

## Quantum Theory of Dispersion

By G. BREIT

*New York University and Carnegie Institution of Washington*

### TABLE OF CONTENTS

Introduction.....	504
I. Discussion of Electromagnetic Field by Means of Hamiltonian Equations.....	506
§1. Introduction of fundamental volume.....	506
§2. Free electromagnetic waves.....	507
§3. Influence of matter.....	510
§4. Meaning of $\text{div } \mathcal{E} = 4\pi e\rho$ and summary of preceding discussion.....	512
§5. Elimination of electrostatic field.....	514
§6. Self energy.....	515
§7. Matter and field treated by one Hamiltonian.....	517
II. The Kramers-Heisenberg Dispersion Formula Derived by Means of the Correspondence Principle.....	521
§1. The Hamiltonian function used.....	521
§2. Transformation to new action variables.....	522
§3. Application of the correspondence principle to coherent scattering.....	525
§4. The $f$ sum rule.....	528
§5. Incoherent scattering.....	529
III. Introductory Remarks about Quantum Theory.....	532
§1. Purpose of remarks.....	532
§2. Schroedinger equation.....	532
§3. The statistical interpretation.....	534
§4. Generalization of Schroedinger's equation to other reference systems.....	537
§5. Operators at different times.....	540
§6. Equations of motion.....	541
IV. The Kramers-Heisenberg Formula as Derived by Born, Heisenberg, Jordan and by Schroedinger.....	543
V. Dirac's Theory of Dispersion.....	555
§1. Introductory.....	555
§2. Quantization of the field and equations of motion.....	557
§3. Elimination of electrostatic field.....	562
§4. Probability amplitudes for numbers of light quanta.....	564
§5. Spontaneous emission.....	566
§6. Absorption and emission.....	568
§7. Scattering and dispersion.....	570
§8. Terms in the square of the vector potential.....	575

### INTRODUCTION

THE theory of optical dispersion has played an important part in the development of quantum theory. It has gone through several stages of evolution. Each of these may be considered as a landmark and an indicator of sophistication in the growth of quantum theory.

The first attempt is that of Kramers and of Kramers and Heisenberg. It lead to the famous Kramers-Heisenberg dispersion formula which represents the point of view of the correspondence principle. It reconciles this principle

as applied to matter with Einstein's mechanism for the interaction of matter and radiation, clarifying the reason for the introduction of negative absorption probabilities. Later on it is of importance in the establishment of the Kuhn, Reiche-Thomas rule which in turn becomes the  $[p, q] = h/2\pi i$  of matrix mechanics.

With the introduction of matrix mechanics we find the problem treated again by Born, Heisenberg and Jordan and at about the same time an independent treatment by Schroedinger appeared from the point of view of his wave mechanics. These discussions lead also to the Kramers-Heisenberg formula. They are more definite than the original derivation. All terms in the formula have a quantitative meaning while the original work of Kramers and Heisenberg there was no way of knowing how the Fourier coefficients in the analysis of the unperturbed conditionally periodic motion had to be averaged. Only in the cases for which some Fourier coefficients vanished identically was it possible to draw a definite conclusion *viz.*, that the average of these coefficients was also zero. The removal of this indefiniteness is a landmark of a fairly satisfactory construction of a quantum theory of matter. It is precise and definite to the same extent to which matter has been treated. But it also has some ambiguities inherent to any method of treatment closing its eyes to the dual (*particle* versus *wave*) nature of electromagnetic radiation. Light is considered as a wave. Its particle nature is ignored. It is natural that no satisfactory understanding of emission of light quanta is obtained and that the phenomena occurring within the natural breadth of a spectral line are obscure. It is also unclear in this form of theory whether the intensity of the Raman radiation should be proportional to the number of atoms present in the final state. The dispersion treatments of Born, Heisenberg, Jordan and of Schroedinger are thus capable of answering correctly all questions about the motion of matter under the action of an externally incident light wave with neglect of the reaction on the matter due to radiation sent out by it. They give no satisfactory account of the phenomenon of spontaneous emission of light quanta. It should nevertheless be remembered that in all questions about radiation which depend unambiguously on the state of matter as for example its electric polarization the version of Born, Heisenberg, Jordan gives answers correct to the same extent to which radiation damping may be neglected.

It remained for Dirac to take into account the dual character of light. In his theory the radiation is treated by means of quantum theory just as matter has been treated in previous attempts. He was able to explain spontaneous emission, to derive anew the Kramers-Heisenberg formula, and to give a definite and correct answer to the other disputable points just mentioned. The development of the theory of radiation and of dispersion thus appeared to be closed.

On closer investigation it became apparent that Dirac's theory is satisfactory only as long as it is made to answer questions about the ordinary range of frequencies. For very high frequencies (of the order  $10^{23}$  sec.<sup>-1</sup>) it predicts an infinite accumulation of energy and is thus also untenable in its

present form. Apart from these essential difficulties the original formulation of Dirac contained another more minor defect inasmuch as no account was taken of the electrostatic field produced by the electron. It so happens that the omission makes no difference in the discussion of radiation problems. The construction of theories free from this defect and satisfying the requirements of relativistic invariance we owe to Heisenberg and Pauli and an alternative procedure to Fermi. The essential difficulties of Dirac's theory of radiation become very pronounced in the treatment of Heisenberg and Pauli. The proof of their reality and the practical consequence of an infinite displacement of spectral lines has been given by Oppenheimer and has been followed by the investigations of Waller, Heisenberg, Rosenfeld, Peierls and Landau. We will return to these later. For the present we must be content with the statement that Dirac's theory gives good answers as long as its troublesome character in the region of high frequencies is treated sympathetically. It is the best available theory at the present time.

In this section of the review we wish to show the interconnections of the three generations of quantum theories as well as their relationship to the great grandfather—the classical electrodynamics. The Kramers, Heisenberg and the Born, Heisenberg, Jordan—Schroedinger theories are very closely related to the ordinary treatments of classical dispersion discussed in the review preceding this. Dirac's treatment is related more intimately to another less familiar way of discussing the interaction between radiation and matter by means of a system of classical Hamiltonian equations. We begin with presenting this point of view.

## I. DISCUSSION OF ELECTROMAGNETIC FIELD BY MEANS OF HAMILTONIAN EQUATIONS

### §1. Introduction of fundamental volume

In any quantum theory so far produced it has been found necessary to use the corresponding non-quantum *classical* theory as a guide. The procedure for forming a quantum theory out of a classical one is greatly simplified if the equations of the classical theory are given in a Hamiltonian form. Once these are given it is usually possible to form a *corresponding* quantum theory, free from internal contradictions, by a definite mathematical procedure. As is well known the operators of the quantum theory are determined at all times if all of their commutation rules are known and if the Hamiltonian function is given. [See III, §2 below.]

In the discussion of radiation one must take account of the fact that light of a definite frequency is absorbed or emitted in definite quanta  $h\nu$ . The mathematical analysis must be therefore in terms of frequencies so as to enable us to quantize the energy contained in any frequency  $\nu$ . Our problem thus consists (1) *in describing electromagnetic radiation by means of variables intimately and directly connected with the energy of the monochromatic constituents of the radiation*, and (2) *in finding such a Hamiltonian function that Maxwell's equations and the force equation of classical electrodynamics are identically sat-*

*isfied*. There are various ways of analyzing electromagnetic waves into monochromatic components. One way which has been frequently used is to consider the space inside a hypothetical set of perfectly conducting walls (electrical conductivity =  $\infty$ ). The electromagnetic field inside such an enclosure or *Hohlraum* may be analyzed into standing waves satisfying the boundary conditions of vanishing tangential electric intensity and normal magnetic intensity at the walls. This general picture has already been used by Jeans in the discussion of the law of equipartition in black body radiation and has been frequently employed for quantum generalizations. It is somewhat more convenient for our discussion to analyze the radiation into progressive rather than standing waves.<sup>1-5</sup> By so doing we simplify calculations in the discussion of problems involving unidirectional emission. It is possible to treat all of these problems however by the method of standing waves as well.

In order to make the discussion of progressive waves definite we imagine all three dimensional space subdivided into exactly equal cubical cells of volume  $V$  and side  $L$ . We suppose that  $V$  is so large that one of the cells contains all of the matter and electromagnetic waves that may possibly interest us in our discussion. We then imagine that all the other cells are filled with matter and with waves in exactly the same state as in the first so that a displacement bringing about geometrical congruence of the boundaries of any two cells would bring about also congruence of their physical contents. The mathematical discussion of such an imaginary cellular space is simpler than the discussion of an arbitrary distribution. For practical purposes it should make no difference whether we have filled remote regions of space with matter and waves correctly or not because  $L$  is taken to be so large as to make the influence of anything contained in one cell on any other thing contained in another cell negligible. The device used here is exactly similar to that employed in the discussion of Fourier integrals and is subject to the same limitations. It is more convenient in the present discussion to use Fourier series with very large but not infinite fundamental intervals because by so doing we always have discrete sets of dynamical variables.

## §2. Free electromagnetic waves

Consider first free electromagnetic waves, i.e., electromagnetic waves in the absence of matter. Let the frequency of a monochromatic component be  $\nu$  its wave-length  $\lambda$ , the velocity of light  $c$ , and let  $\mathbf{k}$  be a vector of length  $2\pi/\lambda$  in the direction of propagation of the wave. For brevity  $\mathbf{k}$  will also be called the propagation vector. We must further specify the state of polarization of the wave which means we must assign in addition the direction of its electric vector. [Circular and elliptical polarizations will be treated later: VI §5.] A unit vector in this direction referring to a wave with propagation vec-

<sup>1</sup> W. Heisenberg and W. Pauli, *Zeits. f. Physik* **56**, 1 (1929), I; **59**, 168 (1930), II.

<sup>2</sup> H. Weyl, *Gruppentheorie und Quantenmechanik*, 1st edition p. 89; 2nd edition p. 91.

<sup>3</sup> P. A. M. Dirac, *Proc. Roy. Soc.* **114**, 243 (1927); **114**, 710 (1927).

<sup>4</sup> L. Landau and R. Peierls, *Zeits. f. Physik* **62**, 188 (1930).

<sup>5</sup> F. H. Kennard, *Phys. Rev.* **37**, 458A (1931).

tor  $k$  will be written  $f_{k,1}$  or  $f_{k,2}$  depending on whether we wish to refer to one or another of two mutually perpendicular plane polarized components. We will frequently write instead

$$f_{k,l} \quad (l = 1, 2).$$

It should be specified here that  $f_{k,l}$  is a vector in a fixed direction and does not reverse itself during the periodic time alternations of the electric intensity  $\mathcal{E}$ . We now define in addition two scalar quantities

$$a_{k,l}, \quad a_{k,l}^+$$

which together characterize the amplitude and phase of the wave  $k, l$ . Both  $a$  and  $a^+$  shall depend on the time  $t$ . For free electromagnetic waves the scalar potential may be set = 0 and the vector potential  $\mathbf{A}$  at any point we suppose to be given by

$$\mathbf{A} = \sum_{k,l} f_{k,l} \frac{c}{(V\nu)^{1/2}} [a_{k,l} e^{-ikr} + i a_{k,l}^+ e^{ikr}]. \quad (1)$$

In this summation every vector  $k$  occurs once with  $l=1$  and another time with  $l=2$ . The vector  $-k$  also occurs once each way. We thus have independent variables  $a_{k,l}$  and  $a_{-k,l}$ . In order that  $\mathbf{A}$  be real it is necessary to have

$$\sum_l (f_{k,l} a_{k,l} + i f_{-k,l} a_{-k,l}^+) = \sum_l (f_{-k,l} a_{-k,l} + i f_{k,l} a_{k,l}^+)^* \quad (1')$$

where  $*$  indicates the conjugate imaginary.

We choose initially

$$a_{k,l}^* = i a_{k,l}^+ \quad (1'')$$

so that every term of (1) is automatically real. We will see that (1'') remains satisfied in the course of time. The values of  $k$  consistent with the requirement of the periodicity of  $\mathbf{A}$  with space period  $L$  are

$$\mathbf{k} = (2\pi/L)(n_1, n_2, n_3) \quad (2); \quad k^2 = 4\pi^2\nu^2/c^2 \quad (2')$$

where  $n_1, n_2, n_3$  are positive or negative integers. We suppose that the electric and magnetic intensities  $\mathcal{E}, \mathcal{H}$  are given respectively by

$$\mathcal{E} = -(\partial \mathbf{A} / \partial t), \quad \mathcal{H} = \text{rot } \mathbf{A} \quad (3)$$

We wish every term in (1) to represent a plane wave. This is accomplished by satisfying (7') below. In order that (3) be satisfied we define:

$$\mathcal{E} = - \sum f_{k,l} \frac{2\pi i \nu}{(V\nu)^{1/2}} [a_{k,l} e^{-ikr} - i a_{k,l}^+ e^{ikr}] \quad (4)$$

$$\mathcal{H} = \sum \frac{c}{(V\nu)^{1/2}} [\mathbf{k} \times f_{k,l}] [-i a_{k,l} e^{-ikr} - a_{k,l}^+ e^{ikr}] \quad (5)$$

and introduce such a Hamiltonian function that (7') follows. The energy of the electromagnetic field in the fundamental volume  $V$  is then obtained from (4) and (5) using (2') as follows

$$(1/8\pi) \int \mathcal{E}^2 dV = - \sum'_{k,l} \pi \nu (f_{k,l} a_{k,l} - i f_{-k,l} a_{-k,l}^+) (f_{-k,l} a_{-k,l} - i f_{k,l} a_{k,l}^+)$$

$$(1/8\pi) \int \mathcal{H}^2 dV = \sum'_{k,l} \pi \nu (f_{k,l} a_{k,l} + i f_{-k,l} a_{-k,l}^+) (f_{-k,l} a_{-k,l} + i f_{k,l} a_{k,l}^+).$$

In these formulas the accent on the summation  $\Sigma'$  indicates that terms in  $k$  and  $-k$  have been grouped together and each pair  $(k, -k)$  is taken once.<sup>6</sup> Hence

$$H_R = \frac{1}{8\pi} \int (\mathcal{E}^2 + \mathcal{H}^2) dV = \sum' 2\pi i \nu (a_{k,l} a_{k,l}^+ + a_{-k,l} a_{-k,l}^+)$$

$$= \sum_{k,l} 2\pi i \nu a_{k,l} a_{k,l}^+. \quad (6)$$

The above expression for  $H_R$  may be considered as a Hamiltonian function with  $a_{k,l}$  as a coordinate like and  $a_{k,l}^+$  as a momentum like variable. We shall regard (4) as definitions of  $\mathcal{E}$ ,  $\mathcal{H}$  and we shall find that (3) is satisfied as a consequence of the equations of motion imposed on  $a$ ,  $a^+$  by the Hamiltonian (6). These equations are:

$$\frac{da_{k,l}}{dt} = \frac{\partial H_R}{\partial a_{k,l}^+} = 2\pi i \nu a_{k,l}; \quad \frac{da_{k,l}^+}{dt} = - \frac{\partial H_R}{\partial a_{k,l}} = -2\pi i \nu a_{k,l}^+ \quad (7)$$

so that the first Eq. (3) follows. The second Eq. (3) is also readily verified by direct substitution.

The time dependence of  $a$ ,  $a^+$  is

$$a_{k,l}(t) = a_{k,l}(0) e^{2\pi i \nu t}; \quad a_{k,l}^+(t) = a_{k,l}^+(0) e^{-2\pi i \nu t}. \quad (7')$$

Comparing (7') with (1), (4), (5) we see that  $a_{k,l}$ ,  $a_{k,l}^+$  refer in fact to monochromatic progressive waves of frequency  $\nu$  (see (2')) with propagation vector  $k$ . For each polarization and wave normal direction we need *two* dynamical variables in order to describe both the amplitude and the phase. Eqs. (7) and their solutions (7') show that the reality of (1), (4), (5) is maintained in the course of time on account of (1''), if it is secured initially.

The above Hamiltonian function (6) together with the Eq. (1) for the vector potential and for the field intensities determines the behavior of the field in a vacuum. In this case the solution by means of Eq. (7) is trivially simple. It tells us only that if the field is analyzed in a Fourier series in space each Fourier component propagates with the velocity of light in the direction of the wave normal.

<sup>6</sup> Explicitly  $\Sigma' = \Sigma$  over the range of values  $n_1=0, 1, 2, \dots, n_2, n_3=0, \pm 1, \pm 2, \dots$

### §3. Influence of matter

Let us consider next the case of matter being present and let us content ourselves at first with discussing the dependence of the radiation field on the state of the matter. We recall that plane electromagnetic waves cannot represent the whole field in the vicinity of the electric charges and currents although they are capable of doing so to a good approximation at some distance away. The essential reason for the failure of plane electromagnetic waves is that they cannot be compounded so as to give an electrostatic field. For its description we need to introduce either an electrostatic potential or else we must add to the right side of (1) some terms of an unperiodic character. The analysis with unperiodic terms in  $A$  has been used by Heisenberg and Pauli<sup>1</sup> while that with an electrostatic potential has been employed by Fermi.<sup>7</sup> The manner in which unperiodic terms can represent an electrostatic field can be seen by considering the electrostatic field produced by a point charge. At a distance  $r$  from the charge  $e$  the electric field has components

$$\mathcal{E} = e(x, y, z)/r^3$$

and the magnetic field  $\mathcal{H} = 0$ . This field may be described by means of (3) using

$$A = -ct(e/r^3)(x, y, z).$$

We note that  $\text{div } A \neq 0$  at  $r=0$  but has a singularity at that point. Similarly the field due to two condenser plates perpendicular to the  $x$  axis is given by

$$A = ct |\mathcal{E}| (1, 0, 0)$$

inside the plates and  $A = 0$  outside. The Fourier analysis of this field consists of terms in  $e^{\pm ikx}$ . The vector  $A$  is thus perpendicular to the wave front ( $yz$  plane) of the plane waves into which  $A$  may be analyzed. This suggests adding to the right side of (1) a term

$$\sum'_k \frac{c}{(V\nu)^{1/2}} \frac{k}{k} (a_k e^{-ikr} + a_k^* e^{ikr}).$$

Here as before  $\Sigma'$  is taken by restricting ourselves to half of all the possible directions as may be done for instance by making  $n_1$  in (2) take positive values only, while  $n_2, n_3$  take positive as well as negative values. We have thus

$$A = \sum_{k,l} \frac{c}{(V\nu)^{1/2}} f_{k,l} (a_{k,l} e^{-ikr} + i a_{k,l}^+ e^{ikr}) + \sum'_k \frac{c}{(V\nu)^{1/2}} \frac{k}{k} (a_k e^{-ikr} + a_k^* e^{ikr}) \quad (8)$$

and define  $\mathcal{E}, \mathcal{H}$  by means of (3). The quantities  $a_k, a_k^*$  we consider to be co-ordinate like independent dynamical variables subject to the initial restrict-

<sup>7</sup> E. Fermi, Lincei Rend. 9, 881 (1929); 12, 431 (1930).

ing condition of one being the complex conjugate of the other. We note that the added term in (8) leaves  $\mathcal{H}$  unaltered so that it is still given by (5). The current density we take to be

$$ej(x, y, z; t) = \sum_k ej_k e^{-ikr} V^{-1/2}. \quad (9)$$

The vector  $j_k$  is a function of the time. The immediate problem is to find such a Hamiltonian function that Maxwell's equations should be a consequence of the equations of motion. We obtain a solution by taking the energy of the radiation in the form

$$H_R = (1/8\pi) \int (\mathcal{E}^2 + \mathcal{H}^2) dV \quad (10)$$

and the interaction energy with the matter as

$$H' = - (e/c) \int j A dV. \quad (11)$$

The Hamiltonian function is then

$$H = H_R + H'. \quad (12)$$

The added term in (8) gives new terms in  $\mathcal{E}$  as well. We let

$$\dot{a}_k = 4\pi\nu b_k^*, \quad \dot{a}_k^* = 4\pi\nu b_k \quad (13)$$

which suggests defining

$$\begin{aligned} \mathcal{E} = & - \sum_{k,l} 2\pi i \nu^{1/2} V^{-1/2} \mathbf{f}_{k,l} (a_{k,l} e^{-ikr} - i a_{k,l}^+ e^{ikr}) \\ & - \sum' 4\pi(\nu/V)^{1/2} (k/k) (b_k^* e^{-ikr} + b_k e^{ikr}). \end{aligned} \quad (14)$$

We now form the Hamiltonian function and show later that  $\mathcal{E} = -\partial A / \partial t$ . Then by (10)

$$H_R = \sum_{k,l} 2\pi i \nu a_{k,l} a_{k,l}^+ + \sum' 4\pi\nu b_k b_k^* \quad (15)$$

$$\mathcal{H}' = -e \sum \nu^{-1/2} j_{-k} (\mathbf{f}_{k,l} a_{k,l} + i \mathbf{f}_{-k,l} a_{-k,l}^+) - \sum' (k/k) (j_{-k} a_k + j_k a_k^*) \nu^{-1/2}. \quad (16)$$

From (15), (16) we see that (13) is satisfied if  $b_k, b_k^*$  are taken to be momentum like dynamical variables corresponding respectively to  $a_k, a_k^*$ .

It follows from (15), (16) that in the absence of  $j$  the variables  $a_k, a_k^*$  do not occur in the Hamiltonian function. The corresponding momenta  $b_k, b_k^*$  are constants. By (13) the variables  $a_k, a_k^*$  depend linearly on the time. The field is thus of the type discussed in the examples of a point charge and the field of a condenser. If  $j$  is not zero the interaction energy (16) brings about a time dependence of  $b_k, b_k^*$ . The equations of motion corresponding to the Hamiltonian (12) are in addition to (13)

$$\dot{b}_k = e \nu^{-1/2} k^{-1} (k j_{-k}), \quad \dot{b}_k^* = e \nu^{-1/2} k^{-1} (k j_k) \quad (17)$$

$$\begin{aligned} \dot{a}_{k,l} &= 2\pi i \nu a_{k,l} - i e \nu^{-1/2} \mathbf{f}_{k,l} j_k \\ \dot{a}_{k,l}^+ &= -2\pi i \nu a_{k,l}^+ + e \nu^{-1/2} \mathbf{f}_{k,l} j_{-k}. \end{aligned} \quad (18)$$



From (18) we find

$$\mathbf{f}_{k,l}\dot{a}_{k,l} + i\mathbf{f}_{-k,l}\dot{a}_{-k,l}^+ = 2\pi i\nu(\mathbf{f}_{k,l}a_{k,l} - i\mathbf{f}_{-k,l}a_{-k,l}^+) - ie\nu^{-1/2}[\mathbf{f}_{k,l}(\mathbf{f}_{k,l}j_k) - \mathbf{f}_{-k,l}(\mathbf{f}_{-k,l}j_k)]. \quad (18)'$$

The left side of the above equation (18') when summed over  $l=1, 2$  and multiplied by  $c/(V\nu)^{1/2}$  is the time derivative of the coefficient of  $e^{-ikr}$  in the first summation of (8). In order that the first summation of formula (14) for  $\mathcal{E}$  be consistent with (8) and the equation  $\mathcal{E} = -\partial A/c\partial t$  the second part of (18') involving  $j_k$  must disappear on summation over  $l$ . This condition is in fact fulfilled because

$$\sum_l \mathbf{f}_{k,l}(\mathbf{f}_{k,l}j_k) = \sum_l \mathbf{f}_{-k,l}(\mathbf{f}_{-k,l}j_k) = j_k - (kj_k)k^{-1}. \quad (18'')$$

Using now (8), (18'), (18'') and (13) we see that  $\mathcal{E} = -\partial A/c\partial t$  is satisfied. Therefore, regarding the second Eq. (3) as the definition of  $\mathcal{H}$ ,  $\text{rot } \mathcal{E} = -\partial \mathcal{H}/c\partial t$ . This finishes the discussion of the first triplet of Maxwell's equations. The second triplet is

$$\text{rot } \mathcal{H} = \partial \mathcal{E}/c\partial t + (4\pi e/c)j.$$

From (18) we have by direct substitution into  $\partial \mathcal{E}/c\partial t$  with  $\mathcal{E}$  given by (14)

$$\begin{aligned} \frac{\partial \mathcal{E}}{c\partial t} &= \sum \frac{4\pi^2\nu^{3/2}}{cV^{1/2}} \mathbf{f}_{k,l}(a_{k,l}e^{-ikr} + ia_{k,l}^+e^{ikr}) - \frac{2\pi e}{cV^{1/2}} \sum \mathbf{f}_{k,l}[\mathbf{f}_{k,l}(j_ke^{-ikr} + j_{-k}e^{ikr})] \\ &\quad - \frac{4\pi e}{cV^{1/2}} \sum' (k/k^2) [k(j_ke^{-ikr} + j_{-k}e^{ikr})]. \end{aligned}$$

Remembering the definition of  $j_k$  (Eq. (9)) and formula (18'') this becomes

$$\sum_{k,l} \frac{4\pi^2\nu^{3/2}}{cV^{1/2}} \mathbf{f}_{k,l}(a_{k,l}e^{-ikr} + ia_{k,l}^+e^{ikr}) = \frac{\partial \mathcal{E}}{c\partial t} + 4\pi(e/c)j. \quad (19)$$

The left side of this equation is however by (5) and (2')<sup>8</sup>  $\text{rot } \mathcal{H}$ . It should be observed that  $\text{div } \mathbf{A} \neq 0$  so that  $\text{rot } \mathcal{H} \neq -\Delta \mathbf{A}$  but  $\text{rot } \mathcal{H} = \text{grad div } \mathbf{A} - \Delta \mathbf{A}$ . For the part of  $\mathbf{A}$  containing  $a_{k,l}$ ,  $a_{k,l}^+$  the divergence vanishes while for the second part in  $a_k$ ,  $a_k^*$  the rot and therefore the  $\text{rot}^2$  operators give zero.

#### §4. Meaning of $\text{div } \mathcal{E} = 4\pi e\rho$ and summary of preceding discussion

It remains to be seen what follows for  $\text{div } \mathcal{E}$ . From (14) we see that the only contributions come from the second summation involving  $b_k$ ,  $b_k^*$  because  $(k \cdot f) = 0$ . The initial values of  $b_k$ ,  $b_k^*$  have been arbitrary so far. Their rate of change however has already been fixed by the Hamiltonian and is connected with the current  $j$ . The charge density  $e\rho$  is similarly left undetermined so far as its initial values are concerned. Its rate of change is given by

$$\partial \rho / \partial t + \text{div } j = 0. \quad (20)$$

<sup>8</sup> Use being made also of  $[k \times [k \times f]] = -k^2 f + k(kf) = -k^2 f$  since  $kf = 0$ .

Eq. (14) gives

$$\operatorname{div} \mathcal{E} = \sum' 4\pi i (\nu/V)^{1/2} k (b_k^* e^{-ikr} - b_k e^{ikr}) \quad (21)$$

and using (17)

$$\operatorname{div} \frac{\partial \mathcal{E}}{\partial t} = \sum' 4\pi e i V^{-1/2} k (j_k e^{-ikr} - j_{-k} e^{ikr}). \quad (22)$$

By (9) the right side of this equation is connected with  $\operatorname{div} j$  which in turn is connected with  $\partial \rho / \partial t$  by the continuity Eq. (20). We may thus write instead of (22)

$$(\partial / \partial t)(\operatorname{div} \mathcal{E}) = -4\pi e \operatorname{div} j = 4\pi e (\partial \rho / \partial t). \quad (23)$$

This equation is consistent with

$$\operatorname{div} \mathcal{E} = 4\pi e \rho \quad (24)$$

but the latter equation does not follow in general from (23). On the contrary by assigning suitable initial values to  $b_k, b_k^*$  we can make  $\operatorname{div} \mathcal{E} - 4\pi e \rho$  have arbitrary initial values through all space. If however we choose these initial values so as to satisfy (24) at  $t=0$ , this Eq. (24) will remain satisfied in the course of time on account of (23). The equation  $\operatorname{div} \mathcal{H} = 0$  is, of course satisfied identically on account of the second Eq. (3).

We recapitulate the above discussion. The vector potential  $\mathbf{A}$  is analyzed into a three dimensional Fourier Series (8) which is a sum of  $e^{-ikr}$  multiplied by vector coefficients. The amplitudes of the components of the vector coefficients are chosen as the dynamical variables. One set of variables  $a_k, a_k^*$  refers to components along the wave normal  $k$  and serves to represent the electrostatic field. Another set  $a_{k,i}, a_{k,i}^+$  refers to components perpendicular to  $k$  and serves to represent transverse electromagnetic waves. *It is found that if  $\mathcal{E}$  and  $\mathcal{H}$  are defined by (14), (15) and the Hamiltonian function  $H$  by (12), (11), (10) or else (12), (15), (16) the equations of motion are (13), (17), (18). Maxwell's equations, with the exception of  $\operatorname{div} \mathcal{E} = 4\pi e \rho$ , are a consequence of the equations of motion. The correct value of  $\operatorname{div} \mathcal{E}$  is however maintained in the course of time as a consequence of the equations of motion provided it has the correct value initially (at  $t=0$ ) and provided also the continuity Eq. (20) is satisfied.*

It is natural that having treated time and space unsymmetrically by making the scalar potential zero we have obtained a difference in treatment for  $\operatorname{div} \mathcal{E} = 4\pi e \rho$  and the related triplet of Maxwell's equations  $\operatorname{rot} \mathcal{H} = \partial \mathcal{E} / c \partial t + 4\pi (e/c) j$ . In some problems the above formulation is awkward. It is rather convenient however for a generalization of classical electrodynamics in quantum theory. The dissymmetry between space and time is then not a drawback because the present quantum dynamics is essentially unsymmetrical in these variables. The dissymmetry is of course entirely in formulas involving  $\mathbf{A}$ . The final result (Maxwell's equations) is relativistic and symmetrical. This way of treating the electrostatic field has been introduced by Heisenberg and Pauli.<sup>1</sup>

### §5. Elimination of electrostatic field

If in all of the formulas we omit the variables  $a_k, a_k^*, b_k, b_k^*$  we obtain a set of equations which is still approximately correct. This set of equations and the corresponding Hamiltonian function have been used by Dirac<sup>3</sup> in his first quantum mechanical treatment of radiation. The magnetic field  $\mathcal{H}$  does not involve the omitted variables at all and is unaffected by this omission. The electric field  $\mathcal{E}$  is seen by (14) to involve  $b_k, b_k^*$ . Dirac's theory is thus not altogether logical in this respect. The missing part is

$$\mathcal{E}' = - \sum' 4\pi(\nu/V)^{1/2} (\mathbf{k}/k)(b_k^* e^{-i\mathbf{k}\cdot\mathbf{r}} + b_k e^{i\mathbf{k}\cdot\mathbf{r}}) \quad (25)$$

This is just the part giving rise to  $\text{div } \mathcal{E}$  as is seen in Eq. (21). Thus  $\mathcal{E}'$  satisfies

$$\text{div } \mathcal{E}' = 4\pi e\rho, \text{rot } \mathcal{E}' = 0$$

the last equation being identically satisfied on account of the form of  $\mathcal{E}'$  independently of the values of  $b_k, b_k^*$ . We may thus call  $\mathcal{E}'$  the electrostatic field due to the charge density  $e\rho$  and take it to be given by

$$\mathcal{E}' = - \text{grad } e \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\tau' \quad (25')$$

this being the only solution of the above equations which vanishes at infinity. It should be remembered that  $\rho$  is also a function of the time. Hence  $\mathcal{E}'$  varies with the time. By calling  $\mathcal{E}'$  the electrostatic field we do not wish to imply constancy but to emphasize the fact that at any time  $t_0$  the field  $\mathcal{E}'$  is identical with an electrostatic field which would exist at the time  $t_0$  if the charge density  $\rho(\mathbf{r}, t)$  had the value  $\rho(\mathbf{r}, t_0)$  for a long time before the instant  $t_0$ . *The field  $\mathcal{E}'$  thus vanishes at least as rapidly as  $(1/r^2)$  at large distances and cannot be responsible therefore for the emission of radiation.* The energy of the radiation field  $H_R$  consists of two parts. The second of these is  $\sum' 4\pi\nu b_k b_k^*$  and is simply

$$W(\mathcal{E}') = (1/8\pi) \int \mathcal{E}'^2 d\tau \quad (25'')$$

with  $\mathcal{E}'$  given by (25'). Consider the field  $\mathcal{E}'$  due to two particles the dimensions of both particles being supposed small compared to their distance apart. For simplicity the particles may be taken to be spheres of very small radii  $a_1, a_2$ . Then (25'') may be replaced by

$$e_1^2/2a_1 + e_2^2/2a_2 + e_1 e_2 / r_{12}$$

where  $r_{12}$  is the distance between the two particles, neglecting for the present the Lorentz-Fitzgerald contraction. The first two terms give a constant contribution to the energy and the only term of interest is  $e_1 e_2 / r_{12}$ . This term was of course included by Dirac in his theory. If the Lorentz contraction is taken into account the terms in  $e^2/2a$  are no longer correct. Since, however, we do not possess a consistent theory of electromagnetic mass the best that can be

done is to consider their variation to be properly taken care of by the relativistic expression for the kinetic energy

$$mc^2/(1 - \beta^2)^{1/2} \quad (\beta = v/c).$$

We see therefore that, *aside from questions dealing with the interaction energy of a particle with its own field, the energy of the field may be correctly calculated by omitting all terms in  $a_k, a_k^*, b_k, b_k^*$  and introducing instead the mutual energies of the particles in the form  $e_1 e_2 / r_{12}$* . This is precisely what has been done by Dirac. As will be seen later quite similar considerations hold if matter and radiation together are treated by a single Hamiltonian function. An analogous discussion can be carried through also in the quantum theory [V §3].

### §6. Self energy

We must now return to the question of self energy. The electromagnetic momentum of a spherical distribution of charge having a radius  $a$  (as measured by an observer moving with the particle) a charge  $e$ , and moving with velocity  $v$  is

$$\mathbf{G} = \frac{2e^2}{3ac^2} \frac{\mathbf{v}}{(1 - \beta^2)^{1/2}}. \quad (26)$$

The energy of the field in the reference system of the particle is

$$W_0 = e^2/2a$$

so that

$$\mathbf{G} = \frac{4(W_0/c^2)\mathbf{v}}{3(1 - \beta^2)^{1/2}}. \quad (26')$$

The energy of the electromagnetic field is on the other hand

$$W = W_0 \frac{1 + (\beta^2/3)}{(1 - \beta^2)^{1/2}} \quad (27)$$

of which the part

$$(1/8)\pi \int \mathcal{E}^2 d\tau = W_0 \frac{1 - (\beta^2/3)}{(1 - \beta^2)^{1/2}} \quad (27')$$

is the energy of the electric field, and the magnetic field contributes a part

$$(1/8\pi) \int \mathcal{H}^2 d\tau = \frac{2W_0\beta^2}{3(1 - \beta^2)^{1/2}}.$$

According to the classical investigations of v. Laue and Einstein the occurrence of the factor  $(4/3)$  in  $(26')$  and of the term  $\beta^2/3$  is attributable to the influence of the material stress energy tensor. The energy of the electrostatic field  $\mathcal{E}'$  is not  $(27')$ , because for the field  $\mathcal{E}'$ , we have  $\text{rot } \mathcal{E}' = 0$ , while  $\mathcal{E}$  is the total electric field for which  $\text{rot } \mathcal{E} = -\partial \mathcal{H}/\partial t \neq 0$ . The energy of the field  $\mathcal{E}'$  as

given by the expression (25'') is easily found if it is remembered that the spherical distribution of charge appears to a stationary observer to be a very thin layer of charge contained between two concentric and *similar* oblate ellipsoids of revolution obtained by contracting two adjacent concentric spheres in the direction of motion by the Lorentz-Fitzgerald factor  $(1-\beta^2)^{1/2}$ . The small axis of the ellipsoid is  $a(1-\beta^2)^{1/2}$  and the equatorial radius is  $a$ . Since the amount of charge is conserved in a Lorentz transformation the amounts of charge in corresponding sections of the spheres and the ellipsoids are equal. It is well known in electrostatics that such an ellipsoidal distribution of charge is identical with the charge distribution on the surface of a charged conducting ellipsoid.<sup>9</sup> This reasoning is quite similar to the one used by Abraham in his prerelativistic discussion of the electromagnetic mass. Only in our case the ellipsoid is oblate while in his case it is prolate. The electrostatic energy (25'') is now obtained from the familiar expression for the electric capacity of an oblate spheroid<sup>9</sup> which in our case is  $\alpha\beta/\sin^{-1}\beta$ . Thus

$$W(\mathcal{E}') = \frac{e^2}{2a} \frac{\sin^{-1}\beta}{\beta} = W_0 \frac{\sin^{-1}\beta}{\beta}. \quad (28)$$

Comparing this with (27) we see that for low velocities ( $\beta \ll 1$ )  $W(\mathcal{E}')$  is the main part of  $W$ , the difference between them being of the order  $\beta^2 W_0$ . This difference is to within terms in  $v^2/c^2$

$$W - W(\mathcal{E}') = (2/3)W_0\beta^2 = \frac{1}{2}(4W_0/3c^2)v^2. \quad (28')$$

Formula (28') gives the contribution to the self energy of the electron due to the first summation in (15). It is somewhat gratifying to see that to this order omitting the electrostatic field brings about an energy change (28') agreeing exactly with the expression for the electromagnetic mass

$$m_0 = 4W_0/(3c^2) \quad (28'')$$

which is required by (26'). This fact is at present of little value in quantum theory applications because for a point electron ( $a=0$ ),  $W_0$  and  $m_0 = \infty$ . It is further seen that for velocities approaching the velocity of light ( $\beta=1$ ) the expression (28) remains finite

$$(W(\mathcal{E}'))_{\beta \rightarrow 1} = (\pi/2)W_0$$

while (27) becomes infinite. In this limiting case

$$W - W(\mathcal{E}') \cong (4/3)W_0/(1-\beta^2)^{1/2} - \pi W_0/2 + \text{terms in } (1-\beta^2)^{1/2} \quad (28''')$$

which again happens to agree with the electromagnetic mass formula. For intermediate values of the velocity the omission of the electrostatic field does not lead to a simple result for the self energy. If the electrostatic field is not omitted we must use the whole expression (27). It does not represent correctly the energy changes of a particle with velocity. It again happens to agree with (28'') for  $\beta \rightarrow 1$ . This as well as the agreement of (28'), (28''') with (28'')

<sup>9</sup> Jeans, *Electricity and Magnetism*, 4th edition, p. 248.

must be regarded as accidental because the total energy of the electromagnetic field given by (27) varies with velocity in a different way from the kinetic energy of a particle of mass  $m$ , viz.:

$$mc^2/(1 - \beta^2)^{1/2}.$$

We must be prepared to find, therefore, that the solution of all the equations following from the Hamiltonian function (12) will lead to total energies of the radiation field which cannot be interpreted as changes in the kinetic energy with velocity. We must also be prepared to find that the corresponding energy for a point electron will be infinite. Neither of these faults should be attributed to the Hamiltonian treatment. It is intrinsic to any theory attempting to discuss the energy of the electromagnetic field without taking into account the stresses holding the electron together. It should be borne in mind that all of the preceding discussion refers to the energy of the field of radiation and not to the total energy, which has no direct meaning so far, since we have not introduced the material mass and since we have not discussed yet the behavior of matter.

#### §7. Matter and field treated by one Hamiltonian

So far matter has been supposed to be in a known state. We proceed to set up a Hamiltonian function to describe the behavior of matter as well. For simplicity we consider one particle of mass  $m$  and charge  $e$ . The customary nonrelativistic Hamiltonian function for such a particle moving in an *external* field of force described by the vector potential  $\mathbf{A}$  is

$$\mathbf{p}^2/(2m) + V(\mathbf{r}) - (e/mc)\mathbf{p}\mathbf{A}(\mathbf{r}) + e^2(\mathbf{A}(\mathbf{r}))^2/(2mc^2)$$

where  $\mathbf{p}$  is the momentum of the particle.

Here  $V(\mathbf{r})$  is the potential energy which may be made to include forces of other than electromagnetic origin. If we attempt to add this function to the energy of the field ( $H_R$ ) we see that the result has many of the characteristics necessary for the combined treatment of field and matter. It is not satisfactory, however, because the vector potential in the term  $\mathbf{p}\mathbf{A}$  has to be taken in different ways for the two treatments. In order to obtain correct equations of motion,  $\mathbf{A}$  must be taken as the external potential, i.e., the total potential minus the potential due to the particle itself while for the treatment of the radiation  $\mathbf{A}$  must include the potential due to the particle. Furthermore, the field due to a point charge of infinitesimal dimensions is infinite in its immediate vicinity and even has a singularity. Its infinite value depends in fact on the direction from which the particle is approached. We see, therefore, that the above expression must be reinterpreted so as to have a meaning and so as to represent correctly the physical behavior of the system. To accomplish this purpose we begin by endowing the particle with a small but finite space extension. We set up a Hamiltonian function representing its behavior and then pass to the limit of a very small space extension of the charge. In discussing matter nonrelativistically we may neglect the Lorentz-Fitzgerald contraction so that the particle may be said to have a shape and size independ-

ent of velocity. We single out a point in the charge which we call its center and to which we refer as

$$\mathbf{r} = (x, y, z).$$

The charge density at any other point  $\mathbf{r}'$  in space we then take to be given by

$$eD(\mathbf{r}' - \mathbf{r}). \quad (29)$$

The function  $D$  describes the shape and size of the charge distribution. We suppose it to be spherically symmetrical about the center  $\mathbf{r}$  so that  $D$  is a function only of  $[(x' - x)^2 + (y' - y)^2 + (z' - z)^2]^{1/2}$ . In order to have the total charge contained in the particle equal to  $e$  we must take

$$\int D(\mathbf{r}' - \mathbf{r}) dV' = 1. \quad (29')$$

We note that if the spatial extension of  $D$  becomes very small this function approaches Dirac's  $\delta$  function as a limit. The Hamiltonian function for matter and radiation together may now be taken as:

$$\begin{aligned} H = H_R - (e/mc) \int \mathbf{p} \mathbf{A}(\mathbf{r}') D(\mathbf{r}' - \mathbf{r}) dV' \\ + (e^2/(2mc^2)) \left[ \int \mathbf{A}(\mathbf{r}') D(\mathbf{r}' - \mathbf{r}) dV' \right]^2 + \mathbf{p}^2/(2m) + V(\mathbf{r}) \end{aligned} \quad (30)$$

where  $H_R$  is given by (10) and (15) and  $\mathbf{A}$  is given by (8). The dynamical variables are  $a_{k,i}, a_{k,i}^+, a_k, a_k^*, b_k, b_k^*$  for radiation and  $\mathbf{r} = (x, y, z)$ ,  $\mathbf{p} = (p_x, p_y, p_z)$  for matter. Eq. (14) for  $\mathcal{E}$  and the relation  $\mathcal{H} = \text{rot } \mathbf{A}$  we again regard as definitions of the field strengths  $\mathcal{E}$ ,  $\mathcal{H}$ . Hamilton's equations of motion for  $x, y, z$  are from (30)

$$\dot{\mathbf{r}} = (\mathbf{p}/m) - (e/(mc)) \int \mathbf{A}(\mathbf{r}') D(\mathbf{r}' - \mathbf{r}) dV'. \quad (31)$$

For  $dp_x/dt$  we obtain, using (31)

$$\frac{dp_x}{dt} = - \frac{\partial V}{\partial x} + (e/c) \int D(\mathbf{r}' - \mathbf{r}) \left( \dot{\mathbf{r}} \frac{\partial \mathbf{A}(\mathbf{r}')}{\partial x'} \right) dV' \quad (32)$$

This Eq. (32) is related to but is not the force equation. For we see from (31) that  $\mathbf{p}$  and  $m\dot{\mathbf{r}}$  are not identical. In order to obtain  $m\dot{\mathbf{r}}$  with the aid of (31) and (32) we need also to know the time derivative of  $\mathbf{A}(\mathbf{r}')$ . This time derivative we write  $\partial \mathbf{A}(\mathbf{r}')/\partial t$  in order to emphasize the fact that  $\mathbf{r}'$  is a fixed point in space. This notation is convenient in order to have exact correspondence with the customary notation of hydrodynamics but it is not necessary since  $\partial \mathbf{A}(\mathbf{r}')/\partial t = d\mathbf{A}(\mathbf{r}')/dt$ . We have

$$\begin{aligned} \frac{d}{dt} \int A(\mathbf{r}') D(\mathbf{r} - \mathbf{r}') dV' &= \int \frac{\partial A(\mathbf{r}')}{\partial t} D(\mathbf{r}' - \mathbf{r}) dV' \\ &+ \int A(\mathbf{r}') \sum_{x,y,z} \dot{x} \frac{\partial D(\mathbf{r} - \mathbf{r}')}{\partial x} dV'. \end{aligned}$$

Performing integrations by parts in the second integral

$$\frac{d}{dt} \int A(\mathbf{r}') D(\mathbf{r}' - \mathbf{r}) dV' = \int D(\mathbf{r}' - \mathbf{r}) \left[ \frac{\partial A(\mathbf{r}')}{\partial t} + \sum_{x,y,z} \dot{x} \frac{\partial A(\mathbf{r}')}{\partial x'} \right] dV'. \quad (32')$$

Combining this with (31), (32) we have

$$m\ddot{\mathbf{r}} = -\frac{\partial V}{\partial \mathbf{r}} + e \int D(\mathbf{r}' - \mathbf{r}) \{ \mathcal{E}(\mathbf{r}') + (1/c) [\dot{\mathbf{r}} \times \mathcal{H}(\mathbf{r}')] \} dV' \quad (33)$$

where the validity of Eqs. (3) has been presupposed. It will have to be verified later that the first Eq. (3) is satisfied as a consequence of (30). Eq. (33) may be interpreted as meaning that the charge element  $eD(\mathbf{r}' - \mathbf{r})dV'$  located at  $\mathbf{r}'$  is subjected to a force  $\mathcal{E}(\mathbf{r}') + (1/c) [\dot{\mathbf{r}} \times \mathcal{H}(\mathbf{r}')]$  per unit charge. This is surely correct so far as any external field  $\mathcal{E}$ ,  $\mathcal{H}$  is concerned. It is also fairly satisfactory for the discussion of reactions due to the field produced by the particle itself. It satisfies the conservation of energy and momentum for field and matter together. For it is precisely for this force that the conservation theorems for field and matter are known to hold. Thus the work done on matter by the field is equal to the rate of decrease of the energy of the field as given by Poynting's flux and by (10), and the rate of change of momentum of the matter and radiation together is equal to zero. The Hamiltonian (30) *should* give therefore *Abraham's prerelativistic results for the electromagnetic mass*. It includes the radiation reaction term  $-(2/3)(e^2/c^3)\ddot{\mathbf{r}}$ .

It remains to verify that we have not destroyed the validity of the first Eq. (3) and of other Maxwell's equations. With  $H$  given by (30) we have for any dynamical variable describing the electromagnetic field, for example for  $a_{k,l}$

$$\frac{\partial H}{\partial a_{k,l}} = \frac{\partial H_R}{\partial a_{k,l}} - (e/c) \int \dot{\mathbf{r}} \frac{\partial A(\mathbf{r}')}{\partial a_{k,l}} D(\mathbf{r}' - \mathbf{r}) dV'. \quad (34)$$

We compare this with (11) and (12). It is clear that all the equations following from (30) for the field are exactly the same as those following from (12) the only difference being that instead of  $\mathbf{j}$  we have

$$\mathbf{j} = \dot{\mathbf{r}} D(\mathbf{r}' - \mathbf{r}). \quad (34')$$

The field thus is produced by a current  $e\mathbf{j}$  due to the motion of the charge distribution  $eD(\mathbf{r}' - \mathbf{r})$  with the velocity  $\dot{\mathbf{r}}$ . It is also true that the charge density

$$e\rho = eD(\mathbf{r}' - \mathbf{r})$$

satisfies the continuity Eq. (20). In fact



$$\operatorname{div} j = \left( \dot{x} \frac{\partial}{\partial x'} + \dot{y} \frac{\partial}{\partial y'} + \dot{z} \frac{\partial}{\partial z'} \right) D(\mathbf{r}' - \mathbf{r})$$

while  $\partial \rho / \partial t$  is just the negative of this expression.

The Hamiltonian function (30) does not involve the time explicitly and should be an integral of motion. This is the case, for by (31) the expression (30) is

$$H = (m/2)\dot{\mathbf{r}}^2 + V(\mathbf{r}) + H_R. \quad (35)$$

The constant value of the Hamiltonian function is, as was to be expected, the total energy of matter and radiation. It is the sum of the kinetic energy, the potential energy  $V$  with respect to forces of nonelectromagnetic origin and the energy of the electromagnetic field  $H_R$ . It may not be superfluous to remark that the numerical value of the kinetic energy is not simply related to the term  $\mathbf{p}^2/(2m)$  in (30) and that a separation of the expression (30) into an energy of matter  $\mathbf{p}^2/(2m) + V(\mathbf{r})$ , an energy of radiation  $H_R$  and an interaction energy is highly artificial from a physical point of view. The simpler aspect is given by the Eq. (35).

The general features of the discussion of the value of  $H$  for an electron in a uniform motion of translation contained in Eqs. (26) to (28'') applies to (30), (35). Only everywhere Abraham's rather than the relativistic expressions should be used.

The Hamiltonian (30) is easily generalized to the case of several particles and becomes

$$H = H_R + \sum_i (1/(2m_i)) \left[ \mathbf{p}_i - (e_i/c) \int A(\mathbf{r}') D(\mathbf{r}' - \mathbf{r}_i) dV' \right]^2 + \sum_i V_i(\mathbf{r}_i) \quad (30')$$

where the index  $i$  refers to the different particles.

In order to treat matter relativistically we try the Hamiltonian function

$$H = H_R + c \left[ m^2 c^2 + (\mathbf{p} - (e/c) \int A(\mathbf{r}') D(\mathbf{r}' - \mathbf{r}) dV')^2 \right]^{1/2}. \quad (36)$$

The equations become shorter if we introduce the abbreviation

$$\int B(\mathbf{r}') D(\mathbf{r}' - \mathbf{r}) dV' = \overline{B(\mathbf{r})}$$

meaning by the bar an average taken over the space extension of the charge distribution. Hamilton's equations of motion following from (36) are:

$$\dot{x} = \frac{c(p_x - (e/c)\bar{A}_x)}{\left[ \sum_{x,y,z} (p_x - (e/c)\bar{A}_x)^2 + m^2 c^2 \right]^{1/2}} \quad (36')$$

from which

$$p_x - (e/c)\bar{A}_x = m\dot{x}/(1 - \beta^2)^{1/2}, \quad c^2\beta^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 \quad (36'')$$

$$\frac{dp_x}{dt} = \frac{e}{c} \left\{ \dot{x} \frac{\partial \bar{A}_x}{\partial x} + \dot{y} \frac{\partial \bar{A}_y}{\partial x} + \dot{z} \frac{\partial \bar{A}_z}{\partial x} \right\} \quad (37)$$

$$\frac{e}{c} \frac{d\bar{A}_x}{dt} = \frac{e}{c} \frac{\partial \bar{A}_x}{\partial t} + \frac{e}{c} \left\{ \dot{x} \frac{\partial \bar{A}_x}{\partial x} + \dot{y} \frac{\partial \bar{A}_x}{\partial y} + \dot{z} \frac{\partial \bar{A}_x}{\partial z} \right\} \quad (37')$$

whence by (36'')

$$\begin{aligned} \frac{d}{dt} \frac{m\dot{x}}{(1-\beta^2)^{1/2}} &= \frac{d}{dt} \left( p_x - \frac{e}{c} \bar{A}_x \right) \\ &= -\frac{e}{c} \frac{\partial \bar{A}_x}{\partial t} + \frac{e}{c} (\dot{y} \text{rot}_z \bar{A} - \dot{z} \text{rot}_y \bar{A}). \end{aligned} \quad (38)$$

For the dynamical variables describing radiation Eq. (34) remains correct with  $\mathbf{r}$  given by (36'). The Hamiltonian function (36) is therefore satisfactory for the discussion of the motion of matter under the influence of an external field. It does not represent correctly the action of the electron on itself. To represent this action correctly we would have to take into account the Lorentz contraction by using instead of  $D(\mathbf{r}' - \mathbf{r})$  everywhere

$$D(\mathbf{r}' - \mathbf{r}) \rightarrow (1 - \beta^2)^{-1/2} D \left( \mathbf{r}' - \mathbf{r} + \frac{((\mathbf{r}' - \mathbf{r})\mathbf{v})\mathbf{v}}{v^2} \left( \frac{1}{(1 - \beta^2)^{1/2}} - 1 \right) \right)$$

where  $\mathbf{v}$  is the velocity. Such a Hamiltonian function has not been devised. The necessary changes are more easily made directly in the quantum theory because of the independence of velocity and momentum in Dirac's relativistic equation for the electron.

## II. THE KRAMERS-HEISENBERG DISPERSION FORMULA DERIVED BY MEANS OF THE CORRESPONDENCE PRINCIPLE

### §1. The Hamiltonian function used

Employing the correspondence principle of Bohr, Kramers<sup>10</sup> proposed a view which generalized the classical formula so as to include in it negative dispersion terms and incoherent scattering. The detailed derivation we owe to Kramers and Heisenberg.<sup>11</sup> Their theory concerns itself with the electric polarization of an atom which is under the influence of an externally applied monochromatic electromagnetic wave. The theory has been worked out for the case in which the frequency of the wave is not in resonance with the atomic frequencies. The reaction of the scattered radiation on the atom is neglected. The wave-length of the incident radiation is supposed large in comparison with atomic dimensions. The whole consideration applies to an atom which in the absence of radiation forms a conditionally periodic system.

Kramers and Heisenberg start out by considering the behavior of a conditionally periodic system under the influence of a light wave. It is supposed

<sup>10</sup> H. A. Kramers, *Nature* **113**, 673 (1924); **114**, 310 (1924).

<sup>11</sup> H. A. Kramers and W. Heisenberg, *Zeits. f. Physik* **31**, 681 (1925); See also J. H. Van Vleck, *Phys. Rev.* **24**, 330 (1924); Particularly p. 344.

that in the absence of the perturbing influence of the light wave the atom may be described by a set of angle variables and a corresponding set of action variables

$$J_1, J_2, \dots, J_f$$

where  $f$  is the number of degrees of freedom of the system. The system is supposed to be nondegenerate. The electric polarization of the system may be expressed in terms of the angle and action variables in the form of a Fourier series

$$\mathcal{M}(t) = \sum_{\tau_1, \dots, \tau_f} (1/2) C_{\tau_1, \dots, \tau_f} e^{2\pi i(\tau_1 w_1 + \dots + \tau_f w_f)} \equiv \sum_{\tau} (1/2) C_{\tau} e^{2\pi i(\tau w)}.$$

The polarization and therefore also the coefficients  $C_{\tau}$  are vectors. The components of the coefficients  $C_{\tau}$  may be real or complex. The summation in the above formula is supposed to be extended over all positive and negative *integral* values for each  $\tau_s$  ( $s=1, 2, \dots, f$ ). Instead of writing the suffixes  $\tau_1, \dots, \tau_f$  it is often customary to replace them by the single letter  $\tau$ . Instead of  $\tau_1 w_1 + \dots + \tau_f w_f$  we similarly write  $(\tau w)$ . Each  $C_{\tau}$  is itself a function of the action variables  $J$ . In the absence of perturbations the  $J$ 's are constants and

$$dw_s/dt = \nu_s = \partial H_0 / \partial J_s$$

(where  $H$  is the Hamiltonian) are also constants. Every term of (39) contains the time in the form  $e^{2\pi i(\tau \nu)t}$  and may be said to represent together with its conjugate the Fourier component of the polarization of frequency  $(\tau \nu) = \tau_1 \nu_1 + \dots + \tau_f \nu_f$ . Needless to say, since  $\mathcal{M}(t)$  is real,

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

where the  $*$  represents the complex conjugate of a quantity. If now the wavelength of the incident radiation is so long that variations in the electric field strength of the wave within atomic dimensions may be neglected, and if the atom is supposed to be affected little by the radiative reactions due to the electromagnetic waves which it sends out, the action of the incident light wave may be represented by taking the Hamiltonian function of the perturbed atom to be

$$H = H_0 - \mathcal{E}(t)\mathcal{M}(t) \quad (40)$$

where  $\mathcal{E}(t)$  is the electric intensity of the incident light wave at the atom. The term added to  $H$  in (40) is, of course, a scalar product of two vectors. In order to see that (40) is a satisfactory Hamiltonian function it is sufficient to write out the equations of motion say in Cartesian coordinates. This we leave to the reader.

## §2. Transformation to new action variables

The action and angle variables for  $H_0$  are not action and angle variables for the Hamiltonian function (40). Thus the action variables  $J$  are no longer constants of the motion and the angle variables do not depend on the time

linearly. Nevertheless it is possible to find a new set of angle-action variables  $\bar{w}, \bar{J}$  closely related to  $w, J$  such that under the influence of the light wave the  $\bar{J}$  are constant and the  $\bar{w}$  depend on the time linearly. The transformation  $w, J \rightarrow \bar{w}, \bar{J}$  is made to be canonical by making

$$\sum_s (\bar{J}_s d\bar{w}_s + w_s dJ_s) = d(\sum_s J_s \bar{w}_s + K(\bar{w}, J; t))$$

so that

$$\bar{J}_s - J_s = \partial K / \partial \bar{w}_s, \quad \bar{w}_s - w_s = -\partial K / \partial J_s. \quad (41)$$

So far  $K$  is arbitrary. It is desirable to choose it so that after  $w, J$  are expressed in terms of  $\bar{w}, \bar{J}$  by means of (41) and are substituted into the new Hamiltonian, there should remain a pure function of  $\bar{J}$ . It should be remembered, however, that if  $K$  involves the time the new Hamiltonian function is not (40) but

$$H(J(\bar{J}, \bar{w}), w(\bar{J}, \bar{w})) - \frac{\partial}{\partial t} K(\bar{w}, J(\bar{J}, \bar{w}); t). \quad (40')$$

An accurate calculation for  $K$  would be difficult. It is easily made, however, to the first order in  $\mathcal{E}(t)$ . This means that if  $J, w$  are expanded as Taylor's Series in terms of  $\mathcal{E}(t)$  with coefficients as functions of  $\bar{J}, \bar{w}$ , if then the results are substituted into (40') and this expression (40') is similarly expanded we can make  $\bar{w}$  and  $t$  disappear in all terms of lower order than the second in  $\mathcal{E}(t)$ . For a monochromatic incident wave we may set

$$\mathcal{E}(t) = \text{Re}(\mathcal{E} e^{2\pi i \nu t}) = (1/2)(\mathcal{E} e^{2\pi i \nu t} + \mathcal{E}^* e^{-2\pi i \nu t}). \quad (42)$$

Then by (40), (40') and the first Eq. (41) we have to the first order in  $\mathcal{E}$  for the Hamiltonian (40')

$$H_0(\bar{J}) - \sum_s \frac{\partial K}{\partial \bar{w}_s} \left( \frac{\partial H_0}{\partial J_s} \right)_{J=\bar{J}} - \mathcal{E}(t) \mathcal{M}(t) - \frac{\partial K}{\partial t}. \quad (40'')$$

By making the sum of the last three terms in (40'') vanish we obtain a Hamiltonian depending only on  $\bar{J}$ . The transformation function  $K$  is thus itself of the first order in  $\mathcal{E}(t)$ . A possible  $K$  is

$$K = - (1/2) \text{Re} \left[ \sum_{\tau} \frac{\mathcal{E} C_{\tau}(J) e^{2\pi i [(\tau \bar{w}) + \nu t]}}{2\pi i [(\tau \bar{w}) + \nu]} \right]. \quad (43)$$

In this formula the summation is again extended over all positive and negative integral values of  $\tau$ . What has been shown so far is that under the action of the light wave the quantities  $\bar{J}$  defined by (41), (43) are constants of the motion and that the  $\bar{w}$  depend on the time linearly. Further

$$d\bar{w}/dt = (\partial H_0 / \partial J)_{J=\bar{J}}.$$

Substituting the values of  $J, w$  in terms of  $\bar{J}, \bar{w}$  into (39) one obtains

$$\mathcal{M}(t) = \mathcal{M}_0(t) + \mathcal{M}_1(t) \quad (44)$$

where

$$\mathcal{M}_0(t) = \sum_t (1/2) C_\tau(\bar{J}) e^{2\pi i(\bar{r}\bar{w})t} \quad (44')$$

and

$$\begin{aligned} \mathcal{M}_1(t) = Re \left\{ \sum_{\tau, \tau'} (1/4) \left[ \frac{\mathcal{E} C_{\tau'}(\bar{J})}{(\tau' \nu(\bar{J}) + \nu)} \left( \tau' \frac{\partial}{\partial J} \right) C_\tau(\bar{J}) - \right. \right. \\ \left. \left. - C_\tau \left( \tau \frac{\partial}{\partial J} \right) \left( \frac{\mathcal{E} C_{\tau'}(\bar{J})}{(\tau' \nu(\bar{J}) + \nu)} \right) \right] e^{2\pi i[(\tau + \tau')\bar{w} + \nu t]} \right\} \end{aligned} \quad (44'')$$

The part  $\mathcal{M}_0(t)$  depends on the time exactly in the same way as (39) provided the same values are given to the  $\bar{J}$  in (44') as to the  $J$  in (39). It may be said to represent the unperturbed polarization of the atom. In addition there is the part  $\mathcal{M}_1(t)$  given by (44''). It depends linearly on  $\mathcal{E}$  and may be said to represent the polarization induced by the incident light wave. Formula (44'') may be also rewritten so as to collect all terms with the same exponentials

$$\begin{aligned} \mathcal{M}_1(t) = Re \left\{ \sum_{\tau_0} \sum_{\tau} (1/4) \left[ \frac{(\mathcal{E} C_{\tau'})}{(\tau' \nu) + \nu} \left( \tau' \frac{\partial}{\partial J} \right) C_\tau - \right. \right. \\ \left. \left. - C_\tau \left( \tau \frac{\partial}{\partial J} \right) \left( \frac{(\mathcal{E} C_{\tau'})}{(\tau' \nu) + \nu} \right) \right] e^{2\pi i[(\tau_0 w) + \nu t]} \right\} \end{aligned} \quad (44''')$$

where

$$\tau_s + \tau_s' = \tau_s^0 \quad (44'''\text{a})$$

We have dropped here the bar over the symbols because from now on we work entirely with  $\bar{J}$ ,  $\bar{w}$  so that no confusion with  $J$ ,  $w$  can arise. It should be remembered that *from now on* when we write  $J$ ,  $w$  we mean *the action and angle variables of the perturbed system*. In the formula (44'') the summation over  $\tau$  is supposed to be performed first for fixed  $\tau^0$ . The values of  $\tau'$  to be used are given by (44'''\text{a}). If  $\tau^0 = 0$  we obtain from (44''') a contribution depending on  $t$  only through  $e^{2\pi i \nu t}$  which is synchronous with the incident wave. In addition there are other terms which are not synchronous with it and which arise from  $\tau^0 \neq 0$ . It is seen that the synchronous part does not depend on the phases (additive constants) of the angle variables  $w$ . The opposite is the case for the nonsynchronous part. For this reason the former polarization is said to give rise to *coherent* (i.e., capable of interference) and the latter to *incoherent* scattering.

We discuss the coherent scattering first. This is due to the induced polarization

$$\begin{aligned} \mathcal{M}_1'(t) = Re \left\{ \sum_{\tau} (1/4) \left[ \frac{\mathcal{E} C_{\tau}^*}{-\nu + (\tau \nu)} \left( \tau \frac{\partial}{\partial J} \right) C_\tau + C_\tau \left( \tau \frac{\partial}{\partial J} \right) \left( \frac{(\mathcal{E} C_{\tau}^*)}{(\tau \nu) - \nu} \right) \right] e^{2\pi i \nu t} \right\} \\ = Re \left\{ \sum_{\tau} (1/4) \left( \tau \frac{\partial}{\partial J} \right) \left( \frac{C_\tau (\mathcal{E} C_{\tau}^*)}{(\tau \nu) - \nu} \right) e^{2\pi i \nu t} \right\} \end{aligned} \quad (45)$$

or

$$\mathcal{M}_1'(t) = \text{Re} \left\{ \sum'_{\tau} (1/4) \left( \tau \frac{\partial}{\partial J} \right) \left[ \frac{C_{\tau}(\mathcal{E}C_{\tau}^*)}{(\tau\nu) - \nu} + \frac{C_{\tau}^*(\mathcal{E}C_{\tau})}{(\tau\nu) + \nu} \right] e^{2\pi i \nu t} \right\} \quad (45')$$

where the accent on the summation indicates that the range of values covered by the  $\tau_1, \tau_2, \dots, \tau_f$  is such that  $(\tau\nu) > 0$ . Thus in (43') there is no term in

$$\begin{aligned} & -\tau_1 \frac{\partial}{\partial J_1} - \tau_2 \frac{\partial}{\partial J_2} - \dots - \tau_f \frac{\partial}{\partial J_f} \text{ if a term in} \\ & +\tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \dots + \tau_f \frac{\partial}{\partial J_f} \text{ occurs.} \end{aligned}$$

We remind the reader that  $(\tau\nu) \equiv \tau_1\nu_1(J) + \tau_2\nu_2(J) + \dots + \tau_f\nu_f(J)$  so that  $(\tau\nu)$  is a combination overtone of atomic frequencies. It is clear that the whole derivation applies only if there is no resonance of the incident frequency  $\nu$  with any of the overtones  $(\tau\nu)$  i.e., provided  $\mathcal{M}_1'(t)$  is truly small.

### §3. Application of the correspondence principle to coherent scattering

Only classical dynamics has been used in deriving (45'). Its form is suggestive for a quantum theory generalization. The operator

$$\left( \tau \frac{\partial}{\partial J} \right) = \tau_1 \frac{\partial}{\partial J_1} + \tau_2 \frac{\partial}{\partial J_2} + \dots + \tau_f \frac{\partial}{\partial J_f}$$

always suggests such a connection according to the correspondence principle of Bohr. The essential reason for this is that quantization consists in assigning to any  $J_i$  a set of values differing by integral multiples of Planck's constant  $h$  so that for any two atomic levels the differences

$$\Delta J_i = \tau_i h$$

where  $\tau_i$  are positive or negative integers. If  $A(J)$  is now a function of the  $J$  and if it varies slowly with the  $J$  then

$$\left( \tau \frac{\partial}{\partial J} \right) A(J) = \tau_1 \frac{\partial A}{\partial J_1} + \dots + \tau_f \frac{\partial A}{\partial J_f}$$

may be thought of as the limiting value of

$$\frac{1}{h} \left( \Delta J_1 \frac{\partial A}{\partial J_1} + \dots + \Delta J_f \frac{\partial A}{\partial J_f} \right) = \frac{1}{h} (\Delta A_{qu}(J)). \quad (46)$$

In order to apply the correspondence principle it is necessary to find a  $A_{qu}(J)$  for the right side of (46). The correspondence principle is powerless to fix the form of  $A_{qu}(J)$ . It only restricts the  $A_{qu}(J)$  to be approximately the classical  $A$  and moreover it makes it necessary to have the  $\partial A_{qu}/\partial J$  be asymptotically in agreement with  $\partial A_{\text{class.}}/\partial J$  in view of (46).

In applying this to (45') we must pay attention to the occurrence of the denominators  $(\tau\nu) \pm \nu$ . In the quantum theory the classical quantity  $(\tau\nu)$  be-

comes a possible transition frequency by (46). This suggests replacing the  $(\tau\nu)$  by the possible emission and absorption frequencies after (45') is generalized by the rule (46). We are not forced to do so by the mathematical procedure. The determining factor in our belief that this is the correct generalization is to be found in the experimental facts. It is well known that emission and absorption frequencies correspond to regions of anomalous dispersion. Unless the  $(\tau\nu)$  in the denominators of (45') are generalized into emission and absorption frequencies we will not obtain agreement with this important experimental fact.

The numerators in (45') contain the  $C_\tau$ . According to Bohr these Fourier coefficients are related to the probabilities of emission by an atom in the following way. Consider an atom in a certain state defined by the values of action variables

$$J_i = J'_i.$$

The atom may jump from this level to a lower energy level emitting a quantum of radiation. In its final state action variables may have values

$$J_i = J''_i.$$

The differences in the values of the action variables are

$$J'_i - J''_i = \tau_i h$$

where the  $\tau_i$  are integers by quantization. The probability of such a radiation switch is postulated to be related to the square of the absolute value of  $C_\tau(J)$ . Whether one should use  $J'$  or  $J''$  or some average of the two in  $C_\tau(J)$  or whether  $|C_\tau(J)|^2$  should itself be averaged in some way is left undetermined by the correspondence principle. Aside from the question of averaging a definite relation is required however. For in the classical theory, spontaneous radiation occurs independently for the frequencies of the various overtones  $(\tau\nu)$  provided we confine our attention to times long in comparison with the reciprocals of these frequencies. The rate at which energy is radiated by the overtone  $(\tau\nu)$  is obtained by considering the radiation from a polarization

$$(1/2)(C_\tau e^{2\pi i(\tau\nu)t} + C_\tau^* e^{-2\pi i(\tau\nu)t})$$

in accordance with (39). This rate is by classical electrodynamics

$$(16\pi^4(\tau\nu)^4/3c^3) |C_\tau|^2.$$

On the quantum theory  $(\tau\nu)$  is replaced by a possible emission frequency  $\nu_{J'J''}$  and  $C_\tau(J)$  is replaced by a proper average determined by the terminal values  $J'$ ,  $J''$ . This we call

$$2e\mathbf{r}_{J'J''}$$

With these substitutions the rate of radiation of energy may be equated to the Einstein spontaneous emission coefficient  $A_{J'J''}$  multiplied by  $h\nu_{J'J''}$  where  $\nu_{J'J''}$  is the emission frequency corresponding to the radiation switch from the level  $J'$  to the level  $J''$ . The resulting relation is

$$A_{J',J''} = \frac{64\pi^4 e^2 \nu^3 J',J''}{3hc^3} |\mathbf{r}_{J',J''}|^2 \quad (\nu_{J',J''} > 0) \quad (47)$$

where

$$2e\mathbf{r}_{J',J''} = \overline{C_\tau(J)}^{(J',J'')}$$

is the connection between the quantum quantity  $\mathbf{r}_{J',J''}$  and the classical  $C_\tau(J)$ . The symbol  $\overline{\quad}^{(J',J'')}$  indicates that the averaging is to be performed in some way between  $J'$  and  $J''$ .

According to Einstein's theory of emission and absorption probabilities there is a universal relation between the  $A_{J',J''}$  and the absorption probability. It is natural to suspect that the quantum generalization of the numerators in (45') are also related to  $A_{J',J''}$  by a universal formula particularly since strong absorption lines give strong anomalous dispersion. The simplest hypothesis to make is that (47') gives the correct averaging also for a generalization of (45'). Eq. (47) may on this view be regarded as a definition of  $|\mathbf{r}_{J',J''}|^2$  in terms of the experimentally obtainable  $A_{J',J''}$ .

Having thus decided on the meaning of the proper generalizations of the numerators and denominators in (45') we use Bohr's rule (46). We suppose the atom to be in a state  $J'$  i.e., in a state in which the variables  $J$  have these values  $J'$ . We pick out two states  $J''$ ,  $J'''$  which correspond to  $J' - J''' = J'' - J' = \tau h$ . This is possible only if all the quantum numbers are high. We have already supposed that Eq. (45') is written so as to have  $(\tau\nu) > 0$ . Since asymptotically

$$(\tau\nu) = (1/h)(H(J') - H(J''')) = (1/h)(H(J'') - H(J'))$$

it follows that

$$H(J'') > H(J') > H(J''').$$

We adopt the convention

$$\nu_{J'',J'} = (1/h)(H(J'') - H(J'))$$

so that  $\nu_{J'',J'} > 0$  if the transition  $J' \rightarrow J''$  is an absorption of radiation. We thus obtain

$$\begin{aligned} \mathcal{M}_1'(t)_{qu} = Re \left\{ \sum_{H(J'') > H(J')} \frac{e^2}{h} \left[ \frac{\mathbf{r}_{J',J''}(\mathcal{E}\mathbf{r}_{J',J''}^*)}{\nu_{J'',J'} - \nu} + \frac{\mathbf{r}_{J',J''}^*(\mathcal{E}\mathbf{r}_{J',J''})}{\nu_{J'',J'} + \nu} \right] e^{2\pi i \nu t} \right. \\ \left. - \sum_{H(J''') < H(J')} \frac{e^2}{h} \left[ \frac{\mathbf{r}_{J',J'''}(\mathcal{E}\mathbf{r}_{J',J'''}^*)}{\nu_{J',J'''} - \nu} + \frac{\mathbf{r}_{J',J'''}^*(\mathcal{E}\mathbf{r}_{J',J'''})}{\nu_{J',J'''} + \nu} \right] e^{2\pi i \nu t} \right\}. \quad (48) \end{aligned}$$

The first summation refers to possible absorptions and the second to emissions. If we suppose that  $\mathbf{r}_{J',J''}^* = \mathbf{r}_{J',J''}$  (suggested by  $C_\tau^* = C_{-\tau}$ ) the distinction between absorptions and emissions becomes unnecessary and

$$\mathcal{M}_1'(t)_{qu} = Re \left\{ \sum \frac{e^2}{h} \left[ \frac{\mathbf{r}_{J',J''}(\mathcal{E}\mathbf{r}_{J',J''})}{\nu_{J'',J'} - \nu} + \frac{(\mathcal{E}\mathbf{r}_{J',J''})\mathbf{r}_{J',J''}}{\nu_{J'',J'} + \nu} \right] e^{2\pi i \nu t} \right\}. \quad (48')$$



#### §4. The $f$ sum rule

Let there be now a number of atoms oriented in a random way with respect to the incident electric intensity as is the case, for example, in a gas. The average of the component of  $\mathcal{M}_1'(t)_{qu}$  in the direction of  $\mathcal{E}$  is

$$\begin{aligned}\overline{\mathcal{M}_1'(t)_{qu}} &= Re \left\{ \sum_{J''} \frac{2e^2 \mathcal{E} \nu_{J''J'} |\mathbf{r}_{J''J'}|^2}{3h(\nu_{J''J'}^2 - \nu^2)} e^{2\pi i \nu t} \right\} \\ &= \frac{2e^2}{3h} \mathcal{E}(t) \sum_{J''} \frac{\nu_{J''J'} |\mathbf{r}_{J''J'}|^2}{\nu_{J''J'}^2 - \nu^2}.\end{aligned}\quad (48'')$$

This quantity divided by  $\mathcal{E}(t)$  multiplied by the number of atoms per unit volume  $N$  and by  $4\pi$  should give  $n^2 - 1$  where  $n$  is the refractive index, for low gas densities. Defining the effective number of isotropic oscillators per atom by

$$n^2 - 1 = \frac{e^2}{\pi m} \sum_{J''} \frac{N f_{J''J'}}{\nu_{J''J'}^2 - \nu^2} \quad (49)^{12}$$

and comparing with (48'') we obtain

$$f_{J''J'} = (8\pi^2 m / 3h) \nu_{J''J'} |\mathbf{r}_{J''J'}|^2. \quad (50)$$

The Ladenburg  $f$  number thus defined is seen to be positive if  $\nu_{J''J'}$  is positive i.e., if the transition  $J' \rightarrow J''$  corresponds to absorption. Otherwise it is negative.

Eq. (47) determines the connection of the Ladenburg  $f$  numbers with the spontaneous emission probabilities, viz.,

$$\begin{aligned}f_{J''J'} &= (A_{J''J'}/3) \tau_{cl}(\nu_{J''J'}) \\ \left( \tau_{cl}(\nu) &= \frac{3mc^3}{8\pi^2 e^2 \nu^2} \right)\end{aligned}\quad (51)$$

where  $\tau_{cl}(\nu)$  is the time during which the energy of a classical oscillator of mass  $m$ , charge  $e$ , frequency  $\nu$  falls to  $(1/2.7182 \dots)$  of its initial value. The proper generalization of formula (51) to the case of degenerate states has already been given as formula (31) of the preceding report.

The sum rule of Kuhn, Reiche-Thomas<sup>12a</sup> is seen once more from (49) to be equivalent to the requirement of having the refractive index of an atom with one electron become the same as that of a free electron at very high frequencies. This requirement makes it necessary to suppose that

$$\sum_{J''} f_{J''J'} = 1$$

for such an atom. The expression (50) for  $f_{J''J'}$ , when substituted into the sum rule, gives a restriction on the numbers

$$|\mathbf{r}_{J''J'}|^2 = |x_{J''J'}|^2 + |y_{J''J'}|^2 + |z_{J''J'}|^2.$$

<sup>12</sup> Cf. Eq. (37) of preceding review.

<sup>12a</sup> W. Thomas, Naturwiss. **13**, 627 (1925); F. Reiche and W. Thomas, Zeits. f. Physik **34**, 510 (1925); W. Kuhn, Zeits. f. Physik **33**, 408 (1925).

If it is supposed that

$$\sum_{J''} \nu_{J''J'} |x_{J''J'}|^2 = \sum_{J''} \nu_{J''J'} |y_{J''J'}|^2 = \sum_{J''} \nu_{J''J'} |z_{J''J'}|^2$$

this restriction means

$$m \sum [(2\pi i \nu_{J''J'} x_{J''J'}) x_{J''J'} - x_{J''J'} (2\pi i \nu_{J''J'} x_{J''J'})] = h/2\pi i \quad (52)$$

which is essentially the set of equations for the diagonal elements of the quantum condition

$$[p_x, x] = p_x x - x p_x = h/2\pi i.$$

### §5. Incoherent scattering

We return to a consideration of formula (44'''). In addition to the synchronous terms used in (45) there are the nonsynchronous terms of frequency

$$(\tau^0 \nu) + \nu = \tau_1^0 \nu_1 + \tau_2^0 \nu_2 + \cdots + \tau_f^0 \nu_f + \nu$$

which give rise to incoherent scattering. They also have a quantum theoretical counterpart. The classical frequency  $(\tau^0 \nu)$  is interpretable in the quantum theory as a transition frequency between two states the quantum numbers of which differ by amounts  $\tau_1^0, \tau_2^0, \cdots \tau_f^0$ . The natural extension of classical incoherent scattering is thus scattering of radiation with frequencies

$$\nu_{J'J'''} + \nu$$

where  $\nu_{J'J'''}$  is an emission or an absorption frequency. To fix ideas we suppose the atom to be in the state  $J'$ . The state  $J'''$  is then any other energy level. We again use the convention

$$h\nu_{J'J'''} = H(J') - H(J''').$$

Thus if  $H(J') > H(J''')$  the frequency  $\nu_{J'J'''}$  is positive and is a possible emission frequency from the state  $J'$ . If  $H(J') < H(J''')$  the number  $\nu_{J'J'''}$  is negative and  $\nu_{J''J'}$  is a possible absorption frequency. We note that (44''') includes a summation over  $\tau^0$  which extends over all positive and negative  $\tau^0$ . The classical incoherent radiation deals therefore with frequencies  $\nu \pm (\tau^0 \nu)$  which may be numerically greater or smaller than  $\nu$ . If the number  $\nu - |(\tau^0 \nu)| < 0$  we must interpret the corresponding frequency as  $|(\tau^0 \nu)| - \nu$ . As long, however, as  $\nu \pm |(\tau^0 \nu)|$  the quantum theory extension is supposed to be the frequency  $\nu + \nu_{J'J'''}$ . This includes both the frequencies greater and smaller than  $\nu$ . The translation of the  $C_r, C_{r'}$  into the quantum theory can be made similarly to that for the coherent scattering formula (48'). The result of Kramers and Heisenberg is:

$$\mathcal{M}_1''(t)_{J'} = Re \left\{ \sum_{J''} \sum_{\substack{J''' \\ (\nu > \nu_{J''J'''})}} \frac{e^2 [\mathbf{r}_{J''J'''}(\mathcal{E} \mathbf{r}_{J''J'''})]}{h [\nu_{J''J'''} + \nu]} - \frac{\mathbf{r}_{J''J'''}(\mathcal{E} \mathbf{r}_{J'J''})}{\nu_{J'J''} + \nu} \right] e^{2\pi i [(\nu_{J'J'''} + \nu)t + \delta_{J'J'''}]} \right\}$$

$$\begin{aligned}
& + \sum_{J''} \sum_{\substack{J''' \\ (\nu_{J'J'''} > \nu)}} \frac{e^2 \left[ \mathbf{r}_{J''J'''} (\mathcal{E} \mathbf{r}_{J''J'}) \right]}{h \left[ \nu_{J''J'} + \nu \right]} \\
& - \frac{\mathbf{r}_{J''J'} (\mathcal{E} \mathbf{r}_{J''J'''})}{\nu_{J''J'''} + \nu} \left[ e^{2\pi i [(\nu_{J''J'} + \nu)t + \delta_{J''J'}]} \right] \Bigg\}. \quad (53)
\end{aligned}$$

Here the phases  $\delta_{J'J''}$  are introduced to account for the undetermined additive constants in the angle variables  $w$ .

The first term in this formula takes account of processes in which an atom may be pictured to have been initially in a state  $J'$  a light quantum  $h\nu$  having then been absorbed to raise the atom temporarily to a state  $J''$  and a light quantum  $h(\nu + \nu_{J'J''})$  having then been reemitted to leave the atom in the final state  $J'''$ . The conservation of energy holds in the form

$$\begin{aligned}
h\nu + H(J') &= h\nu' + H(J''') \quad (\nu > \nu_{J'J''}) \\
\nu' &= \nu + \nu_{J'J''}. \quad (54)
\end{aligned}$$

The restriction  $\nu > \nu_{J'J''}$  insures a positive value of the scattered frequency  $\nu'$ . The frequency  $\nu_{J'J''}$  may be positive or negative giving respectively the antistokes and the normal types of frequency changes. If there is a state  $J''$  with  $H(J'') > H(J')$  such that the absorption frequency  $\nu_{J'J''}$  is approximately  $\nu$  the denominator of the second term of the first line in (53) becomes small. The incoherent scattering of frequency  $\nu + \nu_{J'J''}$  becomes large. In this case incoherent scattering merges with the phenomenon of fluorescence which may take place if  $\nu = \nu_{J'J''}$ . For in this case the quantum can be actually absorbed and the atom then raised to the state  $J''$ . From the state  $J''$  it then falls to  $J'''$  emitting light of frequency  $\nu_{J''J'''}$  which is the same as  $\nu + \nu_{J'J''}$  for  $\nu = \nu_{J'J''}$ . The first term of the first line of (53) can also become small now for  $\nu \cong \nu_{J''J'''}$ , i.e., for the case when the incident frequency is approximately correct for raising the atom from a state  $J''$ , by absorption, to the final state  $J'''$ . It is also to be noted that the amplitudes  $r_{J'J''}$ ,  $r_{J''J'''}$  are connected with the Einstein emission and absorption probabilities through (47) and that therefore only those terms in formula (53) can occur for which spontaneous radiative transitions are possible between  $J'$ ,  $J''$  and  $J''$ ,  $J'''$ . This fact, together with the behavior of the resonance denominators just discussed indicates that the picture of the absorption of  $h\nu$  via the intermediate state  $J''$  has a certain degree of reality even though in the temporary condition  $J''$  the conservation of energy is violated. Considerable light has been thrown since on the reason for this temporary lack of conservation of energy by the more modern quantum theory.

The second term of (53) is subject to the restriction  $\nu_{J'J'''} > \nu$ . The final state  $J'''$  has here a lower energy than the initial. The interpretation of  $e^{2\pi i(\nu_{J''J'} + \nu)t}$  is more easily seen if this exponential is written as  $e^{-2\pi i(-\nu + \nu_{J'J''})t}$  which is interpreted as the emission of a quantum of frequency

$$\nu'' = -\nu + \nu_{J'J''} \quad (\nu_{J'J''} > \nu).$$

The conservation of energy for the initial and final states holds then in the form

$$h\nu + H(J') = h\nu'' + 2h\nu + H(J'''). \quad (55)$$

In the final condition the radiation contains a quantum of frequency  $\nu''$  and an additional quantum of frequency  $\nu$ . The first term of the second line of (53) becomes large if  $\nu \cong \nu_{J'J''}$  i.e., if the incident frequency approaches the condition of inducing emissions of frequency  $\nu$  (Einstein's negative absorptions). In the limiting case  $\nu = \nu_{J'J''}$  we deal with an induced emission in which the atom jumps from  $J'$  to  $J''$  and a quantum  $h\nu$  is emitted. This is followed by an emission of a quantum  $h\nu_{J''J'''} = h(\nu_{J''J'''} - \nu_{J'J''}) = h(\nu_{J''J'''} - \nu)$ . The second term of the second line of (53) becomes large for  $\nu \cong \nu_{J''J'''}$ , i.e., for the case when induced emission is probable from the intermediate to the final state.

For  $J' \equiv J'''$  the second line in (53) disappears and the formula then becomes equivalent to (48'). [As written we must also take the complex conjugate to bring about agreement. This is immaterial in the present consideration.]

The scattering of the type (54) has been hypothesized by Smekal<sup>13</sup> from energetic considerations previously to the theory of Kramers and Heisenberg. Scattering of the type (55) and the quantitative expectations for both kinds are met with first in the work of Kramers and Heisenberg. A number of interesting considerations have been based on the Kramers-Heisenberg formula by Pauli.<sup>14</sup> The development of modern quantum theory makes it advisable to treat the same questions in greater detail in subsequent sections of the report. Nevertheless a reader not familiar with the above reference is strongly recommended its perusal.

For purposes of comparison with more modern treatments we finish the discussion of the above derivation of the Kramers-Heisenberg formula by a recapitulation of the physical picture and results. A conditionally periodic material system is considered under the action of incident radiation. The wave-length of radiation is supposed to be long in comparison with atomic dimensions. In the calculation the damping of the electronic motion due to the scattering is neglected. Under these conditions it is found that for a weak incident monochromatic radiation there is an induced electric moment given by (44'''). By a judicious application of the correspondence principle the formula has been translated into quantum theory by means of (48'), (53). When an atom is in a state  $J'$  it is supposed to scatter radiation as though it has an electric moment given by (53). The case (48') may be included in (53) by making  $J' \equiv J'''$ .

We now note that in (44'') the variables  $\bar{J}, \bar{w}$  are only approximately the variables  $J, w$  of the atom. In (44''') and from there on we have dropped the distinction between the  $\bar{J}$  and the  $J$ . It will be well to remember, therefore,

<sup>13</sup> A. Smekal, *Naturwiss.* **11**, 873 (1923).

<sup>14</sup> W. Pauli, Jr., *Handbuch der Physik*, **XXIII**, pp. 86-108.

that when in (53) we speak of an atom as being in a certain state  $J'$  we really should, to be quite accurate, speak of the variables  $\bar{J}$  having these values  $J'$ . This means that the atom has been subjected to quantization, under the influence of the incident radiation.

It has been shown by Van Vleck<sup>15</sup> and also by Niessen<sup>16</sup> that Bohr's postulate (47) together with Einstein's relation between the emission and absorption probabilities give, for large quantum numbers, an average absorption of energy which is in agreement with that required by the classical theory. By large quantum numbers are meant such quantum numbers that  $(\tau\partial/\partial J)A$  and  $(1/h)\Delta A$  differ by amounts negligible in comparison with  $(\tau\partial/\partial J)A$ . For such large quantum numbers there is, therefore, complete agreement between the classical and quantum theories both for the dispersion of light and for the absorption and emission of energies.

### III. INTRODUCTORY REMARKS ABOUT QUANTUM THEORY

#### §1. Purpose of remarks

In most of the applications of quantum theory made in this review it is convenient to deal with Schroedinger's equation. In some cases, however, the matrix and operator methods are more direct. The equivalence of the wave and matrix methods is well known and many excellent discussions of it are found in the literature. Nevertheless we have to discuss some features of the general quantum theory which are of special importance for the applications made below. Among the points which we wish to emphasize are: (1) the difference between operators representing dynamical variables at different times and the operators occurring in Schroedinger's equation; (2) the fact that the equations used in the method of variations of constants are simply the components of a Schroedinger equation in a new reference system; (3) that the equations between operators which follow from a Hamiltonian function are mathematical consequences of the Schroedinger equation; (4) that all methods available at present which satisfy the requirements of the statistical interpretation reduce to a Schroedinger equation with one time.

The last point (4) is of interest in view of the fact that there are still unsolved difficulties in the treatment of electromagnetic radiation in quantum theory. One hears time and again the opinion that the problem of the interaction of two particles should be attacked by means of wave equations involving two times: a separate time for each particle. It may well be that some of the failures of present quantum theory are attributable to the use of one time. However, it should be realized that the only general form of theory available at present will require revision in order to make this change.

We will refer to this section in the succeeding ones whenever the radiation theory requires the use of a general relation of quantum theory.

#### §2. The Schroedinger equation

The physical question which are answered by applications of quantum theory are always about *probabilities* of observing one or another value of a

<sup>15</sup> J. H. Van Vleck, Phys. Rev. **24**, 330 (1924); Journ. Opt. Soc. **9**, 27 (1924).

<sup>16</sup> K. F. Niessen, Ann. d. Physik **72**, 743 (1924).

quantity in the course of time. In particular the probabilities of observing one or another value of a positional coordinate of a particle is given by the square of the absolute value of the Schroedinger wave function. This wave equation has been supposed by Schroedinger to obey an equation called the Schroedinger equation. The choice of its form is sometimes difficult and is of course crucial for the success or failure of a theory. No general and absolute rule can be given for setting up such equations. We suppose for the present that we know its form. This means that it is given to us as

$$(h/2\pi i)(\partial\psi/\partial t) + H(p, q, \dots; t)\psi = 0. \quad (60)$$

Here the Hamiltonian function  $H$  contains the dynamical variables of the system indicated by  $p, q, \dots$ ; in addition it may also contain the time  $t$ . Thus for instance in the case of the nonrelativistic hydrogen atom

$$H(p, q, \dots; t) = (p_x^2 + p_y^2 + p_z^2)/(2m) - e^2/(x^2 + y^2 + z^2)^{1/2}$$

where  $x, y, z$  are the Cartesian coordinates of the electron;  $p_x, p_y, p_z$  are the corresponding momenta;  $m, e$  are respectively the mass and the charge of the electron. The reader will recall that Schroedinger's original method of treating the problem consists in using (60) with the special expression for  $H$  just written. For  $p_x, p_y, p_z$  Schroedinger substituted  $(h/2\pi i) (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ . We take this opportunity to emphasize the fact that in Schroedinger's equation the dynamical variables  $p, q, \dots$  are replaced by *definite operators which do not involve the time*.

To be quite clear about what is meant by (60) we must say that:

(1) The  $\psi$  function is an ordinary (not a matrix) function of a complete set of simultaneously measurable quantities. In the above example the quantities are  $x, y, z$ .

(2) The dynamical variables  $p, q, \dots$  are to be replaced in  $H$  by definite operators, which operate on the variable in  $\psi$ . The *form of the operator replacing a given physical quantity in the Hamiltonian function is determined by the physical nature of the quantity* and not by the problem which is treated by the Schroedinger Eq. (60). Thus, for example, the momentum  $p_x$  is replaced by  $(h/2\pi i) (\partial/\partial x)$  whether one is treating the hydrogen atom or any other problem.

Let the variables entering  $\psi$  be denoted collectively by  $x'$ . The operators replacing the dynamical variables in  $H$  of the Schroedinger equation are then definite matrices

$$p(x', x''), q(x', x'') \dots \text{etc.}$$

By definition

$$H(x', x''; t)$$

the matrix obtained by substituting the matrix quantities  $p, q$  into the form  $H(p, q, \dots; t)$ . The matrices  $p(x', x'')$  etc. may be singular or nonsingular. Thus in the above example the dynamical variable  $y$  is replaced by the

singular matrix  $y'\delta(x'-x'')\delta(y'-y'')\delta(z'-z'')$  where  $\delta$  is Dirac's  $\delta$  function. The matrix which replaces  $p_y$  is similarly  $\hbar \cdot \delta'(y'-y'')\delta(x'-x'')\delta(z'-z'')/2\pi i$ .

It is important to note that although the dynamical variables  $x, y, z$ ;  $p_x, p_y, p_z$  are function of the time this is not true of the operators which replace them in the Schroedinger equation.

(3) With the above meaning of  $H(x', x''; t)$  the Schroedinger equation (60) means in more detail

$$\frac{\hbar}{2\pi i} \frac{\partial \psi(x'; t)}{\partial t} + \int H(x', x''; t) \psi(x'') dx'' = 0. \quad (60')$$

The above formula applies as written only to the case of continuous ranges of values of  $x'$ . If the  $x'$  are discrete the integration sign must be replaced by a summation.

The justification of the physical plausibility of (60') for any possible set of simultaneously observable physical quantities  $x'$  has been given by Dirac and Jordan. We do not enter into this question fully in this report. We suppose that (60') is the proper equation to use. We use it as a starting point because its practical applications (hydrogen atom, linear oscillator, etc.) have made it popular among a large group of physicists.

### §3. The statistical interpretation

It will be recalled that the  $\psi$  function in the case of the hydrogen atom has a simple interpretation. The square of its absolute value  $|\psi(x, y, z; t)|^2$  represents the probability of finding the electron in the volume element  $dx dy dz$  at the time  $t$ . This interpretation is generalized to mean that  $|\psi(x'; t)|^2$  is the probability of finding any system in the volume element  $dx'$  at the time  $t$ . The most immediate physical context of (60) is the prediction of  $|\psi(x'; t)|^2$  at the time  $t$  for a condition of the physical system described by  $\psi(x'; t)_{t=0}$  at the time  $t=0$ .

We now note that in the example of the hydrogen atom this interpretation definitely implies that at the time  $t$  the average value of  $x$  or  $y$  or  $z$  or any function  $f(x, y, z)$  is

$$\int |\psi(x, y, z; t)|^2 f(x, y, z) dx dy dz.$$

For the momenta  $p_x, p_y, p_z$  there can be no quite so simple interpretation of the average value of let us say  $p_x$  or  $p_x^2$  at the time  $t$ . Nevertheless it is natural to use for any quantity  $Y$  the matrix for which is  $Y(x', x''; t)$  the following extension

$$\bar{Y}(t) = \iint \psi^*(x'', t) Y(x'', x'; t) \psi(x', t) dx' dx''. \quad (61)$$

The matrix  $Y(x'', x'; t)$  is supposed to be computed in exactly the same way from the function  $Y(p, q, \dots; t)$  as has been done for  $H(x'', x'; t)$ .

The plausibility of this postulate may be explained using the measurement of the momentum as an example. According to the direct experimental evidence of Davisson and Germer, G. P. Thomson and others a moving particle should be represented by a wave with a definite wave-length.

$$h/|p|$$

where  $|p|$  is the absolute value of the momentum. The wave function describing such a particle is taken to be within a constant factor

$$e^{2\pi i(\mathbf{p}'\mathbf{r})/h}[(\mathbf{p}'\mathbf{r}) = p'_x x + p'_y y + p'_z z]$$

where  $p'_x, p'_y, p'_z$  are ordinary numbers equal to the known components of momentum. The square of the absolute value of this wave function is the same at all points in space. The particle has then an equal chance of being found anywhere. To make its location somewhat more definite we allow ourselves to superpose a number of such waves and to form

$$\psi(x', y', z') = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h^{-3/2} C_{p'_x, p'_y, p'_z} e^{2\pi i(\mathbf{p}'\mathbf{r})/h} dp'_x dp'_y dp'_z. \quad (61')$$

In optics such a wave is interpreted as a mixture of monochromatic waves. Experimentally this means that the diffraction patterns due to such waves are diffuse on account of the superposition of the patterns due to the different wave-lengths. Similarly in wave-mechanics the diffraction pattern due to the above  $\psi$  is supposed to be diffuse. The probability of finding a particle in a diffraction maximum corresponding to the range  $dp'_x dp'_y dp'_z$ , i.e., the probability of finding a momentum in that range is supposed to be proportional to  $|C_{p'}|^2 dp'_x dp'_y dp'_z$ . We have<sup>17</sup>

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\psi|^2 dx' dy' dz' = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |C_{p'}|^2 dp'_x dp'_y dp'_z = 1$$

so that  $|C_{p'}|^2 dp'_x dp'_y dp'_z$  is not only the relative but also the absolute probability. Inverting the relation between  $\psi$  and  $C$  we have<sup>18</sup>

$$C_{p'} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h^{-3/2} \psi(x', y', z') e^{-2\pi i(\mathbf{p}'\mathbf{r})/h} dx' dy' dz'.$$

The mean of  $p_x$  is now

$$\bar{p}_x = \int |C_{p'}|^2 p'_x dp'_x dp'_y dp'_z$$

which on substituting  $C_{p'}$  in terms of  $\psi$  becomes

$$\bar{p}_x = \iiint p'_x \{ h^{-3} \psi^*(x'', y'', z'') e^{2\pi i(\mathbf{r}'' - \mathbf{r}')\mathbf{p}'/h} \psi(x', y', z') dx' dy' dz' dx'' dy'' dz'' \} dp'_x dp'_y dp'_z.$$

<sup>17</sup> Courant, Hilbert 2nd edition, p. 84.

<sup>18</sup> Idem, p. 69.



We have

$$\begin{aligned} & \iint h^{-3} p_x' e^{2\pi i(\mathbf{r}'' - \mathbf{r}') \cdot \mathbf{p}'} / \hbar \psi(x', y', z') dx' dy' dz' dp_x' dp_y' dp_z' \\ &= \frac{\hbar}{2\pi i} \frac{\partial}{\partial x''} \iint h^{-3} e^{2\pi i(\mathbf{r}'' - \mathbf{r}') \cdot \mathbf{p}'} / \hbar \psi(x', y', z') dx' dy' dz' dp_x' dp_y' dp_z' = \frac{\hbar}{2\pi i} \frac{\partial \psi(x'')}{\partial x''} \end{aligned}$$

so that

$$\bar{p}_x = \int \psi^*(x'') \frac{\hbar}{2\pi i} \frac{\partial \psi(x'')}{\partial x''} dx'' dy'' dz''$$

provided it is permissible to differentiate outside the integral sign.

Similarly by ordinary calculation

$$\overline{p_x^n} = \int (p_x')^n |C_{p'}|^2 dp_x' dp_y' dp_z' = \int \psi^*(x'') \left( \frac{\hbar}{2\pi i} \frac{\partial}{\partial x''} \right)^n \psi(x'') dx'' dy'' dz''.$$

These quantum means are seen to be special cases of (61) with the operator  $p_x^n$  being represented by the matrix  $p_x^n(x', y', z'; x'', y'', z'')$  obtained by multiplying the matrices for  $p_x$   $n$  times.

The matrix operators representing physical quantities are supposed to be Hermitian. It follows that

$$-\frac{\hbar}{2\pi i} \frac{\partial \psi^*(x'; t)}{\partial t} + \int \psi^*(x''; t) H(x'', x'; t) dx'' = 0. \quad (60'')$$

The  $\psi$  function is subjected to the normalization condition

$$\int \psi^*(x'; t) \psi(x'; t) dx' = 1$$

which is maintained in the course of time by (60'), (60''). The rate of change of the mean value of  $Y$  is obtained by differentiating (61) and substituting  $\partial \psi / \partial t$ ,  $\partial \psi^* / \partial t$  from (60'), (60'')

$$(d/dt) \bar{Y}(t) = \overline{(\partial Y / \partial t)} + (2\pi i / \hbar) [\bar{H}, \bar{Y}].$$

Here the first term is the mean of  $\partial Y(p, q, \dots; t) / \partial t$  taken using  $\psi(x; t)$  by means of (61) and the second term is similarly obtained from  $(2\pi i / \hbar) [H, Y]$  where

$$[H, Y] = HY - YH$$

and the expression  $[H, Y]$  is treated as a new  $Y$  in (61). If now the formula (61) is to have a useful meaning we must secure the equation

$$(d/dt) \bar{Y} = \overline{(dY/dt)} \quad (62)$$

for otherwise we would have the physically unsatisfactory condition of having the rate of change of the mean differ from the mean of the rate of change. We

therefore *define* the physical quantity  $dY/dt$  so as to satisfy (62) under all conditions. By so doing we will have a consistent physical nomenclature in which the rate of change of a quantity  $Y$  is such that its mean at any time is  $(d/dt)\bar{Y}$ . Thus the matrix operator representing  $dY/dt$  is taken to be

$$(dY/dt)(x', x''; t) = \partial Y(x', x''; t)/\partial t + (2\pi i/h)[H, Y](x', x''; t). \quad (63)$$

In particular

$$(dH/dt)(x', x''; t) = \partial H(x', x''; t)/\partial t$$

which disappears if  $t$  is not contained explicitly in the Hamiltonian function.

The Schroedinger equation and the requirement (62) lead to (63). Quite similarly by applying (63) repeatedly the proper definitions of the operators for  $d^2Y/dt^2, d^3Y/dt^3, \dots$  follow. Applying (63) to  $Y = Y_1 + Y_2$  and  $Y = Y_1 Y_2$  it is seen that the operators  $dY/dt$  are for these cases

$$\frac{dY}{dt} = \frac{dY_1}{dt} + \frac{dY_2}{dt} \quad \text{and} \quad \frac{dY}{dt} = \frac{dY_1}{dt} Y_2 + Y_1 \frac{dY_2}{dt}.$$

Thus for functions representable by sums, differences, and products the ordinary differentiation rules with respect to the time will hold.

#### §4. Generalization of Schroedinger's equation to other reference systems

At a given time, say at  $t=0$ , we may assign any set of values to  $\psi(x'; t)$  for all values of the variables  $x'$  provided the normalization condition is satisfied. In particular  $\psi$  may be made equal to zero except in the neighborhood of an arbitrary point  $x'_0$  in the  $x'$  space. If such is the case then at  $t=0$  the values of the dynamical variables  $x'$  are  $x'_0$  with certainty because then by hypothesis we can measure  $\bar{x}, \bar{x}^2, \bar{x}^3$  etc., and we will find by (61) that these measured values are  $x_0, x_0^2, x_0^3, \dots$ . This is not true in general of other dynamical variables as is seen again by (61). However, it may also be possible to assign at  $t=0$  such values to  $\psi$  that some other set of dynamical variables is similarly known with absolute certainty from statistical measurements on the system. Thus, for example, in the one electron problem the states described by (61') with  $C_p$  distinct from zero only in a narrow range, correspond by (61) to the momenta  $p_x, p_y, p_z$  having almost definite numerical values  $\bar{p}_x'', \bar{p}_y'', \bar{p}_z''$  because with a proper treatment of convergence (61) yields  $\bar{p}_x^n = \bar{p}_x''^n$  for arbitrary  $n$ .

Let us suppose then that we have any complete set of normal orthogonal functions each of which corresponds to a definite value of some dynamical variables. Let these functions be  $u_n(x')$  and let us suppose that the time is not involved in these functions. We may then quite generally expand any

$$\psi(x'; t) = \sum_n c_n(t) u_n(x'). \quad (64)$$

The indices  $n$  stand collectively for the values of the quantities which the new dynamical variables have if  $\psi = u_n$ . In the special case just discussed [Eq. (61')] the variables  $n$  are the momenta  $p$ .

Substituting (64) into (60') we obtain

$$\frac{\hbar}{2\pi i} \frac{dc_n(t)}{dt} + \sum_m H_{nm}(t) c_m(t) = 0 \quad (65)$$

where

$$H_{nm}(t) = \int \int u_n^*(x') H(x', x''; t) u_m(x'') dx' dx''. \quad (66)$$

We compare (65) with (60') and we see that their form is essentially the same. Whatever difference there is, is purely notational for the context of (65) does not depend on whether we write  $c_n(t)$  or  $c(n; t)$ ; the difference between the summation sign in (65) and the integral in (60') is also only notational because it was agreed in (60') that the integral should be replaced by a summation if the values of the  $x'$  should be discrete. It is important to note, however, that the matrix operators representing the Hamiltonian function are related by the formula (66).

The inverse to (64) is by well-known properties<sup>19</sup> of complete sets of normal orthogonal functions

$$c_n(t) = \int \psi(x'; t) u_n^*(x') dx' \quad (64')$$

and the inverse to (66) is

$$H(x', x''; t) = \sum_{n,m} u_n(x') H_{nm}(t) u_m^*(x''). \quad (66')$$

Write now

$$c_{n'}(t) = \Psi(n'; t), \quad u_{n'}(x') = S(x', n') \quad (64'')$$

we may then also write

$$u_{n'}^*(x') = S^{-1}(n', x')$$

because by the orthogonality relations

$$(SS^{-1})_{x'x''} = \delta(x' - x''), \quad (S^{-1}S)_{n'm'} = \delta_{n'm'}$$

where the products  $SS^{-1}$  and  $S^{-1}S$  are meant to be matrix products. By the definition of  $S$  and  $S^{-1}$  it follows that

$$\tilde{S}^* = S^{-1}$$

<sup>19</sup> The properties used may be summarized in the formulas

$$\begin{aligned} \sum_n u_n^*(x') u_n(x'') &= \delta(x' - x'') \text{ (Dirac's } \delta \text{ function)} \\ \int u_n^*(x') u_m(x') dx' &= \delta_{nm} \text{ where } \delta_{nm} = 1 \text{ } n=m \\ &\quad \delta_{nm} = 0 \text{ } n \neq m. \end{aligned}$$

Applying the first of these to the case of the momentum [Eq. (61')] it will be seen that the  $\delta$  function is in this case meaningless. Nevertheless correct results are obtained by writing (incorrectly)  $\int \hbar^{-3} e^{2\pi i(\mathbf{r}' - \mathbf{r}'') \cdot \mathbf{p}'} / \hbar d\mathbf{p}_x' d\mathbf{p}_y' d\mathbf{p}_z' = \delta(\mathbf{r}' - \mathbf{r}'')$  in the computation of  $\int \{ \hbar^{-3} e^{2\pi i(\mathbf{r}'' - \mathbf{r}') \cdot \mathbf{p}'} / \hbar \psi(\mathbf{r}') dx' dy' dz' \} d\mathbf{p}_x' d\mathbf{p}_y' d\mathbf{p}_z' = \psi(\mathbf{r}'')$ .

where  $\sim$  indicates transposition of indices  $n', x'$ . Formulas (66), (66') are now

$$\begin{aligned} H_{n'n''}(t) &= \iint S^{-1}(n', x') H(x', x''; t) S(x'', n'') dx' dx'' \\ &= (S^{-1} H^{(x)} S)_{n'n''} \\ H(x', x''; t) &= \sum_{n', n''} S(x', n') H(n', n''; t) S^{-1}(n'', x'') \\ &= (S H^{(n)} S^{-1})_{x' x''}. \end{aligned} \quad (66'')$$

Here  $H^{(x)}$  means that we use the matrix  $H(x', x''; t)$  and  $H^{(n)}$  that we use  $H(n', n''; t)$ . Similarly (60') and (65) may be rewritten as

$$\begin{aligned} (h/2\pi i)(\partial \Psi(n'; t)/\partial t) + (H^{(n)} \Psi)_{n'} &= 0 \\ (h/2\pi i)(\partial \psi(x'; t)/\partial t) + (H^{(x)} \psi)_{x'} &= 0 \end{aligned} \quad (67)$$

where

$$\Psi(n'; t) = (S^{-1} \psi)_{n'}; \psi(x'; t) = (S \Psi)_{x'} \quad (68)$$

by (64). Formulas (66''), (67), (68) are simply transcriptions of (66), (66'), (60'), (66'), (64), (64') in the notation defined by (66''). They contain no new physical information. They are somewhat more elegant because the orthogonality relations of complete sets of normal orthogonal functions are accentuated by the notation  $S, S^{-1}$ . The equations may be remembered easily by stating that the form of the Schrodinger equation is unchanged if we transform  $\psi \rightarrow$  (into)  $S^{-1}\psi$  and simultaneously the matrix operator  $H$  into  $S^{-1}HS$ .

The transformation  $\Psi(n'; t) = (S^{-1}\psi)_{n'}$  implies also

$$\Psi^*(n'; t) = \int (S^{-1}(n', x'))^* \psi^*(x'; t) dx' = \int \psi^*(x'; t) S(x', n') dx'$$

or

$$\Psi^*(n'; t) = (\psi^* S)_{n'}. \quad (69)$$

According to (61) we may write for the mean value of any quantity

$$\bar{Y}(t) = (\psi^* Y^{(x)} \psi) = (\Psi^* S^{-1} Y^{(x)} S \Psi) = (\Psi^* Y^{(n)} \Psi)$$

where

$$Y^{(n)} = S^{-1} Y^{(x)} S; \quad (70)$$

is the matrix equation representing the matrix operator  $Y^{(n)}(n', n''; t)$  in terms of the matrix  $Y^{(x)}(x', x''; t)$ . The transformation (70) is seen to be designed so as to leave formula (61) invariant in form whether we work with  $\psi(x', t)$  or  $\Psi(n'; t)$ . Transformations having the form (70) are called *canonical transformations*.

If for any time  $t$  only one  $c_n$  say  $c_{n_0}$  is different from zero the variables  $n$  have certainly the values  $n_0$  at that time. This follows from our definition of

the functions  $u_n(x')$  while we worked with  $\psi(x'; t)$ . In (64'') we have denoted  $c_n(t)$  by  $\Psi(n; t)$ . We see that if  $\Psi(n; t)$  is distinct from zero only for  $n = n_0$  the variables  $n$  have certainly the values  $n_0$ . We may interpret, therefore, the new function  $\Psi(n'; t)$  in the same way as we had interpreted  $\psi(x'; t)$ . Formula (70) shows this conclusion to be general provided all operators are transformed according to it. By (70) and (66'') the form of (63) is invariant under the transformation  $H \rightarrow S^{-1}HS$ ,  $Y \rightarrow S^{-1}YS$  associated with  $\psi \rightarrow S^{-1}\psi$ .

The dynamical variables  $n$  may themselves be represented by matrices in the  $(x', x'')$  reference system. Suppose one of them  $n^{(s)}$  is represented by the matrix operator  $n^{(s)}(x', x'')$ . The condition which makes it sufficient for the functions  $u_n(x')$  to correspond to definite values of the variables  $n$  is then

$$\int n^{(s)}(x', x'') u_{n'}(x'') dx'' = n'^{(s)} u_{n'}(x')$$

or

$$(n^{(s)} u_{n'})_{x'} = n'^{(s)} u_{n'}(x') \quad (71)$$

for all  $x'$ . By (64'') this is

$$(n^{(s)} S)_{x' n'} = n'^{(s)} S(x', n'). \quad (71')$$

This is often called the Schroedinger equation for the variable  $n^{(s)}$  and the transformation function  $S$ . If  $\psi(x'; t) = S(x', n')$  the variables  $n$  have the values  $n'$  at the time at which this equation is satisfied but not necessarily at other times.

### §5. Operators at different times

Return now to (65) and the expansion (64). Start out with the system in the state  $\psi = u_{n'}(x')$  at  $t = 0$ . The Schroedinger equation determines the future state of the system. The values of  $\psi$  for  $t > 0$  we denote by  $\psi_{n'}(x'; t)$  the defining equation being

$$\psi_{n'}(x'; t)_{t=0} = u_{n'}(x'). \quad (72)$$

Define

$${}^tY(n', n''; t) = \iint \psi_{n'}^*(x'; t) Y(x', x''; t) \psi_{n''}(x''; t) dx' dx''. \quad (73)$$

This matrix differs from  $Y(n', n''; t)$  of (70). For  $t = 0$  however they are the same

$${}^tY(n', n''; t)_{t=0} = Y(n', n''; 0). \quad (73')$$

We expand

$$\psi_{n'}(x'; t) = \sum_m c_m^{n'}(t) u_m(x'); c_m^{n'}(0) = \delta_{n'm}$$

and get

$${}^tY(n', n''; t) = \sum_{m', m''} c_{m'}^{*n'}(t) Y(m', m''; t) c_{m'', n''}(t).$$

It is proved without difficulty from (65) that

$$\sum_{n'} c_{m'}^*{}^{*n'}(t) c_{m''}{}^{n'}(t) = \delta_{m'm''}; \quad \sum_{m'} c_{m'}^*{}^{*n'}(t) c_{m'}{}^{n''}(t) = \delta_{n'n''}.$$

This suggests rewriting the above formulas by giving emphasis to these orthogonality relations:

$$\begin{aligned} c_{m'}{}^{n'}(t) &= U_t(m', n'); \quad c_{m'}^*{}^{*n'}(t) = U_t^{-1}(n', m') \\ {}^tY(n', n''; t) &= (U_t^{-1}Y(t)U_t)_{n', n''} \end{aligned} \quad (74)$$

where

$$\begin{aligned} (h/2\pi i)(dU_t(m', n')/dt) + \sum_{m''} H_{m'm''}(t)U_t(m'', n') &= 0 \\ (U_t(m', n'))_{t=0} &= \delta_{m'n'} \end{aligned} \quad (74')$$

determines  $U_t$ . The matrix elements  $H_{m'm''}(t)$  in (74') are the same as  $H_{m'm''}(t)$  of the first line of (66'').

Eqs. (74) apply to all operators  $Y$ . The matrix  $H(n', n''; t)$  can be transformed by a similar equation into  ${}^tH(n', n''; t)$ . The time derivative Eq. (63) becomes in consequence

$${}^t(dY/dt)(n', n''; t) = {}^t(\partial Y/\partial t)(n', n''; t) + (2\pi i/h)[{}^tH(t), {}^tY(t)](n', n''; t). \quad (75)$$

An arbitrary initial state of the system may be described by assigning arbitrary initial values  $c_n(0)$  to the  $c_n(t)$ . Then at a later time

$$\Psi_n(t) = c_n(t) = \sum_m c_m(0)c_n^m(t) = \sum_m U_t(n, m)\Psi(m; 0)$$

or more concisely

$$\Psi(n; t) = (U_t\Psi(0))_n. \quad (76)$$

The mean value of  $Y$  is by (70), (76) and (74)

$$\bar{Y}(t) = (\Psi^*(t)Y(t)\Psi(t)) = (\Psi^*(0){}^tY(t)\Psi(0)). \quad (77)$$

The matrix operators  ${}^tY(t)$  defined by (73) have thus the connection (77) with the supposedly statistically observable quantum theoretical means  $\bar{Y}(t)$ . This connection is often useful because it enables one to determine  $\bar{Y}(t)$  for all  $t$  if the operator  ${}^tY(t)$  is known.

## §6. Equations of motion

The operators  ${}^tY(t)$  have an immediate connection with classical dynamics *via* the correspondence principle. For if there is a quantum equation

$$d{}^tY(t)/dt = f({}^tZ(t), {}^tX(t), \dots)$$

between operators a corresponding classical equation must hold, because the above equation has as a consequence (See also (62))

$$d\bar{Y}(t)/dt = \overline{f(t)}.$$

For those cases in which quantum theory merges with the classical theory  $\bar{Y}(t)$  becomes identical with the classical  $Y(t)$ .

Since by (74) the operator  ${}^tY(t)$  is obtained from  $Y(t)$  by a canonical transformation the equations between different operators  $Y(t)$  are correct also if the  $Y(t)$  are replaced by  ${}^tY(t)$ . In particular the commutation relations are unchanged. We recall that this has as a consequence a formal agreement of the classical Hamiltonian equations

$$\dot{q}_r = \partial H(p, q; t) / \partial p_r, \quad \dot{p}_r = -\partial H(p, q; t) / \partial q_r$$

with corresponding equations between corresponding operators  ${}^tq_r(t)$ ,  ${}^tp_r(t)$  viz.

$$\frac{d{}^tq_r(t)}{dt} = \frac{\partial H({}^tp, {}^tq; t)}{\partial {}^tp_r}, \quad \frac{d{}^tp_r(t)}{dt} = -\frac{\partial H({}^tp, {}^tq; t)}{\partial {}^tq_r}. \quad (78)$$

It must be borne in mind, however, that in the quantum theory these equations are only special applications of (75), and do not contain any more information than (75) or Schroedinger's equation + statistical interpretation.

Instead of analyzing  $\psi(x, t)$  into the normal orthogonal set  $u_n(x)$  as has been done in (64) we could have used some other set of normal orthogonal functions, say  $\bar{u}_{\bar{n}}(x)$ . The whole theory applies as before. But if the functions  $\bar{u}_{\bar{n}}$  themselves contain the time a modification is introduced. We let

$$\bar{u}_{\bar{n}'}(x'; t) = \bar{S}(x', n'; t). \quad (79)$$

Then defining

$$\bar{H}^{(\bar{n})} = \bar{S}^{-1}H^{(x)}\bar{S}; \quad Y^{(\bar{n})} = \bar{S}^{-1}Y^{(x)}\bar{S}$$

and

$$\bar{K}^{(\bar{n})} = \bar{H}^{(\bar{n})} + \frac{h}{2\pi i} \left( \bar{S}^{-1} \frac{d\bar{S}}{dt} \right) \quad (80)$$

the Schroedinger equation (60') transforms itself into

$$(h/2\pi i)(\partial \bar{\Psi} / \partial t) + \bar{K}^{(\bar{n})}\bar{\Psi} = 0 \quad (81)$$

where

$$\bar{\Psi} = \bar{S}^{-1}\psi.$$

The mean value of  $Y$  is as before

$$\bar{Y} = \bar{\Psi}^* Y^{(\bar{n})} \bar{\Psi} \quad (82)$$

and just as before the validity of (62) is secured by an equation similar to (63) with  $\bar{K}$  replacing  $H$ . Similarly  $\bar{K}$  replaces  $H$  in Eq. (74').

Denoting  $S^{-1}\bar{S}$  by  $\dot{S}$  we also have a direct relationship between the  $n$  and the  $\bar{n}$  systems of reference. For then

$$\bar{\Psi} = \dot{S}\Psi. \quad (83)$$

The transformation functions  $\bar{U}(t)$  and  $U(t)$  are related by

$$\bar{U}(t) = \hat{S}(t) U(t) \hat{S}^{-1}(0). \quad (83')$$

Formula (83') shows the connection between Eq. (75) and the corresponding formula in the  $\bar{n}$  reference system. It is easily verified that in this case  ${}^tK(t)$  takes place of  ${}^tH(t)$ .

Formula (80) for  $K$  is useful in those cases in which  $H$  contains the time explicitly in the form of a small additive term. By choosing  $\bar{S}$  properly such a term may be removed.

#### IV. THE KRAMERS-HEISENBERG FORMULA AS DERIVED BY BORN, HEISENBERG, JORDAN AND BY SCHROEDINGER

The correspondence principle derivation of the Kramers-Heisenberg formula suffers from two defects. Firstly the passage from the classical formula (44'') to the quantum formula (53) involved a judicious replacement of the classical Fourier coefficients  $C_r$  by the  $r_{J,J''}$ . This substitution is convincing enough inasmuch as it introduces empirically observable intensities of spectral lines. It is impossible however to give a definite formula for the calculation of the  $r_{J,J''}$  and the theory has therefore succeeded only in establishing a connection between phenomena of emission and those of dispersion. Secondly the process of emission was not understood at the time of the original derivation. Thus it was not clear under what circumstances the light emitted by an atom has coherence properties. An atom falling from one energy level to another and emitting a quantum appears to have nothing to do with any other atom going through the same process. Nevertheless it has been shown experimentally by Wood that in resonance radiation the light quanta emitted by different atoms interfere with each other as shown by the appearance of specular reflection at high vapor densities. [VII §1 below.]

We owe to Born, Heisenberg, Jordan and to Schroedinger a precise interpretation of the quantities  $r_{J,J''}$  in the dispersion formula. Schroedinger used his wave-mechanics and calculated the changes in the charge density distribution produced by an incident light wave. The light wave was taken into account as a periodic perturbing force acting on the atom. The Hamiltonian of the system was exactly the same as that used by Kramers and Heisenberg [(40) above]. Using this Hamiltonian function in the Schroedinger Eq. (60), solutions for the wave function  $\psi$  were worked out correct to the first power of the electric intensity  $\mathcal{E}$ . An atom which at the time  $t=0$  is in the state  $n$  described by the proper function  $u_n(x, y, z)$  is then found to be described at the time  $t$  by the wave function

$$\begin{aligned} u_n e^{-2\pi i \nu_n t} + \sum_m \frac{\mathcal{M}_{mn}}{2h} e^{-2\pi i \nu_n t} \left[ \mathcal{E} \frac{e^{2\pi i \nu t} - e^{-2\pi i \nu_{mn} t}}{\nu + \nu_{mn}} \right. \\ \left. + \mathcal{E}^* \frac{e^{-2\pi i \nu t} - e^{-2\pi i \nu_{mn} t}}{\nu_{mn} - \nu} \right] u_m \end{aligned} \quad (84)$$



as is found by using the expression (42) for  $\mathcal{E}(t)$  and solving (65). In the above expressions as usual  $\nu_n = E_n/h$  where  $E_n$  is the energy of the state  $n$  and  $\nu_{nm} = -\nu_{mn} = \nu_n - \nu_m$ . From this solution matrix elements of all quantum mechanical quantities may be worked out. In particular the matrix elements of the electric polarization may also be obtained.

Born, Heisenberg, Jordan have started out with matrix equations equivalent to the above Schroedinger equation and obtained a solution equivalent to the above wave function expression in all essential respects. If the reader is interested in following the mathematical connection of the treatments he will find it convenient to denote the transformation function of formula (74) by  $U_t(n, m) = (U_t^0 + U_t^0 V)(n, m)$  where

$$U_t^0(n, m) = \delta_{nm} e^{-2\pi i \nu_n t}$$

$$V(n, m) = (1/(2h)) \left[ \mathcal{E} \frac{e^{2\pi i(\nu + \nu_{nm})t} - 1}{\nu_{nm} + \nu} + \mathcal{E}^* \frac{e^{2\pi i(\nu_{nm} - \nu)t} - 1}{\nu_{nm} - \nu} \right] \mathcal{M}_{nm}.$$

Then any operator  $Y$  at the time  $t$  is represented by the matrix

$${}^tY = (1 - V)(U_t^0)^{-1} Y U_t^0 (1 + V) = {}^tY^0 - [V, {}^tY^0]$$

to the first power of  $\mathcal{E}$ . The matrices of Born, Heisenberg and Jordan are obtained by replacing  $e^{2\pi i(\nu_{nm} \pm \nu)t} - 1$  in  $V$  by  $e^{2\pi i(\nu_{nm} \pm \nu)t}$  and thus dealing with wave functions which do not reduce to  $u_n$  for  $t=0$  but contain atomic frequencies only as  $e^{-2\pi i \nu_n t}$ . These functions differ from the above function worked out by Schroedinger by not containing the terms in  $e^{-2\pi i \nu_{nm} t}$  in the square brackets. The matrix elements are with this modification

$$({}^tY - {}^tY^0)_{nm} = \sum_{m'} \frac{1}{2h} \left\{ Y_{nm'}^0 \left[ \mathcal{E} \frac{e^{2\pi i(\nu + \nu_{nm})t}}{\nu_{m'm} + \nu} + \mathcal{E}^* \frac{e^{2\pi i(\nu_{nm} - \nu)t}}{\nu_{m'm} - \nu} \right] \mathcal{M}_{m'm}^0 \right. \\ \left. - \left[ \mathcal{E} \frac{e^{2\pi i(\nu + \nu_{nm})t}}{\nu_{nm'} + \nu} + \mathcal{E}^* \frac{e^{2\pi i(\nu_{nm} - \nu)t}}{\nu_{nm'} - \nu} \right] \mathcal{M}_{n'm}^0 Y_{m'm}^0 \right\}. \quad (84')$$

The matrix elements  $\mathcal{M}_{m'm}^0 Y_{nm'}^0$  are independent of  $t$ . It is customary to interpret the terms containing  $(\nu \pm \nu_{nm})t$  in the exponent as having to do with the emission of frequencies<sup>23</sup>  $\nu \pm \nu_{nm}$ . This procedure is definite for the coherent radiation, but not convincing for Raman lines.

With a very slight modification the line of thought begun in these treatments can be made to lead to a definite result. In fact the paper of Born, Heisenberg and Jordan starts out by discussing the coupling of the atom to radiation precisely from the point of view of stationary states which we will carry through. This point of view is abandoned later in their paper having been used to justify in a sense the treatment of the light wave as an external perturbing force varying periodically with the time.

In ordinary dispersion phenomena at frequencies lying in the optical region the coupling of light to the atom occurs through the electric intensity

<sup>23</sup> The average absorption is easily derivable from (84). One obtains both ordinary absorption and *induced emission* in agreement with (47).

of the incident light wave. In classical physics the electric intensity of a monochromatic wave could be set equal to  $A \cos(2\pi\nu t - \epsilon)$  where  $\epsilon$  is a constant phase. In quantum theory the analogon to such a function may be obtained by observing that the coordinate and the momentum of a harmonic oscillator of frequency  $\nu$  are also represented classically by formulas of the type  $\cos(2\pi\nu t - \epsilon)$ . We imagine a harmonic oscillator coupled to the dispersing atom thinking of the two classically for the present. The force on the atom due to the oscillator is the more nearly harmonic the looser the coupling between them. We attempt therefore to describe the effect of the light wave as that of an oscillator of the same frequency coupled very weakly to the atom. We suppose that the Hamiltonian function is

$$H = \pi\nu(P^2 + Q^2) - \mathbf{C}P\mathcal{M} + H_0 \quad (85)$$

where  $P, Q$  are respectively the momentum and the coordinate of the oscillator.  $\mathbf{C}$  is a small constant vector in the direction of the incident electric intensity.  $H_0$  is the Hamiltonian of the atom unperturbed by radiation. In the classical theory there is no difference between using (85) with very small  $\mathbf{C}$  and using (40). In the quantum theory there is of course a difference inasmuch as the system described by (85) has definite energy states while for (40) we have no way of even defining energy levels. Nevertheless there is also a similarity. The energy levels of a harmonic oscillator are equally spaced, the spacing between adjacent levels being  $h\nu$ . An easy calculation shows that the quantum mean of any operator which is a function of only  $Q$  and  $P$  is a Fourier series in the time with period  $1/\nu$ , provided the oscillator is not perturbed. The state of the unperturbed oscillator is therefore necessarily periodic.

In particular the quantum means of  $Q$  and  $P$  themselves are under these conditions of the form  $A \cos(2\pi\nu t - \epsilon)$ . If  $\mathbf{C}$  is made very small the quantum mean of  $P$  must be taken large in order that the quantum mean of  $\mathbf{C}P$  be comparable with  $\mathcal{E}$ . Thus large quantum numbers of the oscillator should be used. It is then possible to work with a  $\psi$  concentrated into a small region so that the uncertainties of  $P$  and  $Q$  are both negligible when multiplied by  $\mathbf{C}$ . It has in fact been shown by Schroedinger that the eigenfunctions of a harmonic oscillator may be superposed to form such a wave packet.<sup>21</sup> We regard, therefore, the system described by (85) as the proper quantum mechanical counterpart of a classical periodic field of force.

In order to treat the system described by (85) we introduce action and angle variables of the oscillator by means of

$$P + iQ = (J/\pi)^{1/2} e^{2\pi iw}, \quad P - iQ = e^{-2\pi iw} (J/\pi)^{1/2}. \quad (85')$$

<sup>21</sup> E. Schroedinger, *Naturw.* **14**, No. 28, 664 (1926). This paper was written soon after the invention of wave mechanics. The interpretation of the wave function is not in agreement with the one accepted today. Using Schroedinger's formula (8), however, we obtain  $\psi^*\psi = \exp [A^2/2 - (x - A \cos 2\pi\nu_0 t)^2]$  which represents a Gauss error distribution of charge having the mean position  $A \cos 2\pi\nu_0 t$ .

According to the well-known treatment of Dirac the proper values of  $J$  are then of the form

$$J_s = N_s h + \text{const.}$$

where the constant is independent of  $s$ , and  $N$  is an integer. The well-known discussion of a harmonic oscillator shows that the constant may be made zero and the  $N$  take values 0, 1, 2, . . . . The Hamiltonian (85) may be then rewritten

$$H = (J + (h/2))\nu - 2^{-1}\pi^{-1/2}(J^{1/2}e^{2\pi i w} + e^{-2\pi i w}J^{1/2})\mathcal{CM} + H_0. \quad (85'')$$

The matrices of  $J^{1/2}e^{2\pi i w}$  and of  $e^{-2\pi i w}J^{1/2}$  are now defined to within a canonical transformation. They may be chosen so that

$$(J^{1/2}e^{2\pi i w})_{J', J'-h} = (J')^{1/2}; \quad (e^{-2\pi i w}J^{1/2})_{J', J'+h} = (J' + h)^{1/2}$$

are the only nonvanishing matrix elements. Then

$$\begin{aligned} Q_{J', J'-h} &= (1/2i)(J'/\pi)^{1/2}, \quad Q_{J', J'+h} = -(1/(2i))((J' + h)/\pi)^{1/2} \\ P_{J', J'-h} &= (\tfrac{1}{2})(J'/\pi)^{1/2}, \quad P_{J', J'+h} = (\tfrac{1}{2})((J' + h)/\pi)^{1/2} \end{aligned} \quad (86)$$

are the only nonvanishing matrix elements of  $Q$  and  $P$  by (85').

So far we have discussed the matrices for  $Q, P$  as though the oscillator were not coupled to the atom. For the coupled system consisting of the oscillator and the atom the matrices must be referred not only to the energy levels of the unperturbed oscillator but also to the energy levels of the unperturbed atom. The operators  $Q, P$  are supposed to operate only on the oscillator variables. The matrices  $Q, P$  are therefore diagonal in the energy levels of the atom. Thus

$$P_{n', N'; n'', N''} = (h/4\pi)^{1/2} \delta_{n' n''} \{ (N')^{1/2} \delta_{N', N''+1} + (N' + 1)^{1/2} \delta_{N', N''-1} \} \quad (86')$$

where  $n', n''$  denote atomic energy levels and  $N' = J'/h, N'' = J''/h$  are the vibrational quantum numbers of the oscillator. Similarly  $\mathcal{M}$  is diagonal in the oscillator quantum numbers. The matrix elements of  $\mathcal{M}$  referred to the energy levels of the unperturbed atom will be written  $\mathcal{M}_{n' n''}$ . The operator Eq. (85'') gives now directly for the matrix of the total energy

$$\begin{aligned} H_{n', N'; n'', N''} &= [(N' + \tfrac{1}{2})h\nu + h\nu_{n'}] \delta_{n' n''} \delta_{N' N''} \\ &\quad - C(h/4\pi)^{1/2} [(N')^{1/2} \delta_{N', N''+1} + (N' + 1)^{1/2} \delta_{N', N''-1}] \mathcal{M}_{n' n''}. \end{aligned} \quad (87)$$

Eq. (65) may now be used. Attention must be paid to the fact that the single index  $n$  or  $m$  in (65) denoted a definite state of the whole system. The state of the composite system atom + oscillator needs for its specification the quantum numbers  $n$  of the atom, as well as the vibrational quantum numbers  $N$  of the oscillator. The single suffix  $n$  of (65) must be therefore replaced by the double suffix  $n, N$  of (87). The whole discussion is *similar to* that of a *diatomic molecule* with electronic excitation  $n$  and vibration quantum numbers  $N$ . The Schroedinger equation is

$$\frac{\hbar}{2\pi i} \frac{dc_{n,N}}{dt} + \sum_{m,M} H_{n,N;m,M} c_{m,M} = 0.$$

The system may be started in the state  $n', N'$  at  $t=0$ . Its state for  $t>0$  is determined by the above equation. In agreement with the convention of formula (74) the values of  $c_{n,N}$  corresponding to this condition are written  $c_{n,N}^{n',N'}(t) = U_t(n, N; n', N')$ . The solution for  $c_{n,N}^{n',N'}$  is obtained by a direct and simple calculation using (87) and (65). The only  $c_{n,N}^{n',N'}$  which do not vanish to the first power of the small constant  $C$  are

$$\begin{aligned} c_{n',N'}^{n',N'} &= e^{-2\pi i[(N'+1/2)\nu + \nu_{n'}]t} \\ c_{n,N'-1}^{n',N'} &= \left(\frac{N'}{4\pi\hbar}\right)^{1/2} \mathcal{CM}_{nn'} \frac{e^{-2\pi i[(N'+1/2)\nu + \nu_{n'}]t} - e^{-2\pi i[(N'-1/2)\nu + \nu_n]t}}{\nu_{nn'} - \nu} \\ c_{n,N'+1}^{n',N'} &= \left(\frac{N'+1}{4\pi\hbar}\right)^{1/2} \mathcal{CM}_{nn'} \frac{e^{-2\pi i[(N'+1/2)\nu + \nu_{n'}]t} - e^{-2\pi i[(N'+3/2)\nu + \nu_n]t}}{\nu_{nn'} + \nu}. \end{aligned} \quad (87')$$

The reason for the occurrence of the exponential terms of frequencies  $(N'-1/2)\nu + \nu_n$  and  $(N'+3/2)\nu + \nu_n$  is that we have supposed the system to be in the state  $n', N'$  at the time  $t=0$  and that this state is not a stationary state of the coupled system. These terms may be omitted without destroying the agreement with (65). Then at  $t=0$  the system is not exactly in the state  $n', N'$  but only approximately so. There is a probability of the order  $C^2$  of its being in a state  $n, N'-1$  or  $n, N'+1$ . We notice however, that omitting these exponential terms leaves only terms in

$$e^{-2\pi i[(N'+1/2)\nu + \nu_{n'}]t}.$$

The resultant state may be interpreted therefore as a stationary stage of the coupled system having an energy

$$\hbar[(N' + \frac{1}{2})\nu + \nu_{n'}].$$

The fact that the energy of stationary states is independent of  $C$  to the first order follows also from the well-known theory of perturbations of stationary non-degenerate states and formula (87). The fact about (87) which is important in this connection is that the perturbation energy contains no diagonal matrix elements. The resultant formulas

$$\begin{aligned} c_{n',N'}^{n',N'} &= e^{-2\pi i[(N'+(1/2))\nu + \nu_{n'}]t} \\ c_{n,N'-1}^{n',N'} &= \left(\frac{N'}{4\pi\hbar}\right)^{1/2} (\mathcal{CM}_{nn'}) \frac{e^{-2\pi i[(N'+(1/2))\nu + \nu_{n'}]t}}{\nu_{nn'} - \nu}, \\ c_{n,N'+1}^{n',N'} &= \left(\frac{N'+1}{4\pi\hbar}\right)^{1/2} (\mathcal{CM}_{nn'}) \frac{e^{-2\pi i[(N'+(1/2))\nu + \nu_{n'}]t}}{\nu_{nn'} + \nu} \end{aligned} \quad (88)$$

may also be derived by means of the theory of perturbations of stationary states.

The stationary state described by (88) may be still designated by  $n', N'$ . This designation is unique for small  $C$  because the probability of the atom being in  $n'$  and the oscillator simultaneously in  $N'$  is much larger than the simultaneous probability of any two other states. The matrix element of any operator  $Y$  not involving  $t$  explicitly referred to these stationary states is then in accordance with (74)

$$\sum_{m', M'; m'', M''} c_{m', M'; m'', M''}^{*n', N'} Y_{m', M'; m'', M''} c_{m'', M''}^{n'', N''} = {}^t Y_{n', N'; n'', N''}. \quad (89)$$

The matrix elements  $Y_{m', M'; m'', M''}$  are independent of the time and are referred to the stationary states  $m', m''$  of the atom and the stationary states  $M', M''$  of the oscillator. Let now  $Y$  be an operator which acts only on the atomic variables. Then

$$Y_{m', M'; m'', M''} = Y_{m', m''} \delta_{M', M''}$$

is diagonal in the oscillator quantum numbers. If  $C=0$  then by (88)  $c_{n', N' \mp 1}^{n', N'} = 0$  and by (89)  ${}^t Y_{n', N'; n'', N''}$  is also diagonal in  $N', N''$ . Under these conditions there can be no radiation jumps in which the vibrational quantum number  $N'$  changes. If  $C$  is not zero Eq. (88) shows that to the first power of  $C$   ${}^t Y$  has nonvanishing matrix elements in which the vibrational quantum number may change by  $\pm 1$ . The physical reason for this lies in the fact that the stationary state  $n', N'$  results by superposition of states in which the atom is in states  $n''$  and the oscillator in states  $N' \pm 1$ .

In order to say something definite about the probability of a jump from  $n', N'$  to  $n'', N''$  we need to make a definite hypothesis. Such a hypothesis has been made in the first introduction of quantum theory and consists in postulating formula (47) above as the relation between the Einstein emission coefficient  $A_{J'J''}$  and the vector matrix of the electric polarization  $e \mathbf{r}_{J'J''}$ . Here it is important that  $J', J''$  be stationary states of the system.

Writing  $U_t = U_t^{(0)} + V U_t^{(0)}$  where

$$\begin{aligned} c_{m', M'}^{n', N'} &= U_t(m', M'; n', N') \\ U_t^{(0)}(m', M'; n', N') &= \delta_{m', n'} \delta_{M', N'} \exp \left[ -2\pi i \left( (N' + \frac{1}{2})\nu + \nu_{n'} \right) t \right] \\ V(m', M'; m'', M'') &= \left( \frac{M'}{4\pi\hbar} \right)^{1/2} \frac{C \mathcal{N}_{m' m''}}{\nu_{m' m''} + \nu} \delta_{M' - 1, M''} \\ &\quad + \left( \frac{M' + 1}{4\pi\hbar} \right)^{1/2} \frac{C \mathcal{N}_{m' m''}}{\nu_{m' m''} - \nu} \delta_{M' + 1, M''} \end{aligned}$$

we have for any operator not involving  $t$  explicitly as in (74)

$${}^t Y = U_t^{(0)-1} (1 - V) Y (1 + V) U_t^{(0)} = {}^t Y^{(0)} + U_t^{(0)-1} [Y, V] U_t^{(0)}$$

where

$${}^t Y^{(0)} = U_t^{(0)-1} Y U_t^{(0)}$$

is the matrix which would exist if there were no perturbation with the oscillator. Writing

$${}^tV^{(0)} = U_t^{(0)-1} V U_t^{(0)}$$

we have for the matrix elements introduced by the perturbation:

$$\begin{aligned} [{}^tY^{(0)}, {}^tV^{(0)}]_{n, N; n', N-1} &= \sum_{n''} \left( \frac{N}{4\pi h} \right)^{1/2} \left( \frac{Y_{n, n''}^{(0)} \mathcal{M}_{n'' n'}}{\nu_{n'' n'} + \nu} - \frac{\mathcal{M}_{nn''} Y_{n'' n'}^{(0)}}{\nu + \nu_{nn''}} \right) e^{2\pi i(\nu_{nn'} + \nu)t} \\ [{}^tY^{(0)}, {}^tV^{(0)}]_{n, N; n', N+1} &= \sum_{n''} \left( \frac{N+1}{4\pi h} \right)^{1/2} \left( \frac{Y_{n, n''}^{(0)} \mathcal{M}_{n'' n'}}{\nu_{n'' n'} - \nu} - \frac{\mathcal{M}_{nn''} Y_{n'' n'}^{(0)}}{\nu_{nn''} - \nu} \right) e^{2\pi i(\nu_{nn'} - \nu)t} \end{aligned} \quad (90)$$

with

$${}^tY_{n' n''}^{(0)} = Y_{n' n''}^0 e^{2\pi i \nu_{n' n''} t}.$$

In particular these formulas may be applied to the electric polarization  $\mathcal{M}$ . The first line of (90) determines by means of (47) the probability of a jump from the stationary state  $n, N$  to the stationary state  $n', N-1$  with the emission of a quantum of energy  $h(\nu_{nn'} + \nu)$  on the supposition that the system is in the state  $n, N$  initially. It must also be supposed that  $\nu_{nn'} + \nu > 0$  because otherwise a quantum  $h(\nu_{nn'} + \nu)$  can not be emitted. The matrix element of the second line of (90) gives similarly the probability of a jump from the state  $n, N$  to the state  $n', N+1$  with the emission of a quantum  $h(\nu_{nn'} - \nu)$ . Again, in order that it be possible to emit the quantum  $h(\nu_{nn'} - \nu)$  it must be supposed that  $\nu_{nn'} - \nu > 0$ . The first line of (90) gives emissions of the Stokes type if  $\nu_{nn'} < 0$ , the emitted quantum being then  $h(\nu - |\nu_{nn'}|)$ ; if  $\nu_{nn'} > 0$  antistokes Raman lines result. The second line of (90) gives rise to transitions identical with those represented by the second line of (53). We have already interpreted the energy relationships for this type of scattering process by (55). The induced emission of  $h\nu$  was used there. We see now that this interpretation is correct. In fact the transition

$$n, N \rightarrow n', N+1$$

may be interpreted by saying that the energy of the incident radiation increased by  $(N+1)h\nu - Nh\nu = h\nu$  while the energy of the atom decreased by  $h\nu_{nn'}$ .

Formulas (90) have been derived as the matrix elements for the coupled system of oscillator and atom. The first line of (90) corresponds to emissions of a quantum  $h(\nu_{nn'} + \nu)$  only if this quantity is positive. If it is negative and if the first line of (90) does not vanish the frequency  $\nu_{n' n} - \nu$  is a possible absorption frequency of the coupled system. The absorption process has for its initial stage the atom in the state  $n$  and the oscillator in the vibrational state  $N$ . When the system is in this condition it may absorb a quantum  $h(\nu_{n' n} - \nu)$

and end in a state  $n'$  of the atom and the state  $N-1$  of the oscillator. This process may be termed *double absorption* because it occurs through the absorption of a quantum  $h\nu$  from the oscillator and a quantum  $h(\nu_{nn'} - \nu)$  from the radiation. This process was first mentioned by M. Göppert Mayer,<sup>22</sup> who also emphasized the importance of the scattered frequencies  $\nu_{nn'} - \nu$  which occurred previously in the paper of Kramers and Heisenberg. The emission of these frequencies occurs even if  $N$  is zero, i.e., if there is no available energy in the incident radiation. The frequency  $\nu_{nn'} - \nu$  is emitted jointly with an increase in the energy of the oscillator by the amount  $h\nu$ . The process is called

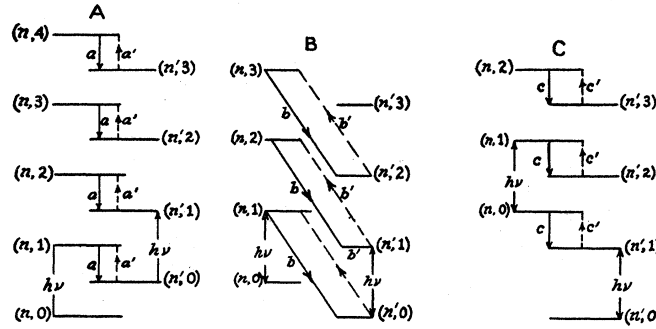


Fig. 1.

$\begin{Bmatrix} a \\ b \end{Bmatrix}$  are  $\begin{Bmatrix} \text{Stokes} \\ \text{antistokes} \end{Bmatrix}$  emissions of  $\begin{Bmatrix} h(\nu - \nu_{n'n}) \\ h(\nu + \nu_{nn'}) \end{Bmatrix}$  from state  $n$  with probability:

$$(128\pi^5\nu'^3/3h^3c^3)\int |\mathbf{m}_1|^2 \rho(\nu) d\nu.$$

$\begin{Bmatrix} a' \\ b' \end{Bmatrix}$  are absorptions of  $\begin{Bmatrix} h(\nu - \nu_{n'n}) \\ h(\nu + \nu_{nn'}) \end{Bmatrix}$  from state  $n'$  and induced emission of  $h\nu$  with probability:  $(16\pi^4/h^4)\int |\mathbf{f}'\mathbf{m}_1|^2 \rho(\nu') [\rho(\nu) + 8\pi h\nu^3/c^3] d\nu$ .

$c$  is emission of  $h(\nu_{nn'} - \nu)$  from  $n$  accompanied by emission of  $h\nu$  with probability:  $(128\pi^5\nu'^3/3h^3c^3)\int |\mathbf{m}_2|^2 [\rho(\nu) + (8\pi h\nu^3/c^3)] d\nu$ .

$c'$  is absorption of  $h(\nu_{nn'} - \nu)$  and of  $h\nu$  from  $n'$  with probability:  $(16\pi^4/h^4)\int |\mathbf{f}'\mathbf{m}_2|^2 \rho(\nu') \rho(\nu) d\nu$

$\mathbf{f}$  is unit vector along electric intensity of  $h\nu$

$\mathbf{f}'$  is unit vector along electric intensity of  $(h(\nu - \nu_{n'n})$  or of  $h(\nu_{nn'} - \nu)) = h\nu'$

$$\mathbf{m}_1 = \sum_{n''} \left( \frac{\mathcal{M}_{nn''}(\mathcal{M}_{n''n'}\mathbf{f})}{\nu_{n''n'} + \nu} - \frac{\mathcal{M}_{n''n'}(\mathcal{M}_{nn''}\mathbf{f})}{\nu + \nu_{nn''}} \right); \quad \mathbf{m}_2 = \sum_{n''} \left( \frac{\mathcal{M}_{nn''}(\mathcal{M}_{n''n'}\mathbf{f})}{\nu_{n''n'} - \nu} - \frac{\mathcal{M}_{n''n'}(\mathcal{M}_{nn''}\mathbf{f})}{\nu_{nn''} - \nu} \right)$$

double emission because it occurs through the simultaneous emission of the light quanta  $h\nu$  and  $h(\nu_{nn'} - \nu)$ . If  $\nu_{nn'} - \nu < 0$  the frequency  $\nu + \nu_{n'n}$  is a possible absorption frequency, the initial state being  $n$ ,  $N$  and the final  $n'$ ,  $N+1$ . We may have, therefore, Stokes or antistokes absorption frequencies depending on whether  $\nu_{n'n} < 0$  or  $\nu_{n'n} > 0$ . The processes involved in these absorptions are bound up with yielding an energy  $h\nu$  to the oscillator, i.e., the emission of a light quantum  $h\nu$ . Theoretically they should occur whether the oscillator is excited above its normal state or not. The presence of  $(N+1)^{1/2}$  in the second line of (90) shows that the probability of such an absorption is

<sup>22</sup> Maria Göppert, Naturwiss. 17, 932 (1929); Maria Göppert Mayer, Ann. d. Physik 9, 401 (1931).

increased by the excitation of the oscillator, i.e., by the presence of light of frequency  $\nu$ . Thus an atom in a state  $n$  exposed to radiation of frequency  $\nu$  shows absorption lines of frequencies  $\nu + \nu_{n'n}$  provided  $\nu + \nu_{n'n} > 0$  and the second line of (90) does not vanish. Fig. 1 shows diagrammatically the Raman transitions the inverse absorptions and the double emissions and absorptions. The probabilities of the processes are given in terms of energy densities. See also V end of §7 below.

Formula (47) enables us to compute the intensities of the Raman lines as well as the intensity of the absorption processes just discussed, by calculating  $[\mathcal{M}^{(0)}, {}^tV^{(0)}]$  in accordance with (90). The matrix elements of this quantity may be identified with  $e\mathbf{r}_{J,J'}$  of (47). In expressing the results for practical applications it is necessary to establish the relation of the energy density of the incident radiation and its electric intensity to the small vector  $\mathbf{C}$  and to the vibrational quantum number  $N$  of the oscillator.

It must be remembered that the condition of the oscillator with which we have been dealing is very similar to a stationary state of the unperturbed oscillator. The smaller  $\mathbf{C}$ , the more nearly is this the case. In a stationary state the quantum means of the coordinate  $Q$  and the momentum  $P$  are zero. Nevertheless by superposing the proper functions of suitably chosen stationary states we can obtain conditions in which these quantum means are not zero. Schroedinger's wave packet is such a condition. Whether we wish to use this wave packet or some other we are limited by the uncertainty relation

$$\Delta P \Delta Q \sim h/2\pi$$

and we further have the approximate relations of quantum means

$$\overline{P^2} - \overline{P}^2 \sim (\Delta P)^2; \quad \overline{Q^2} - \overline{Q}^2 \sim (\Delta Q)^2.$$

By choosing the wave packet suitably we may make  $(\Delta P)^2$  and  $(\Delta Q)^2$  of the order of  $h$ . Thus  $\overline{P^2} - \overline{P}^2$  may be kept constant while  $\overline{P}$  may be increased indefinitely. Further if the total range covered by  $Q$  be  $Q_{\max}$ , the de Broglie wave-length is of the order  $Q_{\max}/N$  so that the "beat wave-length" due to superposition of two proper functions is of the order  $Q_{\max}/(N_1 - N_2)$ . Since for  $\Delta Q \sim h^{1/2}$ ,  $\Delta Q \sim Q_{\max} N^{-1/2}$  and since the "beat wave-length" is of the order of  $\Delta Q$  we have  $N_1 - N_2 \sim N^{1/2}$ . The range of the quantum numbers  $N$  which must be employed in order to construct a proper wave packet is thus of the order  $N^{1/2}$  and can be made negligible in comparison with  $N$  by making  $N^{1/2} \gg 1$ . Under these condition  $(N+1)^{1/2}$  in (90) may be identified with  $N^{1/2}$  and  $N^{1/2}\mathbf{C}$  may be connected with the electric intensity  $\mathcal{E}(t)$  of the incident radiation.

The quantum mean of  $\pi\nu(P^2 + Q^2)$  is the energy of the oscillator  $(N+1/2)h\nu$ . To a sufficient approximation this is  $Nh\nu$ . The quantum mean of  $P^2$  is not constant. Its average over a period  $1/\nu$  is constant. It may be written  $\overline{\overline{P^2}}$  to indicate that the quantum mean has been averaged. Then

$$\overline{\overline{P^2}} = Nh/(2\pi).$$



The time mean of  $\mathcal{E}^2(t)$  where

$$\mathcal{E}(t) = (\frac{1}{2})(\mathcal{E}e^{2\pi i\nu t} + \mathcal{E}^*e^{-2\pi i\nu t})$$

is

$$\overline{\mathcal{E}^2(t)} = (\frac{1}{2})\mathcal{E}\mathcal{E}^*.$$

Since  $\mathcal{CP}$  takes place of  $\mathcal{E}(t)$  we require that  $\mathcal{C}\mathcal{C}\overline{P^2} = \overline{\mathcal{E}^2(t)}$  and we have

$$(h/\pi)N\mathcal{C}\mathcal{C} = \mathcal{E}\mathcal{E}^*. \quad (91)$$

The value of  $N$  obtained from this formula determines to within  $\pm N^{1/2}$  the vibrational quantum numbers of the oscillator which must be used in order that the center of gravity of the wave packet should move through the same range as  $\mathcal{E}(t)$ . The vector  $\mathcal{C}$  is in the same direction as  $\mathcal{E}$ . The vector  $\mathcal{E}$  is complex. In the application of (90) we are concerned only with squares of absolute values of the expressions (90). Thus if we wish we may use instead the expressions

$$\begin{aligned} & [{}^tY^{(0)}, {}^tV^{(0)}]_{n,N;n',N-1} \\ & \rightarrow \sum_{n''} (2h)^{-1} \left( \frac{Y_{nn''}^{(0)} \mathcal{E} \mathcal{M}_{n''n'}}{\nu_{n''n'} + \nu} - \frac{\mathcal{E} \mathcal{M}_{nn''} Y_{n''n'}^{(0)}}{\nu_{nn''} + \nu} \right) e^{2\pi i(\nu_{nn'} + \nu)t} \\ & [{}^tY^{(0)}, {}^tV^{(0)}]_{n,N;n',N-1} \\ & \rightarrow \sum_{n''} (2h)^{-1} \left( \frac{Y_{nn''}^{(0)} \mathcal{E}^* \mathcal{M}_{n''n'}}{\nu_{n''n'} - \nu} - \frac{\mathcal{E}^* \mathcal{M}_{nn''} Y_{n''n'}^{(0)}}{\nu_{nn''} - \nu} \right) e^{2\pi i(\nu_{nn'} - \nu)t} \end{aligned} \quad (90')$$

where  $(Nh/\pi)^{1/2}\mathcal{C}$  of the first line of (90) has been replaced by  $\mathcal{E}$  and  $((N+1)h/\pi)^{1/2}\mathcal{C}$  of the second line has been replaced by  $\mathcal{E}^*$ . Expressions (90') are equal to expressions (90) to within a factor of absolute value = 1. Formula (90') shows that we would have obtained correct results if we had used (84') and associated the terms in  $e^{2\pi i(\nu \pm \nu_{nm})t}$  with the emission of frequencies  $\nu \pm \nu_{nm}$ . So far we are not able to say anything about the phase relations of the emitted radiation which formula (84') implies.

A monochromatic wave is a theoretical idealization which may be realized only approximately in practice. A physical train of waves may be analyzed into monochromatic components. Each component we may represent by an oscillator. An atom exposed to a train of waves may be then pictured as coupled to a number of oscillators. To each oscillator there corresponds a vibrational quantum number. The term in  $\mathcal{C}$  in (87) must then be replaced by a summation of such terms. To the first order in  $\mathcal{C}$  we find then that  $[{}^tY^{(0)}, {}^tV^{(0)}](n, N_1, N_2, \dots; n', N_1-1, N_2, \dots)$  is the same as (90) with  $N$  replaced by  $N_1$ . The "polyatomic molecule" system (atom + several oscillators) has therefore the same probability of changing the vibrational quantum number of one oscillator while leaving the other quantum numbers unchanged as though the other oscillators were not present. The total probability of a radiative transition of the type  $n, N \rightarrow n', N-1$  is then obtained by using instead of the square of the absolute value of (90) a sum of squares each expres-

sion being obtained from (90) by replacing  $N$  by  $N_s$  and  $\nu$  by  $\nu_s$  where  $s$  refers to a given component oscillator. For a nearly monochromatic wave the  $\nu_s$  differ little in comparison with  $\nu_{n'n''}$ . For an assigned polarization of the incident wave the intensity of a given Raman line is therefore proportional to  $\sum_s N_s$  which is proportional to the intensity of the incident radiation.

In the above discussion we have used two physical premises. One of them consists in the supposition that the incident light wave may be treated as a collection of excited oscillators coupled to the atom. The other consists in the supposition that (47) is the correct formula for the calculation of spontaneous emission probabilities. No hypothesis has been made about the phase relationships of waves emitted in such spontaneous emissions. Without Dirac's theory of radiation it is difficult to understand spontaneous emission and the properties of radiation emitted by atoms can be understood only vaguely. Nevertheless we will discuss briefly the question of spontaneous emission from the point of view of the quantum theory of matter.

According to the postulates of quantum theory an atom in a state

$$\psi = \sum c_n(t) u_n$$

has for the observable mean of any quantity  $Y$

$$\bar{Y} = \sum c_n^*(t) Y_{n'n''} c_{n''}(t).$$

If the functions  $u_n$  represent stationary states

$$\bar{Y} = \sum c_n^*(0) Y_{n'n''} c_{n''}(0) e^{2\pi i \nu_{n'n''} t}.$$

One is tempted to apply this formula to a component of electric polarization and to suppose that the electric polarization is connected with the field intensities by the ordinary linear formulas of electrodynamics. It is then necessary to suppose that the field strengths are also matrix quantities. The quantum mean of the electric intensity then contains a Fourier term in  $e^{2\pi i \nu_{n'n''} t}$  with an amplitude proportional to  $c_n^*(0) \mathcal{M}_{n'n''} c_{n''}(0)$ . If only one  $c_n(0)$  is distinct from zero the quantum means of the field intensities as well as their squares, cubes, etc. are independent of the time. The electromagnetic field of an atom in a stationary state would on such a theory be interpretable as a steady field. This fact was pointed out by Schroedinger soon after the inception of wave mechanics, for, according to him  $\psi^* \psi$  is constant in time for stationary states. Once an atom is in a stationary state it has no reason to leave it. This is of course in contradiction with experience. Atoms are known to leave stationary states by spontaneous emission. It is for this reason that we had to make the *ad hoc* hypothesis of formula (47). It is, furthermore, clear that the field calculated on such a theory should be treated with extreme caution because, as has also been pointed out by Schroedinger,<sup>23</sup> its effect would be to spoil the agreement of theory and experiment for the spectrum of hydrogen.

In some cases, however, the electromagnetic field conceived as an ordin-

<sup>23</sup> E. Schroedinger, Ann. d. Physik (4) **82**, 265 (1927).

ary matrix referred to the atomic variables can be used to give correct physical results. Thus, for example by the superposition of proper wave functions we can build up a wave packet and by using large quantum numbers we can make the wave packet move like a classical electron. In this case the field can be computed from the motion of the wave packet by classical formulas. This calculation obviously agrees with the result of calculating the field by quantum formulas because the quantum mean of any power of a field strength is equal in this case to the power of the quantum mean. Furthermore in this limiting case we can determine, by the classical calculation, not only the probability of a certain amount of energy being radiated but also its detailed properties, such as the relationship of the electric intensity to the position of the wave packet, etc. Formula (47) will also agree with the classical calculation but it is not as detailed in its predictions.

In the limit of large quantum numbers the frequency  $\nu_{n'n''}$  approaches asymptotically the classical quantity  $(\tau\nu)$  with  $\tau = n' - n''$ . All terms in  $e^{2\pi i \nu_{n'n''} t}$  combine, and give rise to

$$\sum_{n'', n' = n'' + \tau} c_{n'}^*(0) \mathcal{M}_{n'n''} c_{n''}(0) e^{2\pi i (\tau\nu) t}$$

in the expression for the electric moment. In the expression for the electric moment of the coupled system of atom+oscillator discussed in (87), (88), (90) we also have terms of equal frequency combining and giving rise to

$$\sum_{N' - N'' = \tau} c_{n', N'}^*(0) {}^t\mathcal{M}_{n', N'; n'', N''} c_{n'', N''}(0) e^{2\pi i (\nu_{n'n''} + (N' - N'')\nu) t}$$

where  $\tau = +1$  or  $-1$ . Here  ${}^t\mathcal{M}$  is given by (90).

If it is supposed that

$$c_{n', N'} = a_n b_{N'}$$

the substitution of (90) into (92) gives  $a_n^* a_{n''}$  times the expressions (90') with

$$\mathcal{E} = \sum_N \left( \frac{N\hbar}{4\pi} \right)^{1/2} c b_N^* b_{N-1}.$$

This is in agreement with

$$\mathcal{E}(t) = CP$$

because by (86)

$$\bar{P} = \sum_N \left( \frac{N\hbar}{4\pi} \right)^{1/2} (b_N^* b_{N-1} e^{2\pi i \nu t} + b_{N-1}^* b_N e^{-2\pi i \nu t}).$$

We surmise therefore that (90') gives also a proper account of the phases. In particular if  $n' = n''$  we may regard this as true. Logically this has not been proved. The analogy with the correspondence principle considerations suggests that the combination of terms of equal frequency in (92) is justifiable and for  $n' = n''$  the interpretation of the quantum mean of  $P$  is unambiguous

and independent of supposing  $c_{n',N'} = a_n b_{N'}$ . These questions will be discussed again and in a more satisfactory fashion by means of Dirac's theory. In referring to them above we hope to have conveyed an idea as to the range of validity of the quantum theory of matter when applied to questions of radiation.

## V. DIRAC'S THEORY OF DISPERSION

### §1. Introductory

In the preceding section (IV) we have discussed the interaction of an atom with an electromagnetic wave. In order to understand the action of the wave we found it useful to compare it to that of a harmonic oscillator. When the wave lost the energy  $h\nu$  we said that the oscillator jumped to a lower energy level. The oscillator being harmonic we found it possible to build wave packets performing purely periodic motions which resembled the sinusoidal changes in the electric intensity of a monochromatic light wave.

It must now be pointed out that when we speak in optics of a frequency of light we do so largely to please our imagination. In reality no one has observed the time alternations of the electric and magnetic intensities of a light wave. The reason which made prequantum physics accept the literal significance of frequency is that the consequences of the view were found to be supported by the phenomena of interference, the numerical value of the velocity of light, the continuous transition from electromagnetic waves to light waves through the far infrared and other similar phenomena. With the discovery of the nature of the photoelectric effect, the Compton effect, etc., it became necessary to introduce the hypothesis of light quanta which is not in agreement with the classical wave picture of light. A complete theory should account both for the wave as well as the photon aspects of the phenomenon called light or electromagnetic radiation. It should account for the two aspects in such a way that in the region of Hertzian and radio waves the wave picture should apply without any essential modifications. For light, x-rays and  $\gamma$ -rays the photon aspect should become increasingly pronounced.

For low frequencies the energy in each quantum is small. In order to observe a given energy we need many quanta. The observation of the periodic wave character of an electromagnetic disturbance is possible thus in cases where there are available many quanta of the same frequency; if only one quantum is present we have evidence that this is impossible. For otherwise we would be at a loss to account for the concentration of energy into single electrons in the photoelectric effect. The word *frequency* must nevertheless have some significance even if we deal with a single quantum because it is known that the probability of having a quantum diffracted by a grating is much larger for the diffraction maxima than for other directions. By frequency we should understand the number  $\nu = c/\lambda$  where  $c$  is the velocity of light and  $\lambda$  the wave-length measured in vacuo by interference experiments. The theory should be formulated so that light quanta of energy content  $h\nu$  should be diffracted in the same way as an electromagnetic wave of frequency  $\nu$ . There is no justification however for considering the field intensities as pre-

cisely observable quantities except under special circumstances such as a radio wave.

Qualitatively the electric field of a light wave of frequency  $\nu$  is thus similar to the displacement from equilibrium of a quantized harmonic oscillator. Neither the electric field nor the coordinate of the oscillator can be observed with precision. The energy of radiation of frequency  $\nu$  as well as the energy of the oscillator can change only by the amount  $h\nu$ . To the state of a definitely measureable electric intensity [radio wave] there corresponds a wave package of the oscillator which performs a sinusoidal motion. The large number of light quanta in the radio wave is then seen to correspond to high values of the quantum number of the oscillator. On this analogy the passage from quantum to classical theory occurs for the radiation in the same way as it occurs for mechanical systems *viz.*, through the limit of high quantum numbers.

Again, an explanation of spontaneous emission is readily obtained by the above analogy. If there is one quantum of radiation incident on the atom the oscillator should be thought of as being in the first excited state. Let this quantum be absorbed. Then the oscillator is pictured to have fallen to its normal level. The atom may now fall back to its normal level and the radiation oscillator will become excited. This is much like absorption followed by spontaneous emission. The description is obviously not complete because in general the radiation may consist of all frequencies so that one should suppose that there are oscillators corresponding to every type of monochromatic radiation. The quantitative formulation of such a theory with an oscillator for every frequency polarization and direction of wave normal has been made by Dirac and is known as Dirac's theory of light quanta.<sup>3</sup>

In section I the electromagnetic field has been analyzed into (1) progressive transverse electromagnetic waves represented by the quantities  $a_{k,l}$ ,  $a_{k,l}^+$  and (2) a longitudinal part described by  $a_k$ ,  $a_k^*$ ,  $b_k$ ,  $b_k^*$ . The longitudinal part represents the electrostatic field. The transverse waves represented by  $a_{k,l}$ ,  $a_{k,l}^+$  have each a definite wave-length determined by equation (2'). By (7') these variables depend on the time only through the factors  $e^{\pm 2\pi i\nu t}$  in the absence of matter. Their dependence on the time is the same as that of  $P \mp iQ$  of the simple harmonic oscillator discussed in the preceding section. Dirac's theory consists in supposing that the variables  $a_{k,l}$ ,  $a_{k,l}^+$  may be treated as canonically conjugate quantum variables in such a way as to make each term  $2\pi i\nu_{k,l} a_{k,l}^+ a_{k,l}$  in the expression (6) for the energy of the radiation field have *proper values* differing from each other by integral multiples of  $h\nu_{k,l}$ . By introducing this hypothesis, all the essential features of the treatment which led to Eqs. (90) remain intact. The presence of energy in the form of radiation of a given wave-length, direction of wave normal and state of polarization described by  $k, l$  will still be interpreted as the excitation of the oscillator described by the variables  $a_{k,l}$ ,  $a_{k,l}^+$ . The oscillator will still be able to change its energy by steps  $h\nu_k$  although the quadratic term in  $\mathbf{A}$  in the Hamiltonian (30) introduces a coupling between two different oscillators.

## §2. Quantization of the field and equations of motion

The quantization of  $a_{k,l}$   $a_{k,l}^+$  can be performed quite analogously to that of  $P$  and  $Q$  of the preceding section. Eq. (85') suggests the transformation<sup>24</sup>

$$a_s^+ = -i(J_s/(2\pi))^{1/2} e^{-2\pi i w_s}, \quad a_s = e^{2\pi i w_s} (J_s/(2\pi))^{1/2}; \quad (s = k, l) \quad (93)$$

which leads to

$$[a_s^+, a_s] = a_s^+ a_s - a_s a_s^+ = \hbar/2\pi i$$

and

$$2\pi i \nu_s a_s^+ a_s = J_s \nu_s$$

we treat the  $J_s$  and  $w_s$  as canonically conjugate momenta and coordinates and suppose therefore that  $J_s, w_s$  commute when  $(k, l) \neq (k', l')$ . We have therefore also

$$\begin{aligned} [a_s^+, a_{s'}] &= \hbar \delta_{s,s'} / 2\pi i & [a_s^+, a_{s'}^+] &= [a_s, a_{s'}] = 0. \\ ((k, l) = s) &((k', l') = s'). \end{aligned} \quad (93'')$$

In the classical theory we imposed on the  $a_s, a_s^+$  the additional restriction

$$a_s^* = i a_s^+ \quad (1'')$$

which is equivalent to the requirement that  $a_s + i a_s^+$  and  $(1/(2i)) (a_s - i a_s^+)$  should both be real. In the quantum theory the corresponding operators should be Hermitian. Using (93) and comparing these operators with expression (85') we see that this is equivalent to the requirement that  $P, Q$  be Hermitian. Moreover, classically the quantities  $a_s + i a_s^+$  and  $(1/(2i)) (a_s - i a_s^+)$  cover each the whole range of values from  $-\infty$  to  $+\infty$  because all real values of the vector potential are admitted in classical electrodynamics. This is precisely what is supposed about the range of  $P$  and  $Q$  in the theory of the harmonic oscillator. We conclude therefore that in the transformation (93) the variables  $J, w$  also cover the same range of values as in (85'). If we set

$$J_s = (N_s + 1) \hbar \quad (94)$$

the quantum variable  $N_s$  has proper values

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

There is no physical reason for similarly quantizing the variables  $a_k, b_k$ . They may be treated as canonically conjugate variables subject to the commutation conditions

$$\begin{aligned} [b_k, a_{k'}] &= \frac{\hbar}{2\pi i} \delta_{k,k'}; & [b_k^*, a_{k'}^*] &= \frac{\hbar}{2\pi i} \delta_{k,k'} \\ [a_k^*, a_{k'}] &= [b_k^*, b_{k'}] = 0. \end{aligned} \quad (95)$$

The discussion given in Part I must now be reexamined so as to make sure of the order of factors in the various expressions. We take up for the present

<sup>24</sup> Where this is possible we abbreviate  $k, l$  into  $s$ .

the nonrelativistic description of matter with a correct representation of Maxwell's equations. Classically this was given by (30).

We fix the order in  $H_R$  by writing

$$H_R = \sum_s 2\pi i \nu_s a_s a_s^+ + \sum_k 4\pi \nu_k b_k^* b_k. \quad (96)$$

By (93'), (94) the first summation is zero if all  $N_s = 0$ . Expression (96) does not contain the *zero point energy*. If instead of  $a_s a_s^+$  we had written  $(1/2)(a_s^+ a_s + a_s a_s^+)$  we would also have the zero point energy and we would have satisfied  $H_R = (1/8\pi) \int (\mathcal{E}^2 + \mathcal{H}^2) dV$ . It does not matter whether we use  $H_R$  with zero point energy or not because an additive constant in the expression for the energy does not affect any of the physical results derived. It is nicer to use  $H_R$  without the zero point energy because then the proper values of the first summation of (96) are finite as long as a finite number of oscillators are excited. With the above  $H_R$  we form

$$H = H_R - (e/2mc) \int (\mathbf{p} D_{\mathbf{r}, \mathbf{r}'} + D_{\mathbf{r}, \mathbf{r}'} \mathbf{p}) \mathbf{A}(\mathbf{r}') dV' \\ + (e^2/(2mc^2)) \left[ \int \mathbf{A}(\mathbf{r}') D_{\mathbf{r}, \mathbf{r}'} dV' \right]^2 + \mathbf{p}^2/(2m) + V(\mathbf{r}). \quad (97)$$

The significance of all the symbols is the same as in (30) with the abbreviation

$$D(\mathbf{r} - \mathbf{r}') = D_{\mathbf{r}, \mathbf{r}'}. \quad (97')$$

The linear term in  $\mathbf{A}$  we have symmetrized because  $D_{\mathbf{r}, \mathbf{r}'}$  and  $\mathbf{p}$  do not commute. The vector potential  $\mathbf{A}$  is still given by (8). The quadratic term in  $\mathbf{A}$  could also be symmetrized on introducing (8). This would lead, however, to a complex expression which appears to have no special advantage.

It should be emphasized at this point that we suppose the matter variables  $\mathbf{r}$ ,  $\mathbf{p}$  to be canonically conjugate to each other but to commute with all of the radiation variables. The operator for the velocity is obtained as

$$dx/dt = (2\pi i/\hbar) [H, x] = (1/m) \left[ p_x - (e/c) \int A_x(\mathbf{r}') D_{\mathbf{r}, \mathbf{r}'} dV' \right]. \quad (98)$$

Similarly for the radiation variables

$$\left\{ \begin{array}{l} \frac{da_s}{dt} = 2\pi i \nu_s a_s - (e/2c) \int \left\{ \dot{\mathbf{r}} D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}(\mathbf{r}')}{\partial a_s^+} + D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}(\mathbf{r}')}{\partial a_s^+} \dot{\mathbf{r}} \right\} dV' \\ \frac{da_s^+}{dt} = -2\pi i \nu_s a_s^+ + (e/2c) \int \left\{ \dot{\mathbf{r}} D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}(\mathbf{r}')}{\partial a_s} + D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}(\mathbf{r}')}{\partial a_s} \dot{\mathbf{r}} \right\} dV' \\ \frac{da_k}{dt} = 4\pi \nu_k b_k^*; \quad \frac{da_k^*}{dt} = 4\pi \nu_k b_k^* \\ \frac{db_k}{dt} = \frac{e}{2c} \int \left\{ \dot{\mathbf{r}} D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}}{\partial a_k} + D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}}{\partial a_k} \dot{\mathbf{r}} \right\} dV' \\ \frac{db_k^*}{dt} = \frac{e}{2c} \int \left\{ \dot{\mathbf{r}} D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}}{\partial a_k^*} + D_{\mathbf{r}, \mathbf{r}'} \frac{\partial \mathbf{A}}{\partial a_k^*} \dot{\mathbf{r}} \right\} dV' \end{array} \right. \quad (99)$$

Since  $A$  is linear in  $a_s, a_s^+, a_k$  etc. the derivatives  $\partial A/\partial a_s$ , etc., commute with all the other quantities. The only difference between these equations and the corresponding equations derivable from (30) [See (13), (17), (18)] is that the symmetrized combination  $(1/2) [\dot{\mathbf{r}}D + D\dot{\mathbf{r}}]$  occurs everywhere instead of  $\dot{\mathbf{r}}D$ . We conclude that

$$-\frac{\partial A(\mathbf{r}')}{c\partial t} = \mathcal{E}(\mathbf{r}') = - \sum_s 2\pi i \nu^{1/2} V^{-1/2} \mathbf{f}_s (a_s e^{-i\mathbf{k}\mathbf{r}'} - i a_s^+ e^{i\mathbf{k}\mathbf{r}'}) - \sum_k 4\pi (\nu/V)^{1/2} (\mathbf{k}/k) (b_k^* e^{-i\mathbf{k}\mathbf{r}'} + b_k e^{i\mathbf{k}\mathbf{r}'}) \quad (14)$$

and

$$\mathcal{H} = \sum_s \frac{c}{(V\nu)^{1/2}} [\mathbf{k} \times \mathbf{f}_s] (-i a_s e^{-i\mathbf{k}\mathbf{r}} - a_s^+ e^{i\mathbf{k}\mathbf{r}}) \quad (5)$$

are still valid. Also

$$\text{rot } \mathcal{H} = \partial \mathcal{E}/c\partial t + 4\pi e \mathbf{j}/c \quad (19)$$

holds provided we let

$$\mathbf{j}(\mathbf{r}') = (\tfrac{1}{2})(D_{\mathbf{r},\mathbf{r}'}\dot{\mathbf{r}} + \dot{\mathbf{r}}D_{\mathbf{r},\mathbf{r}'}) \quad (100)$$

which replaces (34'). The derivation of (19) in the quantum theory is exactly the same as in the classical theory equation (99) being used for the computation of  $\partial \mathcal{E}/\partial t$  instead of (17) and (18). Formula (21) for  $\text{div } \mathcal{E}$  also is satisfied identically in the quantum theory. The last two equations (99) prove then also that

$$(\partial/\partial t) \text{div } \mathcal{E} = -4\pi e \text{div } \mathbf{j}$$

as in (23). The identification of  $-\text{div } \mathbf{j}$  with  $\partial \rho/\partial t$  requires a small additional consideration. Let us agree to mean by  $\rho$  the operator

$$\rho = D_{\mathbf{r},\mathbf{r}'}.$$

By the operator  $\partial \rho/\partial t$  we should mean, as in the classical treatment, the time derivative of  $\rho$  if  $\mathbf{r}'$  is kept fixed. This is obtained by the formula

$$\partial \rho/\partial t = (2\pi i/\hbar) [H, \rho].$$

The only dynamical variables involved in  $\rho$  are the coordinates of the electron. For any power of  $x$  we find with  $\dot{x} = (2\pi i/\hbar) [H, x]$  so that

$$dx^n/dt = x^{n-1}\dot{x} + x^{n-2}\dot{x}x + \dots + \dot{x}x^{n-1}.$$

If  $[\dot{x}, x]$  is a  $c$  number this expression simplifies into

$$\frac{dx^n}{dt} = (n/2)(x^{n-1}\dot{x} + \dot{x}x^{n-1}).$$

By (98)  $[\dot{x}, x]$  is a  $c$  number because  $D_{\mathbf{r},\mathbf{r}'}$  commutes with  $x$ . Thus

$$\frac{\partial \rho}{\partial t} = (\tfrac{1}{2}) \sum_{x,y,z} \left( \dot{x} \frac{\partial D_{\mathbf{r},\mathbf{r}'}}{\partial x} + \frac{\partial D_{\mathbf{r},\mathbf{r}'}}{\partial x} \dot{x} \right) = -(\tfrac{1}{2}) \sum_{x',y',z'} \left( \dot{x} \frac{\partial D_{\mathbf{r},\mathbf{r}'}}{\partial x'} + \frac{\partial D_{\mathbf{r},\mathbf{r}'}}{\partial x'} \dot{x} \right)$$



where the derivatives are ordinary classical derivatives. Comparing this with (100) we see that

$$\operatorname{div} j(\mathbf{r}') = \sum_{x'y'z'} \frac{\partial j(\mathbf{r}')}{\partial x'} = - \frac{\partial \rho}{\partial t}.$$

The Eq. (23) is therefore also correct between operators in the form

$$(\partial/\partial t)(\operatorname{div} \mathcal{E} - 4\pi e\rho) = 0. \quad (23)$$

For the proof of this equation it was important to use the symmetrized  $j$  of (100) and therefore also the symmetrized linear term in  $\mathbf{A}$  in (97). Just as in the classical theory equation (23) can be restated by saying that  $\operatorname{div} \mathcal{E} - 4\pi e\rho$  for a given point  $\mathbf{r}'$  is an integral of the motion. It suffices to start with the initial condition

$$\operatorname{div} \mathcal{E} - 4\pi e\rho = 0 \quad (24)$$

in order to have it maintained as a consequence of the equations of motion. The equation  $\operatorname{div} \mathcal{H} = 0$  is, of course, satisfied identically in virtue of the definition  $\mathcal{H} = \operatorname{rot} \mathbf{A}$ . Thus all of Maxwell's equations hold for the operators of the quantum theory.

The classical expression for  $d^2\mathbf{r}/dt^2$  contains the field quantities and matter quantities in the form of products. By (98) the velocity operator contains the vector potentials  $\mathbf{A}$ . It does not commute, with the magnetic field  $\mathcal{H}$ . We should be prepared to find, therefore, that instead of the classical  $[\mathbf{v} \times \mathcal{H}]$  term in the expression for the force a symmetrized expression will occur. By exchanging  $m d\dot{x}/dt$  with  $H$  we obtain the operator for  $m d^2x/dt^2$ . We have

$$\begin{aligned} \frac{2\pi i}{h} [H, p_x] &= - \frac{\partial V}{\partial x} - \frac{e}{2mc} \int \left( \mathbf{p} \frac{\partial D}{\partial x'} + \frac{\partial D}{\partial x'} \mathbf{p} \right) \mathbf{A}(\mathbf{r}') dV' \\ &+ \frac{e^2}{2mc^2} \left\{ \int \mathbf{A}(\mathbf{r}') D dV' \cdot \int \mathbf{A}(\mathbf{r}') \frac{\partial D}{\partial x'} dV' + \int \mathbf{A}(\mathbf{r}') \frac{\partial D}{\partial x'} dV' \cdot \int \mathbf{A}(\mathbf{r}') D dV' \right\} \\ \frac{2\pi i}{h} \left[ H, - (e/c) \int A_x(\mathbf{r}') D dV' \right] \\ &= - (e/c) \int \frac{dA_x(\mathbf{r}')}{dt} D dV' - (e/c) \int A_x(\mathbf{r}') \frac{dD}{dt} dV'. \end{aligned}$$

The index  $\mathbf{r}, \mathbf{r}'$  in  $D$  being understood. The calculation can be simplified by using the fact that  $\mathbf{A}$  and  $d\mathbf{A}/dt$  commute with  $D$ . It follows that

$$[\mathbf{A}(\mathbf{r}'), dD_{\mathbf{r},\mathbf{r}'}]/dt = 0.$$

Hence

$$\begin{aligned} m \frac{d^2x}{dt^2} &= - \frac{\partial V}{\partial x} + e \int \mathcal{E}_x(\mathbf{r}') D dV' \\ &- \frac{e}{2c} \int \left\{ \dot{\mathbf{r}} \frac{\partial D}{\partial x'} \mathbf{A}(\mathbf{r}') + \mathbf{A}(\mathbf{r}') \frac{\partial D}{\partial x'} \dot{\mathbf{r}} + A_x(\mathbf{r}') \frac{dD}{dt} + \frac{dD}{dt} A_x(\mathbf{r}') \right\} dV'. \quad (100') \end{aligned}$$

For  $dD/dt$  we substitute

$$(\frac{1}{2}) \sum_{x,y,z} \left( \dot{x} \frac{\partial D}{\partial x} + \frac{\partial D}{\partial x} \dot{x} \right).$$

We then note that  $\partial D/\partial x'$ ,  $\partial D/\partial y'$ ,  $\partial D/\partial z'$  commute with  $A_x(\mathbf{r}')$ ,  $A_y(\mathbf{r}')$ ,  $A_z(\mathbf{r}')$  and that by (98) the only part of  $\dot{x}$  which does not commute with  $\mathbf{r}$  is  $p_x/m$ . Hence we have identically

$$\begin{aligned} & \dot{y} \frac{\partial D}{\partial y'} A_x(\mathbf{r}') + A_x(\mathbf{r}') \frac{\partial D}{\partial y'} \dot{y} - \frac{A_x(\mathbf{r}')}{2} \left[ \dot{y} \frac{\partial D}{\partial y'} + \frac{\partial D}{\partial y'} \dot{y} \right] \\ & - (1/2) \left[ \dot{y} \frac{\partial D}{\partial y'} + \frac{\partial D}{\partial y'} \dot{y} \right] A_x(\mathbf{r}') \\ & = \frac{1}{2} \left\{ \left[ \dot{y}, \frac{\partial D}{\partial y'} \right] A_x(\mathbf{r}') - A_x(\mathbf{r}') \left[ \dot{y}, \frac{\partial D}{\partial y'} \right] \right\} = \frac{h}{4\pi i m} \left[ \frac{\partial^2 D}{\partial y'^2}, A_x(\mathbf{r}') \right] = 0 \end{aligned}$$

and similarly for all other expressions obtained from this by changing  $\dot{y}$  into  $\dot{x}$  or  $\dot{z}$  or by changing  $A_x$  into  $A_y$  etc. Using these identities and transferring the  $\partial/\partial x'$ ,  $\partial/\partial y'$ ,  $\partial/\partial z'$  operations by partial integration to the  $A_x$ ,  $A_y$ ,  $A_z$  we have

$$\begin{aligned} m \frac{d^2 x}{dt^2} = & - \frac{\partial V}{\partial x} + e \int \left\{ \mathcal{E}_x(\mathbf{r}') D + \frac{1}{2c} [j_y(\mathbf{r}') \mathcal{H}_z(\mathbf{r}') + \mathcal{H}_z(\mathbf{r}') j_y(\mathbf{r}') \right. \\ & \left. - j_z(\mathbf{r}') \mathcal{H}_y(\mathbf{r}') - \mathcal{H}_y(\mathbf{r}') j_z(\mathbf{r}') \right\} dV' \end{aligned}$$

or

$$\begin{aligned} m \frac{d^2 \mathbf{r}}{dt^2} = & - \text{grad } V + e \int \left\{ \mathcal{E}(\mathbf{r}') D_{\mathbf{r}, \mathbf{r}'} \right. \\ & \left. + \frac{1}{2c} [j(\mathbf{r}') \times \mathcal{H}(\mathbf{r}')] - \frac{1}{2c} [\mathcal{H}(\mathbf{r}') \times j(\mathbf{r}')] \right\} dV' \quad (101) \end{aligned}$$

with the significance (100) of  $j$ . Another symmetrized form of the  $x$  component of the magnetic force expression equivalent to the above (and obtainable from (100')) is

$$\frac{e}{2c} \int \{ \dot{y} D \mathcal{H}_z + \mathcal{H}_z D \dot{y} - \dot{z} D \mathcal{H}_y - \mathcal{H}_y D \dot{z} \} dV'. \quad (101')$$

In the above discussion we have confined ourselves to one particle. There is no difficulty in extending the theory to several particles which may have equal or unequal masses and equal or unequal charges. The Hamiltonian function (97) can then be generalized into

$$\begin{aligned} H = & H_R - \sum_i \frac{e_i}{2m_i c} \int (\mathbf{p}_i D_{\mathbf{r}, \mathbf{r}'}^i + D_{\mathbf{r}, \mathbf{r}'}^i \mathbf{p}_i) A(\mathbf{r}') dV' \\ & + \sum_i \frac{e_i^2}{2m_i c^2} \left[ \int A(\mathbf{r}') D_{\mathbf{r}, \mathbf{r}'}^i dV' \right]^2 \\ & + \sum_i \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) \end{aligned} \quad (102)$$

where  $i$  refers to the individual particles and  $n$  is the number of particles. Instead of (98) we have:

$$m_i \frac{dx_i}{dt} = p_{iz} - (e_i/c) \int A_x(\mathbf{r}') D_{zi, \mathbf{r}, \mathbf{r}'}^i dV'. \quad (98')$$

Analogously to (35) we derive from (98') that

$$H = H_R + \sum_i (m_i/2) \dot{\mathbf{r}}_i^2 + V. \quad (103)$$

Similarly for all the other quantities just discussed we derive formulas identical in form with the corresponding classical formulas, the only exception being the force equation (101) in which the symmetrized vector product of  $\dot{\mathbf{r}}$  and  $\mathcal{H}$  must be used just as in the case of a single particle. In particular we have

$$\frac{\partial}{\partial t} \left( \text{div } \mathcal{E} - 4\pi \sum_i e_i \rho_i \right) = 0. \quad (104)$$

### §3. Elimination of electrostatic field

In the classical theory presented in section (I) the electrostatic field determined by  $\text{div } \mathcal{E} = 4\pi e\rho$  and  $\text{rot } \mathcal{E} = 0$ , i.e., by the second sum in (14) may be eliminated from the equations without any difficulty. To do so it is only necessary to observe that using (6) and (1) as well as the Hamiltonian (30) Eqs. (18) follow. Eq. (31) for  $m\dot{\mathbf{r}}$  preserves its form though  $\mathbf{A}$  in it is (1) instead of (18). The force equation (33) is correct so far as the  $[\dot{\mathbf{r}} \times \mathcal{H}]$  part is concerned but needs a correction for the effect of  $\mathcal{E}$ . This correction can be made by introducing terms  $e_1 e_2 / r_{12}$  into the Hamiltonian for any two particles at a distance  $r_{12}$  from each other. With this change we do not have to calculate the variables  $a_k, a_k^*$  and their canonical conjugates  $b_k, b_k^*$  any more. It suffices to compute that part of  $\mathbf{A}, \mathcal{E}, \mathcal{H}$  which may be represented by transverse waves. The electrostatic field can then be also computed at any time from  $\text{div } \mathcal{E}' = 4\pi e\rho$  and  $\text{curl } \mathcal{E}' = 0$ . It should be realized that the Hamiltonian so obtained is different from the value of the original Hamiltonian. Their difference is the electrostatic self energy of the charge distribution  $\sum e_i \rho_i$ . If desired this energy may be added to the new Hamiltonian. Such an addition does not change any of the results. It may be performed as long as the functions  $D_{\mathbf{r}_i^i, \mathbf{r}'}^i$  are not infinitely concentrated. In the limit of infinite concentration of the  $D$  functions the self energy becomes infinite.

On the quantum theory the elimination of the electrostatic field variables can be also performed and leads to very similar results. Eq. (104) means that when the operator  $\text{div } \mathcal{E} - 4\pi \sum_i e_i \rho_i$  is applied to any wave function the result vanishes. By (21)

$$\text{div } \mathcal{E}(\mathbf{r}') = \sum_k' 4\pi i (\nu/V)^{1/2} k (b_k^* e^{-i\mathbf{k}\mathbf{r}'} - b_k e^{i\mathbf{k}\mathbf{r}'}).$$

Writing for the present

$$\rho_i(\mathbf{r}') = \sum_k' i (\nu/V)^{1/2} k (\rho_{ik}^* e^{-i\mathbf{k}\mathbf{r}'} - \rho_{ik} e^{i\mathbf{k}\mathbf{r}'}) \quad (104')$$

we have for any  $k$

$$\left(b_k - \sum_i e_i \rho_{ik}\right) \psi = 0, \left(b_k^* - \sum_i e_i \rho_{ik}^*\right) \psi = 0. \quad (105)$$

Since  $\rho_i(\mathbf{r}')$  is an operator involving the coordinates of the  $i$ th particle, the  $\rho_{ik}$  and the  $\rho_{ik}^*$  also are operators in the  $\mathbf{r}_i$ . The Eqs. (105) are satisfied if we refer  $\psi$  to the variables  $a_k, a_k^*, \mathbf{r}_i$  and let

$$\psi = \phi(\mathbf{r}_i, J_{k,i}) \exp \left\{ \frac{2\pi i}{h} \sum_k \sum_i e_i (\rho_{ik}^* a_k^* + \rho_{ik} a_k) \right\}. \quad (106)$$

In fact substituting (106) into (105) the latter is satisfied identically since when we refer  $\psi$  to  $a_k$  the operator  $b_k$  is  $(h/2\pi i) (\partial/\partial a_k)$ . We substitute (106) into the Schroedinger equation<sup>25</sup> with the Hamiltonian in the form (103). We perform the operations on the exponential factor in (106) and obtain an equation involving only  $\phi$ . It should be noted that in (106) the  $\rho_{ik}, a_k$  are  $c$  numbers. The operation  $\partial/\partial t$  in the Schroedinger equation therefore applies only to  $\phi$ . By (15)  $H_R$  contains the operator  $\sum' 4\pi\nu_k b_k^* b_k$ . This when applied to  $\psi$  operates only on the exponential factor and contributes the term

$$\sum_k' \left( \sum_i e_i \rho_{ik}^* \right) \left( \sum_i e_i \rho_{ik} \right) 4\pi\nu_k. \quad (106')$$

We note that the expression (106') is exactly what we would have obtained if we had performed a classical calculation for  $\sum' 4\pi\nu_k b_k^* b_k$  using a classical Eq. (104'). Thus (106') is simply the classical expression for the electrostatic energy of the charge distributions  $e_i \rho_i$ .

The exponential factor in (106) is transparent to the operator  $V$ . The operators  $m_i dx_i/dt$  which also occur in (103) operate by (98') on the  $\rho_{ik}$ . Each  $m_i dx_i/dt$  applied to (106) gives in addition also a term in  $a_k, a_k^*$  due to the occurrence of these variables in  $A$ . These two terms can be combined to give zero. In fact they are

$$\begin{aligned} & \sum_k e_i \left( \frac{\partial \rho_{ik}^*}{\partial x_i} a_k^* + \frac{\partial \rho_{ik}}{\partial x_i} a_k \right) \\ & - e_i \int (\nu_k/V)^{-1/2} (k_x/k) (a_k e^{-i\mathbf{k}\mathbf{r}} + a_k^* e^{i\mathbf{k}\mathbf{r}}) D_{\mathbf{r}\mathbf{r}'}^i dV'. \end{aligned} \quad (107)$$

We invert (104') and obtain

$$\begin{aligned} i\nu^{1/2} k \rho_{ik}^* &= V^{-1/2} \int e^{i\mathbf{k}\mathbf{r}'} \rho_i(\mathbf{r}') dx' dy' dz'; \\ -i\nu^{1/2} k \rho_{ik} &= V^{-1/2} \int e^{-i\mathbf{k}\mathbf{r}'} \rho_i(\mathbf{r}') dx' dy' dz' \end{aligned}$$

<sup>25</sup> J. R. Oppenheimer, Phys. Rev. **35**, 461 (1930); reference 1, II. The gauge invariance used by these authors is also made use of in the present discussion. On account of this invariance expression (107) below vanishes.

whence differentiating, and remembering that  $\rho_i(\mathbf{r}') = D_{\mathbf{r}_i \mathbf{r}'}^i$ , so that  $\partial \rho_i(\mathbf{r}') / \partial x' = -\partial \rho_i / \partial x_i$  and applying partial integration we find that the above expression (107) disappears. The operator  $m_i dx_i/dt$  contains besides the variables  $a_{k,l}, a_{k,l}^+$ . These produce no effect on the exponential factor in (106). Thus instead of applying  $m_i dx_i/dt$  to  $\psi$  we may equally well omit the  $a_k a_k^*$  in  $\mathbf{A}$ , keep only the terms of Eq. (1), and operate with this remainder of  $m_i dx_i/dt$  on  $\phi$ .

The result is that the Schroedinger equation in  $\psi$  reduces itself to a similar Schroedinger equation in  $\phi$ . Only instead of  $\mathbf{A}$  as given by (8) we must use  $\mathbf{A}$  as given by (1) and to the potential energy  $V$  we must add the classical expression for the electrostatic energy of the charge distribution  $\rho_i e_i$ . We may use therefore the same form of the Hamiltonian function as we have obtained on the classical theory by eliminating the variables of the electrostatic field. As in the classical theory the electrostatic energy of a  $e_i \rho_i$  distribution with itself is a constant and may be combined with the additive constant in the energy expression. Its omission affects only the value of the total energy. The mutual energy of  $e_i \rho_i$  with  $e_j \rho_j$  ( $i \neq j$ ) is the ordinary term  $e_i e_j / r_{ij}$  where  $r_{ij}$  is the distance between the centers of the concentrated distributions  $\rho_i$  and  $\rho_j$ .

It should be noted that it does not follow from the above discussion that  $\phi$  may be used instead of  $\psi$  for all purposes. However, it follows from (106) and the above discussion that this may be done for the calculation of quantum means of functions of  $a_{k,l}, a_{k,l}^+, \mathbf{r}_i, d\mathbf{r}_i/dt$ . For the calculation of such quantities we may also use other wave functions obtained from  $\phi$  by a contact transformation provided this transformation does not involve the variables  $a_k, a_k^*, b_k, b_k^*$ . In particular we may use a contact transformation to the action and angle variables of the system of particles  $e_i$  subject to mutual energies  $e_i e_j / r_{ij}$ . The  $\phi$  function may be then referred also to these action variables. For such purposes we may therefore regard (103) with  $a_k, a_k^*$  omitted and with  $e_i e_j / r_{ij}$  added as the Hamiltonian function operator. In practical applications it will be convenient to incorporate  $\sum e_i e_j / r_{ij}$  in  $V$ .

#### §4. Probability amplitudes for numbers of light quanta

Having eliminated the electrostatic field we have reduced the problem to that of an atom coupled to a system of oscillators each oscillator being associated with the variables  $a_{k,l}, a_{k,l}^+$ . We use the canonical transformation (93) and introduce  $N_{k,l} \equiv N_s$  by (94). The variable  $N_s$  may be interpreted as the number of light quanta travelling in the direction  $\mathbf{k}$ , having a frequency  $\nu_k$  given by  $4\pi^2 \nu_k^2 / c^2 = k^2$ , and a state of polarization (for the present linear) described by  $l$ , the direction of the electric vector being  $\mathbf{f}_s \equiv \mathbf{f}_{k,l}$ . We use the wave function  $\phi$  and have for the Schroedinger equation:

$$\begin{aligned} \frac{\hbar}{2\pi i} \frac{\partial \phi}{\partial t} + \left\{ \sum_s 2\pi i \nu_s a_s a_s^+ + \sum_i \frac{p_i^2}{2m_i} + V \right. \\ \left. - \sum_i \frac{e_i}{2m_i c} \int (\mathbf{p}_i D^i + D^i \mathbf{p}_i) \sum_s \mathbf{f}_s \frac{c}{(V \nu_s)^{1/2}} [a_s e^{-i\mathbf{k}\mathbf{r}'} + i a_s^+ e^{i\mathbf{k}\mathbf{r}'}] dV' \right\} \phi = 0 \quad (108) \end{aligned}$$

$$+ \sum_s \frac{e_i^2}{2m_i c^2} \left[ \int D^i \sum_s f_s \frac{c}{(V_{\nu_s})^{1/2}} [a_s e^{-ikr'} + i a_s^+ e^{ikr'}] dV' \right]^2 \} \phi = 0.$$

We let  $\phi$  be a function of  $N_s$  and of the stationary state quantum numbers  $n$  of the particles  $e_i$  unperturbed by the radiation. The matrix elements of  $e^{\mp 2\pi i w_s}$  must be diagonal in all  $N_{s'}$  for  $(k', l') \neq (k, l)$  and also diagonal in  $n$ . We satisfy all commutation relations by having all but the following matrix elements vanish.

$$(e^{-2\pi i w_s})_{N_s, N_s+1} = 1; (e^{2\pi i w_s})_{N_s, N_s-1} = 1.$$

Remembering (93) the matrices  $a_s, a_s^+$  also are diagonal in  $N_{s'}$ ,  $\mathbf{r}_i$  and have for nonvanishing matrix elements only

$$(a_s^+)_{N_s, N_s+1} = -i(h/2\pi)^{1/2}(N_s + 1)^{1/2}$$

$$(a_s)_{N_s, N_s-1} = (h/2\pi)^{1/2}N_s^{1/2}.$$

The wave function  $\phi$  contains all the  $N_s$  as variables. Using the customary notation

$$\sum_m Y_{nm} \psi_m = (Y\psi)_n$$

where  $Y$  is any operator. We have in our case

$$(a_r^+ \phi)_{n; N_1, N_2, \dots, N_r, \dots} = -i(h/2\pi)^{1/2}(N_r + 1)^{1/2} \phi_{n; N_1, N_2, \dots, N_r+1, \dots}$$

$$(a_r \phi)_{n; N_1, N_2, \dots, N_r, \dots} = (h/2\pi)^{1/2} N_r^{1/2} \phi_{n; N_1, N_2, \dots, N_r-1, \dots}$$

[cf. (65), (67) Section III] It follows from the above formulae that [we write now the suffixes in parentheses alongside the letter to which the suffixes refer]

$$\left( \left( a_r^+ a_r - \frac{h(N_r + 1)}{2\pi i} \right) \phi \right) (n; N_1, N_2, \dots, N_r, \dots)$$

$$= \left( \left( a_r a_r^+ - \frac{hN_r}{2\pi i} \right) \phi \right) (n; N_1, \dots, N_r, \dots) = 0$$

$$(a_r^+ a_r^+ \phi) (n; N_1, N_2, \dots, N_r, \dots)$$

$$= - (h/2\pi) (N_r + 1)^{1/2} (N_r + 2)^{1/2} \phi (n; N_1, \dots, N_r + 2, \dots)$$

$$(a_r a_r \phi) (n; N_1, N_2, \dots, N_r, \dots)$$

$$= (h/2\pi) N_r^{1/2} (N_r - 1)^{1/2} \phi (n; N_1, \dots, N_r - 2, \dots)$$

and for  $r \neq s$

$$(a_r^+ a_s \phi) (n; N_1, N_r, \dots, N_s, \dots) = (a_s a_r^+ \phi) (n; N_1, \dots, N_r, \dots, N_s, \dots)$$

$$= (h/2\pi i) (N_r + 1)^{1/2} N_s^{1/2} \phi (n; N_1, \dots, N_r + 1, \dots, N_s - 1, \dots)$$

$$(a_r^+ a_s^+ \phi) (n; N_1, \dots, N_r, \dots, N_s, \dots)$$

$$= - (h/2\pi) (N_r + 1)^{1/2} (N_s + 1)^{1/2} \phi (n; N_1, \dots, N_r + 1, \dots, N_s + 1, \dots)$$

$$(a_r a_s \phi) (n; N_1, \dots, N_r, \dots, N_s, \dots)$$

$$= + (h/2\pi) N_r^{1/2} N_s^{1/2} \phi (n; N_1, \dots, N_r - 1, \dots, N_s - 1, \dots)$$

We let

$$\left\{ \begin{aligned} \left[ \sum_i (e_i/2m_i) \int (\mathbf{p}_i D_{\mathbf{r}_i, \mathbf{r}'}^i + D_{\mathbf{r}_i, \mathbf{r}'}^i \mathbf{p}_i) \mathbf{f}_s e^{-i\mathbf{k}\mathbf{r}'} dV' \right]_{n', n''} &= (2\pi\hbar)^{1/2} \alpha_{n', n''}^s \\ \left[ \sum_i (e_i/2m_i) \int (\mathbf{p}_i D_{\mathbf{r}_i, \mathbf{r}'}^i + D_{\mathbf{r}_i, \mathbf{r}'}^i \mathbf{p}_i) \mathbf{f}_s e^{i\mathbf{k}\mathbf{r}'} dV' \right]_{n', n''} &= (2\pi\hbar)^{1/2} (\alpha_{n'', n'}^s)^* \\ \sum_i \frac{e_i^2}{m_i} \left[ \int D_{\mathbf{r}_i, \mathbf{r}'}^i D_{\mathbf{r}_i, \mathbf{r}''}^i e^{-i(\mathbf{k}'\mathbf{r}'+\mathbf{k}''\mathbf{r}'')} dV' dV'' \right]_{n', n''} &= \beta_{n', n''}^{k', k''} \end{aligned} \right. \quad (109)$$

$$\beta_{n', n''}^{k', k''} = \beta_{n'', n'}^{k'', k'}; (\beta_{n', n''}^{k', k''})^* = \beta_{n'', n'}^{-k', -k''}$$

Eq. (108) becomes

$$\begin{aligned} & \left( \frac{\hbar}{2\pi i} \frac{\partial}{\partial t} + E_n + \sum_s N_s \hbar \nu_s \right) \phi(n; N_1, N_2, \dots) \\ & - \sum_{n', s} \hbar (\nu_s V)^{-1/2} (N_s^{1/2} \alpha_{nn'}^s \phi(n'; \dots N_s - 1, \dots) \\ & + (N_s + 1)^{1/2} (\alpha_{n', n}^s)^* \phi(n'; \dots N_s + 1, \dots)) \\ & + \sum_{n', s' \neq s''} (h/4\pi V) (\nu_s \nu_{s'})^{-1/2} (\mathbf{f}_s \mathbf{f}_{s'}) \\ & \quad \{ \beta_{nn'}^{k'k''} (N_{s'} N_{s''})^{1/2} \phi(n'; \dots, N_{s'} - 1, \dots, N_{s''} - 1, \dots) \\ & + \beta_{n, n'}^{k', -k''} (N_{s'} + 1)^{1/2} (N_{s''} + 1)^{1/2} \phi(n'; \dots N_{s'} + 1, N_{s''} + 1, \dots) \\ & + \beta_{n, n'}^{k', -k''} N_{s'}^{1/2} (N_{s''} + 1)^{1/2} \phi(n'; N_{s'} - 1, N_{s''} + 1, \dots) \\ & + \beta_{n, n'}^{-k', k''} (N_{s'} + 1)^{1/2} N_{s''}^{1/2} \phi(n'; \dots N_{s'} + 1, N_{s''} - 1) \} \\ & + \sum_{n', s'} (h/4\pi V) \nu_{s'}^{-1} \{ \beta_{n, n'}^{k', k'} N_{s'}^{1/2} (N_{s'} - 1)^{1/2} \phi(n'; N_{s'} - 2, \dots) \\ & + \beta_{n, n'}^{-k', -k'} (N_{s'} + 1)^{1/2} (N_{s''} + 2)^{1/2} \phi(n'; N_{s'} + 2, \dots) \\ & + \beta_{n, n'}^{-k', k'} (2N_{s'} + 1) \phi(n; N_{s'}, \dots) \} = 0. \end{aligned} \quad (110)$$

Here  $s, s', s''$  are written for  $(k, l), (k', l'), (k'', l'')$  respectively.

If an  $N_{s''}$  is not indicated in the Schroedinger function in the above equation it is understood that it has the same value as in the first term. For given initial values of  $\phi(n; N_1, N_2, \dots)$  this equation determines the  $\phi(n; N_1, N_2, \dots)$  for all later times. The terms in  $\alpha_s$  arise from the linear term in  $\mathbf{A}$  while those in  $\beta_{n, n'}^{k'k''}$  arise from  $\mathbf{A}^2$ .

### §5. Spontaneous emission

An atom in an excited state is known to radiate spontaneously. It has already been explained in §1 above that this phenomenon may be understood as a transfer of energy from the atom to the radiation oscillators. In this § we calculate the probability of this transfer by means of the equations of §4. Let the atom be initially ( $t=0$ ) in a definite stationary state  $n_0$  and let us suppose that there are no light quanta at that time. The probability amplitude

for this state of the coupled system of matter and radiation is  $\phi(n_0; 0, 0, \dots)$  so that initially

$$|\phi(n_0; 0, 0, \dots)| = 1 \quad (t = 0)$$

while all other probability amplitudes are zero at  $t=0$ . We wish to find how these probability amplitudes increase.

In this section we *suppose* that the *emission takes place slowly* in comparison with the emission frequencies and furthermore we confine ourselves to the discussion of *times short in comparison with the mean life*. A more complete discussion follows in VI §2. For such times  $\phi(n_0; 0, 0, \dots)$  is larger than any other  $\phi$ . To a first approximation (110) gives then

$$\left( \frac{\hbar}{2\pi i} \frac{\partial}{\partial t} + E_{n_0} \right) \phi(n_0; 0, 0, \dots) = 0$$

so that omitting an arbitrary constant multiplier of absolute value 1

$$\phi(n_0; 0, 0, \dots) = e^{-2\pi i \nu_{n_0} t}.$$

Using this value, (110) may be used to compute the probability amplitude for the atom in the state  $n$  and for a light quantum of type  $s$

$$\phi(n; 0, 0, \dots, 1_s, 0, \dots) \equiv a_{ns}.$$

Supposing that the  $a_{ns}$  are larger than the probability amplitudes for states with two light quanta:  $\phi(n; 0, \dots, 1_{s'}, 0, \dots, 1_{s''}, \dots, 0, \dots)$   $\phi(n; 0, \dots, 2_s, 0, \dots)$ . (See §7 below.) The equation for  $a_{ns}$  is

$$\left( \frac{\hbar}{2\pi i} \frac{d}{dt} + E_n + h\nu_s \right) a_{ns} = h(\nu_s V)^{-1/2} \alpha_{nn_0}^s e^{-2\pi i \nu_{n_0} t}.$$

The solution of this equation subject to the initial condition  $(a_{ns})_{t=0} = 0$  is

$$a_{ns} = (\nu_s V)^{-1/2} \alpha_{nn_0}^s (e^{-2\pi i \nu_{n_0} t} - e^{-2\pi i (\nu_n + \nu_s) t}) / (\nu_s - \nu_{n_0}). \quad (111)$$

By definition  $a_{ns}$  is the probability amplitude for the atom to be in state  $n$  and for one quantum to be present in the form of a progressive wave with direction of propagation and with polarization designated by  $s$ . By (111) the  $a_{ns}$  is largest if  $\nu_s$  approaches  $\nu_{n_0}$ . Thus quanta with frequencies close to  $\nu_{n_0}$  are emitted with more probability than other quanta. The square of the modulus of  $a_{ns}$  is the probability of a quantum  $h\nu_s$  having been emitted while the atom jumped from  $n_0$  to  $n$ . This is

$$|a_{ns}|^2 = [4 (\sin^2 \pi(\nu_s - \nu_{n_0})t) / (\nu_s - \nu_{n_0})^2] |\alpha_{nn_0}^s|^2 / \nu_s V. \quad (111')$$

By (109)  $\alpha_{nn_0}^s$  depends primarily on the direction of  $\mathbf{k}$ . If the wave-length is long in comparison with the dimensions of the atom the  $|\alpha_{nn_0}^s|^2$  varies only slowly with  $\nu_s$ . On the other hand the larger  $t$  in (111') the more sharply does  $(\sin^2 \pi(\nu_s - \nu_{n_0})t) / (\nu_s - \nu_{n_0})^2$  depend on  $\nu_s - \nu_{n_0}$ . If  $|\nu_s - \nu_{n_0}|$  becomes several times  $1/t$  the value of this fraction is quite negligible in comparison with its value for  $\nu_s = \nu_{n_0}$ . We may therefore sum the values of this fraction as



though the factor in front of it were constant and then we can substitute for this factor its value for  $\nu_s = \nu_{n_0 n}$  having averaged  $|\alpha_{s n n_0}|^2$  over all directions  $k$  and states of polarization  $l$ . We have

$$\sum_s \frac{4 \sin^2 \pi(\nu_s - \nu_{n_0 n})t}{(\nu_s - \nu_{n_0 n})^2} = \frac{1}{(\Delta\nu)} \int_{-\infty}^{+\infty} \frac{4 \sin^2 \pi x t}{x^2} dx$$

where  $1/(\Delta\nu)$  is the number of possible distinct values of  $k, l$  per unit frequency range. This number is

$$(1/\Delta\nu) = (8\pi\nu^2/c^3)V.$$

On performing the integration the summation becomes

$$(32\pi^3\nu^2_{n n_0}/c^3)Vt.$$

For very concentrated  $D_{\mathbf{r}_i, \mathbf{r}'_i}$  and for wave-lengths great in comparison with atomic dimensions

$$(2\pi\hbar)^{1/2}\alpha_{s n, n_0} = e^{-ikR} \sum (e_i/m_i)(\mathbf{p}_i \mathbf{f}_s)_{n n_0} \quad (111'')$$

where  $\mathbf{R}$  is the position of the atom. In the present application no confusion will be caused if we write  $\mathbf{p}_i/m_i = \dot{\mathbf{x}}_i$  because we deal now with matrix elements of the atom in its own reference system. In averaging the square of the modulus of the above quantity the  $e^{-ikR}$  disappears. The  $\mathbf{f}_s$  becomes a random direction. We suppose, for simplicity that the states  $n, n_0$  are nondegenerate. The averaging over the directions of  $\mathbf{f}_s$  introduces a factor  $1/3$ . The result is

$$\sum_s |a_{ns}|^2 = \frac{16\pi^2\nu_{n_0 n}}{3\hbar c^3} \left| \left( \sum e_i \dot{\mathbf{x}}_i \right)_{n_0 n} \right|^2 = \frac{64\pi^4\nu^3_{n_0 n}}{3\hbar c^3} \left| \left( \sum e_i \mathbf{r}_i \right)_{n_0 n} \right|^2. \quad (112)$$

If the  $e_i$  are all equal this is the same as (47). The essential restriction in the applicability of these formulas is that the wavelength should be great in comparison with the dimensions of the atom. The step in the calculation at which this restriction was introduced was the use of the approximate value of  $\alpha_{s n n_0}$  with  $e^{-ikr'}$  taken outside of the integral sign. In later applications we will have occasion to use the more accurate form of  $\alpha_{s n, n_0}$  given by (109).

The integrals used in deriving Eq. (112) do not converge if  $1/\Delta\nu$  is not treated as a constant. This question will be treated in VIII below.

## §6. Absorption and emission

More generally consider an initial state in which there are  $N_s^0$  quanta of the kind  $s$  and in which the atom is in the state  $n_0$ . Neglecting for the present the terms in  $\beta$  (their effect will be considered later) we have for short times

$$\phi(n_0; N_1^0, N_2^0, \dots) = e^{-2\pi i(\nu_{n_0} + \sum_s N_s^0 \nu_s)t}$$

and all the other  $\phi$  are small compared to this. These smaller  $\phi$  are determined by:

$$\begin{aligned}
& \left( \frac{\partial}{2\pi i \partial t} + \nu_n + \sum_{s'} N_{s'}^0 \nu_{s'} + \nu_s \right) \phi(n; N_1^0, \dots, N_s^0 + 1, \dots) \\
& \quad = (\nu_s V)^{-1/2} (N_s^0 + 1)^{1/2} \alpha_{nn_0}^s e^{-2\pi i (\nu_{n_0} + \sum N_{s'}^0 \nu_{s'}) t} \\
& \left( \frac{\partial}{2\pi i \partial t} + \nu_n + \sum_{s'} N_{s'}^0 \nu_{s'} - \nu_s \right) \phi(n; N_1^0, \dots, N_s^0 - 1, \dots) \\
& \quad = (\nu_s V)^{-1/2} (N_s^0)^{1/2} (\alpha_{nn_0}^s)^* e^{-2\pi i (\nu_{n_0} + \sum N_{s'}^0 \nu_{s'}) t}
\end{aligned}$$

with the initial conditions that they should vanish at  $t=0$ . The solutions are

$$\begin{aligned}
& \phi(n; N_1^0, \dots, N_s^0 + 1, \dots) \\
& \quad = (\nu_s V)^{-1/2} (N_s^0 + 1)^{1/2} \alpha_{nn_0}^s \frac{e^{-2\pi i (\nu_{n_0} + \sum N_{s'}^0 \nu_{s'}) t} - e^{-2\pi i (\nu_n + \sum N_{s'}^0 \nu_{s'} + \nu_s) t}}{\nu_s - \nu_{n_0 n}} \\
& \phi(n; N_1^0, \dots, N_s^0 - 1, \dots) \\
& \quad = (\nu_s V)^{-1/2} (N_s^0)^{1/2} (\alpha_{nn_0}^s)^* \frac{e^{-2\pi i (\nu_{n_0} + \sum N_{s'}^0 \nu_{s'}) t} - e^{-2\pi i (\nu_n + \sum N_{s'}^0 \nu_{s'} - \nu_s) t}}{\nu_{nn_0} - \nu_s}
\end{aligned} \tag{113}$$

The square of the modulus of the first expression is interpreted as the probability that the atom passed into the state  $n$  with the emission of a quantum  $s$ . Similarly the second expression gives the probability of absorption. If the  $N_s^0$  vary uniformly with respect to  $s$ , as is the case for example in black body radiation, the summation of the squares of the moduli of expressions (113) is entirely similar to that performed in the calculation of spontaneous emission. We note in particular that there is a simple relation between the total probability of emission from  $n_0$  to  $n$  and the probability of absorption from  $n$  to  $n_0$ . To obtain the probability of the latter process we must start at  $t=0$  with the atom in the state  $n$  having again  $N_s^0$  quanta of kind  $s$ . Then

$$\begin{aligned}
& \phi(n_0; N_1^0, \dots, N_s^0 - 1, \dots) \\
& \quad = (\nu_s V)^{-1/2} N_s^{(0)1/2} (\alpha_{nn_0}^s)^* \frac{e^{-2\pi i (\nu_n + \sum N_{s'}^0 \nu_{s'}) t} - e^{-2\pi i (\nu_{n_0} + \sum N_{s'}^0 \nu_{s'} - \nu_s) t}}{\nu_{nn_0} - \nu_s}
\end{aligned} \tag{114}$$

The ratio of the probability of emission from  $n_0$  to  $n$  to the probability (114) of absorption from  $n$  to  $n_0$  in the same field of radiation is then

$$\frac{\sum_s |\phi(n; N_1^0, \dots, N_s^0 + 1, \dots)|^2}{\sum_s |\phi(n_0; N_1^0, \dots, N_s^0 - 1, \dots)|^2} = \frac{\overline{N_s^0} + 1}{\overline{N_s^0}} = \frac{\rho + (8\pi h \nu^3 / c^3)}{\rho} \tag{115}$$

where

$$\rho = \frac{8\pi \nu^2}{c^3} \overline{N_s^0} h \nu_s \tag{115'}$$

is the energy density per unit frequency range. It should be noted that the validity of (115) does not depend on having the wave-length of the radiation great in comparison with atomic dimensions. It depends on the other hand

on having  $N_s^0$  vary uniformly, or else statistically with  $s$ . It is also necessary for its validity that  $t$  be sufficiently large to secure the absorption and emission of light quanta for which  $\nu_s$  is very close to  $\nu_{n_0}$ . Otherwise  $\nu^3$  in (115) has no definite meaning.

Formula (115) may be interpreted by saying that the emission probability may be divided into the *spontaneous* and the *induced* emission probability. The latter is present only if  $\overline{N_s^0} \neq 0$ . Their ratio is  $1/\overline{N_s^0} = 8\pi h\nu^3/c^3\rho$ . The probability of absorption from  $n$  to  $n_0$  is just equal to the probability of induced emission from  $n_0$  to  $n$ . This is in exact agreement with the theory of Einstein.<sup>26</sup> The bearing of Dirac's theory on Einstein's speculation about the unidirectionality of light quanta will be given later (VII §3).

In the above discussion we do not consider contributions to  $\Sigma_s \phi(n; N_1^0 \dots, N_s^0 \pm 1, \dots)$  due to very high values of  $\nu_s$ . In a physically correct theory these should not be of importance. We shall return to this question later in the discussion of the divergence difficulties of the theory.

If the incident radiation does not have random directions the absorption probability does not have to be connected with the spontaneous emission by Einstein's relation. A brief review of the derivation of (112) and a comparison of the form of (111) with (113') shows that the probability of absorption is

$$(8\pi^3/h^2) |(\sum e i \mathbf{f}_s \mathbf{r}_i)_{nn_0}|^2 \rho \quad (116)$$

subject to the same restriction of long wave-length as (112). The bar over the square of the modulus indicates that the latter should be averaged over the polarization vectors  $\mathbf{f}_s$  of the incident radiation.

### §7. Scattering and dispersion

We have found above that if at  $t=0$ ,  $\phi$  is distinct from zero only for  $n=n_0; N_1, N_2, \dots = N_1^0, N_2^0, \dots$  then at a later time  $t$ ,  $\phi(n; N_1^0 \dots N_s^0 \pm 1, \dots)$  is also distinct from zero. We have used an approximate calculation. Clearly the next step in the approximation gives rise to the existence of  $\phi$  in which the number of light quanta either in the state  $s$  or in some other state  $s'$  has changed once more by  $\pm 1$ . The appearance of quanta in states  $s'$  caused by the presence of quanta in states  $s$  is called scattering.

To calculate the probability of scattering we determine the values of  $\phi(n'; N_s^0 \pm 1, N_{s'}^0 \pm 1, \dots)$  on the supposition that  $\phi$  is appreciable only at  $n_0, N_1^0, N_2^0 \dots$  and at  $n; N_1^0, \dots, N_s^0 \pm 1, \dots$ . The equations which determine these values of  $\phi$  are by (110) [remembering that in (110) the  $\Sigma_{s' \neq s''}$  includes every pair  $(s', s'')$  twice]

$$\begin{aligned} & \left( \frac{\partial}{2\pi i \partial t} + \nu_{n'} + \sum N_s^0 \nu_s + \nu_{s'} - \nu_{s''} \right) \phi(n'; N_s^0 + 1, N_{s'}^0 - 1, \dots) \\ &= \sum_n \left\{ \left( \frac{N_{s'}^0 + 1}{\nu_{s'} V} \right)^{1/2} \alpha_{n'n^{s'}} \phi(n; N_s^0, N_{s'}^0 - 1) \right. \\ & \quad \left. + \left( \frac{N_{s'}^0}{\nu_{s'} V} \right)^{1/2} \alpha_{n'n^{s'}} \phi(n; N_s^0 + 1, N_{s'}^0) \right. \\ & \quad \left. + \sum_{s'' \neq s'} \left( \frac{N_{s''}^0}{\nu_{s''} V} \right)^{1/2} \alpha_{n'n^{s''}} \phi(n; N_s^0, N_{s''}^0 + 1, N_{s'}^0 - 1) \right. \\ & \quad \left. + \sum_{s'' \neq s'} \left( \frac{N_{s''}^0}{\nu_{s''} V} \right)^{1/2} \alpha_{n'n^{s''}} \phi(n; N_s^0 + 1, N_{s''}^0, N_{s'}^0 - 1) \right\} \end{aligned} \quad (117.1)$$

<sup>26</sup> A. Einstein, Phys. Zeits. **18**, 121 (1917).

$$\begin{aligned}
& + \left( \frac{N_{s'',0}}{\nu_{s''}V} \right)^{1/2} (\alpha_{nn',s''})^* \phi(n; N_{s',0} + 1, N_{s'',0} \dots) \Big\} \\
& - ((\mathbf{f}_s' \mathbf{f}_{s''})/2\pi V) (\nu_s \nu_{s''})^{-1/2} \beta_{n',n_0}^{k'-k''} (N_{s',0} + 1)^{1/2} (N_{s'',0})^{1/2} \phi(n_0; N_{s',0} N_{s'',0}, \dots) \\
& \quad \left( \frac{\partial}{2\pi i \partial t} + \nu_{n'} + \sum N_s^0 \nu_s + \nu_{s'} + \nu_{s''} \right) \phi(n'; N_{s',0} + 1, N_{s'',0} + 1) \\
& = \sum_n \left\{ \left( \frac{N_{s',0} + 1}{\nu_{s'}V} \right)^{1/2} \alpha_{n'n,s'} \phi(n; N_{s',0}, N_{s'',0} + 1) \right. \\
& \quad + \left( \frac{N_{s'',0} + 1}{\nu_{s''}V} \right)^{1/2} \alpha_{n'n,s''} \phi(n; N_{s',0} + 1, N_{s'',0}) \Big\} \\
& \quad - ((\mathbf{f}_s' \mathbf{f}_{s''})/2\pi V) (\nu_s \nu_{s''})^{-1/2} \beta_{n',n_0}^{k'-k''} (N_{s',0} + 1)^{1/2} (N_{s'',0} + 1)^{1/2} \phi(n_0; N_{s',0}, N_{s'',0}, \dots); \\
& \quad \left( \frac{\partial}{2\pi i \partial t} + \nu_{n'} + \sum N_s^0 \nu_s - \nu_{s'} - \nu_{s''} \right) \phi(n'; N_{s',0} - 1, N_{s'',0} - 1) \\
& = \sum_n \left\{ \left( \frac{N_{s',0}}{\nu_{s'}V} \right)^{1/2} (\alpha_{nn',s'})^* \phi(n; N_{s',0}, N_{s'',0} - 1) \right. \\
& \quad + \left( \frac{N_{s'',0}}{\nu_{s''}V} \right)^{1/2} (\alpha_{nn',s''})^* \phi(n; N_{s',0} - 1, N_{s'',0}) \Big\} \\
& \quad - ((\mathbf{f}_s' \mathbf{f}_{s''})/2\pi V) (\nu_s \nu_{s''})^{-1/2} \beta_{n',n_0}^{-k',-k''} (N_{s',0})^{1/2} (N_{s'',0})^{1/2} \phi(n_0; N_{s',0}, N_{s'',0}).
\end{aligned} \tag{117.2}$$

$$\tag{117.3}$$

In the above formulas, if the value of an  $N_s$  is not specifically indicated, it is understood to be  $N_s^0$ . In addition to the above  $\phi$ 's there are also  $\phi(n'; N_{s'} \pm 2)$  which are of the same order of magnitude. They do not interest us, however, because they represent scattering without change of frequency or direction of the quantum.

We are particularly interested in the appearance of quanta in new directions and with new polarizations and frequencies. Let us agree to refer by  $s'$  to such a quantum. Then  $N_{s',0} = 0$ . The state for which only one  $N_{s'}$  is 1 we write

$$\phi(n'; 1_{s'}, N_{s'',0}).$$

Thus (117.1) gives  $\phi(n'; 1_{s'}, N_{s'',0} - 1)$ . The square of the modulus of this quantity gives the probability that a quantum of the type  $s''$  disappeared and instead a quantum of a new type  $s'$  came into existence. We should note that (117.1) explains the appearance of the  $s'$  quantum as due to the three  $\phi$ 's on the right side of the equation. The last of these is  $\phi(n_0; N_{s',0}, N_{s'',0})$  and refers to the initial state. It comes from the  $A^2$  term in the Hamiltonian. The building up of  $\phi(n'; 1_{s'}, N_{s'',0} - 1)$  due to this term may be called *true scattering* because it arises directly from the initial state. The other contributions come from intermediate states  $\phi(n; N_{s',0} - 1)$  and  $\phi(n; 1_{s'}, N_{s'',0}, \dots)$ . The first of these may be interpreted as a state in which the atom went to a state  $n$  and a light quantum of the type  $s''$  has been absorbed. The scattering due to this term may be said to consist in the absorption of a quantum  $s''$  and the transition of the atom from  $n_0$  to  $n$ , followed by a transition of the

atom from  $n$  to  $n'$  and the emission of  $s'$ . Similarly the second type of intermediate state  $\phi(n; 1_{s'}, N_{s'',0} \dots)$  may be said to participate in scattering by having the atom change from  $n_0$  to  $n$ , a quantum  $s'$  being simultaneously emitted; this being followed by a transition of the atom from  $n$  to  $n'$  simultaneous with the absorption of  $s''$ . A word of caution about this interpretation of (117.1) is appropriate. The description of the role of intermediate states applies to the calculation of the probability amplitude  $\phi(n'; 1_{s'}, N_{s'',0}-1)$  and not to the square of the modulus of this quantity. Exactly the same type of interpretation applies to (117.2). Only where in (117.1) a quantum  $s''$  has been absorbed in (117.2) a quantum  $s''$  is emitted. Eq. (117.3) is not connected with scattering but with the simultaneous absorption of two light quanta  $s', s''$ . Similarly (117.2) for all  $N_{s'',0}$  equal to zero gives the probability of the simultaneous emission of two light quanta even if no light quanta are present initially. Substituting (113) and solving (117.1) subject to the initial condition  $\phi(n'; 1_{s'}, N_{s'',0}-1) = 0$  we obtain

$$\begin{aligned} & \phi(n'; 1_{s'}, N_{s'',0}-1) e^{2\pi i(\nu_{s'} + \sum N_{s''} \nu_{s''})t} \\ &= \frac{(N_{s'',0})^{1/2}}{V(\nu_{s'} \nu_{s''})^{1/2}} \left\{ (\nu_{n'n_0} + \nu_{s'} - \nu_{s''})^{-1} \left[ \left( \sum_n \left( \frac{\alpha_{n'n}^{s'} (\alpha_{n_0 n}^{s''})^*}{\nu_{nn_0} - \nu_{s''}} + \frac{(\alpha_{nn}^{s''})^* \alpha_{nn_0}^{s'}}{\nu_{s'} - \nu_{n_0 n}} \right) \right. \right. \right. \\ & \quad - \frac{(\mathbf{f}_{s'} \mathbf{f}_{s''})}{2\pi} \beta_{n'n_0}^{k', -k''} \left. \right] e^{-2\pi i(\nu_{n_0} - \nu_{s''})t} + \left( \sum_n \left( \frac{\alpha_{n'n}^{s'} (\alpha_{n_0 n}^{s''})^*}{\nu_{n'n} + \nu_{s'}} + \frac{(\alpha_{nn}^{s''})^* \alpha_{nn_0}^{s'}}{\nu_{n'n} - \nu_{s''}} \right) \right. \\ & \quad \left. \left. + \frac{(\mathbf{f}_{s'} \mathbf{f}_{s''})}{2\pi} \beta_{n', n_0}^{k', -k''} \right) e^{-2\pi i(\nu_{n'} - \nu_{s''})t} \right] - \sum_n \left( \frac{\alpha_{n'n}^{s'} (\alpha_{n_0 n}^{s''})^* e^{2\pi i(\nu_{s'} + \nu_{s''})t}}{(\nu_{nn_0} - \nu_{s''})(\nu_{n'n} + \nu_{s'})} \right. \\ & \quad \left. + \frac{(\alpha_{nn}^{s''})^* \alpha_{nn_0}^{s'}}{(\nu_{s'} - \nu_{n_0 n})(\nu_{n'n} - \nu_{s''})} \right) e^{-2\pi i \nu_{n'} t} \left. \right\}. \end{aligned}$$

The combination in square brackets gives large contributions when the factor in front of it is large. This is the case when  $\nu_{n'n_0} + \nu_{s'} - \nu_{s''}$  is small. For quanta  $s'$  satisfying approximately  $\nu_{s'} = \nu_{s''} - \nu_{n'n_0}$  we may as a rule calculate  $\phi(n'; 1_{s'}, N_{s'',0}-1)$  neglecting the last term in the above expression because it does not contain the factor  $1/(\nu_{n'n_0} + \nu_{s'} - \nu_{s''})$ . In order that this be the case the denominators  $(\nu_{nn_0} - \nu_{s''}) \cdot (\nu_{n'n} + \nu_{s'})$  and  $(\nu_{s'} - \nu_{n_0 n}) \cdot (\nu_{n'n} - \nu_{s''})$  should not vanish. This means that (1)  $\nu_{s''}$  may not be either  $\nu_{nn_0}$  or  $\nu_{n'n}$ , i.e., the *incident frequency* may not be a possible *absorption frequency* from the *initial* atomic level, nor may it be an *emission frequency* from the *final* atomic level (2) the *scattered frequency*  $\nu_{s'}$  may not be an *emission frequency* ( $\nu_{n_0 n}$ ) from the *initial* atomic level, nor may it be ( $\nu_{n'n}$ ) an *absorption frequency* from the *final* level. If either one of these conditions is fulfilled we speak of resonance of either the incident or the scattered frequency with the atom. For the case where there is no resonance we may neglect the term in  $\exp[-2\pi i \nu_{n'} t]$ . Further in this case  $|\phi(n'; 1_{s'}, N_{s'',0}-1)|$  has a sharp maximum when  $\nu_{s'}$  approaches  $\nu_{s''} - \nu_{n'n_0}$ . This is due to the factor  $(\nu_{n'n_0} + \nu_{s'} - \nu_{s''})^{-1}$ . The denominators  $\nu_{nn_0} - \nu_{s''}$ ,  $\nu_{s'} - \nu_{n_0 n}$ ,  $\nu_{n'n} + \nu_{s'}$ ,  $\nu_{n'n} - \nu_{s''}$  occurring in the square bracket multiplied by this large factor do not vary critically with  $\nu_{s'}$  unless

there is resonance. The values  $\nu_{s'} = \nu_{s''} - \nu_{n'n_0}$  may be used in them. With these approximations

$$\begin{aligned} \phi(n'; 1_{s'}, N_{s'',0} - 1) e^{2\pi i(\nu_{n_0} + \Sigma N_{s''} \nu_{s'})t} \\ = \frac{(N_{s'',0})^{1/2} [1 - e^{-2\pi i(\nu_{n'n_0} + \nu_{s'} - \nu_{s''})t}]}{V \nu_{s'}^{1/2} \nu_{s'',1/2} (\nu_{n'n_0} + \nu_{s'} - \nu_{s''})} \left[ \sum_n \left( \frac{\alpha_{n'n}^{s'} (\alpha_{n_0 n}^{s''})^*}{\nu_{n n_0} - \nu_{s''}} \right. \right. \\ \left. \left. + \frac{(\alpha_{n n'}^{s''})^* \alpha_{n n_0}^{s'}}{\nu_{s'} - \nu_{n_0 n}} \right) - \frac{(\mathbf{f}_{s'} \mathbf{f}_{s''})}{2\pi} \beta_{n'n_0}^{k', -k''} \right]. \quad (118) \end{aligned}$$

We compare this with the *emission equation* [the first equation (113)]. We see that (118) gives the same value for  $|\phi(n'; 1_{s'}, N_{s'',0} - 1)|^2$  in the neighborhood of  $\nu_{s'} = \nu_{s''} - \nu_{n'n_0}$  as though we had used (113) for the calculation of the probability of appearance of light quanta in the neighborhood of an emission frequency  $\nu_{s''} - \nu_{n'n_0}$  [denoted in (113) by  $\nu_{n_0 n}$ ] with an  $\alpha_{n n_0}^s$  given by

$$\begin{aligned} (\alpha^s)_{\text{effective}} = \left( \frac{N_{s'',0}}{V \nu_{s''}} \right)^{1/2} \left[ \sum_n \left( \frac{\alpha_{n'n}^{s'} (\alpha_{n_0 n}^{s''})^*}{\nu_{n n_0} - \nu_{s''}} \right. \right. \\ \left. \left. + \frac{(\alpha_{n n'}^{s''})^* \alpha_{n n_0}^{s'}}{\nu_{s'} - \nu_{n_0 n}} \right) - \frac{(\mathbf{f}_{s'} \mathbf{f}_{s''})}{2\pi} \beta_{n'n_0}^{k', -k''} \right] \quad (118') \end{aligned}$$

This is correct whether the wave-length is long in comparison with atomic dimensions or not. In the former case the formula may be simplified. For the  $\alpha^s$  we substitute (111'') and write  $(\dot{r}_i)_{n'n}$  for  $(p_i/m_i)_{n'n}$ . Also by (109) if the wave-length is large in comparison with atomic dimensions

$$\beta_{n'n_0}^{k', -k''} = \sum_i (e_i^2/m_i) e^{-i(k' - k'')R} \delta_{n'n_0}. \quad (118'')$$

As before  $\mathbf{R}$  is a vector from the origin to the atom. In the same approximation

$$\begin{aligned} \alpha_{n'n}^{s'} (\alpha_{n_0 n}^{s''})^* &= e^{-i(k' - k'')R} \left( \sum_i e_i \dot{\mathbf{r}}_i \mathbf{f}_{s'} \right)_{n'n} \left( \sum_i e_i \dot{\mathbf{r}}_i \mathbf{f}_{s''} \right)_{nn_0} / (2\pi\hbar) \\ (\alpha_{n n'}^{s''})^* \alpha_{n n_0}^{s'} &= e^{-i(k' - k'')R} \left( \sum_i e_i \dot{\mathbf{r}}_i \mathbf{f}_{s''} \right)_{n'n} \left( \sum_i e_i \dot{\mathbf{r}}_i \mathbf{f}_{s'} \right)_{nn_0} / (2\pi\hbar). \end{aligned}$$

Substituting these values into (118') and letting

$$(\alpha^{s'})_{\text{eff.}} = (2\pi\hbar)^{-1/2} e^{-ik'R} ((\dot{\mathbf{e}}\mathbf{r})_{s'})_{\text{eff.}} = 2\pi i \nu_{s'} (2\pi\hbar)^{-1/2} e^{-ik'R} ((\dot{\mathbf{e}}\mathbf{r})_{s'})_{\text{eff.}} \quad (119)$$

we have

$$\begin{aligned} (\dot{\mathbf{e}}\mathbf{r})_{\text{eff.}} &= e^{ik'R} \left( \frac{N_{s'',0}}{2\pi\hbar V \nu_{s''}} \right)^{1/2} (2\pi i \nu_{s'})^{-1} \left[ \sum_n \left( \frac{(\dot{\mathbf{e}}\mathbf{r})_{n'n} ((\dot{\mathbf{e}}\mathbf{r})_{s'})_{nn_0}}{\nu_{n n_0} - \nu_{s''}} \right. \right. \\ &\quad \left. \left. + \frac{((\dot{\mathbf{e}}\mathbf{r})_{s''})_{n'n} ((\dot{\mathbf{e}}\mathbf{r})_{nn_0})}{\nu_{s'} - \nu_{n_0 n}} \right) - \hbar \mathbf{f}_{s''} \sum_i (e_i^2/m_i) \delta_{n'n_0} \right]; \quad (\dot{\mathbf{e}}\mathbf{r}) \equiv \sum_i e_i \dot{\mathbf{r}}_i. \quad (120) \end{aligned}$$

Here we have abbreviated  $\sum_i e_i \mathbf{r}_i$  into  $(e\mathbf{r})$ . The expression (120) is the effective value of a matrix element which an atom with an emission frequency at  $\nu_{s'}$  would have to possess in order to have the same emission spontaneously as the atom under discussion has under the influence of the incident radiation  $s''$ . Formula (120) may be simplified by using the exchange relations between the  $\mathbf{p}_i$  and  $\mathbf{r}_i$ . Since  $\mathbf{p}_i = m_i \dot{\mathbf{r}}_i$  and  $(\dot{\mathbf{r}}_i)_{nm} = 2\pi i \nu_{nm} (\mathbf{r}_i)_{nm}$  the exchange relations give a number of identities between the components of  $(\mathbf{r}_i)_{nm}$ . It is derived without difficulty from these that

$$- h\mathbf{f}_{s''}(e_i e_j / m_i) \delta_{ij} \delta_{n'n_0} = 4\pi^2 e_i e_j \sum_n [\nu_{n'n}(\mathbf{r}_i)_{n'n}(\mathbf{f}_{s''}\mathbf{r}_j)_{nn_0} + \nu_{n_0n}(\mathbf{r}_j\mathbf{f}_{s''})_{n'n}(\mathbf{r}_i)_{nn_0}].$$

Substituting this for the last term in the square bracket of (120) and replacing the denominator  $\nu_{s'} - \nu_{n_0n}$  by  $\nu_{s''} + \nu_{nn'}$  the square bracket becomes

$$- 4\pi^2 \nu_{s''} \left\{ \frac{(\mathbf{e}\mathbf{r})_{n'n}((\mathbf{e}\mathbf{r})\mathbf{f}_{s''})_{nn_0} \nu_{n'n}}{\nu_{nn_0} - \nu_{s''}} + \frac{((\mathbf{e}\mathbf{r})\mathbf{f}_{s''})_{n'n}(\mathbf{e}\mathbf{r})_{nn_0} \nu_{nn'}}{\nu_{s''} + \nu_{nn'}} \right\}.$$

Observing finally that  $\nu_{n'n}/(\nu_{nn_0} - \nu_{s''}) = -1 - \nu_{s'}(\nu_{nn_0} - \nu_{s''})^{-1}$  and  $\nu_{nn_0}/(\nu_{s''} + \nu_{nn'}) = 1 - \nu_{s'}(\nu_{nn'} + \nu_{s''})^{-1}$ , and also noting that  $\sum_n [(\mathbf{e}\mathbf{r})_{n'n}((\mathbf{e}\mathbf{r})\mathbf{f}_{s''})_{nn_0} - ((\mathbf{e}\mathbf{r})\mathbf{f}_{s''})_{n'n}(\mathbf{e}\mathbf{r})_{nn_0}] = 0$  on account of the commutation relations, (120) becomes

$$\begin{aligned} (\sum e_i \mathbf{r}_i)_{\text{eff.}} = & \frac{e^{ik''R}}{2i\hbar} \left( \frac{8\pi N_{s'',0} \hbar \nu_{s''}}{V} \right)^{1/2} \sum_n \left( \frac{(\mathbf{e}\mathbf{r})_{n'n}((\mathbf{e}\mathbf{r})\mathbf{f}_{s''})_{nn_0}}{\nu_{nn_0} - \nu_{s''}} \right. \\ & \left. + \frac{((\mathbf{e}\mathbf{r})\mathbf{f}_{s''})_{n'n}(\mathbf{e}\mathbf{r})_{nn_0}}{\nu_{s''} + \nu_{nn'}} \right). \end{aligned} \quad (121)$$

In terms of  $\mathcal{E}$  used in section II,  $\mathcal{E}(t) = (1/2)[\mathcal{E}e^{2\pi i \nu t} + \mathcal{E}^*e^{-2\pi i \nu t}]$  we can say that  $\mathcal{E}(t)\mathcal{E}(t)$  when averaged over the time is  $(1/2)\mathcal{E}\mathcal{E}^*$  while the energy density is  $(1/8\pi)\mathcal{E}\mathcal{E}^*$ . This energy density is replaced in the present calculation by  $N_{s'',0}\hbar\nu_{s''}/V$ . A factor of absolute value one does not change the predictions of (121) when we calculate probabilities of emission. We may replace in it, therefore,  $e^{ik''R}(8\pi N_{s'',0}\hbar\nu_{s''}/V)^{1/2}\mathbf{f}_{s''}$  by  $\mathcal{E}^*$ . To within a factor of absolute value one (121) agrees then with the first expression (90') being simply its complex conjugate. In making the comparison we make  $Y$  in (90') equal to  $\mathcal{M} = \sum e_i \mathbf{r}_i$ .

Similarly one can calculate the effective matrix element for the simultaneous emission of  $s'$  and  $s''$  from (117.2). The frequency  $\nu_{s'}$  is then  $-\nu_{s''} + \nu_{n_0n'}$ . The result similarly agrees with the second line of (90'). The above derivation of (121) follows closely Dirac's original treatment.

It will be remembered that in connection with (90') we have derived also the probability of the simultaneous absorption of two light quanta. The same may be done also by means of (117.3). We have also derived the probability of the absorption of one kind of light quantum associated with the induced emission of another. This is contained in (117.1) if both  $N_{s',0}$ ,  $N_{s'',0}$  do not vanish. Then  $\phi(n'; N_{s',0}+1, N_{s'',0}-1)$  is the probability amplitude of a state in which the absorption of  $s''$  occurred simultaneously with the induced

emission of  $s'$ . It is clear that the only difference between the solution of (117.1) for  $N_{s',0} \neq 0$  and that for  $N_{s',0} = 0$  consists in the occurrence of the factor  $(N_{s',0} + 1)^{1/2}$ . The effective matrix element to be used is, therefore, the  $(\sum e_i r_i)_{\text{eff.}}$  given by (121) multiplied by  $(N_{s',0} + 1)^{1/2}$ . This is again in agreement with the conclusion derived in section IV in connection with the discussion of formulas (90').

There is also a simple relation between (117.2) and (117.3). Comparing the equations we see that  $(N_{s',0} + 1)^{1/2}$  of (117.2) becomes replaced by  $(N_{s',0} N_{s'',0})^{1/2}$  in (117.3). The two frequencies can be absorbed when

$$\nu_{s'} + \nu_{s''} = \nu_{n'n_0}$$

while in (117.2) they may be emitted if  $\nu_{s'} + \nu_{s''} = \nu_{n_0 n'}$ . Comparing the right hand sides of (117.2) and (117.3) by means of (113) we see that the effective matrix element for the simultaneous absorption of  $s'$  and  $s''$  differs from the corresponding element for emission simply by having the signs of  $\nu_{s'}$  and  $\nu_{s''}$  changed in the formula. This is again in agreement with the conclusion in section IV in connection with (90'). Only there we had also interchanged the initial and final states of the atom with the result that the matrix element remained unchanged. Thus all of the results are in agreement with (90').

### §8. Terms in the square of the vector potential

In Dirac's treatment it is essential to take into account the quadratic term in the vector potential. In the correspondence principle derivation and in the treatments of Born, Heisenberg, Jordan and of Schroedinger this term was not considered. Nevertheless the two results agree. There is a good reason for this apparent inconsistency.

In the discussions of Born, Heisenberg, Jordan and of Schroedinger the matrix for the atom subjected to the action of the external field is worked out to the first order in the intensity of the field. The Hamiltonian function used is

$$H = (\mathbf{p}^2/(2m)) + V - (\mathcal{E}\mathbf{r}).$$

This Hamiltonian is correct only as long as the action of the magnetic force of the incident light wave is not important. Such is the case whenever the velocity of the particles is small. The same approximation has been made in Dirac's treatment, *viz.*, when the approximate expressions (119) were employed. This is seen if it is remembered that the manner in which the magnetic force enters is through space derivatives of the vector potential. In the transition from (109) to the approximation (119) we have replaced the variable  $e^{-ikr'}$  by local constants  $e^{-ikR}$  and thus have neglected the space derivatives of the vector potential and thereby also the magnetic force. Furthermore with the above Hamiltonian function (containing the electric intensity explicitly) the relation between the velocity and the momentum is

$$\mathbf{p} = m\dot{\mathbf{r}}$$

which does not involve the vector potential at all. This is different from the relations between momentum and velocity in Dirac's treatment. There



$$\mathbf{p} - (e/c)\mathbf{A} = m\dot{\mathbf{r}}.$$

[The symbol  $\mathbf{p}$  thus stands for different things in the two theories.] This brings about the necessity of considering the vector potential in the calculation of the matrix elements of the velocity in addition. For the perturbation of  $\mathbf{p}$  is of the order of  $-(e/c)\mathbf{A}$  so that it would not be consistent to calculate the perturbation of  $m\dot{\mathbf{r}}$  by considering only the perturbation of  $\mathbf{p}$ .

A similar condition exists in calculations of intensity relations for the Compton effect with Schroedinger's quadratic relativistic wave equation by Dirac and Gordon.<sup>27</sup> The current vector in this theory contains the vector potential explicitly. In the calculation of the perturbation of matter by radiation only the first power of the vector potential has to be considered. The retarded potentials are later expressed in terms of the current vector and contain the vector potential once more. This is responsible for the occurrence of the term in  $b_a$  in Gordon's formula (40).

The reason why the term in  $\mathbf{A}$  is so important in the discussion of scattering is thus that  $\mathbf{A}$  enters explicitly into the expression for the current vector.

Dirac's relativistic equation for the electron gives<sup>28</sup> a current vector not involving the vector potential. The square of the vector potential does not occur explicitly. Scattering in this formulation is entirely a second order effect in the interaction energy which is linear in  $\mathbf{A}$ .

<sup>27</sup> P. A. M. Dirac, Proc. Roy. Soc. **111**, 405 (1926); W. Gordon, Zeits. f. Physik **40**, 117 (1926); O. Klein, Zeits. f. Physik **41**, 407 (1927).

<sup>28</sup> I. Waller, Zeits. f. Physik **61**, 837 (1930).