Lectures

on

QUANTUM ELECTRODYNAMICS

by

P. A. M. Dirac

Visiting Professor
Institute for Advanced Study

Fall, 1946

Notes by

S. Kusaka

Palmer Physical Laboratory
Princeton University

Additional copies of the notes may be obtained by writing to The Secretary, Department of Physics, Princeton University at $1.00 a copy.

CONTENTS

Introduction. The theory of quantum electrodynamics which is presented here is not in a completely satisfactory state. However its presentation will bring out the difficulties, and the improvements which will be made in the future will probably change some of the fundamental ideas but will not make the whole theory completely useless.

Relativistic Form of the Hamiltonian Equations.* Since there is a well-defined method of formulating a quantum theory from a classical theory expressed in Hamiltonian form, we shall first put Maxwell's theory in a suitable Hamiltonian form. In ordinary mechanics we have the Hamiltonian function $H=H(p,q,t)$ which is a function of the generalized coordinates $q_1, q_2, \ldots, q_N$, their conjugate momenta $p_1, p_2, \ldots, p_N$, and possibly also of the time t. The equations of motion are then expressed by the equations

$$\frac{dq_r}{dt} = \frac{\partial H}{\partial p_r}, \qquad \frac{dp_r}{dt} = -\frac{\partial H}{\partial q_r} \qquad r = 1, 2, \ldots N$$

and in general the time derivative of any function $\xi(p, q)$ of p_r and q_r is given by

$$\frac{d\xi}{dt} = \sum_{r=1}^{N} \left\{ \frac{\partial \xi}{\partial q_r}\frac{\partial H}{\partial p_r} - \frac{\partial \xi}{\partial p_r}\frac{\partial H}{\partial q_r} \right\} = [\xi, H]$$

where $[\xi, H]$ denotes the Poisson bracket. This usual form of the Hamiltonian equation in which the time variable occurs in a specialized role must be modified since we wish to formulate a relativistic theory. However we need only restrict ourselves to the special theory of relativity since we shall only deal with atomic problems where gravitational effects are negligible.

*P.A.M. Dirac, Ann. L'Inst. Henri Poincare IX, 13 (1939).

In a relativistic theory t must be treated on the same footing as the space coordinates, so we now consider t as another generalized coordinate and introduce -W as its conjugate momentum. Then we have the additional equation, using F as a new Hamiltonian,

$$\frac{dt}{d\tau} = -\frac{\partial F}{\partial W}$$

and since the left side of this equation is unity, the new Hamiltonian F must be chosen so that $F = H - W$.

Furthermore the time derivative of a more general function $\xi(p,q,W,t)$ is

$$\frac{d\xi}{d\tau} = \sum_{r=1}^{N} \left\{ \frac{\partial \xi}{\partial q_r}\frac{\partial F}{\partial p_r} - \frac{\partial \xi}{\partial p_r}\frac{\partial F}{\partial q_r} \right\} - \frac{\partial \xi}{\partial t}\frac{\partial F}{\partial W} + \frac{\partial \xi}{\partial W}\frac{\partial F}{\partial t}$$

and this will be a true equation if

$$\frac{\partial H}{\partial t} = \frac{\partial W}{\partial t}$$

or if $W = H + \text{const.}$

We can then write

$$\frac{d\xi}{d\tau} = [\xi, F]_G$$

where the subscript G on the bracket denotes this generalized form of the Poisson bracket. However in the future we shall deal only with this generalized form of the bracket and the subscript will not be written explicitly.

In the above relation between W and H, the constant is unimportant and can be taken to be zero but the equality sign must be regarded with some reservation. The equation must not be used before the Poisson bracket is evaluated. This results from the fact that the relation between W and H is a consequence of the equations of motion and the initial condition. For this reason we introduce a new symbol

to distinguish between the ordinary equality and this restricted equality. For any two function A and B, $A = B$ if $[A,\xi]=[B,\xi]$ and $C \approx D$ if C and D are equal for the actual motion but $[C,\xi] \neq [D,\xi]$.

With this notation

$$W \approx H$$

The special role of t is still present since the Poisson brackets express the time derivative of the functions. In order to remove this, we replace it by an arbitrary function of all the variables $\tau = \tau(p,q,W,t)$. Then

$$\frac{d\xi}{d\tau} = \frac{d\xi}{dt}\frac{dt}{d\tau} = [\xi,F]\frac{dt}{d\tau} = [\xi, F\frac{dt}{d\tau}] - F[\xi,\frac{dt}{d\tau}]$$

But since $F \approx 0$, the last term vanishes, and hence

$$\frac{d\xi}{d\tau} \approx [\xi,F^*]$$

where

$$F^* = F\frac{dt}{d\tau} \quad \text{and} \quad F^* \approx 0.$$

Our dynamical system is now completely symmetric in the coordinates and the time and is specified by F^*. For the relativistic theory of a single particle we may take for its proper time s. For the case of several particles there is no single function which is suitable, and our theory must be generalized to the case where there are several independent variables and we take them to be the proper times of each, $s_1, s_2, \ldots$. The equations of motion will then be expressed in the form

$$\frac{d\xi}{ds_i} = [\xi,F_i] \qquad i = 1,2,\ldots$$

different values of i referring to different particles. Each

particle has a Hamiltonian $F_i \approx 0$.

The condition for the compatibility of the above equations

$$\frac{\partial^2 \xi}{\partial s_i \partial s_j} = \frac{\partial^2 \xi}{\partial s_j \partial s_i}$$

require that

$$[[\xi, F_j], F_i] = [[\xi, F_i], F_j]$$

and using Poisson's identity we obtain the condition of integrability

$$[\xi, [F_i, F_j]] = 0.$$

This must hold for all ξ so we need

$$[F_i, F_j] = \text{a number}.$$

The condition that $F_i \approx 0$ for all s_j leads to $[F_i, F_j] \approx 0$ showing that this number is zero.

We shall restrict the variability of s_i so that the particle points remain outside each other's light cones. Then the equation $[F_i, F_j] \approx 0$ need hold only for this domain. It is clear that this condition is satisfied if we take a special Lorentz frame in which the same time coordinate is used for all the particles.

The Hamiltonians should be such that

$$\frac{dt_i}{ds_j} = \frac{dx_i}{ds_j} = 0 \qquad \text{for } i \neq j.$$

The transition from the many time theory to the single theory can be made in the following way. The F's are functions of the x's and t's of all the particles and we can unite

$$\frac{d\xi}{dt_i} \approx [\xi, \chi_i F_i]$$

where

$$\chi_i = \frac{ds_i}{dt_i}, \quad \text{since} \quad [\xi, \chi_i F_i] \approx [\xi, F_i]\chi_i.$$

then if we put all $t_i = t$,

$$\frac{d\xi}{dt} \approx \sum_i [\xi, \chi_i F_i] = [\xi, F_T]$$

where $F_T = \sum_i \chi_i F_i$.

Thus the usual non-relativistic form can always be obtained by this specialization.

3. Notation. We use x_μ to denote a point in 4-dimensional space-time with μ taking the values 0, 1, 2, 3, and choose our units so that the velocity of light is unity. The scalar product of two vector A_μ and B_μ in this space is written (A, B) so that

$$(A, B) = A^\mu B_\mu = A_0 B_0 - A_1 B_1 - A_2 B_2 - A_3 B_3$$

and the relation between upper and lower indices are as follows:

$$A^0 = A_0 \ , \ A^1 = - A_1 \ , \ A^2 = - A_2 \ , \ A^3 = - A_3 \ .$$

It is convenient to introduce the 4-dimensional generalized delta function $\Delta(x)$ which is related to the ordinary delta function by the relation

$$\Delta(x) = \begin{cases} 2\,\delta(x^2) & \text{for} \quad x_0 > 0 \\ -2\,\delta(x^2) & \text{for} \quad x_0 < 0 \end{cases}$$

where $x^2 = (x, x) = x_0^2 - |x|^2$. Since we can write

$$\delta(x^2) = \delta\{(x_0 - |x|)(x_0 + |x|)\}$$

$$= \begin{cases} \frac{1}{2|x|}\,\delta(x_0 - |x|) & \text{for} \quad x_0 > 0 \\ \frac{1}{2|x|}\,\delta(x_0 + |x|) & \text{for} \quad x_0 < 0 \end{cases}$$

$$\Delta(x) = \frac{1}{|x|}\left\{\delta(x_0 - |x|) - \delta(x_0 + |x|)\right\}. \qquad (3.1)$$

$\Delta(x)$ is Lorentz invariant. It has the following important

properties. Its Fourier transform is a multiple of itself, since $$\iiint\int \Delta(x)\, e^{-i(a,x)}\, dx_0\, dx_1\, dx_2\, dx_3 = 4\pi^2 i\, \Delta(a)$$ and it satisfies the wave equation

$$\Box\, \Delta(x) = \left(\frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_2^2} - \frac{\partial^2}{\partial x_3^2}\right) \Delta(x) = 0.$$

This can best be proved by applying the operator $\Box$ on $\Delta(x)$ written as the integral over its Fourier transform. Furthermore $$\Delta(-x) = -\Delta(x)$$ so that $$\Delta(0, x_1, x_2, x_3) = 0$$ and

$$\left[\frac{\partial \Delta(x)}{\partial x_0}\right]_{x_0=0} = 4\pi\, \delta(x_1)\, \delta(x_2)\, \delta(x_3)$$

as may easily be verified from the Fourier transform.

4. Fields Associated with an Electron.* Since our plan is to formulate a theory of quantum electrodynamics from a classical theory, our first object is to develop an exact scheme of classical equations based on Maxwell's and Lorentz's work. In relativistic notation the electric and magnetic fields are together represented by an antisymmetrical tensor

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu}$$

where A_μ is the electromagnetic potential which satisfies the condition

$$\frac{\partial A_\mu}{\partial x_\mu} = 0$$

and the wave equation

$$\Box\, A_\mu = 4\pi j_\mu \tag{4.1}$$

where

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

and j_μ is the charge-current vector.

*P.A.M. Dirac, Pro. Roy. Soc. 167 A, 148 (1938).

We make the assumption that there are no charge and current except on the world line of charged particles, and that each particle with charge e gives rise to a charge-current vector given by

$$J_\mu = e \int_{-\infty}^{\infty} \dot{z}_\mu \, \delta(x_0 - z_0)\, \delta(x_1 - z_1)\, \delta(x_2 - z_2)\, \delta(x_3 - z_3)\, ds$$

Where Z_μ defines a point on a world line and is a function of the proper time s measured along it. The dot indicates differentiation with respect to s. For a number of charged particles each with charge e_i and related proper time s_i we sum over the contributions and obtain

$$J_\mu = \sum_i e_i \int_{-\infty}^{\infty} \dot{z}_{\mu i}\, \delta(x_0 - z_{0i})\, \delta(x_1 - z_{1i})\, \delta(x_2 - z_{2i})\, \delta(x_3 - z_{3i})\, ds_i$$

This assumption of the concentration of the charge on a point is rather doubtful physically since it means an infinite density of charge but it is made for the sake of mathematical simplicity. For otherwise, it means that we either have to consider a finite charge distribution which is very difficult to treat in a relativistic way, or assume the breakdown of Maxwell's theory in the neighborhood of the world line which, as has been shown by Born, also leads to great complication. Due to these difficulties, very little progress has been made with these modified theories.

A solution of (4.1) is the well-known Lienard-Wiechert potential, and for a single particle it is

$$A_{\mu\,ret}(x) = \frac{e\, v_\mu}{(v,\, x - z)} \tag{4.2}$$

where $v_\mu = \dot{z}_\mu$ and v_μ and z_μ are evaluated at the retarded point and this is indicated by the subscript "ret" on A_μ.

It can be written in the form of an integral, thus

$$A_{\mu\,ret} = e\int_{-\infty}^{\infty} \frac{v_\mu}{|x-z|}\,\delta(x_0 - z_0 - |x-z|)\,ds$$

where (x-z) denotes the length of the three-dimensional vector, or in terms of the four-dimensional delta function introduced in § 3

$$A_{\mu\,ret} = e\int_{-\infty}^{S(z_0=x_0)} v_\mu\,\Delta(x-z)\,ds$$

the upper limit ensures that only the retarded value comes in. For the general solution we have

$$A_{\mu M} = \sum_i A_{\mu i\,ret} + A_{\mu\,in} \qquad (4.3)$$

where the summation is taken over the different electrons and $A_{\mu\,in}$ is a solution for free space, namely

$$\Box A_{\mu\,in} = 0\,, \qquad \frac{\partial A_{\mu\,in}}{\partial x_\mu} = 0.$$

The subscript M for the general solution indicates that it is the usual solution encountered in Maxwell's theory, and is inserted for future reference. $A_{\mu\,in}$ describes the field of free radiation which arises from the particles at infinity. If the summation included the contribution from all the particles in the universe, the $A_{\mu\,in}$ would vanish.

Since our fundamental equations are completely symmetric in past and future, we can define in the corresponding way

$$A_{\mu\,adv}(x) = -\,e\int_{S(z_0=x_0)}^{\infty} v_\mu\,\Delta(x-z)\,ds$$

where the minus sign comes from the difference in sign between the cases $x_0 > 0$ and $x_0 < 0$ in (3.1), and we can write

$$A_{\mu M} = \sum_i A_{\mu i\, adv} + A_{\mu\, out}$$

as an equation which defines $A_{\mu\, out}$. From (4.3) and (4.4) we obtain

$$A_{\mu\, out} - A_{\mu\, in} = \sum_i e_i \int_{-\infty}^{\infty} v_{\mu i}\, \Delta(x - z_i)\, ds_i \,. \qquad (4.4)$$

5. Equation of Motion for the Electron.* So far we have discussed the electromagnetic fields produced by charged particles which move along a pre-assigned world line. Further conditions are necessary in order to determine how the particles move under the action of the fields. Lorentz first obtained an equation of motion for electrons, which can be written

$$m\dot{v}_\mu - \tfrac{2}{3} e^2 \ddot{v}_\mu - \tfrac{2}{3} e^2 \dot{v}^2 v_\mu = e v_\nu \Big\{ F_\mu^{\nu}{}_{in} + \sum_{others} F_\mu^{\nu}{}_{ret} \Big\}. \qquad (5.1)$$

Since this is a vector equation it consists of 4 equations but actually only 3 of them are independent. The reason is that in our notation $v^2 = 1$ therefore

$$(v, \dot{v}) = 0 \,, \quad (v, \ddot{v}) + \dot{v}^2 = 0$$

and the multiplication of (5.1) by v^μ makes each term vanish identically, the left side due to the above relations and the right side due to the anti-symmetry of F_μ^{ν} . However, this is sufficient since only 3 relations are necessary to determine the world line in 4-dimensional space-time. Lorentz derived equation (5.1) by considering the electron as a finite distribution of charge and in his theory it holds only

*P.A.M. Dirac. Pro. Roy. Soc. 167A, 149 (1938).

approximately. Actually there exist terms in higher derivatives of v_μ with respect to s which depend on the particular assumption made regarding the charge distribution.

The question therefore arises as to whether it is reasonable to use this equation for a point electron. We shall show that by using the laws of conservation of energy and momentum, and making one plausible assumption equation (5.1) holds strictly for a point charge. In Maxwell's theory we can set up a stress tensor

$$T_{\mu\nu} = F_{\mu\alpha} F_{\nu}{}^{\alpha} + \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}$$

whose divergence vanishes

$$\frac{\partial T_{\mu\nu}}{\partial x_\mu} = 0 .$$

It measures the rate of flow of energy and momentum across a unit surface in space-time, and the divergence condition holds everywhere except right on the world line of a charged particle. Consider a tube of radius ϵ around a world line. ϵ is considered much smaller than the physical size of an electron, say something like 10^{-100} cm. The rate of flow of energy and momentum out across the surface of the tube in the proper time interval s_a to s_b is

$$\int_{s_a}^{s_b} \left\{ \frac{1}{2} \frac{e^2}{\epsilon} \dot{v}_\mu - e v_\nu f_\mu^{*\nu} \right\} ds \quad *$$

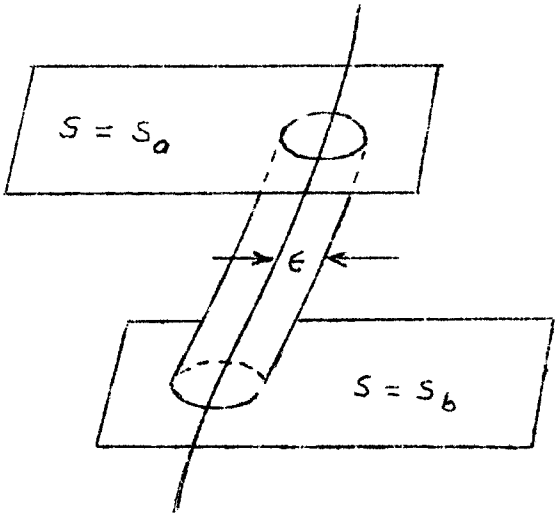

Figure 5.1

*For the derivation of this expression see Dirac, loc. cit.

where

$$f_\mu{}^\nu = F_\mu{}^\nu{}_{in} + \frac{1}{2}\left\{F_\mu{}^\nu{}_{ret} - F_\mu{}^\nu{}_{adv}\right\} + \sum_{others} F_\mu{}^\nu{}_{ret} \quad .$$

If energy and momentum are to be conserved, this must be equal to the difference in the energy and momentum content of the tube at these two times, say

$$B_\mu(s_b) - B_\mu(s_a) \quad .$$

Hence by making the interval infinitesmally small, we have

$$\frac{1}{2}\frac{e^2}{\epsilon} v_\mu - e v_\nu f_\mu{}^\nu = \frac{dB_\mu}{ds} \quad . \tag{5.2}$$

Here it is necessary for us to make some assumption about B_μ since we cannot calculate its value on account of the singularity on the world line. By multiplying (5.2) by v^μ we find

$$v^\mu \frac{dB_\mu}{ds} = 0 \quad \frac{1}{2}\frac{e^2}{\epsilon}$$

and hence the simplest assumption we can make for it is to write

$$B_\mu = k\, v_\mu \quad .$$

Then letting $\epsilon \to 0$ in (5.2) we find that k must be of the form $k = \frac{1}{2}\frac{e^2}{\epsilon} - m$?

and we obtain the equation of motion

$$m\dot{v}_\mu = e v_\nu f_\mu{}^\nu$$

$$= e v_\nu \left\{ F_\mu{}^\nu{}_{in} + \frac{1}{2}\left(F_\mu{}^\nu{}_{ret} - F_\mu{}^\nu{}_{adv}\right) + \sum_{others} F_\mu{}^\nu{}_{ret} \right\} . \tag{5.3}$$

Now the difference between the retarded and advanced fields has no singularity on the world line and actually has the value

$$F_\mu{}^\nu{}_{ret} - F_\mu{}^\nu{}_{adv} = \frac{4e}{3}\left(\ddot{v}_\mu v^\nu - \ddot{v}^\nu v_\mu\right)$$

and inserting this in (5.3) shows that it is identical with Lorentz's equation (5.1). The above method of using a tube

around a world line and applying conservation of energy and momentum thus gives a new way of deriving Lorentz's equation which is justified even for the case when the charge is concentrated on a point. This method is very powerful and has been used to good advantage by Bhabha who also considered the conservation of angular momentum.

The equation of motion (5.1), however, has the grave objection that it contains a second derivative of the velocity. As a consequence there exist more solutions of the equation than are actually needed. That is, there are some solutions which are physically acceptable and others not acceptable; solutions, which we can call physical and non-physical solutions respectively. An example of a non-physical solution is the case where a single electron with no in-going field continuously accelerates and flies off to infinity. In classical theory it is possible to distinguish between physical and non-physical solutions but in going over to quantum theory it gives rise to a fundamental difficulty since there seems to be no way of picking out the admissible solutions in quantum theory. It is a puzzling fact that, though the equations are known experimentally to give correct results when the acceleration is small, the instability shows up even for such cases. Furthermore Eliezer*has shown that there are no physical solutions for the case of an electron moving in the field of a proton, with the electron bound to the proton; that is, where

*C. J. Eliezer, Pro. Camb. Phil. Soc. 39, 173 (1943).

the electron spirals around the proton and eventually falls into it. There are physical solutions for the case where the electron is scattered by a proton, but in cases where the electron starts out with insufficient energy it receives an outward acceleration from the radiation reaction and flies off to infinity in a non-physical solution

6. The Wentzel Field. There exists a well-defined method of going over from classical mechanics to quantum theory, but in order to utilize this correspondence we must first put the classical theory in Hamiltonian form. For this purpose it is convenient to define a new electromagnetic field which was first studied by Wentzel* and shown to be of great importance for this work. We call this field the Wentzel field and distinguish its potentials and field variables by adding the subscript W. It is defined by the relation

$$A_{\mu W}(x) = A_{\mu in}(x) + \sum_i e_i \int_{-\infty}^{s_i} v'_{\mu i}\, \Delta(x - z'_i)\, ds'_i \qquad (6.1)$$

where $v'_\mu = v_\mu(s')$, $z' = z(s')$ are the values of v_μ and z at the arbitrary proper times. The fields are related to the potentials by the usual relation

$$F_{\mu\nu W} = \frac{\partial A_{\nu W}}{\partial x^\mu} - \frac{\partial A_{\mu W}}{\partial x^\nu} .$$

From the form of (6.1) it is clear that the Wentzel field depends on the electron points and has a definite value for specific values of s_i. Also, using the properties of the Δ-function given in §3 we see immediately that $A_{\mu W}$ satisfies the wave equation

$$\Box A_{\mu W} = 0 \qquad (6.2)$$

*G. Wentzel, Zeits. f. Phys. 86, 479 (1933).

and
$$\frac{\partial A_{\mu w}}{\partial x_\mu} = \sum_i e_i \int_{-\infty}^{s_i} v'_\mu \frac{\partial}{\partial x_\mu} \Delta(x - z'_i)\, ds'_i$$
$$= -\sum_i e_i \int_{-\infty}^{s_i} v'_\mu \frac{\partial}{\partial z'_{\mu i}} \Delta(x - z'_i)\, ds'_i$$
$$= -\sum_i e_i \int_{-\infty}^{s_i} \frac{d}{ds'_i} \Delta(x - z'_i)\, ds'_i$$
$$= -\sum_i e_i \Delta(x - z_i) \quad . \tag{6.3}$$

These equations are to be contrasted with those for the Maxwell field A_M

$$\Box A_{\mu M} = 4\pi j_\mu \ , \qquad \frac{\partial A_{\mu M}}{\partial x_\mu} = 0 \ .$$

(6.2) shows that the Wentzel field can always be resolved into waves traveling with the velocity of light, and (6.3) that there are longitudinal as well as transverse waves, though of course they do not play any physical role in radiation.

Considering the contribution of one electron to the Wentzel field, we see from (6.1) that there are at most only two values of s' for which the integrand is not zero. If the field point x is outside the light-cone of s' only the retarded value comes in and

$$A_w = A_{ret} \ ;$$

if x is in the future part of the light-cone, the integrand is always zero and $A_w = 0$;

if x is in the past part of the light-cone, both retarded and

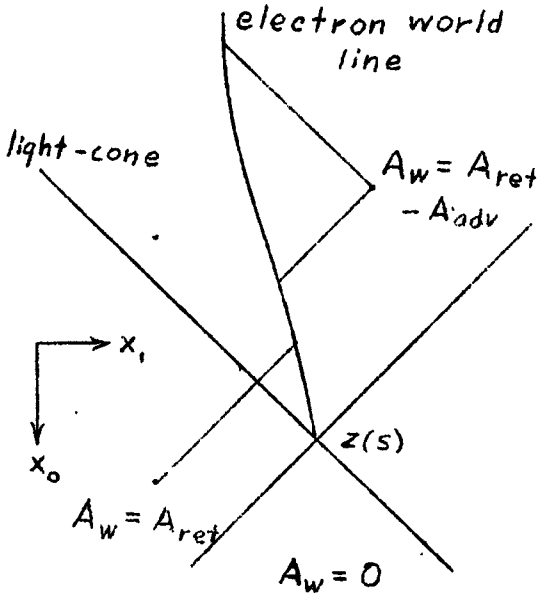

Figure 6.1

advanced values come in and $A_W = A_{ret} - A_{adv}$.
These three cases are illustrated in Figure 6.1.
Thus A_W is discontinuous across the light-cone but it still satisfies the wave equation (6.2) everywhere.

For the case of several electrons, as mentioned in § 2, we choose the proper times such that they all lie outside each other's light-cone, that is

$$\left\{ z_i(s_i) - z_j(s_j) \right\}^2 < 0 . \qquad i \neq j \qquad (6.4)$$

Then we have the three simple cases: for x outside the light-cones of all the electrons

$$A_W = A_{in} + \sum_i A_{i\,ret} = A_M \quad ;$$

for x inside the future light-cone of all the electrons

$$A_W = A_{in} \quad ;$$

and for x inside the past light-cone of all the electrons

$$A_W = A_{in} + \sum_i \left(A_{i\,ret} - A_{i\,adv} \right) = A_{out} .$$

There are, of course, other cases where the field point x lies inside the light-cone of some and outside those of other electrons.

One way of looking at the Wentzel field is to consider it as a field which is defined to be the Maxwell field outside the light-cone and which is defined inside the light-cone by the condition that it satisfy the wave equation (6.2).

We next consider the value of the Wentzel field right at the position of an electron. Since the field has a singularity there, we must specify the path along which we approach the point. For the value of the field at the position of the

i-th electron, we have (i) for a path outside the light-cone of the i-th electron

$$A_w = A_{in} + \sum_j A_{j\,ret} \qquad \text{(with condition (6.4))}$$

(Ii) for a path inside the future light-cone

$$A_w = A_{in} + \sum_{j \neq i} A_{j\,ret}$$

(iii) for a path inside the past light-cone

$$A_w = A_{in} + \sum_{j \neq i} A_{j\,ret} + \left(A_{i\,ret} - A_{i\,adv} \right).$$

We see that (i) contains the usual Coulomb singularity but (ii) and (iii) do not involve any singularity, and half the sum of (ii) and (iii) gives

$$A_w = A_{in} + \sum_{j \neq i} A_{j\,ret} + \frac{1}{2} \left(A_{i\,ret} - A_{i\,adv} \right)$$

which is just the potential needed to give the field entering in the Lorentz equation of motion for the i-th particle (5.3).

Hence if we use the Wentzel field, the field acting on an electron can be given a finite and definite meaning.

7. Equations of Motion in Hamiltonian Form.* We shall now express the equation of motion in hamiltonian form using the generalized formalism developed in §2. For each particle we take the Hamiltonian function

$$F_i = -\frac{1}{2m_i} \left[\left\{ p_i - e_i A(z_i) \right\}^2 - \ell m_i^2 \right] \qquad (7.1)$$

and we shall find from the equations of motion that $F_i \approx 0$. The form of (7.1) is suggested by the simple theory and it

*P.A.M. Dirac, Quantum Electrodynamics. Comm. Dublin Inst. for Adv. Studies, A 1. (1943).

is of the same form as for one electron. The potential $\mathbf{A}(x)$ is not defined as yet, and it is one of our tasks to find its exact expression.

The dynamical variables entering into our problem are $z_{\mu i}$, $p_{\mu i}$ for the particles, and $A_\mu(x)$ for the field, and we need the Poisson brackets between any two of them. For the particle variables we shall make the usual assumptions

$$[z_{\mu i}, z_{\nu j}] = 0 , \quad [p_{\mu i}, p_{\nu j}] = 0 ,$$

$$[z_{\mu i}, p_{\nu j}] = -g_{\mu\nu}\delta_{ij}$$

and in addition we further assume that

$$[z_{\mu i}, A_\mu(x)] = 0 , \quad [p_{\mu i}, A_\nu(x)] = 0 .$$

There is nothing in the elementary theory to tell us what $[A_\mu, A_\nu]$ should be, and for the time being we leave it undetermined.

From the general equation of motion $\frac{d\xi}{ds_i} = [\xi, F_i]$ which holds for any function ξ of all the variables and the above expressions for the brackets, we see immediately that

$$\frac{dz_{\mu i}}{ds_j} = 0 , \quad \frac{dp_{\mu i}}{ds_j} = 0 \qquad i \neq j$$

which tells that the variables of one particle do not depend on the proper times of other particles. For $i = j$ we have

$$\frac{dz_{\mu i}}{ds_i} = [z_{\mu i}, F_i] = -\frac{1}{m_i}\{p_{\nu i} - e A_\nu(z_i)\}[z_{\mu i}, p^\nu_i]$$

$$= \frac{1}{m_i}\{p_{\mu i} - e A_\mu(z_i)\}$$

where $A_\mu(z_i)$ is a well-defined functional of the dynamical variables $A_\mu(x)$ and z_i . We note that we get the

usual connection between momentum and velocity for an electron,

$$p_{\mu i} = m_i v_{\mu i} + e_i A_\mu(z_i) .$$

The condition $v_{\mu i} v^{\mu}{}_i \approx 1$ leads to $F_i \approx 0$.

Furthermore

$$\frac{dp_{\mu i}}{ds_i} = [p_{\mu i}, F_i] = \frac{e_i}{m_i}\left\{p_{\nu i} - e_i A_\nu(z_i)\right\}[p_{\mu i}, A^\nu(z_i)]$$

$$= e_i v_{\nu i}\left(\frac{\partial A^\nu}{\partial x^\mu}\right)_{z_i}$$

and hence

$$m_i \dot{v}_{\mu i} = e_i v_{\nu i}\left(\frac{\partial A^\nu}{\partial x^\mu} - \frac{\partial A^\mu}{\partial x^\nu}\right) . \qquad (7.2)$$

We want this equation to coincide with Lorentz's equation, and from the result of §6 we see that (7.2) will take the required form if we let

$$A(x) = \frac{1}{2}\left\{A_w(x+\lambda) + A_w(x-\lambda)\right\} \qquad (7.3)$$

where λ is a small time-like 4 vector which will ultimately be made to tend to zero, that is,

$$\lambda^2 > 0 , \quad \lambda \to 0 .$$

Also the condition (6.4) on the particle coordinates has to be sharpened to $(z_i - z_j \pm \lambda)^2 < 0$.

This introduction of λ is made for the purpose of insuring that the approach to the position of the particle in evaluating the field is made along a permissible path, and it is important that the limiting process $\lambda \to 0$ is not made too early since otherwise the potential becomes meaningless. On the other hand the theory is not relativistic with finite λ since it

gives a preferred direction for the time axis and only becomes relativistic in the end when λ is made to vanish.

Now the Hamiltonian equations for the field quantities are

$$\frac{dA_\mu(x)}{ds_i} = [A_\mu(x), F_i] \quad , \tag{7.4}$$

which cannot be evaluated since we do not know the Poisson bracket expressions for the potentials. However we do know what the equations should be, and we can define the Poisson brackets to give the correct equations.

We have from (7.3)

$$\frac{dA_\mu(x)}{ds_i} = \frac{1}{2}\left\{\frac{dA_{\mu w}(x+\lambda)}{ds_i} + \frac{dA_{\mu w}(x-\lambda)}{ds_i}\right\}$$

and using the definition (6.1) for the Wentzel field, we obtain

$$\frac{dA_\mu(x)}{ds_i} = \frac{1}{2} e_i V_{\mu i}\left\{\Delta(x+\lambda-z_i) + \Delta(x-\lambda-z_i)\right\} \tag{7.5}$$

In order that (7.4) give rise to (7.5), we see that we must have

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

This completes the scheme of Poisson brackets and the Hamiltonian formulation.

Strictly speaking, the Poisson brackets have been defined so far only for functions of z_i p_i, but it is permissible to define the Poisson brackets of the field quantities in this way. For given any set of independent dynamical variables A, B, C, ... we may postulate any Poisson bracket relations between them, subject to the consistency conditions

$$[[A,B],C] + [[B,C],A] + [[C,A],B] = 0 ,$$

etc. It must then be possible to introduce q's and p's so that

the usual definition of Poisson bracket holds, but as the p's and q's are not needed in the theory there is no need to introduce them.

It is of some interest to mention now we can pass from the many-time scheme to the one-time formalism. The general equations

$$\frac{d\xi}{ds_i} = [\xi, F_i]$$

are equivalent to

$$\frac{d\xi}{dz_{oi}} = [\xi, H_i] \tag{7.6}$$

where

$$H_i = -p_{oi} + e_i A_o(z_i) + \left[m_i^2 + \sum_r \{p_{ri} - e_i A_r(z_i)\}^2\right]^{\frac{1}{2}} \approx 0. \tag{7.7}$$

This expression for H_i is obtained from that of F_i in (7.1) from the conditions that $F_i \approx 0$ shall give $H_i \approx 0$ and that H_i is linear in p_{oi} with coefficient -1. For then (7.6) holds for $\xi = z_{oi}$ which is sufficient to fix the independent variable as z_{oi}. Now putting all the times z_{oi} equal to t, we have for ξ independent of the p_{oi} and z_{oi}

$$\frac{d\xi}{dt} = [\xi, \sum_i H_i] \ . \tag{7.8}$$

However, ξ may involve the field quantities $A_\mu(x_o, x_1, x_2, x_3)$ at arbitrary points in space-time, and putting $x_o = t$ gives an additional variation of ξ with t. We then have instead of (7.8) new equations of motion

$$\frac{d\xi}{dt} = [\xi, \sum_i H_i] + \sum \frac{\partial \xi}{\partial A_\mu(x)} \frac{\partial A_\mu(x)}{\partial t}$$

which can be written in the form

$$\frac{d\xi}{dt} = [\xi, H_T]$$

where $$H_T = \sum_i H_i + H_F$$

and H_F is a function of the field variables such that

$$[A_\mu(x), H_F] = \frac{\partial A_\mu(x)}{\partial t} \qquad \text{for } x_0 = t .$$

H_F is the Hamiltonian of the field in the usual one-time formulation, and $H_F \approx \frac{1}{8\pi}\int (E^2 + H^2)\, dV$ in the limit $\lambda \to 0$. Of course this one-time formulation is not relativistic form.

8. Gauge Transformation. In the last section we expressed the equations of motion for the electron in Hamiltonian form, but this procedure does not give us all the equations which the particle and field variables satisfy. There are, namely, the field equations

$$\Box A_\mu = 0 \tag{8.1}$$

and

$$\frac{\partial A_\mu(x)}{\partial x_\mu} = -\frac{1}{2}\sum_i e_i \left\{ \Delta(x - z_i + \lambda) + \Delta(x - z_i - \lambda) \right\} \tag{8.2}$$

which cannot be derived from the general equation of motion since they do not depend on the proper times s_i. They must therefore be imposed as extra conditions and we must verify that they are consistent with the equations of motion.

First we can readily see that $\Box A_\mu$ has zero Poisson brackets with everything, since

$$[A_\mu(x), A_\nu(x')] = \frac{1}{2} g_{\mu\nu} \left\{ \Delta(x - x' + \lambda) + \Delta(x - x' - \lambda) \right\}$$

and therefore $[\Box A_\mu(x), A_\nu(x')] = 0$.

Thus we can put $\Box A_\mu(x) = 0$ without causing any difficulty

with the equations of motion.

The situation with (8.2) is somewhat different since its Poisson brackets with other variables do not always vanish. Writing (8.2) in the form $R(x) = 0$

with $$R(x) = \frac{\partial A_\mu(x)}{\partial x_\mu} + \frac{1}{2}\sum_i e_i\left\{\Delta(x-z_i+\lambda) + \Delta(x-z_i-\lambda)\right\}$$

we can only have in the Hamiltonian theory the weaker condition

$$R(x) \approx 0 . \tag{8.3}$$

To verify that, if (8.3) holds initially it holds for all time, we have

$$[R(x), p_{\mu i} - e_i A_\mu(z_i)] = -\frac{1}{2} e_i \frac{\partial}{\partial z^\mu_i}\left\{\Delta(x-z_i+\lambda) + \Delta(x-z_i-\lambda)\right\} - \frac{1}{2} e_i \frac{\partial}{\partial x^\mu}\left\{\Delta(x-z_i+\lambda) + \Delta(x-z_i-\lambda)\right\}$$

$$\therefore [R(x), p_{\mu i} - e_i A_\mu(z_i)] = 0$$

and hence $$[R(x), F_i] = 0 .$$

Thus $R(x)$ is a constant of the motion and condition (8.3) is consistent with the equations of motion.

The fact that $R(x)$ is only ≈ 0 means that some of Maxwell's equations also hold only in the curly sense in the Hamiltonian theory. The equations

$$\frac{\partial F_{\mu\nu}}{\partial x^\lambda} + \frac{\partial F_{\lambda\mu}}{\partial x^\nu} + \frac{\partial F_{\nu\lambda}}{\partial x^\mu} = 0$$

still hold in the full sense, since they are deducible directly from the definition of the field

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}$$

but the equations

$$\frac{\partial F_{\mu\nu}}{\partial x_\mu} = \frac{\partial^2 A_\nu}{\partial x_\mu \partial x^\mu} - \frac{\partial^2 A_\mu}{\partial x_\mu \partial x^\nu} \approx \text{a function of the electron variables}$$

hold only in the curly sense, since

$$\frac{\partial^2 A_\nu}{\partial x_\mu \partial x^\mu} = \Box A_\nu = 0$$

but

but $\frac{\partial^2 A_\mu}{\partial x^\nu \partial x_\mu}$ curly equals a function involving z.

From the general dynamical point of view, the condition $R \approx 0$ is on the same footing as $F_i \approx 0$. Now the F's give rise to the equations of motion $\frac{d\xi}{ds_i} = [\xi, F_i]$ and hence we can have a similar equation with $R(x')$

$$\frac{d\xi}{d\tau(x)} = [\xi, R(x')]$$

where τ is the variable canonically conjugate to R. We shall see that this equation generates a gauge transformation.

Considering infinitesimal transformations, we have the difference between the new function $\tilde{\xi}$ and the old ξ

$$\tilde{\xi} - \xi = \epsilon [\xi, R(x')] \qquad (8.3)$$

For $\xi = A_\mu(x)$ we have

$$\tilde{A}_\mu(x) - A_\mu(x) = \epsilon [A_\mu(x), R(x')] = \epsilon \frac{\partial}{\partial x'_\nu} [A_\mu(x), A_\nu(x')]$$

$$= -\frac{1}{2} \epsilon \frac{\partial}{\partial x^\mu} \{ \Delta(x - x' + \lambda) + \Delta(x - x' - \lambda) \}$$

$$= \epsilon \frac{\partial S}{\partial x^\mu}$$

where $S(x) = -\frac{1}{2}\{\Delta(x-x'+\lambda) + \Delta(x-x'-\lambda)\} = \frac{1}{2}\{\Delta(x'-x+\lambda) + \Delta(x'-x-\lambda)\}$

Further $\tilde{z}_i = z_i$

since R(x) does not depend on p_i, and

$$\tilde{p}_{\mu i} - p_{\mu i} = \epsilon [p_{\mu i}, R(x')] = \frac{1}{2} \epsilon e_i \frac{\partial}{\partial z^\mu_i} \{ \Delta(x' - z_i + \lambda) + \Delta(x' - z_i - \lambda) \}$$

$$= \epsilon e_i \frac{\partial S(z_i)}{\partial z^\mu_i}$$

Thus (8.3) generates the most general gauge transformation which preserves the condition $\Box A_\mu(x) = 0$. All quantities of physical significance have zero Poisson brackets with

$R(x')$ for all x', or in other words are gauge invariant, and these conditions correspond to the quantities having zero Poisson brackets with the F_s being constants of the motion. In particular it follows from $[F_i, R]=0$ that the new Hamiltonian after a gauge transformation is the same function of the new variables as the old Hamiltonian is of the old variables.

9. Elimination of the Longitudinal Waves. There exists another kind of a transformation which can be used in electrodynamics to eliminate the longitudinal waves. This transformation is sometimes also called a gauge transformation in the literature but this is not altogether correct. With the above definition of gauge transformation, the longitudinal waves cannot be eliminated since $[A_0(x), A_0(x')] \neq 0$, and this inequality must be preserved in the transformation so that A_0 cannot be brought to zero. The elimination of the longitudinal waves is not relativistic, but it is of interest in spite of this, since it effects a great simplification of the equations.

To carry out this transformation, it is convenient to use the Fourier components of $A_\mu(x)$ instead of $A_\mu(x)$ itself. Thus we write

$$A_\mu(x) = \int \left\{ \eta_{\mu k} e^{i(k,x)} + \bar{\eta}_{\mu k} e^{-i(k,x)} \right\} \frac{d^3k}{k_0} \qquad (9.1)$$

where k is a vector such that

$$k^2 = 0 \ , \quad k_0 > 0 .$$

d^3k is written for $dk_1\ dk_2\ dk_3$ and the k_0 is written in the denominator since d^3k/k_0 is Lorentz invariant. The reason for the convenience of using the Fourier resolution is that the field variables $\eta_{\mu k}$ are functions of the 3-fold infinity of k while $A_\mu(x)$ are functions of the 4-fold infinity of x and satisfy the condition $\Box A_\mu(x) = 0$.

From the Poisson bracket relations for the $A_\mu(x)$ we can readily obtain corresponding relations for the $\eta_{\mu k}$

$$[\eta_{\mu k}, \eta_{\nu k'}] = 0 ,$$

$$[\eta_{\mu k}, \bar{\eta}_{\nu k'}] = -i \frac{g_{\mu\nu}}{4\pi^2} k_0 \cos(k,\lambda)\ \delta_3(k-k') \qquad *$$

and from $R(x) \approx 0$ we have

$$R_k = k^\mu \eta_{\mu k} - \frac{1}{4\pi^2} \cos(k,\lambda) \sum_i e_i\, e^{-i(k,z_i)} \approx 0$$

and its conjugate complex

$$\bar{R}_k \approx 0 .$$

We make the following definition: For a particular k vector taken along the x_3 direction, that is for $k_\mu = (k_0, 0, 0, k_i)$ η_0 and η_3 are defined as giving rise to longitudinal waves and η_1 and η_2 as transverse waves. Thus the elimination of the longitudinal waves means the elimination of η_0 and η_3 and thus the reduction of the number of field variables to half of the original numbers.

It is convenient to build up the contact transformation from infinitesmal contact transformation since in this form the transition to quantum theory can be made directly. We characterize the contact transformation by the variable τ

*See Appendix 1 for the derivation of this expression.

so that each dynamical variable depends on it and transorms according to $\frac{d\xi^\tau}{d\tau} = [\xi^\tau, G^\tau]$, $0 \leq \tau \leq 1$ from its initial value $\xi^0 = \xi$ to its final value ξ^1.

We choose G^τ to have the form

$$G^\tau = i \sum_j e_j \int \left\{ \eta^\tau_{ok} e^{i(k,z_j)} - \bar{\eta}^\tau_{ok} e^{-i(k,z_j)} \right\} \frac{d^3k}{k_o^2} .$$

then

$$z^\tau = z \ , \quad \eta^\tau_{rk} = \eta_{rk} \ , \quad r = 1, 2, 3,$$

$$\frac{d\eta^\tau_{ok}}{d\tau} = - \frac{1}{4\pi^2 k_o} \cos(k,\lambda) \sum_j e_j e^{-i(k,z_j)}$$

Since the left side is independent of τ, we can integrate this equation immediately and obtain

$$\eta^\tau_{ok} = \eta_{ok} - \tau \frac{1}{4\pi^2 k_o} \cos(k,\lambda) \sum_j e_j e^{-i(k,z_j)} \qquad (9.2)$$

Further

$$\frac{dp^\tau_{oi}}{d\tau} = - e_i \int \left\{ \eta^\tau_{ok} e^{i(k,z_i)} + \bar{\eta}^\tau_{ok} e^{-i(k,z_i)} \right\} \frac{d^3k}{k_o}$$

and substituting in the value for η^τ_{ok} of (9.2) we have

$$\frac{dp^\tau_{oi}}{d\tau} = - e_i \int \left\{ \eta^o_{ok} e^{i(k,z_i)} + \bar{\eta}_{ok} e^{-i(k,z_i)} \right\} \frac{d^3k}{k_o}$$

$$+ \tau \frac{1}{2\pi^2} e_i \sum_j e_j \int \cos(k,\lambda) \cos(k, z_i - z_j) \frac{d^3k}{k_o^2} . \qquad (9.3)$$

The integral which occures in the second term of (9.3) is a sum of integrals of the type

$$\int \cos(a,k) \frac{d^3k}{k_o^2} = \begin{cases} \frac{2\pi^2}{|a|} & \text{for } a^2 < 0 \\ 0 & \text{for } a^2 > 0 \end{cases}$$

and using this result we find

$$\frac{dp^\tau_{oi}}{d\tau} = - e_i A_o(z_i) + \frac{1}{2} \tau e_i \sum_{j \neq i} e_j \left\{ \frac{1}{|z_i - z_j + \lambda|} + \frac{1}{|z_i - z_j - \lambda|} \right\} \qquad (9.4)$$

In the summation over j, the term for j=i has vanished since $\lambda^2 > 0$ and all terms for $j \neq i$ remain since z's have been chosen so that they are all outside each other's light-

cones. This vanishing of the $i = j$ term is a very important feature arising from the introduction of the λ-limiting process.. Fermi tried to make a similar theory without it and got an infinite term which he had to subtract out arbitrarily. Integrating (9.4) we have

$$p^1_{oi} - p_{oi} = - e_i A_o(z_i) + \frac{1}{4} e_i \sum_{j \neq i} e_j \left\{ \frac{1}{|z_i - z_j + \lambda|} + \frac{1}{|z_i - z_j - \lambda|} \right\}.$$

Corresponding calculations for the other components of p_i lead to expressions which are slightly more complicated, and the whole result can be expressed in the form

$$p_{oi} - e_i A_o(z_i) = p^1_{oi} - \frac{1}{4} e_i \sum_{j \neq i} e_j \left\{ \frac{1}{|z_i - z_j + \lambda|} + \frac{1}{|z_i - z_j - \lambda|} \right\} \tag{9.5}$$

$$p_{ri} - e_i A_r(z_i) = p^1_{ri} - e_i A^{trans}_r(z_i) + \frac{1}{4} e_i \sum_{j \neq i} e_j \left\{ \right.$$

$$\left. \frac{(z_{ri} - z_{rj} + \lambda_r)(z_{oi} - z_{oj} + \lambda_o)}{|z_i - z_j + \lambda|^3} + \frac{(z_{ri} - z_{rj} - \lambda_r)(z_{oi} - z_{oj} - \lambda_o)}{|z_i - z_j - \lambda|^3} \right\} \tag{9.6}$$

Thus when the Hamiltonian is expressed in terms of the new variables, the longitudinal waves do not appear at all. The new terms involving the z's which can be considered as replacing the longitudinal waves, do not seem to have any simple physical significance in this form, but if we pass to the one-time formalism, we see that they give just the Coulomb interaction energy in the total Hamiltonian. Making the transformations (9.5) and (9.6) in the expression for H_i given by (7.7), we obtain in the limit $\lambda = 0$

$$H_i = - p^1_{oi} + \frac{1}{2} e_i \sum_{j \neq i} \frac{e_j}{|z_i - z_j|}$$

$$+ \left[m_i^2 + \sum_r \left\{ p^1_{ri} - e_i A^{trans}_r(z_i) \right\}^2 \right]^{\frac{1}{2}}$$

In this formulation we can just as well leave out the term p^1_{oi} and then the total Hamiltonian is $H_T = \sum_i H_i + H_F$

or $$H_i = \sum_{j<i} \frac{e_i e_j}{|z_i - z_j|} + \sum_i \left[m_i^2 + \sum_r \left\{ p_{ri}^2 - e_i A_r^{trans}(z_i) \right\}^2 \right]^{\frac{1}{2}} + H_F .$$

It is somewhat surprising that this theory, which is still relativistic even though it no longer appears so, gives rise to Coulomb interaction which is a purely non-relativistic concept.

10. Passage to the Quantum Theory. In formulating a quantum theory from the classical theory which we have developed above, the dynamical variables $z_{\mu i}$, $p_{\mu i}$, $A_\mu(x)$ are made into operators satisfying commutation relations corresponding to the classical Poisson bracket relations. Also the supplementary conditions which we have found necessary to impose for consistency of the equations of motion must be reinterpreted.

In classical theory we have the conditions

$$(z_i - z_j \pm \lambda)^2 < 0 \qquad i, j = 1, 2, \ldots \qquad (10.1)$$

which must be satisfied by the z's, and this inequality is also needed in quantum theory. Since the z's are now operators, the condition is to be interpreted in the representation in which the wave function $\Psi(z, \ldots)$ is diagonal in the z's. We consider the wave function to be defined only in the domain of the z's which satisfies the conditions (10.1). If we go to the one-time system where all the z_{oi}'s are equal, we can pass over to any other representation, say where the momenta p's are diagonal, but in general we cannot do this with the

many-time formalism. The physical interpretation of $|\Psi(z,\ldots)|^2$ as the probability density for the electrons to be at a point z in space-time then holds with condition (10.1), and this is reasonable since no disturbance on the measurement of one electron can influence the measurement of another. It conforms to our idea that disturbances travel with the velocity of light.

One wave equation is provided by each Hamiltonian F_i and we have

$$F_i \Psi = 0 \quad , \tag{10.2}$$

the number of wave equations being equal to the number of time variables; or what is the same thing, to the number of electrons present. If we wish to treat electrons with spin, we must take for F_i the expression

$$F_i = \alpha_{oi} \left\{ p_{oi} - e_i A_o(z_i) \right\} + \sum_r \alpha_{ri} \left\{ p_{ri} - e_i A_r(z_i) \right\} + m_i \tag{10.3}$$

instead of the classical expression (7.1). The α's are new variables which describe the new spin degree of freedom, and the time and space parts are written separately in (10.3) since it would not be correct to write it in the form of a 4-dimensional scalar product because the α's do not form a 4-vector. For consistency of equations (10.2), we must have

$$\frac{\partial^2 \psi}{\partial z_{oi} \partial z_{oj}} = \frac{\partial^2 \psi}{\partial z_{oj} \partial z_{oi}} \tag{10.4}$$

and this leads to the condition

$$F_i F_j \psi = F_j F_i \psi \quad . \tag{10.5}$$

To derive these conditions and to see that they hold, we write (10.3) in the form

$$F_i = \left\{ p_{oi} - e_i A_o(z_i) \right\} + \sum_r \alpha_{ri} \left\{ p_{ri} - e_i A_r(z_i) \right\} + \alpha_{mi} m_i$$

$$F_i = i\hbar \frac{\partial}{\partial z_{oi}} + H_i \quad .$$

Then (10.4) leads to $H_i H_j \chi = H_j H_i \chi$

and this is equivalent to (10.5). The fact that (10.5) holds is quite obvious since for $i \neq j$, all the operators in F_i and F_j commute except the field variables and they commute everywhere except in the region a distance λ from the light-cone.

The extra conditions $\Box A = 0$ can be taken over unchanged into the quantum theory, since they are consistent with the commutation relations. The condition $R(x) \approx 0$ with

$$R(x) = \frac{\partial A_\mu(x)}{\partial x_\mu} + \frac{1}{2} \sum_i e_i \left\{ \Delta(x - z_i + \lambda) + \Delta(x - z_i - \lambda) \right\}$$

has to be interpreted in the sense that we only consider wave functions which satisfy the condition $R(x)\psi = 0$ (10.6)

In general the rule for interpreting supplementary conditions in going from classical to quantum theory is that $A \approx 0$ in classical theory becomes $A\psi = 0$ in quantum theory. In the former the curly equal sign means we cannot form the Poisson bracket of A with arbitrary variables, and in the latter we can only multiply the condition on the left by arbitrary operators and still have the equation hold. Thus if χ is any operator $\chi A \psi = 0$

but $A\chi\psi$ is not necessarily $= 0$

so that $[A,\chi]\psi$ is not necessarily $= 0$

Furthermore, in quantum theory if there are more than one condition, say $A\psi = 0$

$B\psi = 0$

then we can deduce a third $[A, B]\psi = 0$

from $AB\Psi = 0$ and $BA\Psi = 0$, and this must not lead to any inconsistency. There is no correspondence of this in classical theory since $A \approx 0$ and $B \approx 0$ does not necessarily imply $[A,B] \approx 0$

Now the wave equations (10.2) are counted on the same footing as supplementary conditions and we must verify that they and (10.6) do not lead to any inconsistency. We have already seen that the F's are consistent among themselves, and the relation $[R(\lambda), F_i] = 0$ holds in quantum theory as well as in classical theory. Finally

$$[R(x), R(x')] = \left[\frac{\partial A_\mu(x)}{\partial x_\mu}, \frac{\partial A_\nu(x')}{\partial x'_\nu} \right] = \frac{\partial^2}{\partial x_\mu \partial x'_\nu} [A_\mu(x), A_\nu(x)]$$

$$= -\frac{1}{2} \Box \left\{ \Delta(x-x'+\lambda) + \Delta(x-x'-\lambda) \right\} = 0,$$

which completes the proof of the consistency of the equations in quantum theory.

We make the definition that a linear operator which represents a physically observable quantity must be such that when it is applied to an allowed wave function it gives rise to another allowed wave function. In other words, if ξ is an observable, we need to impose the conditions that for Ψ satisfying the supplementary conditions $A\Psi = 0$, $B\Psi = 0$, etc.

$$A\xi\Psi\rangle = 0 \quad , \qquad B\xi\Psi\rangle = 0$$

and thus $[A,\xi]\Psi = 0$, $[B,\xi]\Psi = 0$,

or $[A,\xi] \approx 0$, $[B,\xi] \approx 0$.

Thus when there exist supplementary conditions, the observable must satisfy certain conditions.

Applying this general result to our electrodynamic theory with the condition $R(x) \approx 0$, we see that any observable ξ must satisfy the condition $[R(x), \xi] \approx 0$.
Since $R(x)$ generates a gauge transformation, the physical interpretation of this relation is that the observables must be invariant to gauge transformations.

11. Bracket Notation. We have now developed a consistent set of equations for the operators in quantum theory, and our next step is to set up a representation for these operators. We have already done this for the particle variables and it remains to do it for the field variables. We shall first of all describe a system of notation which is most convenient for describing linear operators and vectors whose directions define the quantum states.

The vectors corresponding to the quantum states are denoted by the symbol $|x\rangle$ which we shall call a ket-vector. When it is operated on by a dynamical variable α we write $\alpha|x\rangle$. Each vector has a conjugate imaginary vector and we write these in the form $\langle y|$ and call them bra-vectors. These can be operated on by linear operators on the right and we write the result $\langle y|\beta$.

A bra-vector $\langle y|$ and a ket-vector $|x\rangle$ have a scalar product which is a number and we denote this by $\langle y|x\rangle$. Similarly the scalar product of $\langle y|$ and $\alpha|x\rangle$ is written $\langle y|\alpha|x\rangle$. The associative and distributive laws of algebra hold for these quantities, and we have the two axioms

$$\hbar\bar{\eta}\eta\bar{\eta}|H'\rangle = H'\bar{\eta}|H'\rangle ,$$

and using the commutation relation (12.2), we obtain

$$(H+\hbar)\bar{\eta}|H'\rangle = H'\bar{\eta}|H'\rangle ,$$

or $$H\bar{\eta}|H'\rangle = (H'-\hbar)\bar{\eta}|H'\rangle .$$

Thus if $\bar{\eta}|H'\rangle \neq 0$, then $\bar{\eta}|H'\rangle$ is a new eigenket of H with eigenvalue $H'-\hbar$.

Similar calculation with η replaced by $\bar{\eta}$ yields the result that if $\eta|H'\rangle \neq 0$, then $\eta|H'\rangle$ is an eigenket with eigenvalue $(H'+\hbar)$. Thus if H' is one eigenvalue, then

$H'-\hbar$, $H'-2\hbar$, . . . and $H'+\hbar$, $H'+2\hbar$, . . .

are all eigenvalues except when one of the conditions are not satisfied. When $\bar{\eta}|H'\rangle = 0$ then $\eta\bar{\eta}|H'\rangle = 0$,

or $H|H'\rangle = 0$, $H'|H'\rangle = 0$,

and thus $H' = 0$. When $\eta|H'\rangle = 0$, $\bar{\eta}\eta|H'\rangle = 0$ and $H' = -\hbar$. However this latter equality can never hold since we can show that the eigenvalues of H can never be negative. We have

$$H'\langle H'|H'\rangle = \langle H'|H|H'\rangle$$
$$= \hbar\langle H'|\eta\bar{\eta}|H'\rangle \geq 0,$$

since the right side is a product of $\langle H'|\eta$ by its complex conjugate. Thus the eigenvalues of H are $0, \hbar, 2\hbar, 3\hbar, \ldots, n\hbar, \ldots, \infty$.

Furthermore, if we have one eigenket, we can get another by multiplying by $\bar{\eta}$ or η and in this way build up all the eigenkets by a simple process. For instance, let $|0\rangle$ be the lowest eigenket: then $\bar{\eta}|0\rangle = 0$ and

$$\eta|0\rangle , \eta^2|0\rangle , \eta^3|0\rangle , \ldots , \eta^n|0\rangle , \ldots \qquad (12.3)$$

are all eigenkets. Any general ket $|x\rangle$ can be written as a sum $|x\rangle = \sum_n x_n \eta^n|0\rangle$

so that any state of the oscillator can be expressed as a

power series in η . Thus we can say that $|x\rangle$ is represented by $\sum_n x_n \eta^n = \psi(\eta)$.

The advantage of using Fock's representation is that the eigenfunctions are power series in η while in the ordinary treatment they are Hermite polynomials multiplied by an exponential function.

The eigenvectors given in (12.3) are not normalized, but the constant can easily be calculated. We take

$$\langle 0|0\rangle = 1$$

then

$$\langle 0|\bar{\eta}^n \eta^n|0\rangle = \langle 0|\bar{\eta}^{n-1}(\eta^n \bar{\eta} + n\eta^{n-1})|0\rangle = n\langle 0|\bar{\eta}^{n-1}\eta^{n-1}|0\rangle ,$$

so that by continuing the process

$$\langle 0|\bar{\eta}^n \eta^n|0\rangle = n! \qquad (12.4)$$

and the normalized eigenkets are $\frac{1}{\sqrt{n!}}\eta^n|0\rangle$.

Let us now find the representatives of the bra-vectors. They cannot be obtained by simply taking the complex conjugate of the ket-vectors since the dynamical variable η that fixes the representation is complex. We introduce η^{-1} defined as follows: $\eta^{-1}\eta^n|0\rangle = \eta^{n-1}|0\rangle$ for $n \geq 1$, $\eta^{-1}|0\rangle = 0$. This is sufficient to define η^{-1} completely since any ket-vector can be expressed linearly in terms of the kets $\eta^n|0\rangle$ $(n = 0, 1, 2, 3, \ldots)$. Thus we have $\eta^{-1}\eta = 1$ but $\eta\eta^{-1} \neq 1$ because $\eta\eta^{-1}|0\rangle = 0$.

η^{-1} is not a true reciprocal of η but it is the nearest thing to it since η which has an eigenvalue 0 has no true reciprocal. With this definition we have

$$\langle 0|\eta^{-n}\eta^m|0\rangle = \langle 0|\eta^{m-n}|0\rangle = 0 \qquad \text{for} \quad n < m ,$$

since $\langle 0|$ and $\eta^{m-n}|0\rangle$ belong to different eigenvalues,

$$\langle 0|\eta^{-n}\eta^{m}|0\rangle = \langle 0|(\eta^{-1})^{n-m}|0\rangle = 0 \quad \text{for } n > m$$

from the definition of η^{-1}, and

$$\langle 0|\eta^{-n}\eta^{m}|0\rangle = 1 \quad \text{for } n = m.$$

Therefore $\langle 0|\eta^{-n}$ is orthogonal to all the eigenkets except $\eta^{n}|0\rangle$, so it is the conjugate imaginary of $\eta^{n}|0\rangle$ (except for a numerical factor), and the $\langle 0|\eta^{-n}$'s for all n form a complete system of eigenbras. Thus we have for a general bra-vector $\langle y| = \langle 0|\sum_n y_n \eta^{-n}$ and $\langle y|$ is represented by $\sum_n y_n \eta^{-n} = \varphi(\eta)$.

Hence we see that every ket-vector is represented by an ascending power series in η and every bra-vector is represented by a descending power series in η. Finally the scalar product of two general bra- and ket-vectors can be written

$$\langle y|x\rangle = \langle 0|\sum_n x_n y_n|0\rangle = \sum_n x_n y_n$$

$$= \frac{1}{2\pi i}\oint \varphi(\eta)\,\psi(\eta)\,\frac{d\eta}{\eta} \qquad (12.5)$$

where the contour is taken around the origin. Thus the domain of η can be taken as Fock did as a unit circle around the origin.

From the normalization factor (12.4) the length of a ket-vector $|x\rangle$ has the value $\langle x|x\rangle = \sum_n n!\,|x_n|^2$. Comparing this with the scalar product (12.5) for $\langle y| = \langle x|$ we obtain a relation between the coefficients of the power series $y_n = n!\,\bar{x}_n$.

We see again that the representation of a bra-vector is not the complex conjugate of the representation of a ket-vector, since our representation is not based on real variables.

It may be pointed out here that there is some asymmetry in the representations of the bra- and ket-vectors. The relation $\bar{\eta}|0\rangle = 0$ leads to $\langle 0|\eta = 0$, and this means that we can add any ascending power series to $\varphi(\eta)$ and write

$$\langle y| = \langle 0|\{\ldots + y_{-2}\eta^2 + y_{-1}\eta + y_0 + y_1\eta^{-1} + y_2\eta^{-2} + \ldots\}$$

without changing the value of $\langle y|$. This arbitrariness, however, does not exist in the ket-vector, and the reason lies in the fact that $\eta^{-1}\eta = 1$ but $\eta\eta^{-1} \neq 1$. It is clear that inclusion of arbitrary negative powers series in $\psi(\eta)$ would alter the value of the scalar product integral (12.5). To express it in another way, any state of the oscillator can be written $|x\rangle$ and the ket-vector can be normalized to $\langle x|x\rangle = 1$. It can be expanded in a power series in η and the probability that the oscillator is in the nth state is given by

$$P_n = n!\,|x_n|^2 . \qquad (12.6)$$

Doing the same thing with the bra-vector $\langle y|$, we have

$$P_n = \frac{1}{n!}\,|y_n|^2 .$$

In the latter formula it is possible to let n take on negative values since the result is still reasonable. Due to the fact that $\frac{1}{n!} = 0$ for negative integers, we obtain $P_n = 0$ for these values of n. The same procedure cannot be done with (12.6) since the n! occurs in the numerator.

Also from the normalization condition for $|x\rangle$ we have

$$\sum_{n=0}^{\infty} n!\,|x_n|^2 = 1$$

which shows that $\psi(n)$ converges for all values of η but from

the corresponding condition for $\langle y|$, we have

$$\sum_{n=0}^{\infty} \frac{1}{n!} |y_n|^2 = 1$$

which shows that $\varphi(\eta)$ need not converge for any value of η. For this reason $\psi(\eta)$ is usually the more convenient function to work with.

To find the representation of the unit operator, we write it in the form $\langle \eta' | 1 | \eta'' \rangle$ and we note that it must satisfy the condition

$$\frac{1}{2\pi i} \oint \langle \eta' | 1 | \eta'' \rangle \frac{d\eta''}{\eta''} \psi(\eta'') = \psi(\eta') .$$

If we consider $\psi(\eta)$ as developed in a power series, we see immediately that the equation will be satisfied with

$$\langle \eta' | 1 | \eta'' \rangle = \sum_{n=0}^{\infty} \left(\frac{\eta'}{\eta''} \right)^n = \frac{\eta''}{\eta'' - \eta'} \qquad (12.7)$$

Thus this reciprocal function replaces the δ-function in the usual representation, and it is one of the good features of the theory in avoiding improper functions.

Furthermore, the commutation relations and all the algebraic relations between η and $\bar{\eta}$ are satisfied if we take $\bar{\eta} = \frac{d}{d\eta}$ and hence for the representation of $\bar{\eta}$, we have the condition

$$\frac{1}{2\pi i} \oint \langle \eta' | \bar{\eta} | \eta'' \rangle \frac{d\eta''}{\eta''} \psi(\eta'') = \frac{d}{d\eta'} \psi(\eta')$$

which has the solution

$$\langle \eta' | \bar{\eta} | \eta'' \rangle = \sum_{n=0}^{\infty} \frac{n \eta'^{n-1}}{\eta''^n} = \frac{\eta''}{(\eta'' - \eta')^2} \qquad (12.8)$$

This function replaces the δ'-function in the usual theory.

We note that both (12.7) and 12.8) can have arbitrary terms of the form $n\,\eta''$ (ascending power series in η'')

added on the right-hand side so that the above expressions are not unique.

So far in the development of the representations for a harmonic oscillator we have not made use of the equation of motion. It is $\frac{d\eta}{dt} = [\eta, H] = i\eta$ so that $\eta \propto e^{it}$.

The above representations can still be used for other Hamiltonians. The equation of motion will be different and the above eigenvectors will not describe stationary states, but the representation will still be a convenient one to use if the Hamiltonian and the observables of interest are expressable as power series in p's and q's.

The generalization of the above treatment of one oscillator to the case of a system of oscillators is quite straightforward. We introduce η and $\bar{\eta}$ for each oscillator and distinguish them by adding a subscript a. Thus the variables are η_a, $\bar{\eta}_a$, $a = 1, 2, \ldots$ and they satisfy the commutation relations

$$\eta_a \eta_b - \eta_b \eta_a = 0 ,$$

$$\bar{\eta}_a \eta_b - \eta_b \bar{\eta}_a = \delta_{ab} . \qquad (12.9)$$

The ket vector for a general state is

$$|x\rangle = \{\text{power series in all the variables } \eta_1, \eta_2, \ldots\}|0\rangle$$

$$= \left\{ \sum_{n_1, n_2, \ldots} x_{n_1, n_2 \ldots} \eta_1^{n_1} \eta_2^{n_2} \ldots \right\} |0\rangle$$

where $|0\rangle$ denotes the ket-vector for the normal state with all the oscillators in the lowest energy level. If $\langle x|x\rangle = 1$ then the probability that the first oscillator is in the

state n_1, the second in the state n_2, etc., is given by

$$P_{n_1 n_2 \ldots} = n_1!\, n_2! \ldots |\chi_{n_1 n_2 \ldots}|^2 .$$

13. Assembly of Bosons. We shall next consider a dynamical system composed of a number of particles of the same kind and in which only symmetrical states occur. Such particles we call bosons. First we take one boson and introduce the basic ket-vectors

$$|\alpha^1\rangle , |\alpha^2\rangle , |\alpha^3\rangle , \ldots , \tag{13.1}$$

which are normalized and orthogonal to each other so that

$$\langle \alpha^a | \alpha^b \rangle = \delta_{ab} .$$

Now if we have a set of u^1 bosons, the states of this system can be obtained by taking a direct product of u^1 of the basic vectors in (13.1). Thus

$$|\alpha_1^a\rangle |\alpha_2^b\rangle |\alpha_3^c\rangle \ldots |\alpha_{u'}^q\rangle = |\alpha_1^a \alpha_2^b \alpha_3^c \ldots \alpha_{u'}^q\rangle \tag{13.2}$$

However this ket-vector is not symmetric since any permutation P of the numbers $1, 2, \ldots, u^1$ applied to (13.2) gives a new vector. For instance taking $P = (1,2)$ gives

$$|\alpha_2^a\rangle |\alpha_1^b\rangle \ldots |\alpha_{u'}^q\rangle = |\alpha_2^a \alpha_1^b \ldots \alpha_{u'}^q\rangle .$$

Now P is a linear operator and it can be used as such. Defining a symmetrizing operator

$$S = u'!^{-\frac{1}{2}} \sum_{\text{all perm.}} P$$

we have

$$u'!^{-\frac{1}{2}} \sum_{\text{all perm.}} P |\alpha_1^a\rangle |\alpha_2^b\rangle \ldots |\alpha_{u'}^q\rangle = S |\alpha^a \alpha^b \ldots \alpha^q\rangle \tag{13.3}$$

and we see that S makes any ket-vector of the type (13.2) symmetrical. If the state a, b, . . . , s are all different, the length of the vector (13.3) is unity, but if n_1 members

of the set $\alpha^a, \alpha^b, \ldots, \alpha^q$ are equal to α^1 and if n_2 members are equal to α^2 etc., then

$$\langle \alpha^a \alpha^b \ldots \alpha^q | S^2 | \alpha^a \alpha^b \ldots \alpha^q \rangle = n_1! \, n_2! \ldots .$$

So far the number of bosons in the system was taken to be a fixed number u^1. We now want to generalize to the case where this number can vary and then we shall be able to treat problems of emission and absorption of particles. The basic kets for this general case are

$$|\cdot\rangle , \quad |\alpha^a\rangle , \quad S|\alpha^a \alpha^b\rangle , \quad S|\alpha^a \alpha^b \alpha^c\rangle \ldots .$$

Where the first ket describes the state with no particles, the second with one particle, the third with two, and so on. Thus we can consider the number of bosons present as a dynamical variable u, which has the eigenvalues $u' = 0, 1, 2, \ldots$.

There is a complete correspondence between the assembly of bosons considered above and a set of oscillators of § 12. We saw that the latter was described by dynamical variables $\eta_a, \eta_b, \ldots$ and a general state was represented by a ket-vector of the form $\eta_1^{n_1} \eta_2^{n_2} \ldots |0\rangle$ while an assembly of bosons was described by a series of component wave functions

$$|\cdot\rangle , \quad |\alpha^a\rangle , \quad S|\alpha^a \alpha^b\rangle , \quad S|\alpha^a \alpha^b \alpha^c\rangle , \ldots .$$

The two sets of kets are mathematically equivalent. There are the same number of each, there is a one-to-one correspondence, and corresponding kets have the same length.

$$(length)^2 = n_1! \, n_2! \ldots$$

if we make the identification

$$\eta_1^{n_1} \eta_2^{n_2} \ldots |0\rangle = S|\alpha^a \alpha^b \alpha^c \ldots\rangle$$

where the n's are such that n_1 is the number of α's on the right equal to α^1, n_2 is the number of α's on the right equal to α^2, etc. Thus a set of oscillators is equivalent to an assembly of bosons in the mathematical sense.

14. Relativistic Representation of the Wave Function. We have seen in § 10 that the passage from classical theory to quantum theory can be made in a straight-forward manner, and we introduced the wave function Ψ which is a function of the coordinates of the electrons in a domain satisfying the condition $(z_i - z_j \pm \lambda)^2 < 0$.

The elimination of the longitudinal waves was carried out in classical theory, but it was done in such a manner that it could be taken over to the quantum theory without any change; only the interpretation was altered. This procedure resulted in the fact that only the transverse part of the field A_r^{tr} $(r = 1, 2, 3)$ remained, and its Fourier analysis

$$A_r^{tr}(x) = \int \left\{ \eta_{rk}^{tr} e^{i(k,x)} + \text{c.c.} \right\} \frac{d^3k}{k_0}$$

showed that η_k^{tr} has only two independent components η_{ak} and η_{bk}. These η's satisfied the commutation relations

$$\bar{\eta}_{ok}\,\eta_{ak} - \eta_{ak'}\,\bar{\eta}_{ak} = \frac{k_0}{4\pi^2}\cos(k,\lambda)\,\delta_3(k-k').$$

By a trivial change in definition, this result which holds for continuous variability of k can be replaced by the expression for k taking on a discrete set of values

$$\bar{\eta}_{ak}\,\eta_{ak'} - \eta_{ak'}\,\bar{\eta}_{ak} = \frac{k_0}{4\pi^2}\cos(k,\lambda)\,\delta_{kk'}\,S_k^{-1} \qquad (14.1)$$

where s_K denotes the density of discrete points in K-space. Since the right hand side of (14.1) is positive for ordinary values of K, we can make an identification of the η's occurring in (14.1) and those of (12.9) by multiplying the former η's by a numerical factor. That is, the transverse waves in quantum electrodynamics are equivalent to an assembly of bosons or to a set of oscillators. With this representation the wave function is

$$\psi(z_1, z_2, \ldots, \text{ascending power series in } \eta^{tr}) .$$

This is the most convenient representation to use in practical applications, but its form is not relativistic. For the discussion of Lorentz transformations we must introduce the representation before eliminating the longitudinal waves.

In general the Fourier analysis of the field gives

$$A_\mu(x) = \int \left\{ \eta_{\mu k} e^{i(k,x)} + c.c. \right\} \frac{d^3k}{k_0}$$

where we have four η's for each value of K, which satisfy the commutation relations

$$\bar{\eta}_{\mu k} \eta_{\nu k'} - \eta_{\nu k'} \bar{\eta}_{\mu k} = - \frac{g_{\mu\nu}}{4\pi^2} k_0 \cos(k,\lambda) \delta_{kk'} s_k^{-1} .$$

For the 1, 2, 3 components the right side is still positive and we can as before identify these η's with those of the oscillator, and obtain in the wave function an ascending power series in η_1, η_2, η_3. For η_0, however, the right side is negative and the identification cannot be made with η, but must be made with $\bar{\eta}$. Thus the wave function will contain an ascending power series in $\bar{\eta}_{0k}$, η_{1k}, η_{2k}, η_{3k}, but again such wave functions are not convenient for the discussion of Lorentz transformation.

In order to obtain a suitable wave function expressed in terms of η_{oa}, η_{ak}, η_{kk}, η_{kb}, we note the following:

$|x\rangle$ is a ket-vectorrepresented by $\psi(\eta) = x_0 + x_1\eta + x_2\eta^2 + \ldots$

or if we take the complex conjugate representation, it is

$\chi(\bar{\eta}) = x_0 + x_1\bar{\eta}^{-1} + 2!\,x_2\bar{\eta}^{-2} + \ldots$.

Similarly, the bra-vector $\langle y|$ is represented by

$$\varphi(\eta) = y_0 + y_1\eta^{-1} + y_2\eta^{-2} + \ldots$$

or by the complex conjugate representation

$$y_0 + y_1\bar{\eta} + \frac{1}{2!}y_2\bar{\eta}^2 + \ldots .$$

The descending power series in the complex conjugate representation for the ket-vector is not quite convenient and we make a simple change.

With $|x\rangle$ represented by $\psi(\eta)$, $\bar{\eta}|x\rangle$ is represented by $\frac{\partial\psi}{\partial\eta}$ but with $|x\rangle$ represented by $\chi(\bar{\eta})$, $\eta|x\rangle$ is not represented by $-\frac{\partial\chi}{\partial\bar{\eta}}$. But we can obtain the desired result by noting that

$$\langle y|x\rangle = \frac{1}{2\pi i}\oint \varphi(\eta)\bar{\psi}(\eta)\frac{d\eta}{\eta},$$

$$\langle y|\bar{\eta}|x\rangle = \frac{1}{2\pi i}\oint \varphi(\eta)\frac{\partial\psi(\eta)}{\partial\eta}\frac{d\eta}{\eta} = -\frac{1}{2\pi i}\oint \frac{\partial}{\partial\eta}\left(\frac{\varphi(\eta)}{\eta}\right)\psi(\eta)\,d\eta .$$

The last relation shows that it is more natural to take the representation of $\langle y|$ as $\tilde{\varphi}(\eta) = \frac{\varphi(\eta)}{\eta}$ rather than $\varphi(\eta)$. Then the representation of $\langle y|\bar{\eta}$ is $-\frac{\partial\tilde{\varphi}}{\partial\eta}$. Now taking the complex conjugate of these results, we take $|x\rangle$ to be represented by $\tilde{\chi}(\bar{\eta}) = \frac{\chi(\bar{\eta})}{\eta}$ and then $\eta|x\rangle$ is represented by $-\frac{\partial\tilde{\chi}}{\partial\bar{\eta}}$.

Using the above result, we can set up a representation for the wave function as an ascending power series in η_{ak}, η_{kk}, η_{kb} and descending power series in η_{oa} beginning with η_{ok}^{-1}. This procedure gives for the wave function a new kind of tensor

quantity which has heretofore not appeared in mathematical physics. For the sake of simplicity we restrict ourselves to one value of k and ignore the dependence of the wave function on the z's. Then the wave function can be written

$$Q(\eta_0, \eta_1, \eta_2, \eta_3) = \sum_{n,r,s,t} a_{nrst} \eta_0^n \eta_1^r \eta_2^s \eta_3^t$$

where n, r, s, t take in all integral values from 0 to ∞.

The square of the length of Q is given by

$$|Q|^2 = \sum_{n,r,s,t} |c_{nrst}|^2 \frac{r!\,s!\,t!}{n!}$$

Now suppose we make a Lorentz transformation of the type

$$\eta_\mu = a_\mu{}^\nu \eta'_\nu$$

then

$$Q = \sum_{n,r,s,t} c_{nrst} \frac{(\text{asc. power series in } \eta'_1, \eta'_2, \eta'_3, \eta'_0)}{(a_0{}^\nu \eta'_\nu)^{n+1}}$$

The factor in the denominator is

$$(a_0{}^\nu \eta'_\nu)^{-n-1} = (a_0{}^0 \eta'_0 + a_0{}^r \eta'_r)^{-n-1}$$

and may be expanded by means of the binomial theorem

$$(a_0{}^0 \eta'_0 + a_0{}^r \eta'_r)^{-n-1} = \frac{1}{(a_0{}^0 \eta'_0)^{n+1}} \left\{ \begin{array}{l} \text{asc. power series in } \eta'_1, \eta'_2, \eta'_3 \\ \text{and desc. power series in } \eta'_0 \end{array} \right\}.$$

Putting this in Q, we have

$$Q = \sum_{n,r,s,t} c_{nrst} \left\{ \begin{array}{l} \text{asc. power series in } \eta'_1, \eta'_2, \eta'_3 \\ \text{and asc. and desc. power series in } \eta'_0 \end{array} \right\}.$$

The terms with non-negative powers of η'_0 can simply be discarded since η'_0 is to be identified with $\bar{\eta}$ of p67 and non-negative powers of $\bar{\eta}$ can be neglected in $\bar{\chi}(\bar{\eta})$ because positive powers of $\bar{\eta}$ can be neglected in $\chi(\bar{\eta})$. The squared length for Q given

above can easily be found to be invariant under a Lorentz transformation. This was shown by Dirac* by considering an infinitesmal Lorentz transformation.

The coefficients c_{hrst} of the power series in Q can be looked upon as a new kind of tensor quantity with an infinite number of components, and it is different from ordinary tensors in that its length squared is positive definite. The name expansor has been suggested for these quantities.

Returning to the consideration of the wave function, we see that with the inclusion of the longitudinal waves, the function is an expansor for each k-value. However, the expansors which appear here are not general ones, due to the occurrence of the supplementary conditions. In quantum electrodynamics the field variables satisfy the conditions

$$\frac{\partial A_\mu}{\partial x_\mu} \approx f(z)$$

which, when resolved in Fourier components, are expressed by the relations $k^\mu \eta_{\mu k} \approx b(z)$

and its complex conjugate

$$k^\mu \bar{\eta}_{\mu k} \approx \bar{b}(z).$$

They provide the following conditions on the wave function:

$$(k^\mu \eta_{\mu k} - b)\psi = 0 , \qquad (14.1)$$

$$(k^\mu \bar{\eta}_{\mu k} - \bar{b})\psi = 0. \qquad (14.2)$$

We have these conditions for each value of k, but for the sake of simplicity let us consider just one k-value. A solution of

*P.A.M. Dirac, Pro. Roy. Soc. A183, 284 (1945).

(14.1) can be written

$$\psi = \frac{1}{k^\mu \eta_\mu - b} G \qquad (14.3)$$

where G is an ascending power series in η_1 , η_2, η_3 , and and $1/(k^\mu \eta_\mu - b)$ is to be expanded as an ascending power series in η_1, η_2 , η_3 and a descending power series in η_0. The condition of (14.2) can be written

$$(k^\mu \bar{\eta}_\mu - \bar{b}) \frac{1}{(k^\mu \eta_\mu - b)} G = 0 . \qquad (14.4)$$

Now since $$[k^\mu \bar{\eta}_\mu - \bar{b}, k^\nu \eta_\nu - b] = i k^\mu k^\nu g_{\mu\nu} = 0 ,$$

the order of the factors in (14.4) can be interchanged, and we have $$(k^\mu \bar{\eta}_\mu - \bar{b}) G = 0 ,$$

or using the fact that $\bar{\eta}_\mu = C \frac{\partial}{\partial \eta^\mu}$ where C is a certain constant factor determined by the commutation relations, we have $$(C k_\mu \frac{\partial}{\partial \eta_\mu} - \bar{b}) G = 0 .$$

This equation can be solved and the solution can be written in the form $$G = e^{-\bar{b} k_r \eta_r / k_s^2 C} G_1 \qquad (14.5)$$

where G_1 satisfies the equation

$$k_\mu \frac{\partial G_1}{\partial x_\mu} = 0 \qquad (14.6)$$

and is an ascending power series in η_1 , η_2 , η_3 .

The condition (14.6) on G_1 expresses the fact that G_1 involves only transverse waves.

The expression for the wave function (14.3) displays the relativistic form since $(k^\mu \eta_\mu - b)$ is an invariant and G can be any ascending power series in $\eta_0, \eta_1, \eta_2, \eta_3$. To obtain the previous G, which was an ascending power series in η_1, η_2, η_3 only, we note that due to the condition (14.1) any multiple of $(k^\mu \eta_\mu - b)$ can be added to G without producing any effect and in this way all the negative powers of η_0 can be eliminated.

Further, instead of (14.5) G can be written in the form

$$G = e^{\bar{b} l^\mu \eta_\mu / c} G_1$$

where l_μ has to be chosen so that $l_\mu k^\mu = 1$. G_1 is then a function only of the η's orthogonal to k. That is G_1 involves η^{tr}'s together with $k^\mu \eta_\mu$, but the latter can again be eliminated by expressing G_1 as a multiple of $(k^\mu \eta_\mu - b)$.

We thus see that we cannot use all the information at our disposal without spoiling the relativistic form of the wave function, although the information itself is relativistic.

15. Solution of the Wave Equation. We now turn to the question of whether the wave equation has any solution at all. The natural procedure is to expand the wave function in powers of the electronic charge e, and it is well known that this leads to divergence already in the second order term.

In our discussion we follow another procedure, and for the sake of simplicity restrict ourselves to the case of one electron. This simplest case is sufficient for our problem since the difficulties which arise stem from the interaction between an electron and the field which it produces, and not with the fields

produced by other electrons. Further let us suppose that the longitudinal waves have been eliminated. Then we have the wave equation

$$\{ p_0 + \alpha_r (p_r - e A_r^{tr}) + \alpha_m m \} \psi = 0 \qquad (15.1)$$

where A_r^{tr} is the transverse part of the electromagnetic potential and satisfies the condition

$$\frac{\partial A_r^{tr}}{\partial x_r} = 0 \qquad (15.2)$$

and the commutation relations

$$[A_r^{tr}(x), A_s^{tr}(x')] = \frac{1}{2} \{ G_{rs}(x - x' + \lambda) + G_{rs}(x - x' - \lambda) \} . \qquad (15.3)$$

Here G_{rs} is the transverse part of $g_{rs} \Delta(x)$, its Fourier components being

$$G_{rs\,k} = \left(g_{rs} - \frac{k_r k_s}{k_0^2} \right) \Delta_k$$

where Δ_k is a Fourier component of $\Delta(x)$.

Our procedure is to put

$$A_r^{tr}(x) = M_r(x) + N_r(x + \lambda) \qquad (15.4)$$

where M and N are field variables with the following properties: They describe waves which are propagated with the velocity of light, M and N satisfying the wave equation

$$\Box M = 0 , \quad \Box N = 0 , \qquad (15.5)$$

They describe transverse waves, both satisfying the condition

$$\frac{\partial M_r}{\partial x_r} = 0 , \quad \frac{\partial N_r}{\partial x_r} = 0 . \qquad (15.6)$$

All the components of M commute with each other at all points, and so do the components of N among themselves.

$$[M_r(x), M_s(x')] = 0 , \quad [N_r(x), N_s(x')] = 0 , \qquad (15.7)$$

and the components of M taken with those of N satisfy the commutation relations

$$[M_r(x), N_s(x')] = \tfrac{1}{2} G_{rs}(x-x') . \qquad (15.8)$$

It is readily verified that the expression of A in terms of the M and N variables (15.4) and the commutation relations (15.7) and (15.8) give the required commutators (15.3) for the vector potential A.

Next we define a new field $B_r(x)$ by writing

$$B_r(x) = M_r(x) - N_r(x-\lambda) . \qquad (15.9)$$

Then any B-variable commutes with any A-variable

$$[B_r(x), A_s(x')] = 0$$

and

$$[B_r(x), B_s(x')] = -\tfrac{1}{2}\left\{ G_{rs}(x-x'+\lambda) + G_{rs}(x-x'-\lambda) \right\} . \qquad (15.10)$$

We note that the commutators for the B's differ from those for the A's only by a difference in sign. Since the B's commute with all the dynamical variables occurring in the Hamiltonian, they have no physical importance, and they can be considered as redundant variables.

Let us now consider the role played by a redundant variable in quantum mechanics. If we denote by q the essential variables and by r the redundant variable, we have a wave function $\psi(q,r)$ and the wave equation involves only q and operators that operate on q.

The wave function will be a solution for each value of r, and in order to obtain physical results we must sum or integrate over r. Thus the probability that the operator q have the

value q' is given by

$$\int |\Psi(q',r)|^2 \, dr \, .$$

In fact a single wave function corresponds to a Gibb's ensemble and not to a physical state. To be quite general we should introduce a weight factor $\rho(r)$ in the integration

$$\int |\Psi(q',r)|^2 \rho(r) \, dr \, .$$

For the same wave function, the physical interpretation is different for different weight factors.

The introduction of the B-field brings in great arbitrariness in our theory, and it is permissible to set up new supplementary conditions as long as they satisfy the compatibility relations and do not restrict the physical solutions too much. We shall impose the condition

$$N_r(x)\,\Psi = 0 \qquad \text{for all } x \text{ in the future light-cone of } z \, . \qquad (15.11)$$

These conditions are compatible among themselves because the components of N commute with each other, and they are compatible with the wave equation, since, as is easily verified

$$G_{rs}(x-z) = 0 \quad \text{for} \quad (x-z)^2 > 0 \, .$$

There are no other compatibility conditions since the longitudinal waves have been eliminated. For $z = -\infty$, we have $N_r(x)\,\Psi = 0$ which we may write as $N_r(x) \approx 0$ and from (15.4) and (15.7) this is equivalent to $A^{tr}_r(x) \approx B_r(x)$. That is, the B-field is the initial value of the A-field. If we denote the Fourier components of A^{tr} by η and $\bar{\eta}$ and those of B by ξ and $\bar{\xi}$, then initially

$$\eta \approx \xi \, , \quad \bar{\eta} = \bar{\xi} \, . \qquad (15.12)$$

Now we have seen that the wave function in our theory can be expressed as an ascending power series in η, and its dependence on the B-field can be expressed as an ascending power series in $\bar{\xi}$. We have $\bar{\xi}$ and not ξ for this power series since the commutation relations (15.10) for the B-field have opposite sign to those of the A-field. Then the conjugates $\bar{\eta}$ and ξ can be replaced by the derivatives

$$\bar{\eta} = c\frac{\partial}{\partial \eta} ,$$

$$\xi = c\frac{\partial}{\partial \bar{\xi}} .$$

The conditions (15.12) can now be written in the form

$$(\eta - c\frac{\partial}{\partial \bar{\xi}})\psi = 0 ,$$

$$(\bar{\xi} - c\frac{\partial}{\partial \eta})\psi = 0 . \qquad (15.13)$$

These equations show that the dependence of ψ on η and $\bar{\xi}$ must be of the form

$$\psi :: e^{\eta\bar{\xi}/c} = 1 + \eta\bar{\xi}/c + \eta^2\bar{\xi}^2/2!c + \ldots .$$

The general term $\eta^n\bar{\xi}^n/n!c$ in the power series represents a state with n photons in the A-field and n photons in the B-field. Thus the effect of the supplementary condition

$$N_r(x)\psi = 0$$

is to have the same number of photons in the A-field and in the B-field initially.

Since the conditions (15.11) hold for all x in the future light cone of z, we have $N(z+\lambda)\psi = 0$ provided we take λ in the future light-cones (i.e. $\lambda_0 > 0$) and the wave equation (15.1) reduces to

$$\{ p_0 + \alpha_r(p_r - eM_r(z)) + \alpha_m m \}\psi = 0 .$$

But now the M-field which is left in the wave equation commutes with all the variables in the equation, and the wave equation is exactly the same as for an unquantized field theory. The equation has a solution for any value of the initial momentum of the electron, and it can be written as a functional in the M-field

$$\Psi(\ldots, \sigma_3, \rho_3, \ldots, M(z))$$

where σ_3 and ρ_3 are spin variables. The wave function is now expressed in a representation in which M is diagonal.

The commutation relations (15.7) and (15.8) between the M- and N-fields show that the relation between them is somewhat like that between coordinates and their conjugate momenta, and the $N_r(x)$ in the supplentary conditions (15.11) can be considered as a sort of a differential operator acting on the M-field. In this way (15.11) may be interpreted as a condition which states that Ψ must be unchanged under a certain change in the M-field. On account

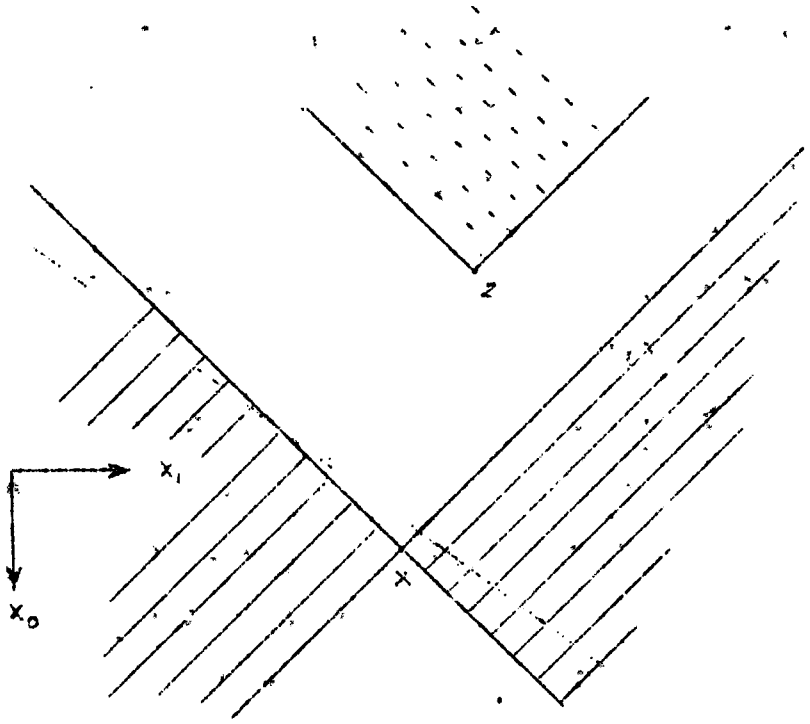

Figure 15.1

of the fact that $G_{rs}(x) = 0$ for $x^2 > 0$, when $N_r(x)$ is applied to $\Psi(z, \sigma_3, \rho_3, M(z))$ it refers to a change in the M-field consisting of only waves in the region outside and on the light-cone of x (shaded region in Fig. 15.1). With x in the future light-cone of z, the M-field in the past light-cone of (the dotted region in Fig. 15.1) is unaffected by this change. It is easily verified that the condition (15.11) holding for any x in the future light-cone of z means that any change in the M-field, satisfying the first of equations (15.5) and (15.6) and such that the M-field in the past light-cone of z is unaffected, leaves Ψ invariant. Thus (15.11) means that Ψ depends only on the M-field in the past light-cone of z.

As mentioned earlier, if we have one solution of the wave equation, we can apply it to different physical situations by choosing different weight factors. However the difficulty is that the wave function $\Psi(M_r)$ in which M is diagonal is not suitable for physical interpretation. It is necessary to transform it to a representation $\Psi(\eta, \bar{\xi})$ in which it appears as a power series in the η's and $\bar{\xi}$'s the various coefficients corresponding to various numbers of photons actually present, and various numbers of photons in the initial field. But this problem of transformation of the representation from $\Psi(M_r)$ to $\Psi(\eta, \bar{\xi})$ is very difficult, and has not been solved yet. If one attempts to do it by using a power series expansion in e, the same divergences which occurred in previous theories appear again.

There is the philosophical objection to the theory that even for the simplest problem of one electron with no ingoing

electromagnetic field, the mathematics is excessively complicated. This probably indicates that the present theory is not the final one.

Another difficulty which exists in the present theory is the separation of physical solutions from the non-physical. A wave function in quantum mechanics corresponds to a family of classical solutions. In the one-electron problem, most of the latter correspond to the case of self-acceleration and are hence non-physical, and the corresponding wave function gives finite probability for emission of radiation even when there is no incident wave. A precise way of separating out the physical from the non-physical solutions has not as yet been formulated.

Appendix 1. Derivation of the Poisson brackets between $\eta_{\mu k}$.

First we obtain the inverse of the Fourier transformation (9.1) for $A_\mu(x)$ which is

$$A_\mu(x) = \int \left\{ \eta_{\mu k}\, e^{i(k,x)} + \bar{\eta}_{\mu k}\, e^{-i(k,x)} \right\} \frac{d^3k}{k_0} .$$

Multiplying by $e^{i\vec{k}\cdot\vec{x}}$ ($\vec{k}\cdot\vec{x}$ denoting a three-dimensional scalar product) and integrating over d^3x, we obtain

$$\int A_\mu(x)\, e^{i\vec{k}\cdot\vec{x}}\, d^3x = \frac{(2\pi)^3}{k_0} \left\{ \eta_{\mu k}\, e^{i k_0 x_0} + \bar{\eta}_{\mu -k}\, e^{-i k_0 x_0} \right\}$$

Differentiating this with respect to x_0, we obtain

$$\int \frac{\partial A_\mu(x)}{\partial x_0}\, e^{i\vec{k}\cdot\vec{x}}\, d^3x = i(2\pi)^3 \left\{ \eta_{\mu k}\, e^{i k_0 x_0} - \bar{\eta}_{\mu -k}\, e^{-i k_0 x_0} \right\} .$$

From these two equations we obtain

$$\eta_{\mu k} = \frac{1}{2(2\pi)^3}\, e^{-i k_0 x_0} \int \left(k_0 - i\frac{\partial}{\partial x_0} \right) A_\mu(x)\, e^{i\vec{k}\cdot\vec{x}}\, d^3x ,$$

and we also check that the expression for $\bar{\eta}_{\mu k}$ obtained in the same way coincides with the complex conjugate of $\eta_{\mu k}$ above which should be the case for $A_\mu(x)$ to be real.

Now for the Poisson brackets, we have

$$[\eta_{\mu k}, \bar{\eta}_{\nu k'}] = \frac{1}{4(2\pi)^6}\, e^{-i(k_0 x_0 - k_0' x_0')} \iint \left(k_0 - i\frac{\partial}{\partial x_0} \right)\left(k_0' + i\frac{\partial}{\partial x_0'} \right) [A_\mu(x), A_\nu(x')]\, e^{i(\vec{k}\cdot\vec{x} - \vec{k}'\cdot\vec{x}')}\, d^3x\, d^3x' .$$

Introducing newspace-time coordinates defined by $X = x - x'$, $Y = \frac{1}{2}(x + x')$ and inserting the value $[A_\mu(x), A_\nu(x')] = \frac{1}{2} g_{\mu\nu} \{ \Delta(X+\lambda) + \Delta(X-\lambda) \}$ we obtain

$$[\eta_{\mu k}, \bar{\eta}_{\nu k'}] = \frac{g_{\mu\nu}}{8(2\pi)^6}\, e^{-i(k_0 - k_0')Y_0 - \frac{i}{2}(k_0 + k_0')X_0} \iint \left(k_0 - i\frac{\partial}{\partial X_0} \right) \left(k_0' - i\frac{\partial}{\partial X_0} \right) \{ \Delta(X+\lambda) + \Delta(X-\lambda) \}\, e^{i(\vec{k}-\vec{k}')\cdot\vec{Y} + \frac{i}{2}(\vec{k}+\vec{k}')\cdot\vec{X}}\, d^3X\, d^3Y$$

the integration over Y can be performed immediately and it yields a δ-function.

$$[\eta_{\mu k}, \bar{\eta}_{\nu k'}] = \frac{g_{\mu\nu}}{8(2\pi)^3}\, \delta_3(k - k')\, e^{-i k_0 X_0} \int \left(k_0 - i\frac{\partial}{\partial X_0} \right)^2 \{ \Delta(X+\lambda) + \Delta(X-\lambda) \}\, e^{i\vec{k}\cdot\vec{X}}\, d^3X$$

Now the Δ-function has the integral representation

$$\Delta(x) = \frac{4\pi}{(2\pi)^3} \int e^{i\vec{K}\cdot\vec{x}} \frac{\sin K_0 x_0}{K_0} \, d^3K$$

so that

$$\int (k_0 - i\frac{\partial}{\partial X_0})^2 \Delta(X+\lambda)\, e^{i\vec{k}\cdot\vec{X}} \, d^3X$$

$$= \frac{4\pi}{(2\pi)^3} \iint (k_0 - i\frac{\partial}{\partial X_0})^2 \, e^{i\vec{K}\cdot(\vec{X}+\vec{\lambda}) + i\vec{k}\cdot\vec{X}} \frac{\sin K_0(X_0+\lambda_0)}{K_0} \, d^3K \, d^3X$$

$$= 4\pi \int (k_0 - i\frac{\partial}{\partial X_0})^2 \delta_3(K+k)\, e^{i\vec{K}\cdot\vec{\lambda}} \frac{\sin K_0(X_0+\lambda_0)}{K_0} \, d^3K$$

$$= 4\pi \, e^{-i\vec{k}\cdot\vec{\lambda}} (k_0 - i\frac{\partial}{\partial X_0})^2 \frac{\sin k_0(X_0+\lambda_0)}{k_0}$$

$$= -i\,8\pi\, k_0\, e^{-i\vec{k}\cdot\vec{\lambda} + ik_0(X_0+\lambda_0)}$$

Hence

$$[\eta_{\mu k}, \bar{\eta}_{\nu k'}] = -i \frac{g_{\mu\nu}}{8\pi^2} k_0 \, \delta_3(k-k') \, e^{-i\vec{k}\cdot\vec{\lambda} + ik_0\lambda_0} + \text{same term with } \lambda \text{ replaced by } -\lambda$$

$$= -i \frac{g_{\mu\nu}}{4\pi^2} k_0 \, \delta_3(k-k') \cos(k,\lambda) \, .$$