

BAKERIAN LECTURE

The physical interpretation of quantum mechanics

BY P. A. M. DIRAC, F.R.S.

(Delivered 19 June 1941.—Received 23 September 1941)

Modern developments of atomic theory have required alterations in some of the most fundamental physical ideas. This has resulted in its being usually easier to discover the equations that describe some particular phenomenon than just how the equations are to be interpreted. The quantum mechanics of Heisenberg and Schrödinger was first worked out for a number of simple examples, from which a general mathematical scheme was constructed, and afterwards people were led to the general physical principles governing the interpretation, such as the superposition of states and the indeterminacy principle. In this way a satisfactory non-relativistic quantum mechanics was established.

In extending the theory to make it relativistic, the developments needed in the mathematical scheme are easily worked out, but difficulties arise in the interpretation. If one keeps to the same basis of interpretation as in the non-relativistic theory, one finds that particles have states of negative kinetic energy as well as their usual states of positive energy, and, further, for particles whose spin is an integral number of quanta, there is the added difficulty that states of negative energy occur with a negative probability.

With electrons the negative-probability difficulty does not arise, and one can get a sensible interpretation of the negative-energy states by assuming them to be nearly all occupied and an unoccupied one to be a positron. This model, however, is excessively complicated to work with and one cannot get any results from it without making very crude approximations. The simple accurate calculations that one *can* make would apply to a world which is almost saturated with positrons, and it appears to be a better method of interpretation to make the general assumption that transition probabilities obtained from these calculations for this hypothetical world are the same as those for the actual world.

With photons one can get over the negative-energy difficulty by considering the states of positive and negative energy to be associated with the emission and absorption of a photon respectively, instead of, as previously, with the existence of a photon. The simplest way of developing the theory would make it apply to a hypothetical world in which the initial probability of certain states is negative, but transition probabilities calculated for this hypothetical world are found to be always positive, and it is again reasonable to assume that these transition probabilities are the same as those for the actual world.

NON-RELATIVISTIC THEORY

Heisenberg, with his discovery of quantum mechanics, introduced a new outlook on the nature of physical theory. Previously, it was always considered essential that there should be a detailed description of what is taking place in natural phenomena, and one used this description to calculate results comparable with experiment. Heisenberg put forward the view that it is sufficient to have a mathematical scheme from which one can calculate in a consistent manner the results of all experiments—a detailed description in the traditional sense is unnecessary and may very well be impossible to establish.

Heisenberg's method focuses attention on to the quantities which enter into experimental results. It was first applied to the theory of spectra, for which these quantities are the energy levels of the atomic system and certain probability coefficients which determine the probability of a radiative transition taking place from one level to another. The method sets up equations connecting these quantities and allows one to calculate them, but does not go beyond this. It does not provide any description of radiative transition processes. It does not even allow one to deduce how the results of a calculation are to be used, but requires one to assume Einstein's laws of radiation (the laws which tell how the probability of a radiative transition process depends on the intensity of the incident radiation), and to assume that certain quantities determined by the calculation are the coefficients appearing in the laws.

Shortly after Heisenberg's discovery, Schrödinger set up independently another form of quantum mechanics, which also enables one to calculate energy levels and probability coefficients and gives results agreeing with those of Heisenberg, but which introduces an important new feature. *It connects together, in one calculation, a set of probability coefficients which act together under certain conditions in Nature;* for example, the set of probability coefficients referring to transitions from one particular initial state to any final state. In this respect it is to be contrasted with Heisenberg's method, which connects together in one calculation all the probability coefficients for a dynamical system, i.e. the probability coefficients from all initial states to all final states.

This feature of Schrödinger's method gives it two important advantages. First, as a consequence of its enabling one to obtain fewer results at a time, it makes the computation much simpler. Secondly, it supplies, in a certain sense, a *description* of what is taking place in Nature, since a calculation leading to results which come into play together under certain conditions

in Nature will be in close correspondence with the physical process that is taking place under those conditions, various points in the calculation having their counterparts in the physical process. A description in this limited sense seems to be the most that is possible for atomic processes. It implies a much less complete connexion between the mathematics and the physics than one has in classical mechanics, and one might be disinclined to call it a description at all, but one may at least consider it as an appropriate generalization of what one usually means by a description. On account of Schrödinger's method allowing a description in this new sense while Heisenberg's allows none, Schrödinger's method introduces an outlook on the nature of physical theory intermediate between Heisenberg's and the old classical one.

When Heisenberg's & Schrödinger's theories were developed it was soon found that they both rested on the same mathematical formalism and differed only with regard to the method of physical interpretation. The formalism is a generalization of the Hamiltonian form of classical dynamics, involving linear operators instead of ordinary algebraic variables, and is so natural and beautiful as to make one feel sure of its correctness as the foundation of the theory. The question of its interpretation, however, which involved unifying Heisenberg's & Schrödinger's ideas into a satisfactory comprehensive scheme, was not so easily settled, and is probably still not finally settled.

The situation of a formalism becoming established before one is clear about its interpretation should not be considered as surprising, but as a natural consequence of the drastic alterations which the development of physics has required in some of the basic physical concepts. This makes it an easier matter to discover the mathematical formalism needed for a fundamental physical theory than its interpretation, since the number of things one has to choose between in discovering the formalism is very limited, the number of fundamental ideas in pure mathematics being not very great, while with the interpretation most unexpected things may turn up.

The best way of seeking the interpretation in such cases is probably from a discussion of simple examples. This way was used for the theory of quantum mechanics and led eventually to a satisfactory interpretation applicable to all phenomena for which relativistic effects are negligible. This interpretation is more closely connected with Schrödinger's method than Heisenberg's, as one would expect on account of the former affording in some sense a description of Nature, and is centred round a *Schrödinger's wave function*, which is one of the things that can be operated on by the linear operators which the dynamical variables have become. The corre-

spondence which the existence of a description implies between the mathematics and the physics makes a wave function correspond to a state of motion of the atomic system, in such a way that, for example, a calculation which gives the transition probabilities from a particular initial state to any final state would be based on that wave function which represents the motion ensuing from this initial state. A wave function is a complex function $\psi(q_1 q_2 \dots q_n t)$ of all the co-ordinates q_1, q_2, \dots, q_n of the system and of the time t , and it receives the interpretation that the square of its modulus, $|\psi(q_1 \dots q_n t)|^2$, is the probability, for the state of motion it corresponds to, of the co-ordinates having values in the neighbourhood of q_1, q_2, \dots, q_n , per unit volume of co-ordinate space, at the time t .

A wave function can be transformed so as to refer to other dynamical variables, for example, the momenta p_1, p_2, \dots, p_n , when it is said to be in another representation. The square of its modulus $|\psi(p_1 \dots p_n t)|^2$ is then the probability, per unit volume of momentum space, of the momenta having values in the neighbourhood of p_1, p_2, \dots, p_n at the time t . A wave function itself never has an interpretation, but only the square of its modulus, and the need for distinguishing between two wave functions having the same squares of their moduli arises only because, if they are transformed to a different representation, the squares of their moduli will in general become different. This brings out the incompleteness of the kind of description which is possible with quantum mechanics.

One may make a slight modification in the wave functions in any representation by introducing a weight factor λ and arranging for the probability to be $\lambda |\psi|^2$ instead of $|\psi|^2$. The weight factor may be any positive function of the variables occurring in the wave function.

Wave functions have to satisfy a certain *wave equation*, namely, the equation

$$i\hbar d\psi/dt = H\psi, \quad (1)$$

where H is a Hermitian (or self-adjoint) linear operator, and is the Hamiltonian of the system expressed in the representation concerned. The wave equation (1) is a generalization of the Hamilton-Jacobi equation of classical mechanics. If S is a solution of the latter equation, then

$$\psi = e^{iS/\hbar} \quad (2)$$

will give a first approximation to a solution of the former.

An important property of the wave equation (1) is that it yields the conservation law—the total probability of the variables occurring in the wave function having any value is constant. The wave function should be normalized so as to make this probability initially unity and then it always

remains unity. This conservation law is a mathematical consequence of the wave equation being linear in the operator d/dt and of H being a self-adjoint operator.

The wave equation is linear and homogeneous in the wave function ψ , and so are the transformation equations. In consequence, one can add together two ψ 's and get a third. The correspondence between ψ 's and states of motion now allows one to infer that there is a relationship between the states of motion, such that one can add or superpose two states to get a third. This relationship constitutes the Principle of Superposition of States, one of the general principles governing the interpretation of quantum mechanics that we are here considering.

Another of these principles is Heisenberg's Principle of Indeterminacy. This is a consequence of the transformation laws connecting $\psi(q)$ and $\psi(p)$, which show that each of these functions is the Fourier transform of the other, apart from numerical coefficients, so that one meets the same limitations in giving values to q and p as in giving values to the position and frequency of a train of waves.

These general principles serve to bring out the departures needed from ordinary classical ideas. They are of so drastic and unexpected a nature that it is not to be wondered at that they were discovered only indirectly, as consequences of a previously established mathematical scheme, instead of being built up directly from experimental facts.

DIFFICULTIES OF THE RELATIVISTIC THEORY

The theory outlined above is not in agreement with the restricted principle of relativity, as is at once evident from the special role played by the time t . Thus, while it works very well in the non-relativistic region of low velocities, where it appears to be in complete agreement with experiment, it can be considered only as an approximation, and one must face the task of extending it to make it conform to restricted relativity. (General relativity need not be considered, since gravitational effects are negligible in atomic theory.) One should be prepared for possible further alterations being needed in basic physical concepts, and hence one should follow the route of first setting up the mathematical formalism and then seeking its physical interpretation.

Setting up the mathematical formalism is a fairly straightforward matter. One must first put classical mechanics into relativistic Hamiltonian form. One must take into account that the various particles comprising the dynamical system interact through the medium of the electromagnetic

field, and one must use Lorentz's equations of motion for them, including the damping terms which express the reaction of radiation. This is done in Appendix I, where, with the help of an action principle, the equations of motion are obtained in the Hamiltonian form (1.40) with the Hamiltonians F_i , one for each particle, given by (1.39). This Hamiltonian formulation may now be made into a quantum theory by following rules which have become standardized from the non-relativistic quantum mechanics. The resulting formalism appears to be quite satisfactory mathematically, but when one proceeds to consider its physical interpretation one meets with serious difficulties.

Take an elementary example, that of a free particle without spin, moving in the absence of any field. The classical Hamiltonian for this system is the left-hand side of the equation

$$p_0^2 - p_1^2 - p_2^2 - p_3^2 - m^2 = 0, \quad (3)$$

where p_0 is the energy and p_1, p_2, p_3 the momentum of the particle, the velocity of light being taken as unity. Passing over to quantum theory by the standard rules, one gets from this Hamiltonian the wave equation

$$(\hbar^2 \square + m^2) \psi = 0, \quad (4)$$

where

$$\square \equiv \frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}.$$

The wave function ψ here is a scalar, involving the co-ordinates x_1, x_2, x_3 and the time x_0 on the same footing, and so it is suitable for a relativistic theory.

If one now tries to use the old interpretation that $|\psi|^2$ is the probability per unit volume of the particle being in the neighbourhood of the point x_1, x_2, x_3 at the time x_0 , one immediately gets into conflict with relativity, since this probability ought to transform under Lorentz transformations like the time component of a 4-vector, while $|\psi|^2$ is a scalar. Also the conservation law for total probability would no longer hold, the usual proof of it failing on account of the wave equation (4) not being linear in $\partial/\partial x_0$.

An important step forward was taken by Gordon (1926) and Klein (1927), who proposed that instead of $|\psi|^2$ one should use the expression

$$\frac{1}{4\pi i} \left| \frac{\partial \bar{\psi}}{\partial x_0} \psi - \bar{\psi} \frac{\partial \psi}{\partial x_0} \right|. \quad (5)$$

This expression is the time component of a 4-vector. Further, it is easily verified that the divergence of this 4-vector vanishes, which gives the conservation law in relativistic form. Thus (5) is evidently the correct mathematical form to use.

This form leads to trouble on the physical side, however, since, although it is real, it is not positive definite like $|\psi|^2$. Its employment would result in one having at times a negative probability for the particle being in a certain place.

This is not the only difficulty. Let us consider the energy and momentum of the particle, and take for simplicity a state for which these variables have definite values. The corresponding wave function will be of the form of plane waves,

$$\psi = e^{-i(p_0x_0 - p_1x_1 - p_2x_2 - p_3x_3)/\hbar}$$

In order that the wave equation (4) may be satisfied, the energy and momentum values p_0, p_1, p_2, p_3 here must satisfy the classical equation (3). This equation allows of negative values for the energy p_0 as well as positive ones and is, in fact, symmetrical between positive and negative energies. The negative energies occur also in the classical theory, but do not then cause trouble, since a particle started off in a positive-energy state can never make a transition to a negative-energy one. In the quantum theory, however, such transitions are possible and do in general take place under the action of perturbing forces.

The wave function may be transformed to the momentum and energy variables. The Gordon-Klein expression (5) then goes over into

$$|\psi(p_0 p_1 p_2 p_3)|^2 p_0^{-1} dp_1 dp_2 dp_3, \quad (6)$$

as the probability of the momentum having a value within the small domain $dp_1 dp_2 dp_3$ about the value p_1, p_2, p_3 , with the energy having the value p_0 , which must be connected with p_1, p_2, p_3 by (3). The weight factor p_0^{-1} appears in (6) and makes it Lorentz invariant, since $\psi(p)$ is a scalar—it is defined in terms of $\psi(x)$ to make it so—and the differential element $p_0^{-1} dp_1 dp_2 dp_3$ is Lorentz invariant. This weight factor may be positive or negative, and makes the probability positive or negative accordingly. Thus the two undesirable things, negative energy and negative probability, always occur together.

Let us pass on to another simple example, that of a free particle with spin half a quantum. The wave equation is of the same form (4) as before, but the wave function ψ is no longer a scalar. It must have two components, or four if there is a field present, and the way they transform under Lorentz transformations is given by the general connexion between the theory of angular momentum in quantum mechanics and group theory. The expression $\Sigma |\psi(x)|^2$, summed for the components of ψ , turns out to be the time component of a 4-vector, and further the divergence of this 4-vector vanishes. Thus it is satisfactory to use this expression as the probability per unit

volume of the particle being at any place at any time. One does not now have any negative probabilities in the theory. However, the negative energies remain, as in the case of no spin.

We may go on and consider particles of higher spin. The general result is that there are always states of negative energy as well as those of positive energy. For particles whose spin is an integral number of quanta, the negative-energy states occur with a negative probability and the positive-energy ones with a positive probability, while for particles whose spin is a half-odd integral number of quanta, all states occur with a positive probability.

Negative energies and probabilities should not be considered as nonsense. They are well-defined concepts mathematically, like a negative sum of money, since the equations which express the important properties of energies and probabilities can still be used when they are negative. Thus negative energies and probabilities should be considered simply as things which do not appear in experimental results. The physical interpretation of relativistic quantum mechanics that one gets by a natural development of the non-relativistic theory involves these things and is thus in contradiction with experiment. We therefore have to consider ways of modifying or supplementing this interpretation.

PARTICLES OF HALF-ODD INTEGRAL SPIN

Let us first consider particles with a half-odd integral spin, for which there is only the negative-energy difficulty to be removed. The chief particle of this kind for which a relativistic theory is needed is the electron, with spin half a quantum. Now electrons, and also, it is believed, all particles with a half-odd integral spin, satisfy the exclusion principle of Pauli, according to which not more than one of them can be in any quantum state. (This principle is obtained in quantum mechanics from the requirement that wave functions shall be antisymmetric in all the particles.) With this principle there are only two alternatives for a state, either it is unoccupied or it is occupied by one particle, and a symmetry appears with respect to these two alternatives.

Some time ago I proposed a way of dealing with the negative-energy difficulty for electrons, based on a theory in which nearly all their negative-energy states are occupied. An unoccupied negative-energy state now appears as a 'hole' in the distribution of occupied negative-energy states and thus has a deficiency of negative energy, i.e. a positive energy. From the wave equation one finds that a hole moves in the way one would expect

a positively charged electron to move. It becomes reasonable to identify the holes with the recently discovered positrons, and thus to get an interpretation of the theory involving positrons together with electrons. An electron jumping from a positive- to a negative-energy state in the theory is now interpreted as an annihilation of an electron and a positron, and one jumping from a negative- to a positive-energy state as a creation of an electron and a positron.

The theory involves an infinite density of electrons everywhere. It becomes necessary to assume that the distribution of electrons for which all positive-energy states are unoccupied and all negative-energy states occupied—what one may call the vacuum distribution, as it corresponds to the absence of all electrons and positrons in the interpretation—is completely unobservable. Only departures from this distribution are observable and contribute to the electric density and current which give rise to electromagnetic field in accordance with Maxwell's equations.

The above theory does provide a way out from the negative-energy difficulty, but it is not altogether satisfactory. The infinite number of electrons that it involves requires one to deal with wave functions of very great complexity and leads to such complicated mathematics that one cannot solve even the simplest problems accurately, but must resort to crude and unreliable approximations. Such a theory is a most inconvenient one to have to work with, and on general philosophical grounds one feels that it must be wrong.

Let us see whether one can modify the theory so as to make it possible to work out simple examples accurately, while retaining the basic idea of identifying unoccupied negative-energy states with positrons. The simple calculations that one can make involve simple wave functions, referring to only one or two electrons, and thus referring to nearly all the negative-energy states being unoccupied. The calculations therefore apply to a world almost saturated with positrons, i.e. having nearly every quantum state for a positron occupied. Such a world, of course, differs very much from the actual world. One can now calculate the probability of any kind of collision process occurring in this hypothetical world (in so far as electrons and positrons are concerned). One can deduce the probability coefficient for the process, i.e. the probability per unit number of incident particles or per unit intensity of the beam of incident particles, for each of the various kinds of incident particle taking part in the process. For this purpose one must use the laws of statistical mechanics, which tell how the probability of a collision process depends on the number of incident particles, paying due attention to the modified form of these laws arising from the exclusion principle of Pauli.

Let us now assume that *probability coefficients so calculated for the hypothetical world are the same as those of the actual world*. This single assumption provides a general physical interpretation for the formalism, enabling one to calculate collision probabilities in the actual world. It does not provide a complete physical theory, since it enables one to calculate only those experimental results that are reducible to collision probabilities, and some branches of physics, e.g. the structure of solids, do not seem to be so reducible. However, collision probabilities are the things for which a relativistic theory is at present most needed, and one may hope in the future to find ways of extending the scope of the theory to make it include the whole of physics.

Comparing the new theory with the old, one may say that the new assumption, identifying collision probability coefficients in the actual world with those in a certain hypothetical world, replaces the old assumption about the non-observability of the vacuum distribution of negative-energy electrons. The approximations needed for working out simple examples in the old theory are equivalent in their mathematical effect to making the new assumption; e.g. these approximations include the neglect of the Coulomb interaction between electron and positron in the calculation of the probability of pair creation and annihilation, and this interaction cannot appear in the new theory, since the calculation there is concerned with a one-electron system. Thus the new theory may be considered as a precise formulation of the old theory together with some general approximations needed for applying it.

The new theory for dealing with the negative-energy states of the electron may be applied to any kind of elementary particle with spin half a quantum, and probably also to particles with other half-odd integral spin values, provided, of course, they satisfy Pauli's exclusion principle. It may thus be applied to protons and neutrons. It requires for each particle the possibility of existence of an antiparticle of the opposite charge, if the original particle is charged. If the original particle is uncharged, one can arrange for the antiparticle to be identical with the original.

PARTICLES OF INTEGRAL SPIN

Most of the elementary particles of physics have half-odd integral spin, but there is the important exception of the light-quantum or photon, with spin one quantum, and there is the new cosmic-ray particle, the meson, also probably with spin one quantum, and other such particles may be discovered in the future. All these kinds of particle, it is believed, satisfy the Bose statistics, a statistics which allows any number of particles to be in

the same quantum state with the same a priori probability. (This statistics is obtained in quantum mechanics from the requirement that wave functions shall be symmetric in all the particles.) For these kinds of particles the previous method of dealing with the negative-energy states is therefore no longer applicable, and there is the further difficulty of the negative probabilities.

When dealing with particles satisfying the Bose statistics, it is useful to consider the operators corresponding to the absorption of a particle from a given state or the emission into a given state. These operators can be treated as dynamical variables, although they do not have any analogues in classical mechanics. If one works out their equations of motion and transformation equations, one finds a remarkable correspondence. The absorption operators from a set of independent states have the same equations of motion and transformation equations as the wave function ψ representing a single particle, and similarly for the emission operators and the conjugate complex wave function $\bar{\psi}$. Thus one can pass from a one-particle theory to a many-particle theory by making the ψ and $\bar{\psi}$ describing the one particle into absorption and emission operators, which must satisfy the appropriate commutation relations. Such a passage is called second quantization.

One can get over the difficulties of negative energies and negative probabilities for Bose particles by abandoning the attempt to get a satisfactory theory of a single particle and passing on to consider the problem of many particles, using a method given by Pauli & Weisskopf (1934) for electrons having no spin and satisfying the Bose statistics. (Such electrons are not known experimentally, but there is no known theoretical reason why they should not exist.) The method of Pauli & Weisskopf is to work entirely with positive-energy states. The operators of absorption from and emission into negative-energy states, arising in the application of second quantization to the one-electron theory, are replaced by the operators of emission into and absorption from positive-energy states of electrons with the opposite charge, respectively. This replacement does not disturb the laws of conservation of charge, energy and momentum. The resulting theory involves spinless electrons of both kinds of charge together, and leads to pair creation and annihilation, as with ordinary electrons and positrons.

The method of Pauli & Weisskopf may be applied in a degenerate form to photons and leads to the quantum electrodynamics of Heisenberg & Pauli (1929 *a* and *b*). To take into account that photons have no charge, one must start with a one-particle theory in which the wave functions are real, so that $\bar{\psi} = \psi$. The part of ψ referring to positive-energy states is then made into the absorption operators from positive-energy states, and the part

referring to negative-energy states into the emission operators into positive-energy states. The resulting scheme of operators, involving only positive-energy photon states, may then be put into correspondence with classical electrodynamics, according to the usual laws governing the correspondence between quantum and classical theory.

It would seem that in this way the difficulties of negative energies and probabilities for Bose particles can be overcome, but a new difficulty appears. When one tries to solve the wave equation (or the wave equations if there are several particles with their respective Hamiltonians) one gets divergent integrals in the solution, of the form, in the case of photons,

$$\int_0^\infty f(\nu) d\nu, \quad f(\nu) \sim \nu^n \text{ for large } \nu, \quad (7)$$

ν being the frequency of a photon. The values 1, 0 and -1 for n are the chief ones occurring in simple examples. Thus the wave equation really has no solutions and the method fails.

I have made a detailed study of the divergent integrals occurring in quantum electrodynamics and have shown (Dirac 1939b) that all those with even values of n can be eliminated by introducing into the equations a certain limiting process, which one can justify by showing that a corresponding limiting process is needed in classical electrodynamics to get the equations of motion into Hamiltonian form. (It appears accordingly in the action principle of Appendix I.) The divergent integrals with odd values of n remain, however, and indicate something more fundamentally wrong with the theory.

Divergent integrals are a general feature of quantum field theories, and it has usually been supposed that they should be avoided by altering the forces or the laws of interaction between the elementary particles at small distances, so as to get the integrals cut off for some high value of ν . However, one can easily see that this is wrong, in the case of electrodynamics at any rate, by referring to the corresponding classical theory. The wave function should have its analogue in the solution of the Hamilton-Jacobi equation, in accordance with equation (2), but already when one tries to solve the Hamilton-Jacobi equation of classical electrodynamics corresponding to the wave equation of Heisenberg & Pauli's quantum electrodynamics, one meets with divergent integrals. Now the classical equations of motion concerned, namely, Lorentz's equations including radiation damping, have definite solutions when treated by straightforward methods and if, on trying to get these solutions by a Hamilton-Jacobi method, one meets with divergent integrals, it means simply that the Hamilton-Jacobi method is

an unsuitable one, and not that one should try to alter the physical laws of interaction to get the integrals to converge. The correspondence between the quantum and classical theories is so close that one can infer that the corresponding divergent integrals in the quantum theory must also be due to an unsuitable mathematical method.

The appearance of divergent integrals with odd n -values in Heisenberg & Pauli's form of quantum electrodynamics may be ascribed to the unsymmetrical treatment of positive- and negative-energy photon states. If instead of using Pauli & Weisskopf's method one keeps to plain second quantization, one can build up a form of quantum electrodynamics symmetrical between positive- and negative-energy photon states, as is done in Appendix II. The new theory leads to similar equations as the old one, but with integrals of the type

$$\int_{-\infty}^{\infty} f(\nu) d\nu \quad (8)$$

instead of (7), and since $f(\nu)$ is always a rational algebraic function, and it is reasonable on physical grounds to approach the upper and lower limits of integration in (8) at the same rate, the divergencies with odd n -values all cancel out.

The work of Appendices I and II shows that the new form of quantum electrodynamics also corresponds to classical electrodynamics in accordance with the usual laws, with the exception that operators corresponding to real dynamical variables in the classical theory are no longer always self-adjoint. This exception is not important, as it rather stands apart from the general mathematical connexion between quantum and classical theory. The Hamilton-Jacobi equation corresponding to the wave equation of the new quantum electrodynamics differs from that of the old one only through being expressed in terms of a different set of co-ordinates, but the new Hamilton-Jacobi equation can be solved without divergent integrals and is connected with a satisfactory action principle, namely, that of Appendix I. Thus the correspondence with classical theory of the new form of quantum electrodynamics is more far-reaching than that of the old form, which provides a strong reason for preferring the new form. It now becomes necessary to find some new physical interpretation to avoid the difficulties of negative energies and probabilities.

Let us consider in more detail the relation between the two forms of quantum electrodynamics. In either form the electromagnetic potentials \mathbf{A} at two points \mathbf{x}' and \mathbf{x}'' must satisfy the commutation relations

$$[A_\mu(\mathbf{x}'), A_\nu(\mathbf{x}'')] = g_{\mu\nu} \Delta(\mathbf{x}' - \mathbf{x}''), \quad (9)$$

obtained from analogy with the classical theory, Δ being the four-dimensional Lorentz-invariant function introduced by Jordan & Pauli (1928), which has a singularity on the light-cone and vanishes everywhere else. In the quantum electrodynamics of Heisenberg and Pauli the \mathbf{A} 's are operators referring to the absorption and emission of photons into positive-energy states. Let us call such operators \mathbf{A}^1 . One could introduce a similar set of operators referring to the absorption and emission of photons into negative-energy states. Let us call these operators \mathbf{A}^2 . They satisfy the same commutation relations (9) and commute with the \mathbf{A}^1 's. One can now introduce a third set of operators

$$\mathbf{A}^3 = 2^{-\frac{1}{2}}(\mathbf{A}^1 + \mathbf{A}^2),$$

which operate on wave functions referring to photons in both positive- and negative-energy states, and which satisfy the same commutation relations (9). The use of this third set gives the new form of quantum electrodynamics arising from plain second quantization.

The three sets of \mathbf{A} 's may be expressed in terms of their Fourier components thus:

$$\mathbf{A}^1(\mathbf{x}) = \left. \int \int \int \{ \mathbf{R}_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})} + \bar{\mathbf{R}}_{\mathbf{k}} e^{-i(\mathbf{k}, \mathbf{x})} \} k_0^{-1} dk_1 dk_2 dk_3 \right\} \quad (10)$$

where $k_0 = \sqrt{(k_1^2 + k_2^2 + k_3^2)},$

$\mathbf{R}_{\mathbf{k}}$ being the emission operator and $\bar{\mathbf{R}}_{\mathbf{k}}$ the absorption operator,

$$\mathbf{A}^2(\mathbf{x}) = \left. \int \int \int \{ \mathbf{R}_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})} + \bar{\mathbf{R}}_{\mathbf{k}} e^{-i(\mathbf{k}, \mathbf{x})} \} k_0^{-1} dk_1 dk_2 dk_3 \right\} \quad (11)$$

where $k_0 = -\sqrt{(k_1^2 + k_2^2 + k_3^2)},$

$$\mathbf{A}^3(\mathbf{x}) = 2^{-\frac{1}{2}} \sum_{k_0=\pm\sqrt{(k_1^2 + k_2^2 + k_3^2)}} \int \int \int \{ \mathbf{R}_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})} + \bar{\mathbf{R}}_{\mathbf{k}} e^{-i(\mathbf{k}, \mathbf{x})} \} k_0^{-1} dk_1 dk_2 dk_3. \quad (12)$$

Since the three sets of \mathbf{A} 's all satisfy the same commutation relations, they must correspond merely to three different representations of the same dynamical variables, and the passage from one to another must be a transformation of the linear type usual in quantum mechanics. Thus, after obtaining the divergency-free solution of the wave equation in the representation corresponding to \mathbf{A}^3 , one could apply a transformation to get the solution in the \mathbf{A}^1 representation. However, the transformation would then introduce the same divergent integrals as appear with the direct solution of the wave equation in the \mathbf{A}^1 representation, so one would not get any further this way.

In working with the \mathbf{A}^3 representation one has redundant dynamical variables. It is as though, in dealing with a system of one degree of freedom with the variables q, p , one decided to treat it as a system of two degrees of freedom by putting

$$q = 2^{-\frac{1}{2}}(q_1 + q_2), \quad p = 2^{-\frac{1}{2}}(p_1 + p_2).$$

This would be quite a correct procedure, but would introduce an unnecessary complication. In the case of quantum electrodynamics, the complication is a necessary one, to avoid the divergent integrals. Let us put

$$\mathcal{B}(\mathbf{x}) = 2^{-\frac{1}{2}}\{\mathbf{A}^1(\mathbf{x}) - \mathbf{A}^2(\mathbf{x})\}. \quad (13)$$

Then the \mathcal{B} 's commute with the \mathbf{A}^3 's, and thus with all the dynamical variables appearing in the Hamiltonian, so they are the redundant variables.

To determine the significance of redundant variables in quantum mechanics one may consider a general case and work in a representation which separates the redundant variables from the non-redundant ones. One then sees immediately that a solution of the wave equation corresponds in general, not to a single state, but to a set of states with a certain probability for each—what in the classical theory is called a Gibbs ensemble. The probabilities of the various states depend on the weights attached to the various eigenvalues of the redundant variables in the wave function, these weights being arbitrary, depending on the weight factor in the representation used. If one works in a representation which does not separate the redundant and non-redundant variables, as is the case in quantum electrodynamics with the representation corresponding to the use of \mathbf{A}^3 , the general result that wave functions represent Gibbs ensembles and not single states must still be valid. Thus one can conclude that there are no solutions of the wave equation of quantum electrodynamics representing single states, but only solutions representing Gibbs ensembles. The problem remains of interpreting the negative energies and probabilities occurring with these Gibbs ensembles.

For any \mathbf{x} , $\mathcal{B}(\mathbf{x})$ commutes with the Hamiltonian and is a constant of the motion. We may give it any value we like, subject to not contradicting the commutation relations. Instead of $\mathcal{B}(\mathbf{x})$ it is more convenient to work with the potential field, $\mathbf{B}(\mathbf{x})$ say, obtained from $\mathcal{B}(\mathbf{x})$ by changing the sign of all the Fourier components containing $e^{ik_0x_0}$ with negative values of k_0 . From (13), (10) and (11),

$$\mathbf{B}(\mathbf{x}) = 2^{-\frac{1}{2}}\Sigma \iiint \{\mathbf{R}_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})} - \overline{\mathbf{R}}_{\mathbf{k}} e^{-i(\mathbf{k}, \mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3, \quad (14)$$

the Σ meaning a summation over both values of k_0 , as in equation (12).

Let us now take \mathbf{B} equal to the initial value of \mathbf{A}^3 , a proceeding which does not contradict the commutation relations since its consequences are self-consistent. Then for the initial wave function ψ ,

$$\{\mathbf{B}(\mathbf{x}) - \mathbf{A}^3(\mathbf{x})\}\psi = 0,$$

or, from (12) and (14),

$$\bar{\mathbf{R}}_{\mathbf{k}} \psi = 0, \quad (15)$$

with k_0 either positive or negative. Thus any absorption operator applied to the initial wave function gives the result zero, which means that the corresponding state is one with no photons present.

The following natural interpretation for the wave function at some later time now appears. That part of it corresponding to no photons present may be supposed to give (through the square of its modulus) the probability of no change having taken place in the field of photons; that part corresponding to one positive-energy photon present may be supposed to give the probability of a photon having been emitted; that corresponding to one negative-energy photon present may be supposed to give the probability of a photon having been absorbed; and so on for the parts corresponding to two or more photons present. *The various parts of the wave function which referred to the existence of positive- and negative-energy photons in the old interpretation now refer to the emissions and absorptions of photons.* This disposes of the negative-energy difficulty in a satisfactory way, conforming to the laws of conservation of energy and momentum. It is possible only because of the redundant variables enabling one to arrange that the initial wave function shall correspond in its entirety to no emissions or absorptions having taken place.

The interpretation is not yet complete, because the theory at present would give a negative probability for a process involving the absorption of a photon, or the absorption of any odd number of photons. To find the origin of these negative probabilities, one must study the probability distribution of the photons initially present in the Gibbs ensemble, which one can do by transforming to the representation corresponding to the A^1 potentials. It is true that one cannot apply this transformation to a solution of the wave equation without getting divergent integrals, as has already been mentioned, but one can apply it to the initial wave function, which is of a specially simple form in the photon variables. This is done in Appendix III, and it is found that the probability of there being n photons initially in any photon state is $P_n = \pm 2$, according to whether n is even or odd. Strictly, to make $\sum_{n=0}^{\infty} P_n$ converge to the limit unity, one must consider P_n as a limit,

$$P_n = 2(\epsilon - 1)^n, \quad (16)$$

with ϵ a small positive quantity tending to zero.

Probabilities 2 and -2 are, of course, not physically understandable, but one can use them mathematically in accordance with the rules for working with a Gibbs ensemble. One can suppose a hypothetical mathematical world with the initial probability distribution (16) for the photons, and one can work out the probabilities of radiative transition processes occurring in this world. One can deduce the corresponding probability coefficients, i.e. the probabilities per unit intensity of each beam of incident radiation concerned, by using Einstein's laws of radiation. For example, for a process involving the absorption of a photon, if the probability coefficient is B , the probability of the process is

$$\sum_{n=0}^{\infty} n P_n B = -\frac{1}{2} B, \quad (17)$$

and for a process involving the emission of a photon, if the probability coefficient is A , the probability of the process is

$$\sum_{n=0}^{\infty} (n+1) P_n A = \frac{1}{2} A. \quad (18)$$

Now the probability of an absorption process, as calculated from the theory, is negative, and that for an emission process is positive, so that, equating these calculated probabilities to (17) and (18) respectively, one obtains positive values for both B and A . Generally, it is easily verified that any radiative transition probability coefficient obtained by this method is positive.

It now becomes reasonable to assume that *these probability coefficients obtained for a hypothetical world are the same as those of the actual world*. One gets in this way a general physical interpretation for the quantum theory of photons. When applied to elementary examples, it gives the same results as Heisenberg & Pauli's quantum electrodynamics with neglect of the divergent integrals, since the extra factor $2^{-\frac{1}{2}}$ occurring in the matrix elements of the present theory owing to the $2^{-\frac{1}{2}}$ in the right-hand side of (12) compensates the factor $\frac{1}{2}$ in the right-hand side of (17) or (18). The present general method of physical interpretation is probably applicable to any kind of particle with an integral spin.

CONCLUSION

It appears that, whether one is dealing with particles of integral spin or of half-odd integral spin, one is led to a similar conclusion, namely, that the mathematical methods at present in use in quantum mechanics are capable of direct interpretation only in terms of a hypothetical world differing very markedly from the actual one. These mathematical methods can be made

into a physical theory by the assumption that results about collision processes are the same for the hypothetical world as the actual one. One thus gets back to Heisenberg's view about physical theory—that all it does is to provide a consistent means of calculating experimental results. The limited kind of description of Nature which Schrödinger's method provides in the non-relativistic case is possible relativistically only for the hypothetical world, and even then is rather more indefinite (e.g. the principle of superposition of states no longer applies), because of the need to use a Gibbs ensemble for describing the photon distribution.

To have a description of Nature is philosophically satisfying, though not logically necessary, and it is somewhat strange that the attempt to get such a description should meet with a partial success, namely, in the non-relativistic domain, but yet should fail completely in the later development. It seems to suggest that the present mathematical methods are not final. Any improvement in them would have to be of a very drastic character, because the source of all the trouble, the symmetry between positive and negative energies arising from the association of energies with the Fourier components of functions of the time, is a fundamental feature of them.

APPENDIX I. THE ACTION PRINCIPLE OF CLASSICAL ELECTRODYNAMICS

There are various forms which the action principle of classical electrodynamics may take, but most of them involve awkward conditions concerning the singularities of the field where the charged particles are situated and are not suitable for a subsequent passage to quantum mechanics. Fokker (1929) set up a form of action principle which does not refer to the singularities of the field and which appears to be the best starting point for getting a quantum theory. Fokker's action integral may conveniently be written with the help of the δ function and is then

$$S = S_1 + S_2,$$

where $S_1 = \sum_i m_i \int ds_i$ (1.1)

and $S_2 = \sum_i \sum_{j \neq i} e_i e_j \iint \delta(\mathbf{z}_i - \mathbf{z}_j)^2 (\mathbf{v}_i, \mathbf{v}_j) ds_i ds_j.$ (1.2)

The scalar product notation

$$(\mathbf{a}, \mathbf{b}) = a^\mu b_\mu = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3$$

is here used, and m_i and e_i are the mass and charge of the i th particle, the 4-vector \mathbf{z}_i gives the four co-ordinates of the point on the world-line of the

*i*th particle whose proper-time is s_i , and \mathbf{v}_i is the velocity 4-vector of the *i*th particle

$$\mathbf{v}_i = d\mathbf{z}_i/ds_i \quad (1.3)$$

satisfying

$$\mathbf{v}_i^2 = 1. \quad (1.4)$$

The integrals are taken along the world-lines of the particles, and the occurrence of the δ function $\delta(\mathbf{z}_i - \mathbf{z}_j)^2$ in S_2 ensures that the only values for \mathbf{z}_i and \mathbf{z}_j contributing to the double integral are those for which $(\mathbf{z}_i - \mathbf{z}_j)^2 = 0$, which means that each of the points \mathbf{z}_i , \mathbf{z}_j is on the past or future light-cone from the other.

The action integral as it stands is not a general one covering all possible states of motion. To make it general one must, as has been pointed out by the author (1938), add to it a term of the form

$$S_3 = \sum_i e_i \int M_\mu(\mathbf{z}_i) v_i^\mu ds_i. \quad (1.5)$$

The 4-vector potential $M_\mu(\mathbf{x})$ may be left for the present an arbitrary function of the field point \mathbf{x} .

For the purpose of deducing the equations of motion one may take the limits of integration in the various integrals to be $-\infty$ and ∞ , as was done by Fokker, but in order to introduce momenta and get the equations into Hamiltonian form one must take finite limits. Let us therefore suppose each s_i goes from s_i^0 to s_i' , and let the corresponding \mathbf{z}_i and \mathbf{v}_i be \mathbf{z}_i^0 , \mathbf{z}_i' and \mathbf{v}_i^0 , \mathbf{v}_i' . It is desirable to restrict the initial values s_i^0 so that the points \mathbf{z}_i^0 all lie outside each other's light-cones, and similarly with the final values s_i' . Thus

$$(\mathbf{z}_i^0 - \mathbf{z}_j^0)^2 < 0, \quad (\mathbf{z}_i' - \mathbf{z}_j')^2 < 0 \quad (i \neq j). \quad (1.6)$$

Before making variations in S , one should replace S_1 by

$$S'_1 = \sum_i m_i \int \sqrt{\mathbf{v}_i^2} ds_i, \quad (1.7)$$

so as to make S homogeneous of degree zero in the differential elements ds_i , \mathbf{v}_i counting as being of degree -1 . The expression for S is then valid with s_i any parameter on the world-line of the *i*th particle, so that \mathbf{v}_i defined by (1.3) does not necessarily satisfy (1.4).

Let us make variations $\partial\mathbf{z}_i(s_i)$ in the world-lines of the particles, $\partial\mathbf{M}(\mathbf{x})$ in the field function $\mathbf{M}(\mathbf{x})$, and Ds'_i in the final values of the s_i , so that the end-points of the world-lines are changed by

$$D\mathbf{z}'_i = \partial\mathbf{z}'_i + \mathbf{v}'_i Ds'_i, \quad (1.8)$$

$\partial\mathbf{z}'_i$ being written for $\partial\mathbf{z}_i(s'_i)$. The initial values of the s_i and the initial points of the world-lines we shall suppose for simplicity to be fixed, since variations

in them would give rise to terms of the same form as those arising from variations in the final values and would not lead to anything new.

Varying S'_1 given by (1.7) and using (1.4) after the variation process, one gets

$$\begin{aligned}\partial S'_1 &= \Sigma_i m_i \left\{ \int_{s_i^0}^{s'_i} (\mathbf{v}_i, \partial \mathbf{v}_i) ds_i + Ds'_i \right\} \\ &= \Sigma_i m_i \left\{ - \int_{s_i^0}^{s'_i} (d\mathbf{v}_i/ds_i, \partial \mathbf{z}_i) ds_i + (\mathbf{v}'_i, \partial \mathbf{z}'_i) + Ds'_i \right\} \\ &= \Sigma_i m_i \left\{ - \int_{s_i^0}^{s'_i} (d\mathbf{v}_i/ds_i, \partial \mathbf{z}_i) ds_i + (\mathbf{v}'_i, D\mathbf{z}'_i) \right\} \quad (1.9)\end{aligned}$$

with the help of (1.8). To get the variation in S_2 given by (1.2) one may, owing to the symmetry between i and j in the double sum, vary only quantities involving i and multiply by 2. The result is

$$\begin{aligned}\partial S_2 &= \Sigma_i \Sigma_{j \neq i} e_i e_j \left\{ \int_{s_i^0}^{s'_i} \int_{s_j^0}^{s'_j} \left[\frac{\partial \delta(\mathbf{z}_i - \mathbf{z}_j)^2}{\partial z_{\mu i}} \partial z_{\mu i} (\mathbf{v}_i, \mathbf{v}_j) + \delta(\mathbf{z}_i - \mathbf{z}_j)^2 (\mathbf{v}_j, \partial \mathbf{v}_i) \right] ds_i ds_j \right. \\ &\quad \left. + \int_{s_j^0}^{s'_j} \delta(\mathbf{z}'_i - \mathbf{z}_j)^2 (\mathbf{v}'_i, \mathbf{v}_j) Ds'_i ds_j \right\} \\ &= \Sigma_i \Sigma_{j \neq i} e_i e_j \left\{ \int_{s_i^0}^{s'_i} \int_{s_j^0}^{s'_j} \left[\frac{\partial \delta(\mathbf{z}_i - \mathbf{z}_j)^2}{\partial z_{\mu i}} (\mathbf{v}_i, \mathbf{v}_j) - \frac{d}{ds_i} \{ \delta(\mathbf{z}_i - \mathbf{z}_j)^2 v_{ji}^\mu \} \right] \partial z_{\mu i} ds_i ds_j \right. \\ &\quad \left. + \int_{s_j^0}^{s'_j} \delta(\mathbf{z}'_i - \mathbf{z}_j)^2 (\mathbf{v}_j, D\mathbf{z}'_i) ds_j \right\}. \quad (1.10)\end{aligned}$$

In varying S_3 given by (1.5), one has to take into account that the total variation in \mathbf{M} at a point $\mathbf{z}_i(s_i)$ on the i th world-line, let us call it $D\mathbf{M}(\mathbf{z}_i)$, consists of two parts, a part $\partial \mathbf{M}(\mathbf{z}_i)$ arising from the variation in the function $\mathbf{M}(\mathbf{x})$ and equal to the value of $\partial \mathbf{M}(\mathbf{x})$ at the point $\mathbf{x} = \mathbf{z}_i$, and a part arising from the variation in \mathbf{z}_i , equal to $\partial \mathbf{M}/\partial x_\nu$ at the point $\mathbf{x} = \mathbf{z}_i$ multiplied into $\partial z_{\nu i}$; thus

$$D\mathbf{M}(\mathbf{z}_i) = \partial \mathbf{M}(\mathbf{z}_i) + (\partial \mathbf{M}/\partial x_\nu)_{\mathbf{z}_i} \partial z_{\nu i}. \quad (1.11)$$

The variation in S_3 is now

$$\begin{aligned}\partial S_3 &= \Sigma_i e_i \left\{ \int_{s_i^0}^{s'_i} [D\mathbf{M}^\mu(\mathbf{z}_i) v_{\mu i} + M^\mu(\mathbf{z}_i) \partial v_{\mu i}] ds_i + M^\mu(\mathbf{z}'_i) v'_{\mu i} Ds'_i \right\} \\ &= \Sigma_i e_i \left\{ \int_{s_i^0}^{s'_i} \left[\partial M^\mu(\mathbf{z}_i) v_{\mu i} + \left(\frac{\partial M^\mu}{\partial x_\nu} \right)_{\mathbf{z}_i} v_{\mu i} \partial z_{\nu i} - \frac{d M^\mu(\mathbf{z}_i)}{ds_i} \partial z_{\mu i} \right] ds_i + M^\mu(\mathbf{z}'_i) Dz'_{\mu i} \right\}. \quad (1.12)\end{aligned}$$

The total variation in S is given by the sum of the three expressions (1·9), (1·10) and (1·12).

By equating to zero the total coefficient of $\partial z_{\mu i}$, one gets the equation of motion of the i th particle. It is

$$-m_i \frac{dv_i^\mu}{ds_i} + e_i \sum_{j \neq i} e_j \int_{s_j^0}^{s'_j} \left[\frac{\partial \delta(\mathbf{z}_i - \mathbf{z}_j)^2}{\partial z_{\mu i}} (\mathbf{v}_i, \mathbf{v}_j) - \frac{d}{ds_i} \{ \delta(\mathbf{z}_i - \mathbf{z}_j)^2 v_j^\mu \} \right] ds_j + e_i \left[\left(\frac{\partial M^\nu}{\partial x_\mu} \right)_{\mathbf{z}_i} v_{\nu i} - \frac{dM^\mu(\mathbf{z}_i)}{ds_i} \right] = 0.$$

Introducing the field function

$$A_i^\nu(\mathbf{x}) = M^\nu(\mathbf{x}) + \sum_{j \neq i} e_j \int_{s_j^0}^{s'_j} \delta(\mathbf{x} - \mathbf{z}_j)^2 v_j^\nu ds_j, \quad (1·13)$$

it may be written

$$m_i \frac{dv_i^\mu}{ds_i} = e_i \left[\left(\frac{\partial A_i^\nu}{\partial x_\mu} \right)_{\mathbf{z}_i} v_{\nu i} - \frac{dA_i^\mu(\mathbf{z}_i)}{ds_i} \right] = e_i \left[\frac{\partial A_i^\nu}{\partial x_\mu} - \frac{\partial A_i^\mu}{\partial x_\nu} \right]_{\mathbf{z}_i} v_{\nu i}. \quad (1·14)$$

It is the correct Lorentz equation of motion of the i th particle, provided A_i^ν is connected with the ingoing and outgoing fields and the retarded and advanced fields of the other particles by the relation, given by the author (1938, equation (42)),

$$A_i^\nu = \frac{1}{2}[A_{\text{in}}^\nu + A_{\text{out}}^\nu] + \frac{1}{2} \sum_{j \neq i} [A_{j \text{ ret}}^\nu + A_{j \text{ adv}}^\nu]$$

or $A_i^\nu(\mathbf{x}) = \frac{1}{2}[A_{\text{in}}^\nu(\mathbf{x}) + A_{\text{out}}^\nu(\mathbf{x})] + \sum_{j \neq i} e_j \int_{-\infty}^{\infty} \delta(\mathbf{x} - \mathbf{z}_j)^2 v_j^\nu ds_j. \quad (1·15)$

According to (1·13) this requires

$$M^\nu(\mathbf{x}) = \frac{1}{2}[A_{\text{in}}^\nu(\mathbf{x}) + A_{\text{out}}^\nu(\mathbf{x})] + \sum_{j \neq i} e_j \left[\int_{-\infty}^{s_j^0} + \int_{s'_j}^{\infty} \right] \delta(\mathbf{x} - \mathbf{z}_j)^2 v_j^\nu ds_j. \quad (1·16)$$

There is no need for equations (1·15), (1·16) to hold throughout space-time. It is sufficient to have them holding in the region which lies inside the future light-cone from \mathbf{z}_i^0 and inside the past light-cone from \mathbf{z}'_i , since it is only the value of $A_i^\nu(\mathbf{x})$ in this region that comes into play in the equation of motion (1·14) of the i th particle. In this region (1·16) may be replaced by

$$M^\nu(\mathbf{x}) = \frac{1}{2}[A_{\text{in}}^\nu(\mathbf{x}) + A_{\text{out}}^\nu(\mathbf{x})] + \sum_j e_j \left[\int_{-\infty}^{s_j^0} + \int_{s'_j}^{\infty} \right] \delta(\mathbf{x} - \mathbf{z}_j)^2 v_j^\nu ds_j, \quad (1·17)$$

with the sum taken over all values of j . By assuming (1·17) to hold throughout space-time one gets an expression for $M^\nu(\mathbf{x})$ independent of i , so that *the equations of motion of all the particles follow from the same action integral*. This is an important feature of Fokker's action principle.

Instead of (1.17) one may take

$$M^\nu(\mathbf{x}) = \frac{1}{2}[A_{\text{in}}^\nu(\mathbf{x}) + A_{\text{out}}^\nu(\mathbf{x})] + \frac{1}{2}\sum_j e_j \left[\int_{-\infty}^{s_j^0} - \int_{s'_j}^{\infty} \right] \Delta(\mathbf{x} - \mathbf{z}_j) v_j^\nu ds_j, \quad (1.18)$$

where $\Delta(\mathbf{y})$ denotes the Jordan & Pauli (1928) Δ function of any 4-vector \mathbf{y} , connected with the δ function of \mathbf{y}^2 by

$$\Delta(\mathbf{y}) = \pm 2\delta(\mathbf{y}^2), \quad (1.19)$$

the + or - sign being taken according to whether y_0 is positive or negative, respectively. (The Δ function used here and in equation (9) has the opposite sign to that of Jordan & Pauli.) The right-hand sides of (1.17) and (1.18) differ only inside the past light-cones from the points \mathbf{z}_j^0 and inside the future light-cones from the points \mathbf{z}'_j , and the field in these places does not enter into the equation of motion of any of the particles, on account of (1.6). (1.18) will be adopted here in preference to (1.17), since the former makes

$$\square M^\nu(\mathbf{x}) = 0, \quad (1.20)$$

on account of the general relation

$$\square \Delta(\mathbf{x}) = 0 \quad (1.21)$$

given by Jordan & Pauli (1928, equation (14)).

A difficulty appears in the work at this stage. In varying the action integral it was implicitly assumed, by the use of (1.11), that $M^\nu(\mathbf{x})$ is a continuous field function in the neighbourhood of the world-lines of the particles, but the expression subsequently obtained for it, (1.18), has discontinuities. One can easily see that those discontinuities on the world-line of any particle due to the terms in the sum in (1.18) referring to the other particles are not serious and do not invalidate the equations of motion (they produce the same sort of effects as a pulse in the incident radiation would). But those at the end-points of the world-line of any particle due to the term in the sum in (1.18) referring to that particle are serious, as they prevent the passage to the Hamiltonian formulation.

To get over the difficulty one must introduce a certain limiting process. Let λ be a small 4-vector whose direction is within the future light-cone, so that

$$\lambda^2 > 0, \quad \lambda_0 > 0, \quad (1.22)$$

and let the $\delta(\mathbf{z}_i - \mathbf{z}_j)^2$ in (1.2) be replaced by

$$\begin{aligned} \delta^*(\mathbf{z}_i; \mathbf{z}_j) &= \delta(\mathbf{z}_i - \mathbf{z}_j + \lambda)^2 \quad \text{for } (z_i - z_j)_0 > 0 \\ &= \delta(\mathbf{z}_i - \mathbf{z}_j - \lambda)^2 \quad \text{for } (z_i - z_j)_0 < 0. \end{aligned} \quad (1.23)$$

This will cause interaction to occur between pairs of particles at points on their world-lines that lie one just outside the light-cone from the other, instead of on it as previously. (It is assumed that two particles never

approach to a distance of the order λ from one another.) One now has to replace δ by δ^* in all the equations as far as (1.17), and must replace (1.18) by

$$M^\nu(\mathbf{x}) = \frac{1}{2}[A_{\text{in}}^\nu(\mathbf{x}) + A_{\text{out}}^\nu(\mathbf{x})] + \frac{1}{2}\sum_j e_j \left[\int_{-\infty}^{s_j^0} \Delta(\mathbf{x} - \mathbf{z}_j + \boldsymbol{\lambda}) v_j^\nu ds_j - \int_{s_j'}^\infty \Delta(\mathbf{x} - \mathbf{z}_j - \boldsymbol{\lambda}) v_j^\nu ds_j \right], \quad (1.24)$$

the restrictions (1.6) having to be sharpened to

$$(\mathbf{z}_i^0 - \mathbf{z}_j^0 \pm \boldsymbol{\lambda})^2 < 0, \quad (\mathbf{z}_i' - \mathbf{z}_j' \pm \boldsymbol{\lambda})^2 < 0 \quad (i \neq j). \quad (1.25)$$

The singularities in $M^\nu(\mathbf{x})$ are now displaced a little from the end-points of the world-lines and no longer come into play, even if the initial points \mathbf{z}_i^0 are also varied, provided the variations in the world-lines are smaller than λ . The equations of motion resulting from the variation principle are no longer exact, but are correct only in the limit $\lambda \rightarrow 0$. The limiting process here introduced appears to be unavoidable for setting up a satisfactory action theory and plays an important role in the later development, leading to the elimination of those divergent integrals (7) in quantum electrodynamics for which n is even.

One can now pass to the Hamiltonian formulation of the equations of motion. For each point in space-time \mathbf{x} , $M^\mu(\mathbf{x})$ may be counted as a coordinate, depending on the proper-times s_i' (it also depends on the proper-times s_i^0 , but this does not here concern us), and will have a conjugate momentum, $K_\mu(\mathbf{x})$ say. These momenta, together with the particle momenta p_i^μ , are defined, as in the general theory of Weiss (1936), by the coefficients of $\partial M^\mu(\mathbf{x})$ and $Dz'_{\mu i}$ in the expression for ∂S , so that

$$\partial S = \Sigma_i p_i^\mu Dz'_{\mu i} + \iiint_{-\infty}^{\infty} K_\mu(\mathbf{x}) \partial M^\mu(\mathbf{x}) dx_0 dx_1 dx_2 dx_3. \quad (1.26)$$

Comparing (1.26) with the sum of (1.9), (1.10) and (1.12), with δ^* written for δ , one obtains

$$K_\mu(\mathbf{x}) = \Sigma_i e_i \int_{s_i^0}^{s_i'} \delta(x_0 - z_{0i}) \delta(x_1 - z_{1i}) \delta(x_2 - z_{2i}) \delta(x_3 - z_{3i}) v_{\mu i} ds_i \quad (1.27)$$

$$\begin{aligned} \text{and } p_i^\mu &= m_i v_i'^\mu + e_i \left[M^\mu(\mathbf{z}_i') + \Sigma_{j \neq i} e_j \int_{s_j^0}^{s_j'} \delta^*(\mathbf{z}_i' - \mathbf{z}_j) v_j^\mu ds_j \right] \\ &= m_i v_i'^\mu + e_i \left[M^\mu(\mathbf{z}_i') + \frac{1}{2} \sum_j e_j \int_{s_j^0}^{s_j} \Delta(\mathbf{z}_i' - \mathbf{z}_j + \boldsymbol{\lambda}) v_j^\mu ds_j \right] \end{aligned} \quad (1.28)$$

when one takes into account (1.23), (1.19) and the fact that a non-vanishing integrand can occur in (1.28) only when $j \neq i$ and $(\mathbf{z}_i' - \mathbf{z}_j + \boldsymbol{\lambda})_0 > 0$.

The momenta satisfy the Poisson Bracket relationships

$$[p_{\mu i}, z_{\nu j}] = g_{\mu\nu} \delta_{ij}, \quad (1.29)$$

$$[K_\mu(\mathbf{x}), M_\nu(\mathbf{x}')] = g_{\mu\nu} \delta(x_0 - x'_0) \delta(x_1 - x'_1) \delta(x_2 - x'_2) \delta(x_3 - x'_3), \quad (1.30)$$

and, of course, the Poisson Bracket of any two momenta or of any two co-ordinates vanishes. Instead of $K_\mu(\mathbf{x})$ it is more convenient to work with the momentum field function $N_\mu(\mathbf{x})$ defined by

$$N_\mu(\mathbf{x}) = \frac{1}{2} \int \int \int \int_{-\infty}^{\infty} \Delta(\mathbf{x} - \mathbf{x}') K_\mu(\mathbf{x}') dx'_0 dx'_1 dx'_2 dx'_3 \quad (1.31)$$

and satisfying

$$\square N_\mu(\mathbf{x}) = 0 \quad (1.32)$$

in virtue of (1.21). Instead of (1.30) one has

$$[N_\mu(\mathbf{x}), M_\nu(\mathbf{x}')] = \frac{1}{2} g_{\mu\nu} \Delta(\mathbf{x} - \mathbf{x}'). \quad (1.33)$$

From (1.31) and (1.27)

$$N_\mu(\mathbf{x}) = \frac{1}{2} \sum_i e_i \int_{s_i^0}^{s_i'} \Delta(\mathbf{x} - \mathbf{z}_i) v_{\mu i} ds_i, \quad (1.34)$$

so that (1.28) may be written

$$\begin{aligned} p_i^\mu &= m_i v_i'^\mu + e_i [M^\mu(\mathbf{z}'_i) + N^\mu(\mathbf{z}'_i + \boldsymbol{\lambda})] \\ &= m_i v_i'^\mu + e_i A^\mu(\mathbf{z}'_i), \end{aligned} \quad (1.35)$$

where

$$A^\mu(\mathbf{x}) = M^\mu(\mathbf{x}) + N^\mu(\mathbf{x} + \boldsymbol{\lambda}). \quad (1.36)$$

From (1.20) and (1.32) the potentials $A_\mu(\mathbf{x})$ satisfy

$$\square A_\mu(\mathbf{x}) = 0, \quad (1.37)$$

showing that they can be resolved into waves travelling with the velocity of light, and from (1.33)

$$[A_\mu(\mathbf{x}), A_\nu(\mathbf{x}')] = \frac{1}{2} g_{\mu\nu} \{ \Delta(\mathbf{x} - \mathbf{x}' + \boldsymbol{\lambda}) + \Delta(\mathbf{x} - \mathbf{x}' - \boldsymbol{\lambda}) \}. \quad (1.38)$$

From (1.4) and (1.35)

$$F_i \equiv \{ \mathbf{p}_i - e_i \mathbf{A}(\mathbf{z}'_i) \}^2 - m_i^2 = 0. \quad (1.39)$$

There is one of these equations for each particle. The expressions F_i may be used as Hamiltonians to determine how any dynamical variable ξ varies with the proper-times s'_i , in accordance with the equations

$$\kappa_i d\xi/ds'_i = [\xi, F_i]. \quad (1.40)$$

Here ξ is any function of the co-ordinates and momenta of the particles and of the fields \mathbf{M} , \mathbf{K} , \mathbf{N} , \mathbf{A} , and the κ 's are multiplying factors not depending on ξ . Taking $\xi = z_i'^\mu$, one finds that

$$\kappa_i = -2m_i, \quad (1.41)$$

to get agreement with (1.35). Taking $\xi = p_i^\mu$ gives one back the equation of motion (1.14) with the λ refinement. Taking $\xi = M_\mu(\mathbf{x})$, one gets

$$\begin{aligned} dM_\mu(\mathbf{x})/ds'_i &= e_i v'_i{}^\nu [M_\mu(\mathbf{x}), A_\nu(\mathbf{z}'_i)] \\ &= \frac{1}{2} e_i v'_{\mu i} \Delta(\mathbf{x} - \mathbf{z}'_i - \boldsymbol{\lambda}), \end{aligned} \quad (1.42)$$

from (1.36) and (1.33). This equation of motion for the field quantities $M_\mu(\mathbf{x})$ does not follow from the variation principle, as it involves only co-ordinates and velocities and not accelerations, and it has to be imposed as an extra condition in the variational method. With the method of development given here, it comes in as a consequence of the choice (1.24) for $M_\mu(\mathbf{x})$.

The above Hamiltonian formulation of the equations of classical electrodynamics may be taken over into the quantum theory in the usual way, by making the momenta into operators satisfying commutation relations corresponding to the Poisson Bracket relations (1.29), (1.30). Equation (1.38) in the limit $\lambda \rightarrow 0$ goes over into the quantum equation (9). The Hamiltonians (1.39) provide the wave equations

$$F_i \psi = 0, \quad (1.43)$$

in which the wave function ψ is a function of the co-ordinates \mathbf{z}'_i of all the particles, subject to the inequalities (1.25), and of the field variables $M_\mu(\mathbf{x})$. One can apply the theory to spinning electrons instead of spinless particles, by modifying the Hamiltonians F_i in the appropriate way.

The present form of quantum electrodynamics differs from that of Heisenberg & Pauli in two respects. First, it involves the λ limiting process. This limiting process may be introduced into the theory of Heisenberg & Pauli, as shown by the author (1939b), and leads to the elimination of the divergent integrals with even n -values. Secondly, it is expressed in terms of a different representation, with the wave functions involving the $M_\mu(\mathbf{x})$ as co-ordinates. This change in the representation eliminates the divergent integrals with odd n -values. The new representation cannot be interpreted in terms of only positive-energy photons, as is the case with the representation of Heisenberg & Pauli, but it can be interpreted in terms of both positive- and negative-energy photons, as will be shown in Appendix II.

APPENDIX II. RELATIVISTIC SECOND QUANTIZATION

The author's original method of second quantization for an assembly of similar Bose particles, consisting in the direct introduction as dynamical variables of the numbers of particles in the various quantum states for

a particle, is not very suitable for the purpose of relativistic generalization. Fock (1934, § I) gave an improved method, which presents the basic ideas in a more direct manner and which allows itself to be easily developed into a relativistic theory. The main features of Fock's method will first be mentioned here, to provide a starting point. The author's (1939a) bracket notation will be used, as it seems to be the neatest and most concise one for this kind of work.

A state of an assembly of u particles satisfying the Bose statistics is represented by a symmetrical wave function $\langle q' q'' \dots q^u | \rangle$, where each value for a variable q represents a state for one particle. Let us suppose for definiteness that the values a q can take on are discrete. Then $|\langle q' q'' \dots q^u | \rangle|^2$ is the probability of the first particle being in the state q' , the second in the state q'' , and so on. Since, however, there is no means of distinguishing one particle from another, the only physically significant interpretation of the wave function is that, if n^a, n^b, \dots are the numbers of particles in the various states q^a, q^b, \dots when the distribution q', q'', \dots, q^u occurs, so that $\Sigma n = u$, then

$$(u!/n^a! n^b! \dots) |\langle q' q'' \dots q^u | \rangle|^2 \quad (2.1)$$

is the total probability of there being these numbers of particles in the various states.

If the number of particles in the assembly is indeterminate, a state of the assembly must be represented by a series of component wave functions

$$\langle \cdot | \rangle, \langle q' | \rangle, \langle q' q'' | \rangle, \langle q' q'' q''' | \rangle, \dots \quad (2.2)$$

referring successively to 0, 1, 2, 3, ... particles, those referring to two or more particles being symmetrical. Each component $\langle q' q'' \dots q^u | \rangle$ determines the probability of there being u particles distributed in any way over the various states, in accordance with (2.1). $\Sigma_{q' \dots q^u} |\langle q' q'' \dots q^u | \rangle|^2$ is the total probability of there being u particles in the assembly, and this quantity summed for all u must equal unity.

Fock's method consists in introducing a set of variables $\xi_{q'}, \xi_{q''}, \dots$, one for each value of q or for each independent state of a particle, and representing a state of the assembly by the function $\langle \xi | \rangle$ of the ξ 's

$$\langle \xi | \rangle = \Sigma_u u!^{-\frac{1}{2}} \Sigma_{q' \dots q^u} \xi_{q'} \xi_{q''} \dots \xi_{q^u} \langle q' q'' \dots q^u | \rangle. \quad (2.3)$$

This function is a polynomial in the ξ 's, the terms of any degree u corresponding to there being u particles in the assembly. The terms of the form $\xi_{q^a}^{n^a} \xi_{q^b}^{n^b} \dots$ correspond to there being n^a particles in the state q^a , n^b in the state q^b , and so on, and hence the operator $\xi_{q^a} \partial / \partial \xi_{q^a}$ gives the number of

particles n^a in any state q^a . From (2.1) it is easily seen that the probability of there being n^a particles in the state q^a , n^b in the state q^b , and so on, is

$$n^a! n^b! \dots | \text{coefficient of } \xi_{q^a}^{n^a} \xi_{q^b}^{n^b} \dots \text{ in } \langle \xi | \rangle |^2. \quad (2.4)$$

The transformation from the set of components (2.2) to the function $\langle \xi | \rangle$ is a linear transformation and is therefore expressible in terms of a transformation function $\langle \xi | q' q'' \dots q^u \rangle$, thus

$$\langle \xi | \rangle = \sum_u \sum_{q' \dots q^u} \langle \xi | q' q'' \dots q^u \rangle \langle q' q'' \dots q^u | \rangle. \quad (2.5)$$

Comparing (2.5) with (2.3), one finds

$$\langle \xi | q' q'' \dots q^u \rangle = u!^{-\frac{1}{2}} \xi_{q'} \xi_{q''} \dots \xi_{q^u}. \quad (2.6)$$

The operators ξ_q , $\partial/\partial\xi_q$, when operating to the right, correspond to the emission of a particle into and the absorption of a particle from the state q , respectively. Let us see what these operators go over into when transformed to the q representation (2.2). Using (2.6) and the laws of the transformation theory, one gets

$$\begin{aligned} \langle \xi | \xi_q | q' \dots q^u \rangle &= \xi_q \langle \xi | q' \dots q^u \rangle = u!^{-\frac{1}{2}} \xi_q \xi_{q'} \dots \xi_{q^u} \\ &= (u+1)^{\frac{1}{2}} \langle \xi | q' \dots q^u \rangle, \end{aligned}$$

so that

$$\xi_q | q' \dots q^u \rangle = (u+1)^{\frac{1}{2}} | q' \dots q^u \rangle. \quad (2.7)$$

This shows that the emission operator ξ_q , operating on one of the basic vectors of the q representation referring to a state of u particles, changes it into a basic vector referring to a state of $(u+1)$ particles, with the coefficient $(u+1)^{\frac{1}{2}}$. Again

$$\begin{aligned} \langle \xi | \frac{\partial}{\partial \xi_q} | q' \dots q^u \rangle &= \frac{\partial}{\partial \xi_q} \langle \xi | q' \dots q^u \rangle = u!^{-\frac{1}{2}} \frac{\partial}{\partial \xi_q} \xi_{q'} \dots \xi_{q^u} \\ &= u!^{-\frac{1}{2}} \sum_{r=1}^u \delta_{q q^r} \xi_{q'} \dots \xi_{q^{r-1}} \xi_{q^{r+1}} \dots \xi_{q^u} \\ &= u^{-\frac{1}{2}} \sum_r \delta_{q q^r} \langle \xi | q' \dots q^{r-1} q^{r+1} \dots q^u \rangle, \end{aligned}$$

so that

$$\frac{d}{d \xi_q} | q' \dots q^u \rangle = u^{-\frac{1}{2}} \sum_r \delta_{q q^r} | q' \dots q^{r-1} q^{r+1} \dots q^u \rangle. \quad (2.8)$$

This shows the effect of the absorption operator $d/d\xi_q$ operating on one of the basic vectors of the q representation.

One can now verify that ξ_q and $\partial/\partial\xi_q$ are adjoint operators when expressed in the q representation. One must make use of

$$\langle q' q'' \dots q^u | q^{u+1} q^{u+2} \dots q^{2u} \rangle = u!^{-1} \delta(q' q'' \dots q^u; q^{u+1} q^{u+2} \dots q^{2u}), \quad (2.9)$$

where $\delta(q' q'' \dots q^u; q^{u+1} q^{u+2} \dots q^{2u})$ equals unity if $q^{u+1}, q^{u+2} \dots q^{2u}$ are any permutation of q', q'', \dots, q^u and zero otherwise. The coefficient $u!^{-1}$ on the

right of (2.9) is needed to make $\langle q' q'' \dots q^u | q^{u+1} q^{u+2} \dots q^{2u} \rangle$ equivalent to the unit matrix when multiplied into symmetrical functions of u variables q by matrix multiplication. From (2.7)

$$\begin{aligned} & \langle q^{u+1} q^{u+2} \dots q^{2u+1} | \xi_q | q' \dots q^u \rangle \\ &= (u+1)^{\frac{1}{2}} \langle q^{u+1} q^{u+2} \dots q^{2u+1} | q q' \dots q^u \rangle \\ &= (u+1)^{\frac{1}{2}} (u+1)!^{-1} \delta(q^{u+1} q^{u+2} \dots q^{2u+1}; q q' \dots q^u), \end{aligned} \quad (2.10)$$

and from (2.8)

$$\begin{aligned} & \langle q' \dots q^u | \partial/\partial \xi_q | q^{u+1} \dots q^{2u+1} \rangle \\ &= (u+1)^{-\frac{1}{2}} \sum_{r=u+1}^{2u+1} \delta_{qq^r} \langle q' \dots q^u | q^{u+1} \dots q^{r-1} q^{r+1} \dots q^{2u+1} \rangle \\ &= (u+1)^{-\frac{1}{2}} u!^{-1} \sum_{r=u+1}^{2u+1} \delta_{qq^r} \delta(q' \dots q^u; q^{u+1} \dots q^{r-1} q^{r+1} \dots q^{2u+1}) \\ &= (u+1)^{-\frac{1}{2}} u!^{-1} \delta(q q' \dots q^u; q^{u+1} \dots q^{2u+1}). \end{aligned} \quad (2.11)$$

The expressions (2.10) and (2.11) are real and equal and the other kinds of matrix elements of ξ_q and $\partial/\partial \xi_q$ all vanish, and hence ξ_q and $\partial/\partial \xi_q$ are adjoint.

The above theory may be made relativistic in a straightforward way. For particles with no spin, one replaces each q by \mathbf{x} , denoting the four co-ordinates of the particle in space-time. Thus one gets, instead of (2.2), the series of component wave functions

$$\langle \cdot | \rangle, \langle \mathbf{x}' | \rangle, \langle \mathbf{x}' \mathbf{x}'' | \rangle, \langle \mathbf{x}' \mathbf{x}'' \mathbf{x}''' | \rangle, \dots, \quad (2.12)$$

those beyond the second being symmetrical. Let us take for simplicity the case of particles of zero rest-mass. Then the various components (2.12) satisfy

$$\square_{\mathbf{x}^r} \langle \mathbf{x}' \mathbf{x}'' \dots \mathbf{x}^u | \rangle = 0, \quad \mathbf{x}^r = \mathbf{x}', \mathbf{x}'', \dots, \mathbf{x}^u. \quad (2.13)$$

One now has to incorporate the Gordon-Klein rule (5) for the probability. This requires one to take for the scalar product $\langle a | b \rangle$ of a wave function $\langle \mathbf{x} | b \rangle$ and a conjugate complex wave function $\langle a | \mathbf{x} \rangle$ referring to a single particle

$$\langle a | b \rangle = \frac{1}{4\pi i} \iiint \left[\frac{\partial}{\partial x_0} \langle a | \mathbf{x} \rangle \cdot \langle \mathbf{x} | b \rangle - \langle a | \mathbf{x} \rangle \frac{\partial}{\partial x_0} \langle \mathbf{x} | b \rangle \right] dx_1 dx_2 dx_3,$$

the integral being taken over any three-dimensional surface $x_0 = \text{constant}$. This result may be written for brevity

$$\langle a | b \rangle = (4\pi i)^{-1} \int \langle a | \mathbf{x} \rangle (\Gamma - \Gamma) \langle \mathbf{x} | b \rangle dx, \quad (2.14)$$

where dx means $dx_1 dx_2 dx_3$ and Γ and Γ' mean $\partial/\partial x_0$ operating on the expression to the left and to the right respectively.

As examples of (2·14), it follows that

$$\langle a|x'' \rangle = (4\pi i)^{-1} \int \langle a|x' \rangle (\Gamma' - \Gamma') \langle x'|x'' \rangle dx'$$

and $\langle x'|b \rangle = (4\pi i)^{-1} \int \langle x'|x'' \rangle (\Gamma'' - \Gamma'') \langle x''|b \rangle dx''.$

These results, considered as equations for the unknown function $\langle x'|x'' \rangle$, have the solution

$$\langle x'|x'' \rangle = -i\Delta(x' - x''), \quad (2·15)$$

as is easily seen by taking the integrals in them to be over the three-dimensional surface $x'_0 = x''_0$ and using the property of the Δ function given by Heisenberg & Pauli (1929a, equation (64))

$$\frac{\partial \Delta(\mathbf{x})}{\partial x_0} = 4\pi \delta(x_1) \delta(x_2) \delta(x_3) \quad \text{for } x_0 = 0. \quad (2·16)$$

Equation (2·15), together with

$$\square \langle \mathbf{x} \rangle = 0, \quad \square |\mathbf{x} \rangle = 0, \quad (2·17)$$

give the relations between the basic vectors $|\mathbf{x}\rangle$, $\langle \mathbf{x}|$ of the present representation for one particle, replacing the conditions of the non-relativistic theory that these vectors are all normalized and orthogonal to one another.

If there are u particles present, equation (2·14) must be generalized to

$$\langle a|b \rangle = (4\pi i)^{-u} \iint \dots \langle a|x' x'' \dots x^u \rangle (\Gamma' - \Gamma') (\Gamma'' - \Gamma'') \dots (\Gamma^u - \Gamma^u) \langle x' x'' \dots x^u | b \rangle dx' dx'' \dots dx^u, \quad (2·18)$$

in which the various Γ and Δ operators all commute with one another. If the number of particles is indeterminate, the right-hand side of (2·18) must be summed for all values of u . If the states a and b are the same and the wave functions are properly normalized, the integrand in (2·18), including the coefficient $(4\pi i)^{-u}$, gives the probability of there being u particles in the neighbourhoods of the points x', x'', \dots, x^u , per unit volume of each neighbourhood. Corresponding to (2·9), equation (2·15) must be generalized to

$$\begin{aligned} \langle x' x'' \dots x^u | x^{u+1} x^{u+2} \dots x^{2u} \rangle \\ = (-i)^u u!^{-1} \Sigma \Delta(x' - x^a) \Delta(x'' - x^b) \dots \Delta(x^u - x^z), \end{aligned} \quad (2·19)$$

in which a, b, \dots, z denotes some permutation of $u+1, u+2, \dots, 2u$ and the summation is over all such permutations.

One can proceed to set up the relativistic analogue of the fundamental transformation (2·3). Introducing a set of variables ξ_x , one for each point in space-time, and assuming the transformation function

$$\langle \xi | x' x'' \dots x^u \rangle = u!^{-1} \xi_{x'} \xi_{x''} \dots \xi_{x^u} \quad (2·20)$$

as the analogue of (2·6), one gets

$$\begin{aligned}\langle \xi | \rangle &= \Sigma_u (4\pi i)^{-u} \int \int \dots \langle \xi | \mathbf{x}' \mathbf{x}'' \dots \mathbf{x}^u \rangle (\Gamma' - \Gamma') \dots (\Gamma^u - \Gamma^u) \langle \mathbf{x}' \dots \mathbf{x}^u | \rangle dx' \dots dx^u \\ &= \Sigma_u (4\pi i)^{-u} u!^{-1} \int \int \dots \xi_{\mathbf{x}'} \dots \xi_{\mathbf{x}^u} (\Gamma' - \Gamma') \dots (\Gamma^u - \Gamma^u) \langle \mathbf{x}' \dots \mathbf{x}^u | \rangle dx' \dots dx^u.\end{aligned}\quad (2·21)$$

The analogue of (2·7) is

$$\xi_{\mathbf{x}} | \mathbf{x}' \dots \mathbf{x}^u \rangle = (u+1)^{\frac{1}{2}} | \mathbf{x} \mathbf{x}' \dots \mathbf{x}^u \rangle, \quad (2·22)$$

showing that $\xi_{\mathbf{x}}$ corresponds to the emission of a particle at the point \mathbf{x} in space-time. Since $\square | \mathbf{x} \mathbf{x}' \dots \mathbf{x}^u \rangle = 0$, as follows from the conjugate imaginary equation to (2·13) applied to $(u+1)$ particles, one has

$$\square \xi_{\mathbf{x}} = 0. \quad (2·23)$$

This shows that the $\xi_{\mathbf{x}}$ are not all independent, as is to be expected physically.

The lack of independence of the $\xi_{\mathbf{x}}$ prevents one from giving a direct meaning to $\partial/\partial \xi_{\mathbf{x}}$. One can, however, get an operator to play the part of the absorption operator by taking the adjoint $\xi_{\mathbf{x}}^*$ to $\xi_{\mathbf{x}}$ in the \mathbf{x} representation. With the help of (2·22) and (2·19)

$$\begin{aligned}\langle \mathbf{x}' \dots \mathbf{x}^u | \xi_{\mathbf{x}} | \mathbf{x}^{u+1} \dots \mathbf{x}^{2u-1} \rangle &= u^{\frac{1}{2}} \langle \mathbf{x}' \dots \mathbf{x}^u | \mathbf{x} \mathbf{x}^{u+1} \dots \mathbf{x}^{2u-1} \rangle \\ &= (-i)^u u^{\frac{1}{2}} u!^{-1} \Sigma \Delta(\mathbf{x}^a - \mathbf{x}) \Delta(\mathbf{x}^b - \mathbf{x}^{u+1}) \dots \Delta(\mathbf{x}^z - \mathbf{x}^{2u-1}),\end{aligned}\quad (2·24)$$

where $\mathbf{x}^a, \mathbf{x}^b, \dots, \mathbf{x}^z$ denotes some permutation of $\mathbf{x}', \mathbf{x}'', \dots, \mathbf{x}^u$ and the summation is over all such permutations. The sum in (2·24) may be replaced by

$$\Sigma_{r=1}^u \Delta(\mathbf{x}^r - \mathbf{x}) \Sigma \Delta(\mathbf{x}^a - \mathbf{x}^{u+1}) \dots \Delta(\mathbf{x}^z - \mathbf{x}^{2u-1}), \quad (2·25)$$

where $\mathbf{x}^a \dots \mathbf{x}^z$ now denotes some permutation of $\mathbf{x}' \dots \mathbf{x}^{r-1} \mathbf{x}^{r+1} \dots \mathbf{x}^u$ and the second Σ refers to all such permutations. Expression (2·25) equals

$$\Sigma_r \Delta(\mathbf{x}^r - \mathbf{x}) i^{u-1} (u-1)! \langle \mathbf{x}' \dots \mathbf{x}^{r-1} \mathbf{x}^{r+1} \dots \mathbf{x}^u | \mathbf{x}^{u+1} \dots \mathbf{x}^{2u-1} \rangle$$

from (2·19). Substituting this for the sum in (2·24) and taking the conjugate complex of both sides of the equation, one obtains

$$\begin{aligned}\langle \mathbf{x}^{u+1} \dots \mathbf{x}^{2u-1} | \xi_{\mathbf{x}}^* | \mathbf{x}' \dots \mathbf{x}^u \rangle &= i u^{-\frac{1}{2}} \Sigma_r \Delta(\mathbf{x}^r - \mathbf{x}) \langle \mathbf{x}^{u+1} \dots \mathbf{x}^{2u-1} | \mathbf{x}' \dots \mathbf{x}^{r-1} \mathbf{x}^{r+1} \dots \mathbf{x}^u \rangle\end{aligned}$$

and hence

$$\xi_{\mathbf{x}}^* | \mathbf{x}' \dots \mathbf{x}^u \rangle = -i u^{-\frac{1}{2}} \Sigma_r \Delta(\mathbf{x} - \mathbf{x}^r) | \mathbf{x}' \dots \mathbf{x}^{r-1} \mathbf{x}^{r+1} \dots \mathbf{x}^u \rangle. \quad (2·26)$$

This is the relativistic analogue of (2·8). It shows, from (1·21), that

$$\square \xi_{\mathbf{x}}^* = 0. \quad (2·27)$$

Let us determine the commutation relation connecting $\xi_{\mathbf{x}}$ and $\xi_{\mathbf{x}}^*$. From (2.22) and (2.26)

$$\begin{aligned}\xi_{\mathbf{x}}^* \xi_{\mathbf{x}^0} |\mathbf{x}' \dots \mathbf{x}^u\rangle &= (u+1)^{\frac{1}{2}} \xi_{\mathbf{x}}^* |\mathbf{x}^0 \mathbf{x}' \dots \mathbf{x}^u\rangle \\ &= -i \sum_{r=1}^u \Delta(\mathbf{x} - \mathbf{x}^r) |\mathbf{x}^0 \mathbf{x}' \dots \mathbf{x}^{r-1} \mathbf{x}^{r+1} \dots \mathbf{x}^u\rangle,\end{aligned}$$

and again

$$\begin{aligned}\xi_{\mathbf{x}^0} \xi_{\mathbf{x}}^* |\mathbf{x}' \dots \mathbf{x}^u\rangle &= -iu^{-\frac{1}{2}} \xi_{\mathbf{x}^0} \sum_{r=1}^u \Delta(\mathbf{x} - \mathbf{x}^r) |\mathbf{x}' \dots \mathbf{x}^{r-1} \mathbf{x}^{r+1} \dots \mathbf{x}^u\rangle \\ &= -i \sum_{r=1}^u \Delta(\mathbf{x} - \mathbf{x}^r) |\mathbf{x}^0 \mathbf{x}' \dots \mathbf{x}^{r-1} \mathbf{x}^{r+1} \dots \mathbf{x}^u\rangle.\end{aligned}$$

Hence $(\xi_{\mathbf{x}}^* \xi_{\mathbf{x}^0} - \xi_{\mathbf{x}^0} \xi_{\mathbf{x}}^*) |\mathbf{x}' \dots \mathbf{x}^u\rangle = -i \Delta(\mathbf{x} - \mathbf{x}^0) |\mathbf{x}' \dots \mathbf{x}^u\rangle,$

so that

$$\xi_{\mathbf{x}}^* \xi_{\mathbf{x}^0} - \xi_{\mathbf{x}^0} \xi_{\mathbf{x}}^* = -i \Delta(\mathbf{x} - \mathbf{x}^0). \quad (2.28)$$

This result may be regarded as the analogue in second quantization of the result (2.15) for a single particle.

The above relativistic theory for spinless particles of zero rest-mass may easily be extended to apply to photons. The wave function for a single particle must be made into a 4-vector with components corresponding to the four components of the electromagnetic potential. The 4-vector wave function may be considered as a function of the four coordinates \mathbf{x} and of another variable μ taking on the four values 0, 1, 2, 3, and may be written $\langle \mathbf{x}_\mu | \rangle$. The μ here is not a suffix attached to \mathbf{x} , but is an independent variable, which is written in the lower position to express that the whole function $\langle \mathbf{x}_\mu | \rangle$ is subject to the contravariant law under Lorentz transformations, and which may be raised by the usual rule

$$\langle \mathbf{x}^\mu | \rangle = g^{\mu\nu} \langle \mathbf{x}_\nu | \rangle. \quad (2.29)$$

The multiplication rule (2.14) must be extended to

$$\langle a | b \rangle = -g^{\mu\nu} (4\pi i)^{-1} \int \langle a | \mathbf{x}_\mu \rangle (\Gamma - \Gamma) \langle \mathbf{x}_\nu | b \rangle dx, \quad (2.30)$$

in which the $-$ sign is inserted with $g^{\mu\nu}$ for convenience, so as to give a factor +1 for μ and ν equal to 1, 2 or 3. Similarly, (2.15) must be extended to

$$\langle \mathbf{x}'_\mu | \mathbf{x}''_\nu \rangle = ig_{\mu\nu} \Delta(\mathbf{x}' - \mathbf{x}''). \quad (2.31)$$

In a corresponding way the wave function for u particles must be made into a tensor of rank u and may be written $\langle \mathbf{x}'_\mu \mathbf{x}''_\nu \dots \mathbf{x}^u_\tau | \rangle$, in which the μ, ν, \dots, τ are independent variables from the \mathbf{x} 's. It must satisfy the symmetry condition that it remains unchanged when any permutation is applied to

the \mathbf{x} 's and the same permutation to μ, ν, \dots, τ . The multiplication rule (2.18) must be extended to

$$\langle a|b\rangle = (-4\pi i)^{-u} \int \cdots \int \langle a|\mathbf{x}'_\mu \mathbf{x}'_\nu \cdots \mathbf{x}'_r\rangle (\Gamma' - \Gamma') (\Gamma'' - \Gamma'') \cdots (\Gamma^u - \Gamma^u) \langle \mathbf{x}'^\mu \mathbf{x}''^\nu \cdots \mathbf{x}''^r|b\rangle dx' dx'' \cdots dx^u. \quad (2.32)$$

The further development of these extensions is obvious and need not be mentioned in detail. The result is that the emission operator becomes a 4-vector $\xi_{\mathbf{x}\mu}$, and satisfies with its adjoint $\xi_{\mathbf{x}\mu}^*$ in the \mathbf{x}_μ representation the commutation relation

$$\xi_{\mathbf{x}\mu}^* \xi_{\mathbf{x}'\nu} - \xi_{\mathbf{x}'\nu} \xi_{\mathbf{x}\mu}^* = ig_{\mu\nu} \Delta(\mathbf{x} - \mathbf{x}'), \quad (2.33)$$

corresponding to (2.28).

To establish a connexion between the present theory of photons and the quantum electrodynamics of the end of Appendix I, one must put

$$M_\mu(\mathbf{x}) = (\frac{1}{2}\hbar)^{\frac{1}{2}} \xi_{\mathbf{x}\mu}, \quad N_\mu(\mathbf{x}) = (\frac{1}{2}\hbar)^{\frac{1}{2}} \xi_{\mathbf{x}\mu}^*, \quad (2.34)$$

so as to make the commutation relation (1.33) go over into (2.33). The wave function ψ of Appendix I, which is of the form of a power series in the variables $M_\mu(\mathbf{x})$, may now be identified with $\langle \xi | \rangle$. The interpretation of ψ that one would expect from the rules of non-relativistic quantum mechanics, namely, that $|\psi|^2$ gives the probability of the variables $M_\mu(\mathbf{x})$ having specified values, is no longer applicable. Instead one must work from (2.4) suitably modified to fit in with the Gordon-Klein expression for the probability for one particle. One may make a Fourier resolution of the $\xi_{\mathbf{x}\mu}$ and replace the continuous range of Fourier components by a discrete set. The wave function then becomes a power series in ξ variables referring to the Fourier components and its coefficients will give, according to the formula (3.24) of Appendix III, the probabilities of there being various numbers of photons in various momentum states, some with positive and some with negative energy. These probabilities are then to be re-interpreted, in accordance with one of the fundamental assumptions of the text, as the probabilities of various photons having been emitted or absorbed.

The real dynamical variables $M_\mu(\mathbf{x})$ appear in the present theory as operators $(\frac{1}{2}\hbar)^{\frac{1}{2}} \xi_{\mathbf{x}\mu}$ which are not self-adjoint. This marks a departure from one of the usual rules of quantum mechanics. However, the Hamiltonians F_i are self-adjoint in the limit $\lambda \rightarrow 0$, since from (1.39) and (1.36) they involve the field variables only through the linear combination $M_\mu(\mathbf{x}) + N_\mu(\mathbf{x})$ or $\xi_{\mathbf{x}\mu} + \xi_{\mathbf{x}\mu}^*$. This result is sufficient to ensure the conservation law—that if the wave function is initially normalized it always remains so—and makes the existence of other real dynamical variables not corresponding to self-adjoint operators unimportant.

APPENDIX III. THE TRANSFORMATION CONNECTING THE
TWO FORMS OF QUANTUM ELECTRODYNAMICS

The new form of quantum electrodynamics set up in Appendices I and II and the old form of Heisenberg & Pauli (amended by the introduction of the λ limiting process) involve two different representations of the same set of operators describing the electromagnetic field, and one can discuss the transformation from one of these representations to the other on the lines of the general transformation theory of quantum mechanics. One cannot, as before mentioned, carry out the transformation for solutions of the wave equations, since such solutions in the old representation contain divergent integrals, but one can carry it out for other representatives of states of the field of a sufficiently general kind to make the methods of the transformation theory applicable.

It is convenient to work with the Fourier components of the various field quantities. With the new form of quantum electrodynamics, the emission operator $\xi_{\mathbf{x}\mu}$ satisfies (2.23) and its Fourier resolution is like that of the wave function representing a single photon and reads

$$\xi_{\mathbf{x}\mu} = (2\pi)^{-1} \Sigma \iiint \xi_{\mathbf{k}\mu} e^{i(\mathbf{k}, \mathbf{x})} k_0^{-1} dk_1 dk_2 dk_3, \quad (3.1)$$

where the Σ means a summation over both values $\pm \sqrt{(k_1^2 + k_2^2 + k_3^2)}$ for k_0 , as in equation (12). The Fourier coefficient $\xi_{\mathbf{k}\mu}$ introduced by (3.1) is defined only for $\mathbf{k}^2 = 0$, and is the operator of emission of a photon with energy and momentum given by the 4-vector $\mathbf{k}\hbar$. Similarly the Fourier resolution of the absorption operator $\xi_{\mathbf{x}\mu}^*$ reads

$$\xi_{\mathbf{x}\mu}^* = (2\pi)^{-1} \Sigma \iiint \xi_{\mathbf{k}\mu}^* e^{-i(\mathbf{k}, \mathbf{x})} k_0^{-1} dk_1 dk_2 dk_3, \quad (3.2)$$

where $\xi_{\mathbf{k}\mu}^*$, the adjoint of $\xi_{\mathbf{k}\mu}$, is the operator of absorption of a photon with energy and momentum $\mathbf{k}\hbar$. The commutation relation (2.33) connecting $\xi_{\mathbf{x}\mu}$ and $\xi_{\mathbf{x}\mu}^*$ leads to the following commutation relation connecting $\xi_{\mathbf{k}\mu}$ and $\xi_{\mathbf{k}\mu}^*$

$$\left. \begin{aligned} \xi_{\mathbf{k}\mu}^* \xi_{\mathbf{k}'\nu} - \xi_{\mathbf{k}'\nu} \xi_{\mathbf{k}\mu}^* &= -g_{\mu\nu} k_0 \delta(k_1 - k'_1) \delta(k_2 - k'_2) \delta(k_3 - k'_3) \\ &\quad \text{when } k_0, k'_0 \text{ have the same sign,} \\ &= 0 \quad \text{when } k_0, k'_0 \text{ have opposite signs.} \end{aligned} \right\} \quad (3.3)$$

The potentials A_μ occurring in the Hamiltonians are, according to (1.36) and (2.34), in the limit $\lambda \rightarrow 0$,

$$A_\mu(\mathbf{x}) = M_\mu(\mathbf{x}) + N_\mu(\mathbf{x}) = (\tfrac{1}{2}\hbar)^{\frac{1}{2}} (\xi_{\mathbf{x}\mu} + \xi_{\mathbf{x}\mu}^*),$$

and thus have the Fourier resolution

$$A_\mu(\mathbf{x}) = (\frac{1}{2}\hbar)^{\frac{1}{2}}(2\pi)^{-1}\Sigma \iiint \{\xi_{\mathbf{k}\mu} e^{i(\mathbf{k}, \mathbf{x})} + \xi_{\mathbf{k}\mu}^* e^{-i(\mathbf{k}, \mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3. \quad (3.4)$$

This agrees with $\mathbf{A}^3(\mathbf{x})$ of equation (12), provided one takes

$$R_{\mathbf{k}\mu} = \hbar^{\frac{1}{2}}(2\pi)^{-1}\xi_{\mathbf{k}\mu}, \quad \bar{R}_{\mathbf{k}\mu} = \hbar^{\frac{1}{2}}(2\pi)^{-1}\xi_{\mathbf{k}\mu}^*. \quad (3.5)$$

The potentials B_μ defined by

$$B_\mu(\mathbf{x}) = M_\mu(\mathbf{x}) - N_\mu(\mathbf{x}) = (\frac{1}{2}\hbar)^{\frac{1}{2}}(\xi_{\mathbf{x}\mu} - \xi_{\mathbf{x}\mu}^*)$$

commute with the A 's, as is easily seen from (2.33), and are thus constants of the motion and also redundant variables. They have the Fourier resolution

$$B_\mu(\mathbf{x}) = (\frac{1}{2}\hbar)^{\frac{1}{2}}(2\pi)^{-1}\Sigma \iiint \{\xi_{\mathbf{k}\mu} e^{i(\mathbf{k}, \mathbf{x})} - \xi_{\mathbf{k}\mu}^* e^{-i(\mathbf{k}, \mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3. \quad (3.6)$$

From (3.5) one sees that the present B_μ is the same as that of equation (14). The $B_\mu(\mathbf{x})$ satisfy a similar commutation relation to the $A_\mu(\mathbf{x})$, equation (9), with the opposite sign on the right.

With the old form of quantum electrodynamics of Heisenberg & Pauli, the Fourier resolution of the potentials A_μ occurring in the Hamiltonians must be effected in the form

$$A_\mu(\mathbf{x}) = \hbar^{\frac{1}{2}}(2\pi)^{-1} \left. \begin{aligned} &\iiint \{\alpha_{\mathbf{k}\mu} e^{i(\mathbf{k}, \mathbf{x})} + \bar{\alpha}_{\mathbf{k}\mu} e^{-i(\mathbf{k}, \mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3, \\ &k_0 = \sqrt{(k_1^2 + k_2^2 + k_3^2)}, \end{aligned} \right\} \quad (3.7)$$

in which $\alpha_{\mathbf{k}\mu}$, $\bar{\alpha}_{\mathbf{k}\mu}$, defined only for positive k_0 , are considered as the operators of emission into and absorption from positive-energy photon states, and satisfy the commutation relation

$$\bar{\alpha}_{\mathbf{k}\mu} \alpha_{\mathbf{k}'\nu} - \alpha_{\mathbf{k}'\nu} \bar{\alpha}_{\mathbf{k}\mu} = -g_{\mu\nu} k_0 \delta(k_1 - k'_1) \delta(k_2 - k'_2) \delta(k_3 - k'_3). \quad (3.8)$$

The redundant variables $B_\mu(\mathbf{x})$, if introduced into this theory, may correspondingly be resolved into Fourier components in the form

$$B_\mu(\mathbf{x}) = \hbar^{\frac{1}{2}}(2\pi)^{-1} \left. \begin{aligned} &\iiint \{\bar{\beta}_{\mathbf{k}\mu} e^{i(\mathbf{k}, \mathbf{x})} + \beta_{\mathbf{k}\mu} e^{-i(\mathbf{k}, \mathbf{x})}\} k_0^{-1} dk_1 dk_2 dk_3, \\ &k_0 = \sqrt{(k_1^2 + k_2^2 + k_3^2)}, \end{aligned} \right\} \quad (3.9)$$

with $\beta_{\mathbf{k}\mu}$, $\bar{\beta}_{\mathbf{k}\mu}$, which are defined only for positive k_0 , commuting with $\alpha_{\mathbf{k}\mu}$, $\bar{\alpha}_{\mathbf{k}\mu}$ and satisfying the commutation relation

$$\bar{\beta}_{\mathbf{k}\mu} \beta_{\mathbf{k}'\nu} - \beta_{\mathbf{k}'\nu} \bar{\beta}_{\mathbf{k}\mu} = -g_{\mu\nu} k_0 \delta(k_1 - k'_1) \delta(k_2 - k'_2) \delta(k_3 - k'_3). \quad (3.10)$$

Comparing (3.7) with (3.4) and (3.9) with (3.6) one finds, for $k_0 > 0$,

$$\left. \begin{aligned} \sqrt{2} \alpha_{\mathbf{k}\mu} &= \xi_{\mathbf{k}\mu} - \xi^*_{-\mathbf{k}\mu}, & \sqrt{2} \bar{\alpha}_{\mathbf{k}\mu} &= -\xi_{-\mathbf{k}\mu} + \xi^*_{\mathbf{k}\mu}, \\ \sqrt{2} \beta_{\mathbf{k}\mu} &= -\xi_{-\mathbf{k}\mu} - \xi^*_{\mathbf{k}\mu}, & \sqrt{2} \bar{\beta}_{\mathbf{k}\mu} &= \xi_{\mathbf{k}\mu} + \xi^*_{-\mathbf{k}\mu}. \end{aligned} \right\} \quad (3.11)$$

It should be remembered here that the bar denotes the conjugate complex of a dynamical variable and the star denotes the adjoint of an operator, which are two different things in the present theory. From equations (3.11), one sees that $\bar{\alpha}_{\mathbf{k}\mu} = \alpha^*_{\mathbf{k}\mu}$ and $\bar{\beta}_{\mathbf{k}\mu} = -\beta^*_{\mathbf{k}\mu}$.

The wave function in the new theory is a function of the $\xi_{\mathbf{k}\mu}$ and $\xi_{-\mathbf{k}\mu}$ (keeping $k_0 > 0$), and in the old theory it is a function of the $\alpha_{\mathbf{k}\mu}$ and of certain variables of the **B** field, which may conveniently be taken to be the $\beta_{\mathbf{k}\mu}$. To set up the transformation connecting the new theory with the old, one must obtain the transformation function connecting the $\xi_{\mathbf{k}\mu}$, $\xi_{-\mathbf{k}\mu}$ representation with the $\alpha_{\mathbf{k}\mu}$, $\beta_{\mathbf{k}\mu}$ representation. For this purpose it is sufficient to deal with only one value for the vector **k** and one direction μ , since different values for these variables correspond to different degrees of freedom, for which the transformations are independent. It is convenient now to put $\xi_{-\mathbf{k}\mu} = \xi_{\mathbf{k}\mu}$ and to drop the suffixes **k** and μ , so that equations (3.11) become

$$\left. \begin{aligned} \sqrt{2} \alpha &= \xi - \xi^*, & \sqrt{2} \bar{\alpha} &= -\xi + \xi^*, \\ \sqrt{2} \beta &= -\xi - \xi^*, & \sqrt{2} \bar{\beta} &= \xi + \xi^*. \end{aligned} \right\} \quad (3.12)$$

The value of **k** concerned may be regarded as one of a discrete set (by enclosing the electromagnetic field in a finite box, or by some other way), and the commutation relations (3.3), (3.8), (3.10) then become

$$\xi^* \xi - \xi \xi^* = c, \quad \xi^* \zeta - \zeta \xi^* = -c, \quad (3.13)$$

$$\bar{\alpha} \alpha - \alpha \bar{\alpha} = c, \quad \bar{\beta} \beta - \beta \bar{\beta} = c, \quad (3.14)$$

where, if we take the case of $\mu = 1, 2$ or 3 , c is some positive number.

The ξ, ζ representation, or any representation of the type of the ξ representation introduced in Appendix II, is of a more general kind than the usual representations of quantum mechanics, since the variables ξ, ζ occurring in the wave functions do not correspond to self-adjoint operators. In dealing with such a representation one must therefore refer back to first principles. The ξ, ζ representation is defined so as to make the representative of ξ or ζ times any $| \rangle$ equal the representative of that $| \rangle$ multiplied by the ξ or ζ variable, respectively, thus

$$\langle \xi' \zeta' | \xi \rangle = \xi' \langle \xi' \zeta' | \rangle, \quad \langle \xi' \zeta' | \zeta \rangle = \zeta' \langle \xi' \zeta' | \rangle. \quad (3.15)$$

From the commutation relations (3.13) one can see that when ξ^* or ζ^* is multiplied into any $| \rangle$, its representative $\langle \xi' \zeta' | \rangle$ gets operated on by $c \partial / \partial \xi'$, $-c \partial / \partial \zeta'$, respectively, thus

$$\langle \xi' \zeta' | \xi^* | \rangle = c \frac{\partial}{\partial \xi'} \langle \xi' \zeta' | \rangle, \quad \langle \xi' \zeta' | \zeta^* | \rangle = -c \frac{\partial}{\partial \zeta'} \langle \xi' \zeta' | \rangle. \quad (3.16)$$

The conjugate complex of a representative $\langle \xi' \zeta' | \rangle$ is a function of the conjugate complex variables, ξ^{**} , ζ^{**} say, corresponding to the adjoint operators ξ^* , ζ^* , and should thus be written $\langle | \xi^{**} \zeta^{**} \rangle$. The conjugate complexes of equations (3.15) are therefore

$$\langle | \xi^* | \xi^{**} \zeta^{**} \rangle = \xi^{**} \langle | \xi^* \zeta^* \rangle, \quad \langle | \zeta^* | \xi^{**} \zeta^{**} \rangle = \zeta^{**} \langle | \xi^* \zeta^* \rangle, \quad (3.17)$$

and those of equations (3.16) are

$$\langle | \xi | \xi^{**} \zeta^{**} \rangle = c \frac{\partial}{\partial \xi^{**}} \langle | \xi^* \zeta^* \rangle, \quad \langle | \zeta | \xi^{**} \zeta^{**} \rangle = -c \frac{\partial}{\partial \zeta^{**}} \langle | \xi^* \zeta^* \rangle. \quad (3.18)$$

Let us now determine the form of the matrix $\langle \xi' \zeta' | \xi^{**} \zeta^{**} \rangle$. From (3.15)

$$\langle \xi' \zeta' | \xi | \xi^{**} \zeta^{**} \rangle = \xi' \langle \xi' \zeta' | \xi^* \zeta^* \rangle, \quad \langle \xi' \zeta' | \zeta | \xi^{**} \zeta^{**} \rangle = \zeta' \langle \xi' \zeta' | \xi^* \zeta^* \rangle,$$

and from (3.18)

$$\begin{aligned} \langle \xi' \zeta' | \xi | \xi^{**} \zeta^{**} \rangle &= c \frac{\partial}{\partial \xi^{**}} \langle \xi' \zeta' | \xi^* \zeta^* \rangle, \\ \langle \xi' \zeta' | \zeta | \xi^{**} \zeta^{**} \rangle &= -c \frac{\partial}{\partial \zeta^{**}} \langle \xi' \zeta' | \xi^* \zeta^* \rangle. \end{aligned}$$

Hence

$$\left(c \frac{\partial}{\partial \xi^{**}} - \xi' \right) \langle \xi' \zeta' | \xi^{**} \zeta^{**} \rangle = 0, \quad \left(c \frac{\partial}{\partial \zeta^{**}} + \zeta' \right) \langle \xi' \zeta' | \xi^{**} \zeta^{**} \rangle = 0.$$

Similarly, working from (3.16) and (3.17), one finds

$$\left(c \frac{\partial}{\partial \xi'} - \xi^{**} \right) \langle \xi' \zeta' | \xi^{**} \zeta^{**} \rangle = 0, \quad \left(c \frac{\partial}{\partial \zeta'} + \zeta^{**} \right) \langle \xi' \zeta' | \xi^{**} \zeta^{**} \rangle = 0.$$

Hence

$$\langle \xi' \zeta' | \xi^{**} \zeta^{**} \rangle = \exp [(\xi' \xi^{**} - \zeta' \zeta^{**})/c], \quad (3.19)$$

apart from a constant factor, which may be taken to be unity without essential loss of generality.

The law of multiplication of representatives must now be chosen so as to make the $\langle \xi' \zeta' | \xi^{**} \zeta^{**} \rangle$ given by (3.19) play the part of the unit matrix.

It is easily seen that this condition is fulfilled if one takes the product of $\langle a|\xi^{*'}\zeta^{*'}\rangle$ and $\langle\xi'\zeta'|b\rangle$ to be

$$\langle a|b\rangle = S_{\xi'} S_{\zeta'} \langle a|\xi^{*'}\zeta^{*'}\rangle \langle\xi'\zeta'|b\rangle, \quad (3.20)$$

where the symbols $S_{\xi'}$, $S_{\zeta'}$ denote certain kinds of summation, which are defined as follows.

If $f(\xi^*)$ and $g(\xi)$ are two power series in ξ^* and ξ respectively, so that

$$f(\xi^*) = \sum_{n=0}^{\infty} f_n \xi^{*n}, \quad g(\xi) = \sum_{n=0}^{\infty} g_n \xi^n,$$

then

$$S_{\xi} f(\xi^*) g(\xi) = \sum_{n=0}^{\infty} c^n n! f_n g_n. \quad (3.21)$$

The definition of S_{ξ} is similar, with $-c$ substituted for c . These definitions lead immediately to

$$S_{\xi''} S_{\zeta''} \langle\xi'\zeta'|\xi^{*''}\zeta^{*''}\rangle \langle\xi''\zeta''|b\rangle = \langle\xi'\zeta'|b\rangle,$$

$$S_{\xi'} S_{\zeta'} \langle a|\xi^{*'}\zeta^{*'}\rangle \langle\xi'\zeta'|\xi^{*''}\zeta^{*''}\rangle = \langle a|\xi^{*''}\zeta^{*''}\rangle,$$

which are the required conditions for $\langle\xi'\zeta'|\xi^{*''}\zeta^{*''}\rangle$ to play the part of the unit matrix. (3.21) is equivalent to the following explicit definition for S_{ξ} ,

$$S_{\xi} f(\xi^*) g(\xi) = (\pi c)^{-1} \int_{-\infty}^{\infty} f(\xi^*) g(\xi) e^{-\xi^*\xi/c} d\xi_R d\xi_I, \quad (3.22)$$

where ξ_R and ξ_I denote the real and pure imaginary parts of the complex variable ξ . The corresponding definition for S_{ζ} is

$$S_{\zeta} f(\zeta^*) g(\zeta) = (\pi c)^{-1} \int_{-\infty}^{\infty} f(\zeta^*) g(-\zeta) e^{-\zeta^*\zeta/c} d\zeta_R d\zeta_I. \quad (3.23)$$

The multiplication rule given by (3.20), (3.21) is connected with the physical interpretation of the wave functions $\langle\xi'\zeta'|\rangle$, namely that the probability of there having been m photons emitted into the momentum and energy state \mathbf{k} (corresponding to the variable ξ) and n photons absorbed from this state is

$$c^m m! (-c)^n n! |\text{coefficient of } \xi'^m \zeta'^n \text{ in } \langle\xi'\zeta'|\rangle|^2. \quad (3.24)$$

this being the quantity whose sum over all m and n must be put equal to unity to normalize the wave function. The form (3.24) is a generalization of (2.4), arising on account of the c and $-c$ on the right-hand sides of the commutation relations (3.13). It gives a negative probability for an odd number of photons having been absorbed, in agreement with what is stated in the text.

Let us now determine the transformation function $\langle \alpha' \beta' | \xi^* \zeta^* \rangle$. From the first of equations (3·12)

$$\sqrt{2} \langle \alpha' \beta' | \alpha | \xi^* \zeta^* \rangle = \langle \alpha' \beta' | \xi - \zeta^* | \xi^* \zeta^* \rangle,$$

which gives, according to (3·18) and (3·17)

$$\sqrt{2} \alpha' \langle \alpha' \beta' | \xi^* \zeta^* \rangle = \left(c \frac{\partial}{\partial \xi^*} - \zeta^* \right) \langle \alpha' \beta' | \xi^* \zeta^* \rangle.$$

Similarly, the other three equations (3·12) give

$$\sqrt{2} c \frac{\partial}{\partial \alpha'} \langle \alpha' \beta' | \xi^* \zeta^* \rangle = \left(c \frac{\partial}{\partial \xi^*} + \xi^* \right) \langle \alpha' \beta' | \xi^* \zeta^* \rangle,$$

$$\sqrt{2} \beta' \langle \alpha' \beta' | \xi^* \zeta^* \rangle = \left(c \frac{\partial}{\partial \zeta^*} - \xi^* \right) \langle \alpha' \beta' | \xi^* \zeta^* \rangle,$$

$$\sqrt{2} c \frac{\partial}{\partial \beta'} \langle \alpha' \beta' | \xi^* \zeta^* \rangle = \left(c \frac{\partial}{\partial \xi^*} + \zeta^* \right) \langle \alpha' \beta' | \xi^* \zeta^* \rangle.$$

The solution of these four equations is

$$\langle \alpha' \beta' | \xi^* \zeta^* \rangle = \sqrt{2} \exp [(\alpha' \beta' + \sqrt{2} \alpha' \xi^* + \sqrt{2} \beta' \zeta^* + \xi^* \zeta^*)/c], \quad (3·25)$$

in which the arbitrary multiplying factor is chosen for convenience to be $\sqrt{2}$. A different choice would only mean a different weight factor in the α, β representation.

It remains to determine the law of multiplication for representatives in the α, β representation, which can best be done by first determining the unit matrix $\langle \alpha' \beta' | \alpha^{**} \beta^{**} \rangle$. We have

$$\begin{aligned} \langle \alpha' \beta' | \alpha^{**} \beta^{**} \rangle &= S_{\xi'} S_{\zeta'} \langle \alpha' \beta' | \xi^* \zeta^* \rangle \langle \xi' \zeta' | \alpha^{**} \beta^{**} \rangle \\ &= 2S_{\xi'} S_{\zeta'} \exp [(\alpha' \beta' + \sqrt{2} \alpha' \xi^* + \sqrt{2} \beta' \zeta^* + \xi^* \zeta^*)/c] \\ &\quad \times \exp [(\alpha^{**} \beta^{**} + \sqrt{2} \alpha^{**} \xi' + \sqrt{2} \beta^{**} \zeta' + \xi' \zeta')/c]. \end{aligned}$$

Performing the operation $S_{\xi'}$ according to (3·22), one gets

$$\begin{aligned} \langle \alpha' \beta' | \alpha^{**} \beta^{**} \rangle &= 2S_{\xi'} \exp [(\alpha' \beta' + \sqrt{2} \beta' \zeta^* + \sqrt{2} \beta^{**} \zeta' + \alpha^{**} \beta^{**})/c] \\ &\quad \times (\pi c)^{-1} \int \int \exp \{[(\sqrt{2} \alpha' + \zeta^*) \xi^* + (\sqrt{2} \alpha^{**} + \zeta') \xi' - \xi^* \xi']/c\} d\xi_R d\xi_I \\ &= 2S_{\xi'} \exp [(\alpha' \beta' + \sqrt{2} \beta' \zeta^* + \sqrt{2} \beta^{**} \zeta' + \alpha^{**} \beta^{**})/c] \\ &\quad \times \exp [(\sqrt{2} \alpha' + \zeta^*) (\sqrt{2} \alpha^{**} + \zeta')/c]. \end{aligned}$$

Again, performing the operation $S_{\xi'}$ according to (3.23), one gets

$$\begin{aligned}
 & \langle \alpha' \beta' | \alpha^{**} \beta^{**} \rangle \\
 &= 2 \exp [(\alpha' \beta' + \alpha^{**} \beta^{**} + 2\alpha' \alpha^{**})/c] \\
 &\quad \times (\pi c)^{-1} \int \int \exp \{[\sqrt{2}(\beta' + \alpha^{**})\xi^{**} - \sqrt{2}(\beta^{**} + \alpha')\xi' - 2\xi^{**}\xi']/c\} d\xi_R d\xi_I \\
 &= \exp [(\alpha' \beta' + \alpha^{**} \beta^{**} + 2\alpha' \alpha^{**})/c] \exp [-(\beta' + \alpha^{**})(\beta^{**} + \alpha')/c] \\
 &= \exp [(\alpha' \alpha^{**} - \beta' \beta^{**})/c]. \tag{3.26}
 \end{aligned}$$

This result is similar to (3.19). It shows that the product of a wave function $\langle \alpha' \beta' | b \rangle$ and a conjugate complex wave function $\langle a | \alpha^{**} \beta^{**} \rangle$ is, corresponding to (3.20),

$$\langle a | b \rangle = S_{\alpha'} S_{\beta'} \langle a | \alpha^{**} \beta^{**} \rangle \langle \alpha' \beta' | b \rangle, \tag{3.27}$$

where S_{α} is defined like S_{ξ} in (3.21) and S_{β} is defined like S_{ζ} .

The multiplication rule (3.27) leads to the physical interpretation of the wave functions $\langle \alpha' \beta' | \rangle$. It gives, for the probability of there being m photons of the **A** field and n photons of the **B** field in the momentum and energy state \mathbf{k} , the result

$$c^m m! (-c)^n n! \text{coefficient of } \alpha'^m \beta'^n \text{ in } \langle \alpha' \beta' | \rangle^2, \tag{3.28}$$

analogous to (3.24). There is now complete mathematical symmetry between the pair of variables α, β and the pair ξ, ζ , though there remains the difference of physical interpretation, that α and β refer to the existence of photons in two fields and ξ and ζ refer to the emission and absorption of photons.

Let us apply the transformation to the initial state of the field, for which, of course, no emissions or absorptions have taken place, so that the ξ', ζ' wave function is

$$\langle \xi' \zeta' | \text{init} \rangle = 1. \tag{3.29}$$

$$\begin{aligned}
 \text{Then } \langle \alpha' \beta' | \text{init} \rangle &= S_{\xi'} S_{\zeta'} \langle \alpha' \beta' | \xi^{**} \zeta^{**} \rangle \langle \xi' \zeta' | \text{init} \rangle \\
 &= \sqrt{2} S_{\xi'} S_{\zeta'} \exp [(\alpha' \beta' + \sqrt{2} \alpha' \xi^{**} + \sqrt{2} \beta' \zeta^{**} + \xi^{**} \zeta^{**})/c]
 \end{aligned}$$

from (3.25). The operators $S_{\xi'}, S_{\zeta'}$ now reduce to putting $\xi^{**} = 0, \zeta^{**} = 0$ and hence

$$\langle \alpha' \beta' | \text{init} \rangle = \sqrt{2} e^{\alpha' \beta' / c}. \tag{3.30}$$

Applying the interpretation (3.28), one finds, firstly, that the probability of there being m photons of the **A** field and n of the **B** field in the state \mathbf{k} is zero unless $m = n$. This is what one would expect from the assumption that the **B** field equals the **A** field initially, an assumption that was made in the

text. Secondly, one finds that the probability of there being n photons of the **A** or the **B** field in the state **k** is $P_n = (-)^n 2$, a result which is used in the text.

REFERENCES

- Dirac, P. A. M. 1938 *Proc. Roy. Soc. A*, **167**, 148–169.
 Dirac, P. A. M. 1939a *Proc. Camb. Phil. Soc.* **35**, 416–418.
 Dirac, P. A. M. 1939b *Ann. Inst. Poincaré*. This paper has not yet appeared in England.
 Fock, V. 1934 *Phys. Z. Sowjet*, **6**, 425–469.
 Fokker, A. D. 1929 *Z. Phys.* **58**, 386–393.
 Gordon, W. 1926 *Z. Phys.* **40**, 117–133.
 Heisenberg, W. & Pauli, W. 1929a *Z. Phys.* **56**, 1–61.
 Heisenberg, W. & Pauli, W. 1929b *Z. Phys.* **59**, 168–190.
 Jordan, P. & Pauli, W. 1928 *Z. Phys.* **47**, 151–173.
 Klein, O. 1927 *Z. Phys.* **41**, 407–442.
 Pauli, W. & Weisskopf, V. 1934 *Helv. Phys. Acta*, **7**, 709–731.
 Weiss, P. 1936 *Proc. Roy. Soc. A*, **156**, 192–220.
-

Molecular structure and rubber-like elasticity

I. The crystal structures of β gutta-percha, rubber and polychloroprene

By C. W. BUNN, B.A., B.Sc.

The Research Department of Imperial Chemical Industries Ltd., Northwich

(Communicated by Sir William Bragg, F.R.S.—Received 13 March 1941)

The crystal structures of β gutta-percha, rubber and polychloroprene have been determined by interpretation of X-ray diffraction photographs.

β Gutta-percha ($-\text{CH}_2-\text{C}(\text{CH}_3)=\text{CH}-\text{CH}_2-$)_n is orthorhombic, with axial lengths $a_0 = 7.78 \text{ \AA}$, $b_0 = 11.78 \text{ \AA}$, $c_0 = 4.72 \text{ \AA}$. Four long-chain molecules pass through this cell parallel to the *c* axis. The space group is $P\ 2_12_12_1$, and the co-ordinates of the five carbon atoms of one structural unit are:

(CH ₂)	0.926	(C)	0.000	(CH)	0.000	(CH ₂)	0.074	(CH ₃)	0.970
	0.110		0.146		0.074		0.110		0.277
	0.676		0.960		0.177		0.462		0.980

The molecules are asymmetric; all the molecules passing through any one crystal are identical—either left-handed or right-handed, not mixed. The carbon chain is a non-planar zigzag; each chain unit C—C=C—C is planar, and has the *trans* configuration, but the connecting links (CH₂—CH₂) lie in a different plane; the plane —C—C=C— makes an angle of 115° with plane —C—C=C—.