of the ∇^2 's is obvious; μ is given by (50). By inserting this in (51) we get an equation which can be split up by supposing ψ to be the product of a function of ξ, η, ζ only (say ϕ) and one of x, y, z only (say χ). In the splitting up an arbitrary constant is introduced, which is represented by E_t in the following equations. For ϕ we get

$$\frac{1}{m+M} \nabla_{\xi, \eta, \zeta}^2 \phi + \frac{8\pi^2 E_t}{h^2} \phi = 0, \quad (52)$$

and for χ

$$\frac{1}{\mu} \nabla_{x, y, z}^2 \chi + \frac{8\pi^2}{h^2} (E - E_t - V) \chi = 0. \quad (53)$$

The former describes the motion of the centre of gravity under no forces, according to wave mechanics; the constant E_t corresponds to its translational energy and can have any non-negative value. $E - E_t$ corresponds to the internal energy. The second equation is exactly that of the one-body problem for a mass-point with mass μ moving in a fixed field V . Hence for the proper values corresponding to the internal energy there will be no

difference other than that m is replaced by μ (see (14')) in the formula for the Rydberg constant. Thus Sommerfeld's important result, mentioned above, is re-stated in wave mechanics. Owing to the analytical simplicity of this deduction, there has not been much ado about it in the literature. But it really is one of the most immediate proofs that there must be something true in the many-dimensional wave-treatment—however irritating the latter may be at first.

13. Perturbation of an arbitrary system.

The theory of the perturbation of an arbitrary system really presents no new features as compared with the perturbation theory of the one-electron atom, a special case of which has been discussed in sections 7–9; but we shall widen our outlook by stating it afresh in a concise form. The general wave equation (44) of section 10 can be written:

$$\dot{\psi} = \frac{2\pi i}{h} \left(-\frac{h^2}{8\pi^2} \nabla^2 \psi + V\psi \right). \quad (54)$$

We will write H for the *operator*

$$H = -\frac{h^2}{8\pi^2} \nabla^2 + V.$$

(V as an *operator* means: “to multiply by V ”.) Then by (43), section 10, the proper functions ψ_k are precisely *those* which are *reproduced* by the operator H , apart from a multiplying constant, which is the proper value:

$$H[\psi_k] = E_k \psi_k. \quad (55)$$

Equation (54) takes the simple form

$$\dot{\psi} = \frac{2\pi i}{h} H[\psi]. \quad (56)$$

Now, adding to V a small perturbing field, which may or may not contain the time explicitly, means altering the operator H slightly. (Of course an alteration of H might also be produced in another way, e.g. by altering one of the masses, &c. It will do no harm if this more general case is included in our treatment.) We shall call the altered operator $H + H'$, bearing in mind that H' is to be a "small" operator. We have to solve

$$\dot{\psi} = \frac{2\pi i}{h} (H[\psi] + H'[\psi]). \quad . \quad . \quad (57)$$

Tentatively substituting

$$\psi = \sum c_k \psi_k e^{2\pi i \nu_k t}; \quad \nu_k = \frac{E_k}{h}, \quad . \quad . \quad (58)$$

with slowly varying time-functions c_k , we obtain in the first instance

$$\sum_k \dot{c}_k \psi_k e^{2\pi i \nu_k t} = \frac{2\pi i}{h} \sum_k c_k H'[\psi_k] e^{2\pi i \nu_k t}.$$

This equation will be satisfied if it is orthogonal to *all* the ψ_l 's*. Multiply by ψ_l and integrate over the whole configuration-space:

$$\dot{c}_l = \frac{2\pi i}{h} \sum_k c_k a_{lk} e^{2\pi i t(\nu_k - \nu_l)}, \quad l = 1, 2, 3, 4, \dots, \quad (59)$$

where
$$a_{lk} = \int dq H'[\psi_k] \psi_l, \quad . \quad . \quad . \quad (60)$$

and $\int dq$ always means a multiple integral over the whole configuration-space. The a_{kl} 's are *small* quantities.

We will suppose the perturbation to be conservative.

* We take it for granted that with respect to the completeness and orthogonality of the proper functions the general case behaves like the simple hydrogen case. That is quite safe. We also, as there, avoid encumbering our formulæ by explicitly taking account of a *continuous* spectrum of proper values.

Then the a_{kl} 's are constants; just as in the special cases treated before, only the exponentials with vanishing exponent will cause appreciable variations of the c_l 's. First take the system to be non-degenerate. Then, dropping the other terms, which only furnish slight oscillations, you get, for every c_l ,

$$\dot{c}_l = \frac{2\pi i a_{ll}}{h} c_l; \quad c_l = c_l^0 e^{\frac{2\pi i a_{ll}}{h} t}, \quad \dots \quad (61)$$

which, if you substitute it in (58), merely means that the frequency is slightly altered by the amount

$$\frac{a_{ll}}{h}.$$

Now take a case of degeneracy. Let the amplitudes $c_l, c_{l+1}, \dots, c_{l+a-1}$ belong to a different proper functions, all belonging to the same proper value E_l , or proper frequency ν_l . Then in each of the equations relating to them you will have not only one, but a , vanishing exponents, which give rise to secular changes. Hence these a amplitudes will be determined by the following set of equations:

$$\dot{c}_{l+\rho} = \frac{2\pi i}{h} \sum_{\lambda=0}^{a-1} c_{l+\lambda} a_{l+\rho, l+\lambda}; \quad \rho = 0, 1, 2, \dots, a-1. \quad (62)$$

These equations show that under the influence of a slight perturbation there will in general be an exchange of amplitudes between degenerate modes of vibration which belong to the same proper value. It is correct to talk of an *exchange*, since it is easily proved from equation

(62) that
$$\sum_{\rho=0}^{a-1} |c_{l+\rho}|^2 = \text{const.}$$

Yet when thinking of this exchange we must remember

that the set of proper functions $\psi_{l+\rho}$ ($\rho = 0, 1, \dots, \alpha - 1$) is arbitrary up to an orthogonal linear substitution of determinant 1. This induces a similar substitution of the amplitudes c_l . Given a definite perturbation, i.e. definite values of the quantities $a_{l+\lambda, l+\rho}$, it is always possible to find at least one orthogonal substitution of the $\psi_{l+\rho}$'s which brings the equations (62) into the simple form (61) of the non-degenerate case. Then these particular proper functions, selected in a way that suits this particular form of perturbation, will under its influence have *constant* amplitude-squares, but will in general belong to slightly different proper frequencies. The α -fold proper value has been split up into α slightly differing proper values; the degeneracy is removed by the disturbing field, and the particularly chosen proper functions of the degenerate problem are the non-degenerate proper functions "in zero approximation" to the single proper values of the perturbed problem. The α slight alterations in proper value can be shown to be the α roots of the "secular" equation

$$\begin{vmatrix} a_{ll} - x, & a_{l, l+1}, & \dots, & a_{l, l+\alpha} \\ a_{l+1, l}, & a_{l+1, l+1} - x, & \dots, & a_{l+1, l+\alpha} \\ \cdot & \cdot & \cdot & \cdot \\ a_{l+\alpha, l}, & a_{l+\alpha, l+1}, & \dots, & a_{l+\alpha, l+\alpha} - x \end{vmatrix} = 0$$

Of course it may happen that these roots are not all different; a certain degeneracy is then retained. We may either say that the members of an arbitrarily chosen set of the degenerate functions all vibrate with the unperturbed frequency, but exchange their amplitudes—or that the members of the appropriately chosen set have constant amplitudes, but that each function has a slightly different frequency: these two assertions are of course

identical. For—as we may put it—either: a vibration of varying amplitude has not really got the frequency which we ascribe to it; or: two or more slightly different frequencies, when superimposed, lead to a “beat phenomenon”, i.e. to a varying amplitude.

14. Interaction between two arbitrary systems.

Take now two arbitrary systems, at first without interaction, one of which is described according to wave mechanics (see equation (56)) by

$$\dot{\psi} = \frac{2\pi i}{h} H[\psi],$$

and the other by

$$\dot{\phi} = \frac{2\pi i}{h} L[\phi].$$

Multiply the first by ϕ , the second by ψ , and add the resulting equations; you get

$$\frac{d}{dt}(\phi\psi) = \frac{2\pi i}{h}(H + L)[\phi\psi],$$

since the operator H does not affect ϕ and L does not affect ψ . The latter equation is the wave equation of the “combined system”, i.e. of the system formed by *mentally* uniting the two systems to form one. (The process is exactly the reverse of what is so often done in “splitting up” an equation by supposing the solution to be the product of two functions, dependent on different individual variables.) The proper functions of the combined system are the *products* of any one of the proper functions of the first system and any one of the second system. The proper value that belongs to such a product is easily seen to be the *sum* of the respective proper values. (This corresponds to the additivity of energy in ordinary

mechanics.) By the addition of proper values a *new degeneracy* may be caused in the combined system, even though the single systems were non-degenerate. (Let us suppose the latter case, for the sake of simplicity.) Let E, E' be two proper values of the first system, F, F' two of the second system, and suppose that

$$E + F' = E' + F = G$$

or
$$E - E' = F - F'.$$

Hence: if a common *difference of proper values* exists between the two systems, it will give rise to a two-fold degenerate proper value G of the combined system. For simplicity's sake, suppose that other relations of the same kind are absent, and *now* suppose that a slight *interaction* of the two systems takes place, changing the operator $H + L$ into $H + L + T$, where T will of course



contain the variables both of the first and of the second system. Then the amplitudes belonging to $E + F'$ and to $E' + F$ will show a slow secular interchange, all the others remaining essentially constant. The sum of the squares of the two amplitudes in question is also constant. Interpreted in the single systems, this cannot very well have any other meaning but that e.g. the amplitude of F increases at the expense of that of F' and, so to speak, to compensate for the amplitude of E' increasing at the expense of that of E . This seems to be the appropriate wave-mechanical description of what in the older

form of the quantum theory was called the transfer of a quantum of energy $E - E'$ ($= F - F'$) from one system to the other.

15. The physical meaning of the generalized ψ -function.

Perhaps the latter conclusions are obscured by the fact that we have hitherto avoided putting forward any definite assumption as to the physical interpretation of the function $\psi(q_1, q_2, \dots, q_n, t)$ relating to a system whose configuration in terms of ordinary mechanics is described by the generalized co-ordinates q_1, q_2, \dots, q_n . This interpretation is a very delicate question. As an obvious generalization of the procedure of spreading out the electronic charge according to a relative density function $\psi\bar{\psi}$ (which furnished satisfactory results in the one-electron problem; see section 5), the following view would present itself in the case of a general mechanical system: the real natural system does not behave like the picture which ordinary mechanics forms of it (e.g. a system of point-charges in a definite configuration), but rather behaves like what would be the result of spreading out the system, described by q_1, \dots, q_n , *throughout its configuration-space* in accordance with a relative density function $\psi\bar{\psi}$. This would mean that, if the ordinary mechanical picture is to be made use of at all, the actual system behaves like the ordinary mechanical picture, present in all its possible configurations at the same time, though "stronger" in some of them than in others.

I maintained this view for some time. The fact that it proves *very* useful can be seen from the one-electron problem (see section 5). *No* other interpretation of the ψ -function is capable of making us *understand* the large

amount of information which the constants a_{kl} furnish about the intensity and polarization of the radiation. Yet this way of putting the matter is surely not quite satisfactory. For what does the expression "to behave like" mean in the preceding sentences? The "behaviour" of the ψ -function, i.e. its development in time, is governed by nothing like the laws of classical mechanics; it is governed by the wave-equation.—

An obvious *statistical* interpretation of the ψ -function has been put forward, viz. that it does not relate to a single system at all but to an assemblage of systems, $\psi\bar{\psi}$ determining the fraction of the systems which happen to be in a definite configuration. This view is a little unsatisfactory, since it offers no explanation whatever why the quantities a_{kl} yield all the information which they do yield. In connexion with the statistical interpretation it has been said that to any physical quantity which would have a definite physical meaning and be in principle (*principiell*) measurable according to the classical picture of the atom, there belong definite proper values (just as e.g. the proper values E_k belong to the energy); and it has been said that the result of measuring such a quantity will always be one or the other of these proper values, but never anything intermediate. It seems to me that this statement contains a rather vague conception, namely that of *measuring* a quantity (e.g. energy or moment of momentum), which relates to the classical picture of the atom, i.e. to an obviously wrong one. Is it not rather bold to interpret measurements according to a picture which we know to be wrong? May they not have quite another meaning according to the picture which will finally be forced upon our mind? For example: let a beam of electronic rays pass through a layer of mercury vapour, and

measure the deflection of the beam in an electric and in a magnetic field before and after the beam has traversed the vapour. According to the older conceptions this is interpreted as a measurement of differences of energy-levels in the mercury atom. The wave-picture furnishes another interpretation, namely, that the frequency of part of the electronic waves has been diminished by an amount equal to the difference of two proper frequencies of the mercury. Is it quite certain that these two interpretations do not interfere with one another, and that the old one can be maintained together with the new one? Is it quite certain that the conception of energy, indispensable as it is in macroscopic phenomena, has *any* other meaning in micro-mechanical phenomena than the number of vibrations in h seconds?

