ANNALS OF PHYSICS: **62**, 44–97 (1971)

Electrodynamics of Direct Interparticle Action II. Relativistic Treatment of Radiative Processes

F. HOYLE AND J. V. NARLIKAR

Institute of Theoretical Astronomy, University of Cambridge, Cambridge, England

Received April 21, 1970

This paper is a sequel to an earlier paper that described nonrelativistic quantum electrodynamics in terms of the time symmetric theory of direct interparticle action. The restriction to a nonrelativistic treatment is removed in the present paper. The path integral approach to quantum mechanics is extended to include relativistic particles with spin as a preliminary to achieving this end.

The response of the Universe can be obtained in two ways, by an extension of the method of the previous paper or by using a general condition analogous to the general condition, $\sum_{\text{particles}} [F_{\text{ret}} - F_{\text{adv}}] = 0$, of Wheeler and Feynman. The second method is used in the present paper since it avoids the need to calculate the detailed properties of the absorber. The rules of quantum electrodynamics in the form stated by Feynman are derived, thereby showing that all the practical results of the usual theory of quantum electrodynamics can also be obtained from the direct particle theory.

Radiation corrections are discussed. The direct particle theory throws some light on the processes of renormalization. The point of view that both mass and charge should be finite before as well as after renormalization is partially developed.

1. INTRODUCTION

The purpose of this paper is to show that electromagnetic interaction can be fully described by the theory of direct interparticle action—that an independent quantized field with photons is unnecessary. In an earlier paper [1, referred to hereafter as I] we discussed the same problem within the framework of a nonrelativistic theory. We found that the path integral approach to quantum mechanics was more convenient than the usual Schrödinger approach. In the present paper we wish to widen the scope of the discussion to relativistic spin $\frac{1}{2}$ particles, i.e., particles described by the Dirac equation, so as to include the full range of phenomena studied in quantum electrodynamics. As in the nonrelativistic case the discussion is divided into two parts: (i) the motion of the Dirac particles and (ii) the nature of the electromagnetic response of the Universe to the behavior of a local system. Of these only the second part is specifically connected with the direct particle theory.

In I the response was obtained by looking at the detailed processes occurring in the future absorber, when the local system makes a quantum transition. This method was the quantum analogue of the derivations I and II of Wheeler and Feynman [2] in the classical absorber theory of radiation. Although this method has the advantage that it stresses the important part played by the absorber in the whole calculation, it leaves the impression that the result is dependent on particular properties of the absorber. To correct this impression, Wheeler and Feynman gave another derivation (cf. IV in [2]) which leads to a general condition to be satisfied by the absorber in order to give the correct response. In the present paper we shall obtain the quantum analogue of this general condition in a relativistic form.

The first part of the discussion—the motion of Dirac particles—presents difficulties, however. Following I, it would seem that a path integral method will be more useful in the second part of the discussion than the straightforward use of the Dirac equation. Here we meet our main difficulty, since the nonrelativistic path integral formulation is not easily generalized to relativistic particles with spin.

In the nonrelativistic case the action for a free particle is given by

$$S = \int_{t_1}^{t_2} \tfrac{1}{2} m \dot{\mathbf{a}}^2 \, dt, \qquad t_2 > t_1 , \tag{1}$$

where $\mathbf{a}(t)$ is a typical path Γ starting at point $1 : (\mathbf{a}_1 , t_1)$ and ending at point $2 : (\mathbf{a}_2 , t_2)$, and m is the mass of the particle. In the path integral approach, the amplitude for the particle to go from 1 to 2 along Γ is given by

$$P(\Gamma) = (\text{const.}) \exp(iS), \tag{2}$$

and the quantum mechanical propagator from 1 to 2 is given by summing $P(\Gamma)$ for all Γ,

$$K(2;1) = \sum_{\Gamma} P(\Gamma) = \int \exp(iS) \, \mathscr{D}^3\mathbf{a}. \tag{3}$$

Here and throughout the paper we take $\hbar = 1$, $c = 1$.

Provided the sum over Γ is a path integral with a suitably defined measure, $K(2; 1)$ can be shown to satisfy the inhomogeneous Schrödinger equation

$$\left[\frac{\partial}{\partial t_2} + \frac{1}{2im} \nabla_2{}^2\right] K(2;1) = \delta_4(2,1). \tag{4}$$

Explicitly, $K(2; 1)$ is given by

$$K(2;1) = \begin{cases} \left[\dfrac{m}{2\pi i(t_2 - t_1)}\right]^{3/2} \exp\left\{\dfrac{im(\mathbf{a}_2 - \mathbf{a}_1)^2}{2(t_2 - t_1)}\right\}, & t_2 > t_1 , \\ 0, \quad t_2 < t_1 . \end{cases} \tag{5}$$

Instead of using (2) we can also define $P(\Gamma)$ as a chain of propagators along Γ. Divide Γ into a large number of small segments, $[\mathbf{a}^{(i)}, t^{(i)}]$, $i = 0, 1,..., N$, denoting the points of division, and $[\mathbf{a}^{(0)}, t^{(0)}]$, $[\mathbf{a}^{(N)}, t^{(N)}]$ corresponding to 1, 2 respectively. $P(\Gamma)$ is given by

$$P(\Gamma) = \prod_{i=1}^{N} A_i^{-1} K[\mathbf{a}^{(i)}, t^{(i)}; \mathbf{a}^{(i-1)}, t^{(i-1)}], \tag{6}$$

where the A_i are constants.

Consider now the Dirac particle of mass m. If we try to describe its behavior by a similar method, starting with the relativistically invariant action

$$S = -\int m\, da, \tag{7}$$

where da is the element of proper time along the world line of a, we run immediately into trouble since the classical action (7) does not contain spin. A path integral like (3), using (7) for S, will not therefore reproduce the properties associated with the Dirac particle. We may follow instead the alternative approach suggested by (6). We look for a solution $K_0(2; 1)$ of the inhomogeneous Dirac equation

$$(\nabla\!\!\!/_2 + im)\, K_0(2; 1) = \delta_4(2, 1), \tag{8}$$

with the boundary condition

$$K_0(2; 1) = 0, \qquad t_2 < t_1 . \tag{9}$$

Such a solution exists and is given by

$$K_0(2; 1) = \frac{1}{2\pi} (\nabla\!\!\!/_2 - im)\Big[\delta(q_{21}^2) - \frac{m}{2q_{21}} J_1(mq_{21})\, \theta(q_{21}^2)\Big], \tag{10}$$

for $t_2 > t_1$, and by (9) for $t_2 < t_1$, where

$$q_{21}^2 = (t_2 - t_1)^2 - (\mathbf{a}_2 - \mathbf{a}_1)^2, \tag{11}$$

$\theta(x)$ is the heaviside function, and J_1 is the Bessel function of order unity. We now define $P(\Gamma)$ by

$$P(\Gamma) = \prod_{i=1}^{N} A_i^{-1} K_0[\mathbf{a}^{(i)}, t^{(i)}; \mathbf{a}^{(i-1)}, t^{(i-1)}], \tag{12}$$

where the A_i are constants. Since for small q_{21} the delta function $\delta(q_{21}^2)$ dominates in (10), we may approximate a continuous path Γ by a zigzag path made up of null segments. We shall return to this point later. With suitable constants A_i (which depend on the measure for the path integral) we can write

$$K_0(2; 1) = \sum_{\Gamma} P(\Gamma) = \int P(\Gamma)\, \mathscr{D}^3\Gamma. \tag{13}$$

The above approach is that used by Feynman [3] in his discussion of the quantum theory of electromagnetic action. Here in this introduction we summarize Feynman's treatment and results. In the following section we give an alternative method of arriving at the same results. The difference is that Feynman's work makes use of second quantization of the particles (cf. [4, Appendix]), whereas the following section uses only the concepts of the path integral method.

Consider an electron of charge e in an external potential B_i. This field is due to other particles and is unquantized—it is a direct particle field. The amplitude (12) is modified to

$$P^B(\Gamma) = P(\Gamma) \exp\left[-ie \int_{\Gamma} B_i\, da^i\right], \tag{14}$$

and (13) is changed to

$$K_0{}^B(2; 1) = \int P^B(\Gamma)\, \mathscr{D}^3\Gamma. \tag{15}$$

By expanding the exponential in (14) and performing the path integrals for each term we get a perturbation expansion for $K_0{}^B(2; 1)$,

$$\begin{aligned} K_0{}^B(2; 1) = {} & K_0(2; 1) + (-ie) \int K_0(2; 3)\, \not{B}(3)\, K_0(3; 1)\, d\tau_3 \\ & + (-ie)^2 \iint K_0(2; 3)\, \not{B}(3)\, K_0(3; 4)\, \not{B}(4)\, K_0(4; 1)\, d\tau_3\, d\tau_4 \\ & + \cdots . \end{aligned} \tag{16}$$

The integrations with respect to $d\tau_3$, $d\tau_4$,... are over the four-dimensional region between the planes $t = t_1$, $t = t_2$. If $2'$ is any point in the interior of this region and $K_0{}^B(2'; 1)$ is defined by replacing point 2 by $2'$ in (16), then it is easily verified that $K_0{}^B(2'; 1)$ satisfies the inhomogeneous Dirac equation

$$[\not{\nabla}_2{}' + ie\not{B}(2') + im]\, K_0{}^B(2'; 1) = \delta_4(2', 1). \tag{17}$$

With the definition

$$\not{\nabla}_2 K_0{}^B(2; 1) = \lim_{2' \to 2} [\not{\nabla}_{2'} K_0{}^B(2'; 1)], \tag{18}$$

$K_0{}^B(2; 1)$ also satisfies the inhomogeneous Dirac equation

$$[\nabla\!\!\!/_2 + ie B\!\!\!/(2) + im]\, K_0{}^B(2; 1) = \delta_4(2, 1). \tag{19}$$

So far the discussion has run parallel to the nonrelativistic case, but now we encounter a new situation. $K_0(2; 1)$ can be written in terms of a complete set of suitably normalized stationary states u_n with energies E_n of the free particle Dirac equation,

$$K_0(2; 1) = \begin{cases} \sum_{E_n} u_n(2)\, \bar{u}_n(1), & t_2 > t_1 , \\ 0 \quad , & t_2 < t_1 . \end{cases} \tag{20}$$

The set $\{E_n\}$ contains negative as well as positive energies. The negative energy states do not affect the propagation of wavefunctions made up entirely of positive energy states. However, such a restriction cannot be maintained if we are dealing with a charged particle in an external potential B_i. As seen in (16), positive energy states can be scattered into negative energy ones by B_i.

To avoid a cascading into negative energy states we need the hole theory and the sea of negative energy states all filled according to the Pauli principle. The problem is therefore that of many electrons, each described by a K_0. In the path integral formulation we now have to deal with a multiple expression

$$\iint \cdots \int P(\Gamma_1, \Gamma_2, ..., \Gamma_n, ...)\, \mathscr{D}^3\Gamma_1 \mathscr{D}^3\Gamma_2 \cdots \mathscr{D}^3\Gamma_n \,.... \tag{21}$$

together with the restrictions imposed by the hole theory—in particular the antisymmetric restriction on the wavefunction.

The greatly increased complexity of this many-particle problem can be ameliorated by Feynman's use of the propagator $K_+(2; 1)$, defined by

$$K_+(2; 1) = \begin{cases} \sum_{E_n > 0} u_n(2)\, \bar{u}_n(1), & t_2 > t_1 , \\ -\sum_{E_n < 0} u_n(2)\, \bar{u}_n(1), & t_2 < t_1 , \end{cases} \tag{22}$$

which also satisfies the inhomogeneous Dirac equation (8). We can similarly build up a propagator $K_+{}^B(2; 1)$ in the presence of B_i by using the perturbation expansion (16) but with K_+ replacing K_0. This new propagator satisfies

$$(\nabla\!\!\!/_2 + ie B\!\!\!/(2) + im)\, K_+{}^B(2; 1) = \delta_4(2, 1), \tag{23}$$

provided the derivatives of $K_+{}^B(2; 1)$ are defined in the same way as those of $K_0{}^B(2; 1)$.

The result proved by Feynman [4, Appendix] is that we can ignore the hole theory in an amplitude calculation provided $K_+{}^B(2;1)$ is used instead of $K_0{}^B(2;1)$, and provided the amplitude is multiplied by C_v, the amplitude for vacuum to remain a vacuum. C_v is given by

$$C_v = \exp(-L), \tag{24}$$

where

$$L = \sum_{n\geqslant 2} L^{(n)}. \tag{25}$$

$L^{(n)}$ is the amplitude for the occurrence of a closed loop in which B_i acts n times. Thus

$$L^{(n)} = \frac{(-ie)^n}{n} \int \cdots \int \mathrm{Tr}[K_+(n;1)\, \not{B}(1)\, K_+(1;2)\, \not{B}(2) \cdots K_+(n-1;n)\, \not{B}(n)]\, d\tau_1\, d\tau_2 \cdots d\tau_n\,. \tag{26}$$

$L^{(n)}$ for n odd vanishes by Furry's theorem [5].

In the treatment outlined above the path integral method is used only in arriving at the hole theory, not directly in arriving at the propagator K_+ or the perturbation expansion in terms of K_+. The difference between the hole theory and the theory using K_+ can be seen as follows. In the hole theory all paths go forward in time and lead to the propagator K_0. In the hole theory we cannot talk of scattering of a single electron by B_i; we must also take into account pair creation and annihilation arising from the negative energy sea. In the Feynman positron theory, on the other hand, we can talk of scattering a single electron, provided we allow the electron to be scattered backwards in time and provided we introduce C_v in the manner described above.

At the outset we stated that the discussion would be divided into two parts: (i) the motion of Dirac particles and (ii) the nature of the electromagnetic response of the Universe to the behavior of a local system. Of these two parts our essential concern is with (ii). The treatment of (i) outlined above is sufficient to enable us to deal with (ii). It turns out, however, that (ii) is more readily treated when (i) is formulated in such a way that the perturbation expansion in K_+ is obtained directly from a path integral. This will be done in Section 2 and the discussion of (ii) will be postponed to Section 3.

To conclude this introduction we note that the theory so far developed refers to particles acted on by an external potential—it does not yet include interactions between pairs of particles directly. Anticipating the results of Section 3, the interaction between particles a, b with charges e_a, e_b is given by the action.

$$R[\mathbf{a}, \mathbf{b}] = -e_a e_b \iint \delta_+(q_{AB}^2)\, da^i\, db_i\,, \tag{27}$$

where da^i, db_i are four-dimensional coordinate elements along the paths $\mathbf{a}(t)$, $\mathbf{b}(t)$, and $\delta_+(x)$ is defined by

$$\delta_+(x) = \frac{1}{\pi} \int_0^\infty e^{-i\omega x}\, d\omega. \tag{28}$$

q_{AB}^2 is the square of the invariant distance between typical points A, B on the paths $\mathbf{a}(t)$, $\mathbf{b}(t)$ of a, b respectively. To consider the interaction of N particles $a, b, \ldots$ we must work with the multiple path integral

$$\int \cdots \int \cdots \int \cdots P(\Gamma_1, \ldots, \Gamma_a, \ldots, \Gamma_b, \ldots) \exp \left\{ \tfrac{1}{2} \sum_a \sum_b R[\mathbf{a}, \mathbf{b}] \right\} \times \mathscr{D}^3\Gamma_1 \cdots \mathscr{D}^3\Gamma_a \cdots \mathscr{D}^3\Gamma_b \cdots . \tag{29}$$

Again (29) can be handled within the framework of the hole theory and the K_0 propagator. Can we use the more elegant method of the K_+ propagator also in this case?

This question has been answered in the affirmative by Feynman [3, section 7]. By considering the electromagnetic quantized field as an independent entity and by eliminating the quantized field oscillators altogether, Feynman arrived at the same expression (27) as we shall obtain in Section 3 for the response of the Universe to the unquantized direct particle field. He then showed that a perturbation expansion in terms of K_+ is still possible, and gave the following rule:

"Suppose that we have a process to order k in e^2 (i.e., having k virtual photons) and order n in the external potential B_μ. Then the matrix element for the process with one more virtual photon and two less potentials is that obtained from the previous matrix by choosing from the n potentials a pair, say $B_\mu(1)$ acting at 1 and $B_\nu(2)$ acting at 2, replacing them by $ie^2\delta_{\mu\nu}\delta_+(s^2{}_{12})$, adding the results for each way of choosing the pair, and dividing by $k + 1$, the present number of photons."[1]

We can apply this rule together with an induction argument starting with $k = 0$, this being the case of external potentials only. Since we already know that a perturbation expansion with the K_+ propagator can be used in the latter case the required result follows for any $k \geqslant 0$.

2. A Path Integral Method Leading to the Perturbation Expansion in Terms of the K_+ Propagator

We begin with the free particle case. Define two propagators $K_0{}^{\pm}$ by

$$\begin{aligned} K_0{}^+(2;1) &= \theta(t_2 - t_1) \sum_n u_n(2)\, \bar{u}_n(1), \\ K_0{}^-(2;1) &= -\theta(t_1 - t_2) \sum_n u_n(2)\, \bar{u}_n(1). \end{aligned} \tag{30}$$

[1] We are using q for the invariant distance instead of s.

Both satisfy the inhomogeneous Dirac equation

$$(\nabla\!\!\!/_2 + im)\, K_0^{\pm}(2; 1) = \delta_4(2, 1), \tag{31}$$

and both are confined within the light cone at point 1, $K_0^+(2; 1)$ being nonzero only for points 2 inside or on the forward half of this cone, and $K_0^-(2; 1)$ being similarly confined to the backward half of the cone. Evidently K_0^+ is the same as the K_0 of the preceding section. Here, however, we have a time symmetric situation and our $\pm$ notation refers to going forwards or going backwards in time.

Corresponding to these propagators we distinguish two types of path, Γ_{21}^+ going forwards in time from point 1 to 2 ($t_2 > t_1$) and a path Γ_{21}^- going backward from 1 to 2 ($t_2 < t_1$). Both are monotonic with respect to time.

To define amplitude along Γ_{21}^+ divide $t_2 - t_1$ into a large number of small subintervals at intermediate points X_i, $i = 0, 1,..., N$. Let $i = 0$ correspond to point 1 and $i = N$ to point 2. Define $P(\Gamma_{21}^+)$ as

$$\prod_{i=1}^{N} A_i^{-1} K_0^+(i; i-1), \tag{32}$$

for N large. We shall discuss whether or not N should tend to infinity in a later section. The amplitude $P(\Gamma_{21}^-)$ for a Γ^- path is similarly defined with a chain of $K_0^-(i; i-1)$ propagators. Clearly the paths Γ_{21}^+ and Γ_{21}^- must be timelike throughout.

The measure for the sum over paths is so defined that

$$K_0^{\pm}(2; 1) = \sum_{\Gamma_{21}^{\pm}} P(\Gamma_{21}^{\pm}) = \int P(\Gamma_{21}^{\pm})\, \mathscr{D}^3 \Gamma_{21}^{\pm} \,. \tag{33}$$

Suppose we study a free particle in the four-dimensional slab $t_1 \leqslant t \leqslant t_2$. At $t = t_1$ paths coming from the past represent an amplitude from some previous history of the particle. We denote this amplitude by ψ_+. Similarly at $t = t_2$ paths coming from the future give an amplitude which we denote by ψ_-. Thus the amplitude at a point 3 within the slab is given by paths Γ_{31}^+ starting at points 1 on $t = t_1$ and weighted by $\psi_+(1)$, together with paths Γ_{32}^- starting at points 2 on $t = t_2$ and weighted by $\psi_-(2)$. That is to say, the amplitude $\psi(3)$ is given by

$$\psi(3) = \int K_0^+(3; 1)\, \gamma_4 \psi_+(1)\, d^3\mathbf{x}_1 - \int K_0^-(3; 2)\, \gamma_4 \psi_-(2)\, d^3\mathbf{x}_2 \,. \tag{34}$$

If ψ_+ is made up of positive energy solutions u_n at $t = t_1$ then ψ_+ for $t > t_1$ is always made up of the same positive energy solutions. Similarly if ψ_- is made up of negative energy solutions at $t = t_2$ then ψ_- for $t < t_2$ is always made up of negative energy solutions. We impose this property as a boundary condition. Then (34) is equivalent to

$$\psi(3) = \int K_+(3; 1)\, \gamma_4 \psi_+(1)\, d^3\mathbf{x}_1 - \int K_+(3; 2)\, \gamma_4 \psi_-(2)\, d^3\mathbf{x}_2 . \tag{35}$$

Suppose now that we are given an external potential B_i nonzero in $t_1 \leqslant t \leqslant t_2$ but zero outside this range. This does not interfere with the boundary condition just discussed. We now admit that within the slab $t_1 \leqslant t \leqslant t_2$ interaction with B_i can turn the time direction of a path. Thus paths within the slab can be zigzag with as many reversals as we like, whereas paths outside the slab are monotonic with respect to t.

The amplitude for a Γ_{21}^+ path is defined by

$$P^B(\Gamma_{21}^+) = P(\Gamma_{21}^+) \exp\left[-ie \int_{\Gamma_{21}^+} B_i\, da^i\right], \tag{36}$$

and that for a Γ_{12}^- path (point 1 on $t = t_1$, point 2 on $t = t_2$) by

$$P^B(\Gamma_{12}^-) = P(\Gamma_{12}^-) \exp\left[-ie \int_{\Gamma_{12}^-} B_i\, da^i\right]. \tag{37}$$

Paths within the slab need not be monotonic with respect to t, however. Suppose we have a path from 1 to 2 with $2n$ reversals. Denoting intermediate points by i, sections between successive reversals are monotonic, and the amplitude is given by

$$P^B(\Gamma_{21}) = \prod_i P^B(\Gamma_{i,i-1}^{\pm}), \tag{38}$$

where the $+$sign holds for forward going sections and the $-$sign for the backward going sections.

Clearly it is also possible to have paths starting at t_1 and ending at t_1. In this case we get a ψ_- at $t = t_1$ even though there may be no ψ_- at $t = t_2$. This does not happen when $B_i = 0$. Similarly we could have a $\psi_+ \neq 0$ at t_2 originating solely from ψ_- at t_2, through reversals by B_i.

We now wish to calculate the following problem. Given ψ_+ on $t = t_1$ and $\psi_- = 0$ on $t = t_2$, what are ψ_+ on $t = t_2$ and ψ_- on $t = t_1$?

Consider the evaluation of $\psi_+(t_2)$. Formally we have

$$\psi_+{}^B(2) = \int\int P^B(\Gamma_{21})\, \gamma_4 \psi_+(1)\, \mathscr{D}^3\Gamma_{12} d^3\mathbf{x}_1 , \tag{39}$$

which includes paths with reversals. To evaluate (39) consider the paths according to the number of reversals, beginning with paths Γ_{21}^+ without reversals.

Divide the time range $t_1 \leqslant t \leqslant t_2$ into a large number of small intervals. Consider one such interval $(t, t + \epsilon)$. Let the path Γ_{21}^+ intersect the time sections at t and $t + \epsilon$ at spatial points $\mathbf{x}$, $\mathbf{y}$ respectively. Since Γ_{21}^+ is timelike, $|\mathbf{y} - \mathbf{x}| \leqslant \epsilon$. Expand the exponential factor in (36) in a power series. The unity term gives the usual free particle contribution. The linear term gives

$$-ie \sum_{\substack{\text{all} \\ \text{intervals}}} \epsilon \iint \{B_4(\mathbf{x}, t) - (\mathbf{y} - \mathbf{x}) \cdot \mathbf{B}(\mathbf{x}, t)/\epsilon\} P(\Gamma_{21}^+) \gamma_4 \psi_+(1) \mathscr{D}^3\Gamma_{21}^+ d^3\mathbf{x}_1 . \quad (40)$$

Now $P(\Gamma_{21}^+)$ can be written as a product of three factors

$$P(\Gamma_{21}^+) = P(\Gamma_{2y}^+) P(\Gamma_{yx}^+) P(\Gamma_{x1}^+). \quad (41)$$

The sum over Γ_{21}^+ is equivalent to sums over Γ_{2y}^+, Γ_{yx}^+, Γ_{x1}^+, together with integrations over $\mathbf{x}$, $\mathbf{y}$. The sum over Γ_{x1}^+ propagates the positive energy part of the wavefunction at 1 to $\mathbf{x}$. This gives $\psi_+(\mathbf{x}, t)$. The sum over $P(\Gamma_{yx}^+)$ gives $K_0{}^+[\mathbf{y}, t + \epsilon; \mathbf{x}, t]$. The sum over $P(\Gamma_{2y}^+)$ similarly gives $K_0{}^+[2; \mathbf{y}, t + \epsilon]$. The latter function propagates only the forward going paths at $t + \epsilon$, and the weight of these paths is given by the positive energy part of the wavefunction at $(\mathbf{y}, t + \epsilon)$. In general, because of B_i in $(t, t + \epsilon)$, there will be a negative energy part in the wavefunction at $t + \epsilon$. This negative energy part gives the weight of backward going paths "reflected" by the step $(t, t + \epsilon)$. The negative energy part does not affect $\psi_+(2)$. We ensure that only the positive energy part of the wavefunction contributes by using $K_+[2; \mathbf{y}, t + \epsilon]$ instead of $K_0{}^+[2; \mathbf{y}, t + \epsilon]$. So we get for (40)

$$-ie \sum \epsilon \iint K_+[2; \mathbf{y}, t + \epsilon] \gamma_4 [B_4(\mathbf{x}, t) - (\mathbf{y} - \mathbf{x}) \cdot \mathbf{B}(\mathbf{x}, t)/\epsilon]$$
$$\cdot K_0{}^+(\mathbf{y}, t + \epsilon; \mathbf{x}, t) \gamma_4 \psi_+(\mathbf{x}, t) d^3\mathbf{x}\, d^3\mathbf{y} . \quad (42)$$

Because $|\mathbf{y} - \mathbf{x}| \leqslant \epsilon$ we can replace $K_+[2; \mathbf{y}, t + \epsilon]$ by $K_+[2; \mathbf{x}, t]$ to within a term that tends to zero with ϵ. Using (30) we have

$$K_0{}^+[\mathbf{y}, t + \epsilon; \mathbf{x}, t] = \sum_n U_n(\mathbf{y}) \bar{U}_n(\mathbf{x}) e^{-i\epsilon E_n}$$
$$= \delta_3(\mathbf{y}, \mathbf{x}) \gamma_4 - i\epsilon \sum_n H(\mathbf{y}) U_n(\mathbf{y}) \bar{U}_n(\mathbf{x}) + O(\epsilon^2), \quad (43)$$

where H is the Hamiltonian and

$$u_n(\mathbf{x}, t) = U_n(\mathbf{x}) e^{-iE_n t}. \quad (44)$$

Thus (42) becomes

$$-ie \sum \epsilon \int K_+[2; \mathbf{x}, t]\, \gamma_4 B_4(\mathbf{x}, t)\, \psi_+(\mathbf{x}, t)\, d^3\mathbf{x}$$

$$-ie \sum i\epsilon \iint K_+[2; \mathbf{x}, t]\, \gamma_4(\mathbf{y} - \mathbf{x}) \cdot \mathbf{B}(\mathbf{x}, t) \tag{45}$$

$$\cdot \left[\sum_n H(\mathbf{y})\, U_n(\mathbf{y})\, \bar{U}_n(\mathbf{x})\right] \gamma_4 \psi_+(\mathbf{x}, t)\, d^3\mathbf{x}\, d^3\mathbf{y}\,.$$

Now for a free particle

$$H \equiv -i\boldsymbol{\alpha} \cdot \boldsymbol{\nabla} + m\gamma_4$$

and the $m\gamma_4$ term makes no contribution to the second part of (45). The only portion of this second part that survives as $\epsilon \to 0$ is that which arises on differentiating $\mathbf{y} - \mathbf{x}$, using integration by parts with respect to $\boldsymbol{\nabla}\mathbf{y}$ from $H(\mathbf{y})$. Since $\sum_n U_n(\mathbf{y})\, \bar{U}_n(\mathbf{x}) = \delta_3(\mathbf{y}, \mathbf{x})\, \gamma_4$ it is easy to see that (45) is just

$$-ie \sum \epsilon \int K_+[2; \mathbf{x}, t]\, \not{B}(\mathbf{x}, t)\, \psi_+(\mathbf{x}, t)\, d^3\mathbf{x}. \tag{46}$$

For small ϵ the summation can be replaced by integration with respect to t. Denoting $(\mathbf{x}, t)$ as point 3 we therefore obtain

$$-ie \int K_+(2; 3)\, \not{B}(3)\, \psi_+(3)\, d\tau_3\,, \tag{47}$$

the integration with respect to $d\tau_3$ being over the four dimensional slab $t_1 \leqslant t \leqslant t_2$.

The second order term in the expansion of the exponential in (36) can be worked out in the same way and gives

$$(-ie)^2 \iint_{t_4 > t_3} K_+(2; 4)\, \not{B}(4)\, K_+(4; 3)\, \not{B}(3)\, \psi_+(3)\, d\tau_3\, d\tau_4\,, \tag{48}$$

and similarly for higher order terms. The appearance of K_+ propagators after each $\not{B}$ can be understood in the following terms. After each $\not{B}$ scattering the wavefunction appears in a general $\psi_+ + \psi_-$ form. Since we are propagating along a Γ^+ path we must use only the ψ_+ part. That is to say we need to act on ψ_+. But this is the same as K_+ acting on $\psi_+ + \psi_-$.

Consider now paths with reversals. Evidently to arrive at $t = t_2$ from $t = t_1$ we must have an even number of reversals. It will serve to bring out the main features if we treat the simplest case of two reversals. Consider a path that goes from t_1 to $t_3 > t_1$, then reverses to $t_4 < t_3$, and reverses again to $t_2 > t_4$. In order for paths of this kind to contribute to the amplitude the potential must be nonzero at t_3 and t_4. Again dividing $t_1 \leqslant t \leqslant t_2$ into small steps, the integral $\int B_i\, da^i$ is replaced by a sum over steps. This integral appears squared in the second order of the perturbation expansion, so that in this order we have a double sum over steps. The double sum contains a member that refers to the steps $(t_3 - \epsilon, t)$ and $(t_4, t_4 + \epsilon)$. This member gives the lowest order contribution that the path in question makes to the total amplitude.

We have to sum over all paths from t_1 to t_2 with reversals at t_3 and t_4. Since for such paths the particle is free (in the second order of the perturbation expansion) between t_1 and $t_3 - \epsilon$ it follows that ψ_+ at t_1 is carried into a ψ_+ wavefunction at $t_3 - \epsilon$. Within the thin slab $(t_3 - \epsilon, t_3)$ scattering by B_i changes the wavefunction into a $\psi_+ + \psi_-$ combination at t_3. Since the paths are now backwards and the particle is again free (second order of the perturbation expansion) between t_3 and $t_4 + \epsilon$ only the ψ_- part of the wavefunction at t_3 is carried to $t_4 + \epsilon$. Between $t_4 + \epsilon$ and t_4 there is a further scattering by B_i, and this again sets up a $\psi_+ + \psi_-$ combination at t_4. The paths now go forward to t_2 and since the particle is once again free only the ψ_+ part of the wavefunction at t_4 is carried to t_2. This "accounting" procedure is rather complicated when the $K_0{}^{\pm}$ propagators are used, but the K_+ propagator automatically keeps a correct record at each reversal. It is therefore preferable to use K_+ for the sections in which the particle is free.

Summing over all paths that reverse at times t_3, t_4 $(t_3 > t_4)$ in $[t_1, t_2]$ we obtain

$$(-ie)^2 \iint_{t_3>t_4} K_+(2;4)\, \not{B}(4)\, K_+(4;3)\, \not{B}(3)\, \psi_+(3)\, d\tau_3\, d\tau_4\,, \tag{49}$$

and (48) and (49) together give

$$(-ie)^2 \iint K_+(2;4)\, \not{B}(4)\, K_+(4;3)\, \not{B}(3)\, \psi_+(3)\, d\tau_3\, d\tau_4 \tag{50}$$

as the second order term of the perturbation expansion.

In this last step we are assuming that the case $t_3 = t_4$ gives no contribution to the amplitude. This case can be dealt with in the following way. When the path first reverses at t_3 the ψ_- part of the wavefunction is of order ϵ, being proportional to the thickness of the slab $(t_3 - \epsilon, t_3)$. When this ψ_- part is scattered again by the same slab the resulting ψ_+ part of the wavefunction is of order ϵ^2. This ψ_+ part then goes to t_2. We have a similar contribution for each thin slab, but since there are only of the order of ϵ^{-1} such slabs the total "double scattering" is of order ϵ and tends to zero with ϵ.

The above discussion is easily generalized to give

$$\psi_+(2) = \int K_+{}^B(2;1)\, \gamma_4 \psi_+(1)\, d^3\mathbf{x}_1\,, \tag{51}$$

where

$$K_+^B(2; 1) = K_+(2; 1) - ie \int K_+(2; 3) \not{B}(3) K_+(3; 1) d\tau_3$$

$$+ (-ie)^2 \iint K_+(2; 4) \not{B}(4) K_+(4; 3) \not{B}(3) K_+(3; 1) d\tau_3 d\tau_4 + \cdots . \quad (52)$$

A similar analysis leads to

$$\psi_-(1') = \int K_+^B(1'; 1) \gamma_4 \psi_+(1) d^3\mathbf{x}_1 , \quad (53)$$

where $K_+^B(1'; 1)$ is given by replacing point 2 in (52) by point 1′, the latter point being on $t = t_1$.

The same procedure for (52) as that used for (16) leads easily to

$$[\not{\nabla}_2 + ie\not{B}(2) + im] K_+^B(2; 1) = \delta_4(2, 1). \quad (54)$$

Equations (51) and (53) give the solution to the problem we set out to solve, and (52) is the required perturbation expansion. The discussion following equation (27) applies equally well here.

The factor C_v in (24) has not yet appeared in the present discussion. This is hardly surprising since we have been dealing only with one particle. After considering the many-particle problem in the next section we shall obtain C_v in Section 4.

3. Many-Particle Interactions and the Generalized Response of the Universe

We begin by restating the conclusions of the preceding section in a slightly different notation and in a somewhat more general form.

We now describe the slab under investigation by $0 \leqslant t \leqslant T$ instead of by $t_1 \leqslant t \leqslant t_2$. The planes $t = 0, t = T$ will be referred to as the surface S and points 1, 2,..., 1′, 2′,... may be anywhere on S. The work of the previous section is described by

$$\psi_{\text{out}}(1') = \int K_+^B(1'; 1) \not{n}_1 \psi_{\text{in}}(1) dS_1 . \quad (55)$$

Here we introduce the concept of ingoing and outgoing parts of the wavefunction at the surface S. In the notation of the previous section ψ_{in} is a ψ_+ wavefunction on $t = 0$ and is a ψ_- wavefunction on $t = T$, while ψ_{out} is a ψ_+ wavefunction on $t = T$ and a ψ_- wavefunction on $t = 0$. n^i is the unit inward normal to S and dS_1 is the surface element at point 1.

Most one-electron problems in quantum electrodynamics are concerned with calculating ψ_{out} when ψ_{in} and the external field B_i are given.

Consider the following generalization of (55) to the many-particle problem

$$\Psi_{\text{out}}(1', 2',...) = \int\int \cdots \int K_+{}^B(1'; 1)\,\not{n}_1 K_+{}^B(2'; 2)\,\not{n}_2 \cdots \Psi_{\text{in}}(1, 2,...)\, dS_1\, dS_2 \cdots . \quad (56)$$

The wavefunction Ψ_{in} now has four components for each particle and has a point 1, 2,... for each particle. Similarly Ψ_{out} has four components and a member of 1′, 2′,... for each particle. The matrix $K_+{}^B(1'; 1)\not{n}_1$ acts on the components of Ψ_{in} for the particle at point 1, $K_+{}^B(2'; 2)\not{n}_2$ acts on the components for the particle at point 2, and so on. The order in which these matrices are written is irrelevant.

This generalization is valid for particles acted on by an external field but is invalid when there are mutual interactions between the particles. Then we write

$$\Psi_{\text{out}}(1', 2',...) = \int\int \cdots \int K[1', 2',...; 1, 2,...]\, \not{n}_1\not{n}_2 \cdots \Psi_{\text{in}}(1, 2,...)\, dS_1\, dS_2 \cdots \quad (57)$$

in which the many-particle propagator $K[1', 2',...; 1, 2,...]$ is determined in the following way.

We recall that

$$K_+{}^B(1'; 1) = \int P(\Gamma_{1'1}) \exp\left[-ie_a \int_{\Gamma_{1'1}} B_i\, da^i\right] \mathscr{D}^3\Gamma_{1'1} . \quad (58)$$

The multiple path integral for a many-particle system in an external field B_i is

$$K[1', 2',...; 1, 2,...] = \int \cdots \int P(\Gamma_{1'1})\, P(\Gamma_{2'2}) \cdots \exp\left[-ie_a \int_{\Gamma_{1'1}} B_i\, da^i - ie_b \int_{\Gamma_{2'2}} B_i\, db^i - \cdots\right] \times \mathscr{D}^3\Gamma_{1'1}\mathscr{D}^3\Gamma_{2'2} \cdots . \quad (59)$$

and this separates into the product $K_+{}^B(1'; 1)\ K_+{}^B(2; 2)$... as stated above. But when the particles are mutually interacting (59) is changed to

$$K[1', 2',...; 1, 2,...] = \int\int \cdots \int P(\Gamma_{1'1})\, P(\Gamma_{2'2}) \cdots \exp(iR) \cdot \exp\left[-i\left\{e_a \int_{\Gamma_{1'1}} B_i\, da^i + \cdots\right\}\right] \mathscr{D}^3\Gamma_{1'1}\mathscr{D}^3\Gamma_{2'2} \cdots, \quad (60)$$

in which the factor $\exp(iR)$ represents interparticle action. In general $\exp(iR)$ takes such a form as to prevent (60) from being separable. The particles are here denoted by a, b,... and da_i, db_i,... are coordinate displacements along the paths $\Gamma_{1'1}$, $\Gamma_{2'2}$ ··· respectively.

We now approach part (ii) of our discussion (cf. Introduction). We aim to prove that all the results of the usual theory of quantum electrodynamics follow from choosing the classical time-symmetric action for R,

$$R = -\tfrac{1}{2} \sum_a \sum_b e_a e_b \int_{\Gamma_{1'1}} \int_{\Gamma_{2'2}} \delta(q_{AB}^2)\, da_i\, db^i. \tag{61}$$

Before proceeding, however, we pause to note that if $\Psi_{\text{in}}(1, 2,...)$ is antisymmetric with respect to every pair of particles then so is $\Psi_{\text{out}}(1', 2',...)$. This can be seen from (57), (60), and (61). The exclusion principle has a rather mysterious role in the usual formulation of quantum electrodynamics. The present path integral approach throws some light on its meaning. It prevents the paths of different particles from sharing common points. The paths $\Gamma_{1'1}$, $\Gamma_{2'2}$ cannot cross, for example—more accurately, paths with common points make zero total contribution to the amplitude. Paths for the same particle, on the other hand, can have common points or common segments, because paths for the same particle are alternatives—the particle follows one of its possible paths. It is indeed this property that permits us to say whether two paths belong to the same particle or to different particles.

Returning to the multiple path integral (60), the electromagnetic factor is the exponential of i times

$$-\sum_a e_a \int B_i\, da^i - \tfrac{1}{2} \sum_a \sum_b e_a e_b \iint \delta(q_{AB}^2)\, da^i\, db_i\,, \tag{62}$$

the line integrals being over the paths $\Gamma_{1'1}$, $\Gamma_{2'2}$,... for the particles a, b,... respectively. The paths are segments lying within the slab $0 \leqslant t \leqslant T$. There are interactions with paths outside the slab, and it is these interactions that give B_i. There are contributions to B_i from both $t < 0$ and $t > T$. The interaction with the future arises because paths interact through the time symmetric function $\delta(q^2)$. The external field from $t < 0$ has already been discussed in detail in *I*, so we shall not repeat the discussion here. That is to say

$$B^i = B^i_{t<0} + B^i_{t>T}\,, \tag{63}$$

and we take

$$B^i_{t<0} = 0. \tag{64}$$

The first term of (62) is then the response of the Universe.

Apart from containing terms with $a = b$, which we shall discuss at the end of the present section, (62) is the same as the classical action in the direct particle theory. It is interesting to recall the classical treatment of the action, since the classical treatment and the quantum treatment will turn out to have similar features. First we write

$$-\tfrac{1}{2}\sum_{a\neq b}\sum e_a e_b \int\int \delta(q_{AB}^2)\, da^i\, db_i = -\tfrac{1}{2}\sum_a e_a \int A^i_{(a)}\, da_i\,,$$

where

$$A^i_{(a)}(X) = \sum_{b\neq a} e_b \int \delta(q_{XB}^2)\, db^i. \tag{65}$$

Separating $\delta(q_{\mathrm{AB}}^2)$ into advanced and retarded parts,

$$\delta(q_{\mathrm{AB}}^2) = \frac{1}{2\mid \mathbf{x}_A - \mathbf{x}_B \mid}[\delta(t_A - t_B - \mid \mathbf{x}_A - \mathbf{x}_B \mid) + \delta(t_A - t_B + \mid \mathbf{x}_A - \mathbf{x}_B \mid)], \tag{66}$$

and writing

$$A^i_{(a)}(X) = \tfrac{1}{2}[A^i_{(a)}(X)_{\mathrm{ret}} + A^i_{(a)}(X)_{\mathrm{adv}}], \tag{67}$$

$$A^i_{(a)}(X)_{\mathrm{ret}} = \sum_{b\neq a} A^{i(b)}(X)_{\mathrm{ret}} = \sum_{b\neq a} e_b \int \frac{\delta(t_X - t_B - \mid \mathbf{x} - \mathbf{x}_B \mid)}{\mid \mathbf{x} - \mathbf{x}_B \mid} db^i, \tag{68}$$

$$A^i_{(a)}(X)_{\mathrm{adv}} = \sum_{b\neq a} A^{i(b)}(X)_{\mathrm{adv}} = \sum_{b\neq a} e_b \int \frac{\delta(t_X - t_B + \mid \mathbf{x} - \mathbf{x}_B \mid)}{\mid \mathbf{x} - \mathbf{x}_B \mid} db^i, \tag{69}$$

enables the classical action to be written in the form

$$-\sum_a e_a \int B^i_{t>T}\, da_i - \tfrac{1}{2}\sum_a e_a \int [A^i_{(a)\mathrm{ret}} + A^i_{(a)\mathrm{adv}}]\, da_i\,. \tag{70}$$

The classical response condition is

$$B_i(X)_{t>T} = \tfrac{1}{2}\sum_a [A_i^{(a)}(X)_{\mathrm{ret}} - A_i^{(a)}(X)_{\mathrm{adv}}]. \tag{71}$$

Eliminating $B^i{}_{t>T}$ between (70) and (71) we get

$$-\sum_a e_a \int \{A^i_{(a)\mathrm{ret}} + \tfrac{1}{2}[A^{i(a)}_{\mathrm{ret}} - A^{i(a)}_{\mathrm{adv}}]\}\, da_i\,. \tag{72}$$

The first term in the curly brackets is the usual retarded classical field of all particles other than a, and the second term is the usual radiation reaction on a.

Our aim is to obtain a generalized response condition for the quantum case, analogous to (71), that together with (62) leads to an electromagnetic phase factor in (60) giving all the results of the usual theory of quantum electrodynamics. The required form has already been noted in (27) and (29).

The quantum case differs from classical theory in that conjugate paths, $\mathbf{a}'(t)$ for particle a, $\mathbf{b}'(t)$ for particle b,... must be introduced. These conjugate paths go from points 1*, 2*,... on S to the same points 1', 2',..., as do the paths $\mathbf{a}(t)$, $\mathbf{b}(t)$ $\cdots$. Instead of (62) we now have (62) minus the contribution of the conjugate paths. It is easy to see that this combination of paths and conjugate paths gives

$$-\sum_a e_a \left[\int \{\tfrac{1}{2}[A^i_{\text{ret}} + A^i_{\text{adv}}] + B^i_{t>T}\}\, da_i - \int \{\tfrac{1}{2}[A'^i_{\text{ret}} + A'^i_{\text{adv}}] + B'^i_{t>T}\}\, da_i'\right] \tag{73}$$

in which

$$A^i(X)_{\text{ret}} = \sum_b e_b \int \frac{\delta(t_X - t_B - |\mathbf{x} - \mathbf{x}_B|)}{|\mathbf{x} - \mathbf{x}_B|}\, db^i, \tag{74}$$

$$A^i(X)_{\text{adv}} = \sum_b e_b \int \frac{\delta(t_X - t_B + |\mathbf{x} - \mathbf{x}_B|)}{|\mathbf{x} - \mathbf{x}_B|}\, db^i, \tag{75}$$

and where the primed fields refer to the conjugate paths and are given by expressions similar to (74), (75), and $B'^i_{t>T}$ is the response to the conjugate paths. Multiplying (73) by i and taking the exponential gives the electromagnetic phase factor.

Our point of view concerning the conjugate paths was discussed at length in I. The first integral in (73) cannot be expressed in a form suited to practical calculation except is association with the second integral. That is to say, we do not know the response of the universe except in a probability calculation. The usual theory of quantum electrodynamics, on the other hand, allows a complete calculation of the amplitude alone, irrespective of whether a probability is taken. Since amplitudes do not appear in physical statements the usual theory seems unsatisfactory in this respect. It is true that physical statements can be made by discarding the phase of the amplitude, but then why go to the trouble to calculate the phase if it is not required?

The slab $0 \leqslant t \leqslant T$ now has an explicit meaning. It is the slab in which an experiment is performed, i.e., for which we require a physical statement.

Paths and conjugate paths together permit a separation of positive and negative frequencies. In classical theory there is one path for each particle and this one path contributes both positive and negative frequencies. In quantum electrodynamics paths give positive frequencies and conjugate paths give negative frequencies. Negative and positive frequencies are introduced in the following way. Define

$$A^i(X)_{\text{ret}\pm} = \sum_b e_b \int \frac{\delta_\pm(t_X - t_B - |\mathbf{x} - \mathbf{x}_B|)}{2\,|\mathbf{x} - \mathbf{x}_B|}\, db^i, \tag{76}$$

$$A^i(X)_{\text{adv}\pm} = \sum_b e_b \int \frac{\delta_\pm(t_X - t_B + |\mathbf{x} - \mathbf{x}_B|)}{2\,|\mathbf{x} - \mathbf{x}_B|}\, db^i, \tag{77}$$

where

$$\delta_{\pm}(x) = \frac{1}{\pi}\int_0^\infty e^{\mp i\omega x}\,d\omega = \delta(x) \mp \frac{i}{\pi}\frac{\mathscr{P}}{x}. \tag{78}$$

It may be noted that whereas the sum of $A^i_{\text{ret}+}$, $A^i_{\text{ret}-}$ is a vector, namely, the classical retarded field

$$A^i(X)_{\text{ret}} = A^i_{\text{ret}+} + A^i_{\text{ret}-}, \tag{79}$$

$A^i_{\text{ret}+}$, $A^i_{\text{ret}-}$ are not vectors by themselves. Similarly

$$A^i(X)_{\text{adv}} = A^i_{\text{adv}+} + A^i_{\text{adv}-} \tag{80}$$

is the classical advanced field but $A^i_{\text{adv}+}$, $A^i_{\text{adv}-}$ are not vectors by themselves.

This separation into positive and negative frequencies permits two new vectors to be defined, $A^i_{\text{ret}+} - A^i_{\text{adv}+}$ and $A^i_{\text{ret}-} - A^i_{\text{adv}-}$. These vectors appear in quantum electrodynamics but not in classical theory. It is straightforward to show that

$$A^i(X)_{\text{ret}+} - A^i(X)_{\text{adv}+} = \sum_b e_b\left[\int_{t_X>t_B}\delta_+(q^2_{XB})\,db^i - \int_{t_B>t_X}\delta_-(q^2_{BX})\,db^i\right], \tag{81}$$

$$A^i(X)_{\text{ret}-} - A^i(X)_{\text{adv}-} = \sum_b e_b\left[\int_{t_X>t_B}\delta_-(q^2_{XB})\,db^i - \int_{t_B>t_X}\delta_+(q^2_{BX})\,db^i\right]. \tag{82}$$

Although $\delta_{\pm}(q^2)$ are invariant functions it is not immediately clear that (81) and (82) are vectors, because of the time restrictions. However, the contributions of the principal parts in $\delta_{\pm}(q^2)$ have opposite signs. It is easy to see therefore that the time restrictions do not apply to the principal parts, which accordingly give vectors. The time restrictions continue to apply to the $\delta(q^2)$ part of $\delta_{\pm}(q^2)$, but since $\delta(q^2)$ is confined to the light cone the restrictions are invariant, and the $\delta(q^2)$ part also gives vectors.

The quantum response corresponding to (71) is

$$B_i(X)_{t>T} = \tfrac{1}{2}\sum_b \{[A_i^{(b)}(X)_{\text{ret}+} - A_i^{(b)}(X)_{\text{adv}+}] + [A_i'^{(b)}(X)_{\text{ret}-} - A_i'^{(b)}(X)_{\text{adv}-}]\}. \tag{83}$$

As stated above, the paths carry the positive frequencies and the conjugate paths carry the negative frequencies. If we imagine a coalescence of paths and conjugate paths (83) becomes the same as (71). To obtain an equivalent condition for $B_i'(X)_{t>T}$ we interchange paths with conjugate paths and positive frequencies with negative frequencies. This double interchange leaves (83) unaltered, so that

$$B_i'(X)_{t>T} = B_i(X)_{t>T}. \tag{84}$$

To complete the work of the present section we calculate the contribution to (73) from particles a, b. For $a \neq b$ we have four paths **a**, **b**, **a**′, **b**′ which combine in the pairs (**a**, **b**), (**a**′, **b**′), (**a**′, **b**), (**a**, **b**′). The contribution of (**a**, **b**) has two terms,

$$-e_a e_b \iint \delta(q_{AB}^2)\, da_i\, db^i, \tag{85}$$

from the time symmetric interparticle action and

$$-\tfrac{1}{2} e_a e_b \left\{ \iint_{t_A > t_B} [\delta_+(q_{AB}^2) - \delta_-(q_{AB}^2)]\, da_i\, db^i \right.$$

$$\left. + \iint_{t_B > t_A} [\delta_+(q_{BA}^2) - \delta_-(q_{BA}^2)]\, da_i\, db^i \right\} \tag{86}$$

from B_i. Since

$$2\delta(q_{AB}^2) = \delta_+(q_{AB}^2) + \delta_-(q_{AB}^2), \qquad \delta_\pm(q_{AB}^2) = \delta_\pm(q_{BA}^2), \tag{87}$$

the sum of (85) and (86) is

$$-e_a e_b \iint \delta_+(q_{AB}^2)\, da_i\, db^i, \tag{88}$$

which is the same as (27). The pair (**a**′, **b**′) similarly give

$$+e_a e_b \iint \delta_-(q_{A'B'}^2)\, da_i'\, db'^i. \tag{89}$$

Only the B_i, B_i' terms in (73) contribute to (**a**′, **b**). It is not hard to show that these terms combine to give

$$e_a e_b \left\{ \iint_{t_{A'} > t_B} \delta_+(q_{A'B}^2)\, da_i'\, db^i - \iint_{t_B > t_{A'}} \delta_-(q_{BA'}^2)\, da_i'\, db^i \right\}. \tag{90}$$

The pair (**a**, **b**′) gives

$$e_a e_b \left\{ \iint_{t_{B'} > t_A} \delta_+(q_{B'A}^2)\, da^i\, db_i' - \iint_{t_A > t_{B'}} \delta_-(q_{AB'}^2)\, da^i\, db_i' \right\}. \tag{91}$$

The total contribution of the particles a, b to the electromagnetic phase factor is given by summing (88), (89), (90), (91), multiplying by i and taking the exponential,

$$\exp\Big\{ie_a e_b\Big[\iint_{t_{A'}>t_B}\delta_+(q^2_{A'B})\,da_i'\,db^i - \iint_{t_B>t_{A'}}\delta_-(q^2_{BA'})\,da_i'\,db^i$$
$$+\iint_{t_{B'}>t_A}\delta_+(q^2_{B'A})\,da^i\,db_i' - \iint_{t_A>t_{B'}}\delta_-(q^2_{AB'})\,da^i\,db_i'$$
$$-\iint\delta_+(q^2_{AB})\,da^i\,db_i + \iint\delta_-(q^2_{A'B'})\,da_i'\,db'^i\Big]\Big\}. \tag{92}$$

This is the influence functional $F[\mathbf{a}, \mathbf{b}; \mathbf{a}', \mathbf{b}']$ contributed by the particles a, b. It can be expressed in terms of three-dimensional Fourier integrals using

$$\delta_\pm(q^2_{21}) = \pm 4\pi i \int \frac{1}{2K}\exp\{\mp iK\,|\,t_2 - t_1\,|\pm i\mathbf{k}\cdot(\mathbf{x}_2 - \mathbf{x}_1)\}\,\frac{d^3\mathbf{k}}{(2\pi)^3}, \tag{93}$$

where $K = |\,\mathbf{k}\,|$ and q^2_{21} is the invariant distance between points 1 and 2. We get

$$F[\mathbf{a}, \mathbf{b}; \mathbf{a}', \mathbf{b}']$$
$$= \exp\Big[-\frac{e_a e_b}{4\pi^2}\int d\Omega\int K\,dK\,\Big\{\iint \exp[-iK\,|\,t_A - t_B\,| + i\mathbf{k}\cdot(\mathbf{x}_B - \mathbf{x}_A)]\,da_i db^i$$
$$+\iint \exp[iK\,|\,t_{A'} - t_{B'}\,| + i\mathbf{k}\cdot(\mathbf{x}_{A'} - \mathbf{x}_{B'})]\,da_i'\,db'^i$$
$$-\iint \exp[iK(t_A - t_{B'}) + i\mathbf{k}\cdot(\mathbf{x}_{B'} - \mathbf{x}_A)]\,da_i\,db'^i$$
$$-\iint \exp[iK(t_B - t_{A'}) + i\mathbf{k}\cdot(\mathbf{x}_{A'} - \mathbf{x}_B)]\,da_i'\,db^i\Big\}\Big]. \tag{94}$$

The different terms in (94) can be interpreted as follows. For $t_A > t_B$ the positive terms in the curly bracket contribute to a downward transition of particle b and an upward transition of particle a, and *vice versa* for $t_B > t_A$, whereas the negative terms contribute to downward transitions of both a and b. Thus (94) refers to absorption and stimulated emission. Since the coefficient in front of all terms in the exponential is the same, the probabilities for these two processes are the same. The influence functional (94) leads to all the usual results for the interaction of two charged particles. There is a similar factor in the multiple path integral (60) for every pair of particles.

Consider next the case $a = b$. There are contributions from three path combinations $(\mathbf{a}, \mathbf{a}')$, $(\mathbf{a}, \mathbf{a})$, and $(\mathbf{a}', \mathbf{a}')$. These may be obtained in the same way as in the above discussion, or by identifying $\mathbf{a}$ and $\mathbf{b}$, $\mathbf{a}'$ and $\mathbf{b}'$ in the above formulas.

However, if we follow the latter course a factor $\frac{1}{2}$ must be introduced, because of the $\frac{1}{2}$ in (61) and because **a** and **a**′ interact only once. The result is

$$F[\mathbf{a}, \mathbf{a}'] = \exp \Big\{ ie_a{}^2 \Big[\iint_{t_{A'} > t_A} \delta_+(q^2_{A'A})\, da_i{}'\, da^i - \iint_{t_A > t_{A'}} \delta_-(q^2_{AA'})\, da_i{}'\, da^i$$

$$- \tfrac{1}{2} \iint \delta_+(q^2_{A\tilde{A}})\, da^i\, d\tilde{a}_i + \tfrac{1}{2} \iint \delta_-(q^2_{A'\tilde{A}'})\, da'^i\, d\tilde{a}_i{}' \Big] \Big\}. \tag{95}$$

In terms of Fourier integrals the corresponding result is

$$F[\mathbf{a}, \mathbf{a}'] = \exp \Big[\frac{e_a{}^2}{4\pi^2} \int d\Omega \int_0^\infty K\, dK \Big\{ - \iint \exp[iK(t_A - t_{A'}) + i\mathbf{k} \cdot (\mathbf{x}_{A'} - \mathbf{x}_A)]\, da^i da_i{}'$$

$$+ \iint_{t_A > t_{\tilde{A}}} \exp[iK(t_{\tilde{A}} - t_A) + i\mathbf{k} \cdot (\mathbf{x}_{\tilde{A}} - \mathbf{x}_A)]\, da^i\, d\tilde{a}_i$$

$$+ \iint_{t_{A'} > t_{\tilde{A}'}} \exp[iK(t_{A'} - t_{\tilde{A}'}) + i\mathbf{k} \cdot (\mathbf{x}_{\tilde{A}'} - \mathbf{x}_{A'})]\, da'^i\, d\tilde{a}_i{}' \Big\} \Big]. \tag{96}$$

This is the same result as was obtained in I. The first term in the curly bracket gives spontaneous transitions whereas the second and third terms contribute "radiative correction effects." There is a factor of the form (96) in the multiple path integral (60) for every particle.

For many purposes it is preferable to express the $\delta_\pm(q^2)$ functions in (92) and (95) as four-dimensional Fourier integrals, using

$$\delta_\pm(q^2) = -4\pi \int_{\epsilon \to 0+} \frac{1}{k^2 \pm i\epsilon}\, e^{-ik\cdot(x_2 - x_1)} \frac{d^4k}{(2\pi)^4}\,. \tag{97}$$

We have given the three-dimensional Fourier form of the influence functionals in order to show that our present results, using the generalized response condition (83), are the same as the results of the explicit calculation in I.

We come now to the problem of the self-action contributions to (62). The contribution of particle a is

$$-\tfrac{1}{2} e_a{}^2 \iint \delta(q^2_{A\tilde{A}})\, da^i\, d\tilde{a}_i\,, \tag{98}$$

where A, $\tilde{A}$ both lie on the path $\mathbf{a}(t)$. In classical theory (98) is omitted entirely; particles do not act on themselves. It is actually sufficient in classical theory to make a weaker hypothesis—that sufficiently small steps of a path do not act on themselves. For classical paths are wholly timelike and $q^2_{A\tilde{A}} > 0$ giving $\delta(q^2_{A\tilde{A}}) = 0$ when A and $\tilde{A}$ are separated.

In quantum theory, on the other hand, there is a difference between omitting (98) and making the hypothesis that individual path steps do not act on themselves, since the latter hypothesis still permits different Γ^+ and Γ^- sections of a path to interact, whereas a complete omission of (98) would remove all self-interaction. In the above work we included (98) in order to obtain the usual electron–positron interaction. If there were no self-action, the response $B^i_{t>T}$ would contribute the exponential of

$$-\tfrac{1}{2}ie_a \int [A^{(a)}_{i\text{ret}+} - A^{(a)}_{i\text{adv}+}]\, da^i$$
$$= -ie_a \int A^{(a)}_{i\text{ret}+}\, da^i + \tfrac{1}{2}ie_a \int [A^{(a)}_{i\text{ret}+} + A^{(a)}_{i\text{adv}+}]\, da^i \qquad (99)$$

to the (**a**, **a**) term in (96). The second term in (99) is minus the self-action contribution (98). It is cancelled therefore when (98) is included, so that the first term of (99) yields the (**a**, **a**) part of (96). If (98) is taken to be zero for timelike paths the second term of (99) is also zero, and we still obtain (96), even though self-action is absent. This explains why in I we were able to obtain the correct influence functional, since in I we were concerned only with electrons, i.e., with Γ^+ paths treated non-relativistically, the same as classical paths.

It seems possible that in quantum theory (98) may vanish for Γ^+ or Γ^- paths individually, even without requiring the hypothesis that individual path steps do not act on themselves. Suppose A and $\tilde{A}$ are on the same step, $da^i \equiv d\tilde{a}^i$. Since quantum paths are made up of null segments, $q^2_{A\tilde{A}} = 0$, but this does not cause difficulty so long as A and $\tilde{A}$ are distinct, since $da_i\, d\tilde{a}^i$ is proportional to $q^2_{A\tilde{A}}$ and $q^2_{A\tilde{A}}\delta(q^2_{A\tilde{A}}) = 0$. In general for an individual Γ^+ or Γ^- path, either we have $q^2_{A\tilde{A}} > 0$, $\delta(q^2_{A\tilde{A}}) = 0$, or the path between A and $\tilde{A}$ is a straight null line. So long as A and $\tilde{A}$ are taken to be interior prints within da^i, $d\tilde{a}^i$ we have $da_i\, d\tilde{a}^i$ proportional to $q^2_{A\tilde{A}}$ and $\delta(q^2_{A\tilde{A}})\, da_i\, d\tilde{a}^i$ is again zero.

The self-action integral becomes ambiguous, however, when A and $\tilde{A}$ are permitted to coincide. This ambiguity is removed if we omit self-action on a sufficiently small scale. We shall find in Section 5 that this question is intimately connected with divergences in radiation correction processes.

4. Vacuum Loops

In Section 1 we described Feynman's method of introducing the K_+ propagator. This method leads to the result that the amplitude for any process in quantum electrodynamics can be calculated from a perturbation expansion, provided a factor $C_v = \exp(-L)$, with L defined by (25), (26), is included to represent vacuum loops. In Section 2 we obtained the perturbation expansion for a single particle in an external field. This was done by the path integral method. In Section 3 this method was extended to a many-particle system, and the mutual interactions of the particles were correctly represented by combining the classical electromagnetic action with the response of the Universe. This is the position we have reached. To complete our program it remains to return to the problem of vacuum loops and and to recover the factor C_v.

We show first that C_v can be obtained by applying an antisymmetrization process to the perturbation expansion (52) for a particle in an external field. Since (52) was obtained from a path integral, it is necessary to show next how the extra terms required by antisymmetrization can also be represented by a path integral. This one-particle treatment must then be generalized to many particles and must include interparticle action. We proceed in these three stages.

The idea underlying the antisymmetrization of (52) is illustrated in general terms in Fig. 1. In (a) we have two particles going forward in time, one particle going from 1 to 3 by a Γ^+ path, the other from 4 to 2 also by a Γ^+ path. The amplitude for (a) does not represent the total amplitude for there to be a particle at 2 and a particle at 3. The exclusion principle requires subtraction of (b), with 1 going to 2 and 4 to 3.

In (c) we have a single particle path with two time reversals, and with an external field B_i acting at the points 3 and 4 of reversal. The configuration of points is the same in (c) as in (a), and the electron paths from 1 to 3, 4 to 2 are the same. If now we follow the suggestion of the preceding section—that time reversals generate "new" particles—we require the Γ^+ sections from 1 to 3, 4 to 2 in (c) to be on the same footing as the same two paths in (a). We should then subtract the amplitude for the possibility that the Γ^+ paths starting from 1 and 4 go to 2 and 3, as they do in (b) instead of to 3 and 2. The effect of this redirection of the paths is shown in (d). We now have a single Γ^+ path 1 to 2 and a disconnected closed loop with a Γ^+ path from 4 to 3 and a Γ^- path from 3 to 4. The external field acts at the two points of reversal of the loop, not on the Γ^+ path from 1 to 2. We shall next give quantitative expression to this idea.

The amplitude for Fig. 1(c) is given by restricting the second order term in the perturbation expansion to $t_3 > t_4$:

$$(-ie)^2 \iint_{t_3 > t_4} K_+(2; 4)\, \not{B}(4)\, K_+(4; 3)\, \not{B}(3)\, K_+(3; 1)\, d\tau_3\, d\tau_4 . \tag{100}$$

Since 1 to 3 is a Γ^+ path, (100) operates on a ψ_+ wavefunction at point 1. Writing the wavefunction as $v_+(1)$ and integrating spatially with respect to 1, the amplitude for the double scattering of Fig. 1(c) is

$$(-ie)^2 \iint_{t_3 > t_4} K_+(2; 4)\, \not{B}(4)\, K_+(4; 3)\, \not{B}(3)\, v_+(3)\, d\tau_3\, d\tau_4 . \tag{101}$$

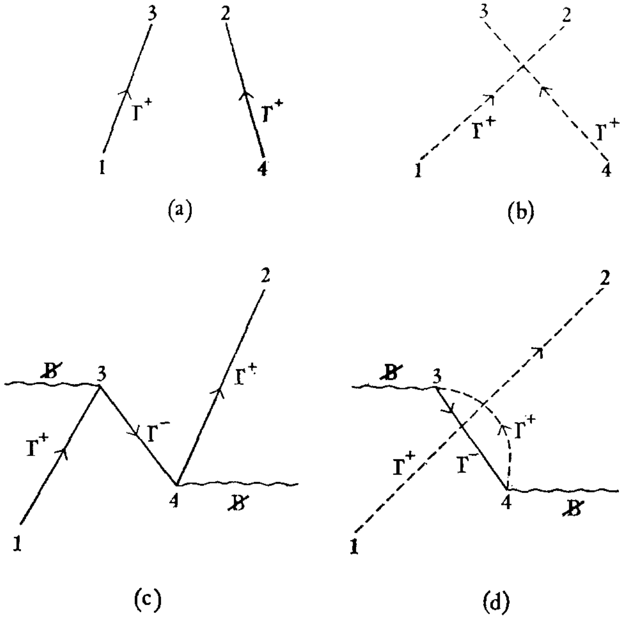

FIG. 1. In two-particle scattering the exclusion principle requires the amplitude for (b) to be subtracted from that for (a). The same principle when applied to multiple scattering of a "single" particle by electromagnetic potentials leads to closed loops. This is shown in diagrams (c) and (d).

The spatial integration with respect to 1 is over the $t = 0$ face of the slab $0 = t_1 \leqslant t \leqslant t_2 = T$, the integrations with respect to $d\tau_3$, $d\tau_4$ are taken through the slab, and point 2 is on $t = T$. Since $t_2 > t_4$ we have

$$K_+(2;4) = \sum_{E_n>0} u_n(2)\,\bar{u}_n(4), \tag{102}$$

and we can write (101) in the form

$$(-ie)^2 \sum_{E_n>0} u_n(2) \iint_{t_3>t_4} \bar{u}_n(4)\,\not{B}(4)\,K_+(4;3)\,\not{B}(3)\,v_+(3)\,d\tau_3\,d\tau_4\,. \tag{103}$$

Next we have to subtract the amplitude for Fig. 1(d). This is

$$-(-ie)^2 \sum_{E_n>0} v_+(2) \iint_{t_3>t_4} \bar{u}_n(4)\,\not{B}(4)\,K_+(4;3)\,\not{B}(3)\,u_n(3)\,d\tau_3\,d\tau_4\,. \tag{104}$$

The sum of (103) and (104) is antisymmetric with respect to interchange of u_n and v_+, as was required by the qualitative discussion based on Fig. 1.

Now we use the matrix relation

$$\sum_{\text{Spins}} \bar{u}_n(4)\, \not{B}(4)\, K_+(4;3)\, \not{B}(3)\, u_n(3) = \operatorname{Tr}\left[\sum_{\text{Spins}} u_n(3)\, \bar{u}_n(4)\, \not{B}(4)\, K_+(4;3)\, \not{B}(3)\right] \tag{105}$$

together with

$$K_+(3;4) = \sum_{E_n>0} u_n(3)\, \bar{u}_n(4), \qquad t_3 > t_4 \tag{106}$$

and obtain for (104),

$$-(-ie)^2\, v_+(2) \iint_{t_3>t_4} \operatorname{Tr}[K_+(3;4)\, \not{B}(4)\, K_+(4;3)\, \not{B}(3)]\, d\tau_3\, d\tau_4\,. \tag{107}$$

The restriction $t_3 > t_4$ can be removed provided $\frac{1}{2}$ is introduced, and the resulting coefficient of $v_+(2)$ is

$$-\tfrac{1}{2}(-ie)^2 \iint \operatorname{Tr}[K_+(3;4)\, \not{B}(4)\, K_+(4;3)\, \not{B}(3)]\, d\tau_3\, d\tau_4\,. \tag{108}$$

Apart from the unity term, this is the lowest order term in C_v. Asymmetrization of higher order terms in the perturbation expansion proceeds in the same way. In third order there are new terms in which the potential scatters v_+ and also acts twice in a closed loop. In fourth order the new terms scatter v_+ twice and act twice in a closed loop, act four times in a closed loop and act twice in each of two loops. The net effect of all terms of the perturbation expansion is to give $C_v = \exp(-L)$ with L defined by (25) and (26).

Next we see how (108) and other terms arising from the perturbation expansion can be represented by path integrals. The loop Γ^0 is made up of $\Gamma^{\pm}$ sections, as was the path Γ_{21} in (38). The amplitude for a loop can be obtained from (38) by identifying points 1 and 2, and taking the trace of the right hand side,

$$P^B(\Gamma^0) = \operatorname{Tr}\left[\prod_i P^B(\Gamma^{\pm}_{i,i-1})\right]. \tag{109}$$

By summing loops in the slab $0 \leqslant t \leqslant T$ we obtain the path integral

$$\int P^B(\Gamma^0)\, \mathscr{D}^3\Gamma^0. \tag{110}$$

The field B_i is a simple exponential factor in each $P^B(\Gamma^{\pm}_{i,i-1})$ factor on the right hand side of (109). Taken together these factors give the integral of B_i round Γ^0.

We write

$$P^B(\Gamma^0) = \exp\left[-ie \oint_{\Gamma^0} B_i\, dl^i\right] P(\Gamma^0). \tag{111}$$

where

$$P(\Gamma^0) = \mathrm{Tr}\left[\prod_i P(\Gamma^{\pm}_{i,i-1})\right]. \tag{112}$$

It is remarkable that $-L$ in $\exp(-L) = C_v$ is just minus the simple path integral (110),

$$-L = -\int P(\Gamma^0) \exp\left[-ie \oint_{\Gamma^0} B_i dl^i\right] \mathscr{D}^3\Gamma^0. \tag{113}$$

dl^i is a coordinate displacement along Γ^0.

The terms $L^{(n)}$ in (25), (26) can be obtained from (113) by expanding the electromagnetic phase factor. The method is the same as that used to obtain the perturbation expansion (52) from (39). Equation (113) is the path integral, with the correct sign, for one loop.

The path integral for n closed loops, $l, m,\ldots$ is

$$\frac{1}{n!}(-1)^n \int \cdots \int P(\Gamma_l^0)\, P(\Gamma_m^0)$$

$$\cdots \exp\left[-ie \oint_{\Gamma_l^0} B_i\, dl^i - ie \oint_{\Gamma_m^0} B_i\, dm^i - \cdots\right] \mathscr{D}^3\, \Gamma_l^0 \mathscr{D}^3\, \Gamma_m^0 \cdots . \tag{114}$$

This may be compared with (59) for the particles $a, b,\ldots$ with the open paths $\Gamma_{1'1}, \Gamma_{2'2},\ldots$. The factor $1/n!$ arises because the set of loops $l, m,\ldots$ is counted $n!$ times in the summations.

Before continuing with (114) we pause to note that the amplitudes $P(\Gamma_l^0)$ are determined in terms of the $\Gamma^{\pm}_{ii-1}$ paths which were defined in Section 2 in terms of the $K_0^{\pm}$ propagators, whereas the $L^{(n)}$ terms in (25) involve the K_+ propagators. This use of the K_+ propagators is due to the trace property of $P(\Gamma^0)$. Thus the trace can be represented by a summation with respect to wavefunctions,

$$P(\Gamma^0) = \sum \bar{u}[\Pi P(\Gamma^{\pm}_{i,i-1})]\, u, \tag{115}$$

where the summation is over normalized states for any momentum $\mathbf{p}$. We require both $E = \pm\sqrt{\mathbf{p}^2 + m^2}$. The space-time dependence of u cancels that of $\bar{u}$ because the wavefunctions are taken at the "beginning" and "ending" of the loop, i.e., at the same point. Remembering that the Γ^+ sections of Γ^0 must be weighted by a ψ_+

wavefunction, and the Γ^- sections by a ψ_- wavefunction, the K_+ propagator is needed to keep correct accounting, exactly as in the derivation of (52).

We have to generalize our path integral so that it includes both particles a, b,... and loops l, m,... and also so that mutual interactions are present as well as an external field. The path integral is to give the many particle propagator $K[1', 1; 2', 2;...]$. The expression (60) for K applies when there are no loops. When there are n loops l, m,... there is a contribution to K given by

$$\frac{1}{n!}(-1)^n \int \cdots \int P(\Gamma_{1'1})\, P(\Gamma_{2'2}) \cdots P(\Gamma_l^0)\, P(\Gamma_m^0) \cdots \exp\{i(R+S)\}$$
$$\times \exp\Big[-i\Big\{e_a \int_{\Gamma_{1'1}} B^i\, da_i + e_b \int_{\Gamma_{2'2}} B^i\, db_i + \cdots + e \oint_{\Gamma_l^0} B^i\, dl_i + e \oint_{\Gamma_m^0} B^i\, dm_i + \cdots\Big\}\Big]$$
$$\cdot\, \mathscr{D}^3\Gamma_{1'1}\mathscr{D}^3\Gamma_{2'2} \cdots \mathscr{D}^3\Gamma_l^0\mathscr{D}^3\Gamma_m^0 \cdots . \qquad (116)$$

The factor $\exp(iR)$ is the same as in Section 3. The factor $\exp(iS)$ is new and gives the interactions between particles and loops and between loops and loops, through

$$S = -\sum_a \sum_l e_a e \iint \delta(q_{AL}^2)\, da_i\, dl^i - \tfrac{1}{2}e^2 \sum_l \sum_m \iint \delta(q_{LM}^2)\, dl^i\, dm_i . \qquad (117)$$

The many-particle propagator is given by summing (116) for $n = 0, 1, 2,...$. The factor half appears in the second term of (117) because the pair of loops (l, m) is counted twice. The particle a and the loop l are only counted once in the first term. In the above formulas e is the electron charge, which implies that we are considering the electron vacuum, although even powers of e always appear in practical calculations, so the sign of e is irrelevant.

The discussion now proceeds as it did in Section 3. For every pair of particles we had paths **a**, **b** and conjugate paths **a**′, **b**′ leading to the contributions F[**a**, **b**; **a**′, **b**′], F[**a**, **a**′], F[**b**, **b**′] to the influence functional. For every pair of loops (l, m) we now have closed paths ***l***, **m** and closed conjugate paths ***l***′, **m**′ giving contributions F[***l***, **m**; ***l***′, **m**′], F[***l***, ***l***′], F[**m**, **m**′]. These lead to vacuum diagrams—they have the effect of multiplying amplitudes by a phase factor exp iC, C real. Such factors disappear in the present theory because these contributions to the influence functional appear only after the response of the Universe has been used, and we are then concerned with the calculation of a probability, not an amplitude. Loops by themselves are of no relevance. But the mixed contributions F[**a**, ***l***; **a**′, ***l***′] are of physical importance. They produce a scattering of particle **a** and contribute $-27Mc/\text{sec}$ to the Lamb shift. Using the method of Section 3 we obtain

$$F[\mathbf{a}, l; \mathbf{a}', l'] = \exp\Big[iee_a \Big\{ \iint_{t_{A'} > t_L} \delta_+(q^2_{A'L})\, da'^i\, dl_i - \iint_{t_L > t_{A'}} \delta_-(q^2_{LA'})\, da'^i\, dl_i$$

$$+ \iint_{t_{L'} > t_A} \delta_+(q^2_{L'A})\, da^i\, dl_i' - \iint_{t_A > t_{L'}} \delta_-(q^2_{AL'})\, da^i\, dl_i'$$

$$- \iint \delta_+(q^2_{AL})\, da^i\, dl_i + \iint \delta_-(q^2_{A'L'})\, da'^i\, dl_i' \Big\}\Big]. \tag{118}$$

This completes the program set out at the beginning of this section. It actually completes the program of this series of papers, since we have now obtained the required influence functionals between particles and particles, between particles and loops, and the influence functional representing the action of a particle on itself—the radiation reaction. All the usual results of quantum electrodynamics follow by applying the perturbation expansion to these influence functionals.

We are now in an equivocal position, however, for just as we can obtain all the usual correct practical results of quantum electrodynamics so we can obtain all the unwanted divergences in radiation correction processes. The critical question we now have to ask is whether these divergences are inevitable, or does the present direct particle theory provide an escape from them, as it does in classical theory? We shall consider this question in the next two sections, and will end the paper by a general discussion of the present position of the direct particle theory.

5. The Electron Self-Energy and the Radiative Correction to Scattering

We come now to the analogue of the self-energy divergence in field theory. For a particle we have

$$-\tfrac{1}{2}e_a{}^2 \iint \delta(q^2_{A\tilde{A}})\, da^i\, d\tilde{a}_i\,, \tag{119}$$

which together with the response of the Universe leads to the term

$$-\tfrac{1}{2}e_a{}^2 \iint \delta_+(q^2_{A\tilde{A}})\, da^i\, d\tilde{a}_i \tag{120}$$

in the influence functional associated with particle a. We already discussed the interpretation of (119) at the end of Section 3 and saw that to avoid ambiguity in the meaning of the line integrals points A and $\tilde{A}$ should be distinct. In Section II we defined the path steps by a set of points X_i, $i = 0, 1, \ldots, N$. We left open the question of whether N should tend to infinity. If we decide affirmatively and arrange X_i so that all steps tend to zero, $A \to \tilde{A}$ cannot be avoided and the outcome is completely equivalent to the usual theory. We obtain all the usual divergences and require all the usual apparatus of renormalization theory. Here, however, we propose to examine the effect of preventing $A \to \tilde{A}$. To preserve a four-dimensional notation we require

$$| X_A{}^i - X_{\tilde{A}}{}^i | \geqslant | \epsilon^i |, \qquad i = 1, 2, 3, 4, \tag{121}$$

where the $| \epsilon^i |$ are to be regarded as small compared to the Compton wavelength. To make (121) physically meaningful ϵ^i must be a vector associated with the particle. For a free particle we can try

$$\epsilon^i = \epsilon \cdot \frac{p^i}{m}, \tag{122}$$

where ϵ is to be a length very small compared to m^{-1}. Thus we can choose

$$\epsilon = 2Gm, \tag{123}$$

the gravitational radius of the particle, giving $m\epsilon \approx 10^{-40}$. Although we have not permitted the path steps to tend strictly to zero, the scale of the steps is nevertheless exceedingly small compared to any physical structures that have hitherto been considered.

We emphasise that we are not seeking here to justify (121), (122), (123), but to investigate their effect and to compare the resulting situation with the cutoff procedures of the usual theory. If indeed gravitation plays a critical role then modification of the space-time structure on the scale ϵ is required. Only through such an investigation could one expect to understand ϵ^i more deeply.

We show in this section that (121), (122), (123) lead to a finite self-energy even though the frequency in the three-dimensional Fourier integral of the $\delta_+(q^2)$ function tends to infinity. And in Section 6 we shall show that the divergences in vacuum polarization, which are usually thought to be of a different nature to the divergence of the self-energy are removed by the same procedure.

For simplicity of presentation we calculate first the self-energy for an electron with $\mathbf{p} = \mathbf{0}$. Suppose at $t = 0$ we are given the wavefunction $\mathscr{U}_0$ corresponding to positive energy m. We wish to examine the effect of the term (120) in the influence functional for the slab $0 \leqslant t \leqslant T$. The wavefunction at point 2 on $t = T$ is usually calculated from the perturbation expansion

$$\mathscr{U}_0 e^{-imT} - ie^2 \iint K_+(2, 3)\, \gamma_i K_+(3, 4)\, \gamma^i\, \delta_+(q_{34}^2)\, \mathscr{U}_0 e^{-imt_4}\, d\tau_3\, d\tau_4 + \cdots, \tag{124}$$

where the integrations with respect to $d\tau_3$, $d\tau_4$ are over the slab. This permits points 3 and 4 to coincide, which is the situation we now wish to avoid. From (121) and (122) with $p^i = (-\mathbf{p}, E) = (0, m)$ the evaluation of (124) is now subject to

$$| t_3 - t_4 | \geqslant \epsilon. \tag{125}$$

To deal with the restriction (125) we separate the second term of (124) into two terms Q_1, Q_2 such that $t_3 > t_4$ in Q_1 and $t_3 < t_4$ in Q_2.

For $T \geqslant t_3 > t_4 \geqslant 0$ write

$$K_+(2, 3) = \int_{E_2>0} \frac{1}{2E_2} (\gamma_4 E_2 - \boldsymbol{\gamma} \cdot \mathbf{p}_2 + m) \, e^{-iE_2(T-t_3)+i\mathbf{p}_2\cdot(\mathbf{x}_2-\mathbf{x}_3)} \frac{d^3\mathbf{p}_3}{(2\pi)^3}, \tag{126}$$

$$K_+(3; 4) = \int_{E>0} \frac{1}{2E} (\gamma_4 E - \boldsymbol{\gamma} \cdot \mathbf{p} + m) \, e^{-iE(t_3-t_4)+i\mathbf{p}\cdot(\mathbf{x}_3-\mathbf{x}_4)} \frac{d^3\mathbf{p}}{(2\pi)^3}, \tag{127}$$

where

$$E_2^2 = \mathbf{p}_2^2 + m^2, \qquad E^2 = \mathbf{p}^2 + m^2. \tag{128}$$

The momentum representation for $\delta_+(q^2{}_{34})$ is given in (93). Performing integrations with respect to $d^3\mathbf{x}_3$, $d^3\mathbf{x}_4$, $d^3\mathbf{p}$, $d^3\mathbf{p}_2$ we get

$$\mathbf{p}_2 = 0, \qquad \mathbf{p} = -\mathbf{k}, \tag{129}$$

so that $E_2 = m$, $E = \sqrt{K^2 + m^2}$, and Q_1 becomes

$$Q_1 = (-ie^2) \cdot (4\pi i) \cdot \frac{1}{2^3} \iiint_{t_3>t_4} (\gamma_4 + 1)\, \gamma_i(\gamma_4 E + \boldsymbol{\gamma} \cdot \mathbf{k} + m)\, \gamma^i \mathscr{U}_0 \cdot \frac{1}{EK}$$
$$\cdot\, e^{-imT+i(m-E-K)(t_3-t_4)} \frac{d^3\mathbf{k}}{(2\pi)^3} \, dt_3 \, dt_4 . \tag{130}$$

Using symmetrical integration to remove terms linear in $\mathbf{k}$, Q_1 is simplified to

$$Q_1 = \xi_1 \mathscr{U}_0 e^{-imT}, \tag{131}$$

where

$$\xi_1 = \frac{e^2}{\pi} \iiint_{t_3>t_4} \frac{2m - \sqrt{K^2 + m^2}}{\sqrt{K^2 + m^2}} \, e^{i(m-E-K)(t_3-t_4)} \, K \, dK \, dt_3 \, dt_4 . \tag{132}$$

Similarly we have

$$Q_2 = \xi_2 \mathscr{U}_0 e^{-imT}, \tag{133}$$

where

$$\xi_2 = \frac{e^2}{\pi} \iiint_{t_4>t_3} \frac{2m + \sqrt{K^2+m^2}}{\sqrt{K^2+m^2}} e^{i(m+E+K)(t_3-t_4)} K\, dK\, dt_3\, dt_4 . \qquad (134)$$

It is usually argued that divergences in Q_1, Q_2 arise from high frequency quanta, i.e., when $K \to \infty$. This is based on a calculation in which the integrations with respect to time are performed first. Now because of the time restrictions we integrate first with respect to frequency. For $t_3 \neq t_4$ the integrals with respect to K converge. To see this we examine the behavior of (132) and (134) at large K. Thus for $K \gg m$,

$$\xi_1 \sim \frac{e^2}{\pi} \iiint_{t_3>t_4} (2m - K)\, e^{i(m-2K)(t_3-t_4)}\, dK\, dt_3\, dt_4$$

$$= \frac{e^2}{\pi} \iint_{t_3>t_4} e^{im(t_3-t_4)} \int_0^\infty (2m - K)\, e^{-2iK(t_3-t_4)}\, dK\, dt_3\, dt_4 , \qquad (135)$$

and

$$\xi_2 \sim \frac{e^2}{\pi} \iint_{t_3<t_4} e^{im(t_3-t_4)} \int_0^\infty (2m + K)\, e^{2iK(t_3-t_4)}\, dK\, dt_3\, dt_4 . \qquad (136)$$

The actual values of ξ_1, ξ_2 differ from these asymptotic values by finite quantities.

Equations (135) and (136) can be simplified further by using the results

$$\lambda \neq 0, \qquad \int_0^\infty e^{-i\lambda x}\, dx = \frac{1}{i\lambda}, \qquad \int_0^\infty x e^{-i\lambda x}\, dx = -\frac{1}{\lambda^2} . \qquad (137)$$

Thus

$$\xi_1 \sim \frac{e^2}{\pi} \iint_{t_3>t_4} e^{im(t_3-t_4)} \left\{ \frac{-im}{t_3 - t_4} + \frac{1}{4(t_3 - t_4)^2} \right\} dt_3\, dt_4 , \qquad (138)$$

$$\xi_2 \sim \frac{e^2}{\pi} \iint_{t_3<t_4} e^{im(t_3-t_4)} \left\{ \frac{im}{t_3 - t_4} - \frac{1}{4(t_3 - t_4)^2} \right\} dt_3\, dt_4 . \qquad (139)$$

Interchanging the naming of t_3 and t_4 in ξ_2 we have

$$\xi_1 + \xi_2 \sim \frac{2e^2}{\pi} \iint_{t_3>t_4} \left\{ -im \frac{\cos m(t_3 - t_4)}{t_3 - t_4} + i \frac{\sin m(t_3 - t_4)}{4(t_3 - t_4)^2} \right\} dt_3\, dt_4$$

$$= \frac{2ie^2}{\pi} \int_0^{mT} dx \int_{m\epsilon}^{x} \left\{ -\frac{\cos y}{y} + \frac{\sin y}{4y^2} \right\} dy, \qquad (140)$$

where

$$x = mt_3, \qquad y = m(t_3 - t_4), \tag{141}$$

and we have now taken account of the restriction $| t_3 - t_4 | \geqslant \epsilon$.

For a slab with $mT \gg 1$, which is the situation usually considered, we have

$$\xi_1 + \xi_2 \sim \frac{2ie^2}{\pi} \int_0^{mT} dx \int_{m\epsilon}^{\infty} \frac{1}{y} \left\{ \frac{\sin y}{4y} - \cos y \right\} dy. \tag{142}$$

For $m\epsilon \ll 1$, the main contribution to (142) comes from $y \ll 1$ when the trigonometrical factor $\sim -3/4$, and

$$\xi_1 + \xi_2 \sim -iAT, \tag{143}$$

where

$$A = -\frac{3e^2}{2\pi} m \ln(m\epsilon). \tag{144}$$

Since we are considering paths to have finite time steps ϵ, it would be a stricter procedure to divide the whole time range $[0, T]$ into many steps of length ϵ and to work out $\xi_1 + \xi_2$ as a sum over steps rather than as an integral. It turns out when this is done that the result is the same as (144).

The wavefunction at $t = T$ to order e^2 is therefore given by

$$\mathscr{U}_0 e^{-imT} \left\{ 1 + i \cdot \frac{3e^2}{2\pi} m \ln(m\epsilon) T \right\} \approx \mathscr{U}_0 e^{-im'T} \tag{145}$$

for T not too large, where

$$m' = m \left\{ 1 - \frac{3e^2}{2\pi} \ln(m\epsilon) \right\} = m + A. \tag{146}$$

The effect of the response of the universe is to "change" the mass m to m' given by (146). The usual theory gives

$$A = \frac{3e^2}{2\pi} m \ln \frac{K_0}{m}, \tag{147}$$

where K_0 is a cutoff applied to the K integration, which is carried out after integrations with respect to t_3, t_4. Although (144) and (147) are similar in form, the cutoff in (144) can be given a physical basis, for example, in terms of the gravitational radius of the particle, whereas that in (147) is a mathematical variable from a Fourier integral and should not play a role in the theory.

We wish next to consider the corresponding problem for a free particle with $p^i = (-\mathbf{p}, +E)$, $\mathbf{p} \neq \mathbf{0}$, $E > 0$ and 4-spinor $\mathscr{U}$. According to (121), (122) there are restrictions now on all the coordinates. This leads to a more complicated situation. It is possible, however, to simplify the work by noting first that if, as before, we restrict only the time coordinate, by $| t_3 - t_4 | \geqslant \epsilon^4$, the outcome for the wavefunction at point 2 on $t = T$ is

$$\mathscr{U} e^{i\mathbf{p}\cdot\mathbf{x}_2 - i(E+\Delta E)T}, \tag{148}$$

where

$$E\Delta E = -\frac{3e^2m^2}{2\pi}\ln(E\epsilon^4). \tag{149}$$

The result (149) is obtained in the Appendix. Since we are integrating invariant functions over a four-dimensional region it is clear that if we had used

$$| X_3{}^{\mu} - X_4{}^{\mu} | \geqslant \epsilon^{\mu} \tag{150}$$

to restrict one of the space coordinates, $\mu = 1, 2, 3$, instead of the time coordinate, we should have obtained $\ln(p_\mu \epsilon^\mu)$ instead of $\ln(E\epsilon^4)$, and that if we had used

$$| X_3{}^{i} - X_4{}^{i} | \geqslant \epsilon^{i}, \qquad i = 1, 2, 3, 4, \tag{151}$$

for all coordinates we should have obtained $\ln(p_i \epsilon^i)$. This is simply $\ln(m\epsilon)$, so that in place of (149) we have the invariant result

$$E\,\Delta E = -\frac{3e^2}{2\pi}\, m^2 \ln(m\epsilon) = m\,\Delta m. \tag{152}$$

Unlike the usual theory the so-called "wavefunction renormalization," represented by the constant

$$B = -\frac{e^2}{2\pi}\ln(m\epsilon), \tag{153}$$

did not appear in the above work. We shall return to this point in Section 7. The constant B does appear, however, if we seek the wavefunction at an intermediate time t, $0 < t < T$. We again consider the case of a particle with wavefunction $\mathscr{U}_0$ corresponding to energy m at $t = 0$. If we were to restrict our time integrals to the interval $(0, t)$ the analysis would be the same as that given above, but with t replacing T and we should obtain $\mathscr{U}_0 e^{-im't}$ for the wavefunction at time t. But we must permit integrations up to T. This does not change Q_1, which has $t \geqslant t_3 > t_4 \geqslant 0$. The restriction $t_3 \leqslant t$ here arises because $\delta_+(q_{34}^2)$ leaves the particle in a positive energy state after the response is completed at point 3, and hence we must go forward in time from t_3 to t. But for Q_2, in addition to $t \geqslant t_4 > t_3 \geqslant 0$, we have also to consider

$$T \geqslant t_4 > t > t_3 \geqslant 0. \tag{154}$$

It is this additional time range that makes a contribution to the wavefunction at time t, over and above $\mathscr{U}_0 e^{-im't}$.

The extra contribution is

$$\xi_2' \mathscr{U}_0 e^{-imt}, \tag{155}$$

where

$$\xi_2' \sim \frac{e^2}{\pi} \iint_{T \geqslant t_4 > t > t_3 \geqslant 0} e^{im(t_3 - t_4)} \int (2m + K)\, e^{2iK(t_3 - t_4)}\, dK\, dt_3\, dt_4$$

$$\sim \frac{e^2}{\pi} \iint_{T \geqslant t_4 > t > t_3 \geqslant 0} e^{im(t_3 - t_4)} \left\{ \frac{-im}{t_4 - t_3} - \frac{1}{4(t_4 - t_3)^2} \right\} dt_3\, dt_4 . \tag{156}$$

The intervals for t_3, t_4 have been kept open at t in the above formulas to indicate that $t_4 - t_3$ has a minimum value ϵ. We require

$$t_3 \leqslant t - \alpha\epsilon, \qquad t_4 \geqslant t + \beta\epsilon, \tag{157}$$

where α, β are positive numbers satisfying $\alpha + \beta = 1$. All possible choices for α, β lead to the same final result. Taking $\alpha = 0$, and making the substitutions

$$x = m(t_4 - t), \qquad y = m(t_4 - t_3), \tag{158}$$

we obtain

$$\xi_2' \sim -\frac{e^2}{\pi} \int_{m\epsilon}^{m(T-t)} dx \int_x^{x+mt} e^{-iy} \left\{ \frac{i}{y} + \frac{1}{4y^2} \right\} dy. \tag{159}$$

Provided t is not close to zero or to T the $1/4y^2$ in (159) gives a large contribution when both x and y are close to $m\epsilon$. After integrating with respect to y for this term we get

$$\xi_2' \sim -\frac{e^2}{4\pi} \int_{m\epsilon} \frac{dx}{x} \sim \frac{e^2}{4\pi} \ln(m\epsilon) \sim -\frac{1}{2} B. \tag{160}$$

Provided t is not close to zero or to T, the required wavefunction to order e^2 is therefore

$$(1 - iAT - \tfrac{1}{2}B)\, \mathscr{U}_0 e^{-imt}. \tag{161}$$

The appearance of B within the slab is of importance when an external potential acts on the particle. In general and external potential[2] can be represented by a four-dimensional Fourier integral

$$B_i(X) = \int b_i(q)\, e^{-iq\cdot x} \frac{d^4q}{(2\pi)^4}\,. \tag{162}$$

The scattering problem is to determine the amplitude for a state of momentum $\mathbf{p}_2$ at $t_2 = T$ given that the state has momentum $\mathbf{p}_1$ at $t_1 = 0$. From the path integral point of view we have to consider the effect of the phase factor

$$\exp\left[-ie \int_{\Gamma_{21}} B_i\, dx^i\right] \tag{163}$$

on this problem.

Expanding (163) as a power series, the unity term leads to (145), which involves no scattering. The linear term in B_i produces scattering, however. In most cases of physical interest higher order terms in B_i give smaller contributions to the scattering than the linear term, so it is usual to restrict the problem to determining the effect of the linear term. It is found that only those components in the Fourier integral for which

$$\mathbf{q} \cong \mathbf{p}_2 - \mathbf{p}_1\,, \qquad q_4 \cong E_2 - E_1\,, \tag{164}$$

where $E_2 = (\mathbf{p}_2{}^2 + m^2)^{1/2}$, $E_1 = (\mathbf{p}_1{}^2 + m^2)^{1/2}$, make a contribution to the scattering into states of momentum $\mathbf{p}_2$. For finite T the Fourier components making effective contribution can depart from the strict equalities of (164) by small terms of order T^{-1}.

The scattering amplitude can be developed in a series of ascending powers of q^i. Since b^i, q^i appear in relativistically invariant forms, and since $b(q)\cdot q = 0$ because of the Lorentz condition, the b^i components can only appear in the combination $\not{b}$, while the q^i components can appear as q^{2n+2} or as $q^{2n}\not{q}$, $n = 0, 1, 2,\ldots$. The most general form for the scattering amplitude is therefore seen to be

$$\not{b}(q) \sum_{n=0}^{\infty} C_n q^{2n} + [\not{b}(q)\not{q} - \not{q}\not{b}(q)] \sum_{n=0}^{\infty} D_n q^{2n}, \tag{165}$$

where the C_n, D_n are constants. Detailed calculations in quantum electrodynamics show that none of C_n, $n = 1, 2,\ldots$; D_n, $n = 0, 1, 2,\ldots$ involve the cutoff $\ln(K_0/m)$. Because the present theory differs from the usual theory only in the form of the cutoff, not in the places where it occurs, the present theory leads to (165) with the same coefficients—except possibly for C_0. It is well-known that the second series

[2] The reader will note the clash between our notation B_i for the external potential and B for the constant $-e^2 \ln (m\epsilon)/2\pi$.

in (165) gives the anomalous magnetic moment of the electron, and the entire series without the C_0 term gives the Lamb shift. The observable effects of including

$$\exp\left[-\tfrac{1}{2}e^2 \iint \delta_+(q^2_{A\tilde{A}})\, da^i\, d\tilde{a}_i\right] \tag{166}$$

in scattering are therefore the same in the present theory as they are in the usual theory.

Although we are working only to the first order in b_i it is important to notice that (165) includes all orders in the expansion of (166)—in the usual language we can have any number of virtual photons. Each order in e^2 is handled by Feynman–Dyson graphs. It is found that while individual graphs give cutoff dependent contributions to C_0 the sum of the contributions of all graphs does not involve the cutoff. The well known situation to order e^2 is shown in Fig. 4. The cutoff dependent term from I cancels those from II and III.

Once again we expect the same result for C_0 in the present theory. To emphasire the part played by the constant B within the slab we demonstrate in the Appendix that this is so for the graphs of Fig. 4. The calculation is similar to those of the usual theory. Here we use a much simpler method, depending on the closed form of the electromagnetic phase factor in the path integrals, to show that C_0 must be cutoff independent to all orders in e^2, and hence it follows that the detailed calculations based on Feynman–Dyson graphs must lead to the cancellations described above.

The C_0 term in (165) survives as the components q^i tend to zero. Let the field B_i such that $b_i(q)$ is zero except near $q^i = 0$ for all i: That is to say, we collapse the Fourier coefficients b^i on to the origin in q space. As the region in which $b_i(q)$ is nonzero shrinks, B_i becomes effectively constant, and

$$\exp\left[-ie\int_{\Gamma_{21}} B_i\, dx^i\right] \to \exp[-ieB\cdot(x_2 - x_1)]. \tag{167}$$

The electromagnetic phase factor arising from B_i becomes path independent. The gauge transformation

$$B_i \to B_i - \frac{\partial\chi}{\partial x^i}, \qquad \chi = B_i x^i, \tag{168}$$

reduces the field to zero. Since a gauge transformation produces no physical effect the situation is the same as for a zero field. There can be no scattering and therefore no cutoff dependent terms in C_0.

Consider the case $B^i = (-\mathbf{B}, 0)$ in further detail, writing $\mathbf{p} = -e\mathbf{B}$. Let the four-component spinor at $t = 0$ again be $\mathscr{U}_0$ corresponding to a free particle state of energy m. Suppose that the wavefunction on $t = 0$ has no spatial dependence.

Then we have to consider $\mathscr{U}_0 e^{i\mathbf{p}\cdot\mathbf{x}_1}$ on $t = 0$ because of the $e^{i\mathbf{p}\cdot\mathbf{x}_1}$ factor in (167). Before propagation to $t = T$ the product $\mathscr{U}_0 e^{i\mathbf{p}\cdot\mathbf{x}_1}$ must be separated into positive and negative energy components for $\pm E = \pm\sqrt{p^2 + m^2}$. Write

$$\mathscr{U}_0 = \mathscr{U}_+ + \mathscr{U}_- , \tag{169}$$

where $\exp[i(\mathbf{p}\cdot\mathbf{x} \mp Et)]\mathscr{U}_\pm$ are solutions of the free particle Dirac equation. Only $\mathscr{U}_+$ is propagated forward to $t = T$. Propagation subject to (166) was studied earlier in this section. After multiplying by $e^{-i\mathbf{p}\cdot\mathbf{x}_2}$ from (167) the wavefunction at point 2 on $t = T$ is

$$\mathscr{U}_+ e^{-i(E+\Delta E)T}, \tag{170}$$

where ΔE is given by $E\,\Delta E = m\,\Delta m$, ΔE being defined in (152). If the field had been zero the wavefunction on $t = T$ would have been

$$\mathscr{U}_0 e^{-i(m+\Delta m)T}. \tag{171}$$

The difference between (170) and (171) appears to contradict what was said in the previous paragraph. The difference has arisen from the specification of the wavefunction at $t = 0$. Propagation from $t < 0$ cannot give a wavefunction on $t = 0$ with spinor $\mathscr{U}_0$ and without an $\mathbf{x}_1$ dependence unless $B_i = 0$ for $t < 0$. We have therefore introduced a discontinuity in the field at $t = 0$, a discontinuity that cannot be removed by a gauge transformation. Indeed the discontinuity generates components in the Fourier integral (162) for b_i away from the origin in q-space, which is just the situation we set out to avoid. It is therefore necessary, if we wish to have no $\mathbf{x}_1^-$ dependence on $t = 0$ to specify $\mathscr{U}_+$ as the 4-spinor. Alternately if we wish to specify $\mathscr{U}_0$ we must have an $\mathbf{x}_1^-$ dependence on $t = 0$ given by the factor $e^{-i\mathbf{p}\cdot\mathbf{x}_1}$. In both these cases the constant field has no effect.

6. Vacuum Polarization

The amplitude for the process shown in Fig. 2 is

$$-(-ie^2)(-ie) \iiint \bar{u}_2(5)\, \gamma_i u_1(5)\, \delta_+(q_{54}^2)\, \mathrm{Tr}[\gamma^i K_+(4;3)\, \not{B}(3)\, K_+(3;4)]\, d\tau_3\, d\tau_4\, d\tau_5 , \tag{172}$$

the integrations with respect to $d\tau_3$, $d\tau_4$, $d\tau_5$ being over the slab $0 \leqslant t \leqslant T$. The expression (172) is the lowest order contribution of the vacuum in the presence of an external field B_i to the scattering of a particle from state u_1 into state u_2. This amplitude follows without difficulty from the general path integral of Section 4. Since the integral of a constant potential around a closed path is zero, it is always possible to subtract an artibrary constant potential from B_i. We shall make use of this possibility at a later stage.

The scattering is the same as that produced by a potential

$$A^i(5) = ie^2 \iint \delta_+(q_{54}^2)\,\mathrm{Tr}[\gamma^i K_+(4;3)\,\not{B}(3)\,K_+(3;4)]\,d\tau_3\,d\tau_4\,. \tag{173}$$

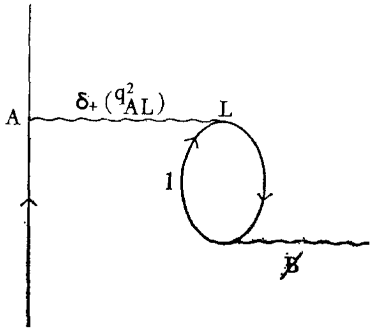

FIG. 2. The simplest vacuum diagram leading to an observable result (−27 Mc/sec in the Lamb shift).

This is the potential produced by a current $j^i(5)$ given by

$$j^i(5) = \frac{1}{4\pi}\,\Box_5 A^i(5) = ie^2 \int \mathrm{Tr}[\gamma^i K_+(5;3)\,\not{B}(3)\,K_+(3;5)]\,d\tau_3\,. \tag{174}$$

In this section we show that (174) can be evaluated in such a way that no quadratic divergence is present, and we show that the cutoff dependent logarithmic term is of the same form as that encountered in the preceding section. Given j^i we can construct A^i and hence obtain the scattering from

$$-ie \int \bar{u}_2(5)\,\not{A}(5)\,u_1(5)\,d\tau_5\,. \tag{175}$$

With a slight change of notation we write

$$j_k(1) = ie^2 \int \mathrm{Tr}[K_+(1;2)(\not{B}_2 - \not{B}_1)\,K_+(2;1)\,\gamma_k]\,d\tau_2 \tag{176}$$

where $\not{B}_2$, $\not{B}_1$ are the values of $\not{B}$ at points 2 and 1. The potential $\not{B}_1$ is constant and we are now making use of the possibility of subtracting a constant potential from the variable potential $\not{B}_2$. This step is useful in respect of the Taylor expansion

$$\not{B}_2 - \not{B}_1 = \gamma^i[B_{i,l}q^l + \tfrac{1}{2}B_{i,lm}q^l q^m + \cdots], \tag{177}$$

where the derivatives $B_{i,l}$, $B_{i,lm}$ are evaluated at point 1, and $q^i = x_2{}^i - x_1{}^i$. We write the propagators in the form

$$K_+(2;1) = (-\nabla\!\!\!/_2 + im)\, I_+(q^2) = -2\not{q} I_+'(q^2) + imI_+(q^2), \tag{178}$$

$$K_+(1;2) = 2\not{q} I_+'(q^2) + imI_+(q^2), \tag{179}$$

where

$$I_+'(q^2) = \frac{d}{dq^2}\, I_+(q^2), \tag{180}$$

and

$$I_+(q^2) = -\frac{1}{4\pi}\,\delta(q^2) + \frac{m}{8\pi q}\, H_1^{(2)}(mq). \tag{181}$$

In the Hankel function q is the positive square root of q^2 when $q^2 > 0$ and is the negative square root when $q^2 < 0$. The limiting form of $I_+(q^2)$ as $q^2 \to 0$ is

$$I_+(q^2) \to -\frac{1}{4\pi}\,\delta_+(q^2).$$

Inserting (177), (178), (179) in (176) we get

$$j_k(1) = ie^2 \int \mathrm{Tr}[(2\gamma^p q_p I_+' + imI_+)(B_{i,l}q^l + \tfrac{1}{2}B_{i,lm}q^l q^m + \cdots)\,\gamma^i \cdot (-2\gamma^r q_r I_+' + imI_+)\,\gamma_k]\, d^4q. \tag{182}$$

To simplify (182) we note that the integrations remove terms of odd order in q_l, and that the trace removes terms that contain an odd number of γ matrices. Thus

$$j_k(1) = (ie^2)\cdot B_{i,lm} \int \mathrm{Tr}[-2\gamma^p\gamma^i\gamma^r\gamma_k q_p q^l q^m q_r I_+'^2 - \tfrac{1}{2}m^2 I_+{}^2 \gamma^i\gamma_k q^l q^m + \cdots]\, d^4q. \tag{183}$$

Next we use

$$\mathrm{Tr}(\gamma^i\gamma_k) = 4\eta^i{}_k,$$
$$\mathrm{Tr}(\gamma^p\gamma^i\gamma^r\gamma_k) = 4(\eta^{pi}\eta^r{}_k + \eta^p{}_k\eta^{ir} - \eta^{pr}\eta^i{}_k), \tag{184}$$

to give

$$j_k(1) = -8ie^2 B_{i,lm} \int (2q^i q_k - q^2\eta^i{}_k)\, I_+'^2 q^l q^m d^4q - 2ie^2m^2 B_{i,lm} \int \eta^i{}_k I_+{}^2 q^l q^m d^4q + \cdots. \tag{185}$$

Only the first term in (185) is cutoff dependent. The rest remain finite as $\epsilon \to 0$. Since it is the cutoff dependent term we are seeking we shall confine further attention to this first term.

By symmetry we can write

$$\int I_+'^2 q^l q^m q^i q^k d^4q = \lambda[\eta^{lm}\eta^{ik} + \eta^{li}\eta^{mk} + \eta^{lk}\eta^{mi}]. \tag{186}$$

Contracting l against m and i against k gives

$$\int \{I_+'(q^2)\, q^2\}^2\, d^4q = 24\lambda. \tag{187}$$

The integral on the left hand side of this equation is worked out in the Appendix. Using (A.56) we have

$$\lambda = -\frac{i}{192\pi^2}\ln(m\epsilon). \tag{188}$$

Returning to the first term of (185) we note that

$$\begin{aligned} B_{i,lm} &\int (2q^i q_k - q^2\eta^i{}_k)\, I_+'^2 q^l q^m d^4q \\ &= 2\lambda B_{i,lm}(\eta^{lm}\eta^i{}_k + \eta^{li}\eta^m{}_k + \eta^l{}_k\eta^{mi} - 3\eta^{lm}\eta^i{}_k) \\ &= -4\lambda\Box B_k + 2\lambda B^l_{,lk} + 2\lambda B^m_{,km} \\ &= -4\lambda\Box B_k\,, \end{aligned} \tag{189}$$

since $B^m_{,km} = B^m_{,mk} = 0$ and $B^l_{,lk} = 0$. Using (189) we get

$$j_k(1) \sim 32ie^2\lambda\Box B_k\,. \tag{190}$$

With (188) for λ,

$$j_k(1) \sim \frac{e^2}{6\pi^2}\ln(m\epsilon)\,\Box\, B_k\,. \tag{191}$$

The d'Alembertian of B_k is taken at point 1. Hence there is no vacuum current except at the sources of the external field B_k. In the direct particle theory this means there is no polarization current except at the particles. If J_k is the particle current we have $\Box B_k = 4\pi J_k$ and

$$j_k(1) = \frac{2e^2}{3\pi}\ln(m\epsilon)\, J_k(1). \tag{192}$$

The result that the polarization current occurs only at the particles holds also for the finite terms not considered explicitly above.

The usual theory gives

$$j_k(1) \sim -\frac{2e^2}{3\pi} \ln\left(\frac{M}{m}\right) J_k(1), \tag{193}$$

in terms of a momentum cutoff M. In Section 5 we saw that ϵ in the present theory replaces the photon cutoff K_0 of the usual theory. Now we see that ϵ also replaced the momentum cutoff of the usual theory. There is no requirement for two distinct modes of cutoff.

We also note that the quadratic divergence, encountered in most treatments of the usual theory, did not appear in the above work.

7. General Discussion

We now review the position that has been reached. Infinities are removed in classical theory by requiring that points on paths do not act on themselves—this is sufficient to remove all self-action since classical paths are timelike. Infinities are removed in quantum theory provided there is no self-action on a scale less than ϵ.

In Section 5 we discussed the possibility that ϵ is defined by the gravitational radius $2Gm$. If so, then ϵ is smaller than the Compton wavelength m^{-1} by a factor $\sim 10^{-40}$, and $-\ln(m\epsilon) \sim 10^2$. The "experimental mass" m' obtained in Section 5 is then about $4m/3$. Since the vacuum current (192) is proportional to the particle current we can think of the effect of the vacuum as changing the charge e to an "experimental" charge e'. For $-\ln(m\epsilon) \sim 10^2$ we get $e' \sim 0.9e$. Even for ϵ as small as the gravitational radius the modifications produced by self-action are not so great as to make the power series expansion of

$$\exp\left[-ie^2 \iint \delta_+(q^2_{A\tilde{A}})\, da^i\, d\tilde{a}_i\right] \tag{194}$$

an unreasonable procedure.

This raises an issue concerning renormalization theory. The aim is to show that the effects of (194) are equivalent to a change of m, e to some m', e', which are then identified with the experimental values, by which is meant the interpretation of experiment in accordance with a theory that omits the cutoff dependent terms arising from (194). This is done by an induction method working in terms of the expansion of (194) (see [5, p. 330]). The required result is first proved to order e^2 using cutoff values for K_0, M that are small enough for the e^2 term to appreciably exceed the e^4, e^6,... terms. Under the same convergence condition the result is next assumed to order e^{2n} and is thence proved to order e^{2n+2}.

This renormalization demonstration is then frequently followed either by the statement or the implication that K_0, M can be permitted to tend to infinity without vitiating the result. This may be true. A process of analytic continuation may exist that extends the result to high values of K_0, M for which the expansion in e^2 cannot be used, but there is no good reason for this supposition. A similar situation would arise in the present theory if we let $\epsilon \to 0$. Instead we prefer to keep ϵ finite. It does not strain credulity to suppose that something new happens on the scale of the gravitational radius of a particle. Rather would it be surprising if this were not so. With ϵ finite, it is then important that we need neither K_0 nor M. Both K_0, M can tend to infinity—indeed both must tend to infinity since these quantities are mathematical variables in Fourier integrals.

The expression (194) is only a part of the general influence functional. The forms $F[\mathbf{a}, \mathbf{a}']$, $F[\mathbf{a}, \mathbf{b}; \mathbf{a}', \mathbf{b}']$, given in (94), (96) of Section 3, show a clear distinction between scattering problems and decay problems. The normal procedure in scattering is to add amplitudes for different frequencies before taking the square of the modulus to give a probability. In decay problems, on the other hand, the square of the modulus for each separate frequency is taken first and the resulting moduli are then added for the total probability. This is necessary to give an appropriate time dependence for the decay of a quantum state. It is usually argued that, while the frequencies for scattering are determined by the δ_+ function and are therefore systematically phased with respect to each other, the frequencies in decay problems arise from randomly phased field oscillators. The total amplitude should always be obtained by adding all frequencies but in decay problems cross-products between different frequencies average to zero when the square of the modulus is taken. No doubt a consistent theory can be formulated in this way, but the issue is treated scantily if at all in most quantum theory texts, so the issue would seem to cause some embarrassment.

The situation is clear-cut in the present theory. The influence functionals contain couplings of paths with paths, conjugate paths with conjugate paths, and couplings of paths with conjugate paths. From (94) and (96) it is seen that when a path interacts with a conjugate path it is the same frequency that operates on both. There is no case in which a frequency K interacting with a path is combined with a different frequency K' acting on a conjugate path. This is the decay problem case. When paths interact with paths, on the other hand, the interaction frequencies have no correlation with those that operate in the coupling of conjugate paths with conjugate paths. This is the scattering case.

Vacuum loops were introduced in Section 4 in two ways, through antisymmetrization of the perturbation expansion and through a path integral. Although the two methods are equivalent the path integral taken alone appears rather as a cumbersome irrelevancy. One can ask: Why bother to put in vacuum loops at all? That the Lamb shift would otherwise be 27 Mc/sec wrong hardly seems a satisfactory answer. An answer is needed that shows the theory with vacuum loops to be more elegant than the theory without them.

It is possible for paths with time reversals to intersect themselves. Such an intersecting path, belonging to the amplitude for Compton scattering, is shown in Fig. 3(a). Remembering that the paths of different particles do not intersect each other—this is the meaning of the exclusion principle—it is natural to take the view that individual paths should not intersect themselves. We must then subtract the amplitude for Fig. 3(a) from the amplitude given by the theory without loops.

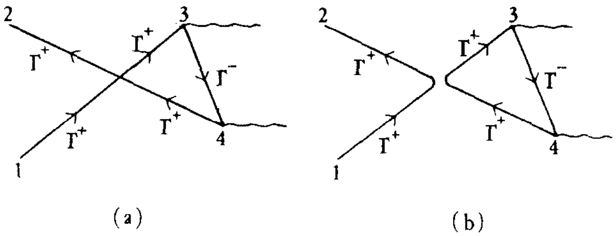

FIG. 3. Closed loops are necessary to take account of the fact that paths do not cross themselves. The self-intersecting path in (a) is redrawn in (b) to represent a closed loop touching a path without time reversal. The amplitude for (b) must be subtracted from that given by the theory without closed loops.

Figure 3(a) is redrawn in 3(b) as a path without time reversal and a closed loop in which there are two external "photon lines." There is a strong presumption that what we did in Section 4 amounted to just this subtraction. It is true that the formulation of Section 4 permitted loops to be disjoint from the paths of particles. But we already noted in Section 4 that the interaction of loops with themselves leads only to a phase factor e^{iC} in the amplitude, C real, and this does not appear in the present theory since we are concerned throughout with probabilities. Moreover it was noted in Section 6 that it is only when loops touch particle paths—the case of Fig. 3(b)—that there is a contribution to the vacuum current, and hence to scattering. The case of Fig. 3(b) is the only one that is of physical relevance.

It seems then that closed loops are not an addition at all. It is rather that they represent the removal of something unwanted. That this is indeed a correct interpretation is shown by the sign of the vacuum current. Both the path and the loop in Fig. 3(b) are directed the same way at their point of contact. Consequently Fig. 3(b) represents a current in the same sense as other paths in the same direction and in the same time direction. This means that for a beam of particles of one kind, electrons say, the vacuum current has a sign opposite to that of the electrons themselves—opposite because we have to subtract Fig. 3(b).

It remains an open question as to whether the discussion of Section 4 is the most effective way to formulate this idea. We suspect that either some further mathematical results remain to be discovered or a better procedure can be found.

We started these papers with the intention of discussing electromagnetic interaction. We had little thought in the beginning of investigating the nature of free particles. In the event, the electromagnetic part of the work, given in Section 3, turned out to be comparatively straightforward, whereas the free particle formulation has given us considerable trouble. The difficulties can be said to arise from the path integral method, but we think this is not so much a defect in the path integral method as an indication of omission in the theory as a whole.

We already noted in Section 1 that Feynman's method of arriving at the propagator is somewhat roundabout. In Section 2 we proceeded in a more direct way, but even so an awkward question remains. We introduced the concept of $\Gamma^{\pm}$ paths, with the hypothesis that Γ^{+} paths are weighted by wavefunctions ψ_{+} of positive energy and Γ^{-} paths by wavefunctions ψ_{-} of negative energy. There was then little difficulty in showing that the K_{+} propagator was required to give proper accounting in the perturbation expansion. But what lies behind this hypothesis?

We hope to consider this question in future work. Here we simply indicate the direction in which we think the answer may lie. The K_{+} propagator is $(-\nabla\!\!\!/ + im)I_{+}$, and for small q^2 we have $I_{+} = -\delta_{+}(q^2)/4\pi$. The appearance of $\delta_{+}(q^2)$ in analogy to the situation in respect of the electromagnetic field is an indication that some form of response of the Universe is involved, and the following remarkable circumstance points also to this conclusion. The propagators $K_0^{\pm}(2; 1)$ introduced in Section 2 satisfy the inhomogeneous Dirac equation. So does

$$\tfrac{1}{2}[K_0^{+}(2; 1) + K_0^{-}(2; 1)]. \tag{195}$$

On the other hand

$$\tfrac{1}{2}\left[\sum_{E_n>0} u_n(2)\,\bar{u}_n(1) - \sum_{E_n<0} u_n(2)\,\bar{u}_n(1)\right] \tag{196}$$

satisfies the homogeneous Dirac equation. The sum of (195) and (196) is the propagator $K_{+}(2; 1)$. Expressed in this way there is an analogy with the Wheeler–Feynman formulation of classical electromagnetism, the retarded and advanced electromagnetic fields of the particles being replaced by K^{+} and K^{-}. The presumption is that (195) is the free particle propagator and that (196) is a response of the Universe.

When in (83) of Section 3 we called for the response of the Universe we limited the theory to answering probability questions concerning outgoing particles from the slab $0 \leqslant t \leqslant T$. Of course one need not call for the response of the Universe but then nothing useful can be said because we cannot use the influence functionals of Section 3 in an amplitude calculation. In this respect the present theory is different from the usual theory. However, since probabilities are always taken at the end in the usual theory, the difference is not noticed in practical calculations.

We end with some remarks on the interpretive aspects of quantum theory. Intermediate steps are used in most calculations, the steps at which a perturbation acts, for example. Such steps must be dealt with in terms of the amplitude—this is implicit in the path integral evaluation. We work with amplitudes inside the slab defined by $t = 0$, $t = T$. The probabilities apply on the boundary.

In the usual theory taking probabilities is described as an "experiment." In our terminology the experimental situation lies within the slab and calling for the response of the Universe constitutes the experiment. Our calculations do not tell us what will happen in an individual case, however; the probabilities we have obtained represent an average of the different things that happen when many experiments of a similar kind are carried out. In an individual experiment we proceed in a different way.

After an individual experiment one is faced by a "new deal" situation. One can calculate the probability of a track of a specified kind appearing in a bubble chamber. But once it has been determined in an actual experiment whether such a track is present or not we no longer continue to describe the situation in terms of probabilities, or in terms of the original amplitude. It has always been a mysterious feature of quantum theory that by actual observation our description of the world is sharpened. The experimenter is in a better position to determine the future state ($t > T$) of his bubble chamber after he has examined it ($t = T$) then if he had simply extended the time range of his original calculations. This situation arises because we can only specify the response of the Universe statistically in our calculations. The response of Section 3 does not give the response in a particular case, it represents the average response for an ensemble of similar cases. By actual experiment, on the other hand, we discover information about the response even in a particular case.

Can our theoretical description of the world be improved to include the sharpening achieved by experiment? Attempts to do this have been made from time to time in terms of "hidden variables." Such attempts have been unsuccessful and the present discussion shows why. We pay the penalty of being restricted to statistical calculations because of our insistence on localizing problems. We cannot expect to know more about the interior of our slab until we know more about the exterior. Quantum uncertainty arises from separating restricted volumes from the rest of the Universe.

Appendix

1. In Section 5 we quoted (148), (149) for the propagation of a state of momentum $\mathbf{p}$, going from $\mathscr{U}e^{i\mathbf{p}\cdot\mathbf{x}_1}$ on $t = 0$ to $\mathscr{U}e^{i\mathbf{p}\cdot\mathbf{x}_2 - i(E+\Delta E)t_2}$ for $t_2 = T$. Here $E = (\mathbf{p}^2 + m^2)^{1/2}$ and $\mathscr{U}$ satisfies

$$(\gamma_4 E - \boldsymbol{\gamma}\cdot\mathbf{p})\mathscr{U} = \not{p}\mathscr{U} = m\mathscr{U}. \tag{A1}$$

We establish (149) for ΔE, starting from (125). In place of (130) we get

$$Q_1 = (-ie^2)(4\pi i)\frac{1}{2^3}\iiint_{t_3>t_4} \frac{\not{p}+m}{E}\gamma_i \frac{\gamma_4 E_1 - \boldsymbol{\gamma}\cdot(\mathbf{p}-\mathbf{k}) + m}{E_1}\gamma^i\mathscr{U}\cdot\frac{1}{K}$$
$$\cdot\, e^{-iET + i(E-E_1-K)(t_3-t_4)}\frac{d^3\mathbf{k}}{(2\pi)^3}\,dt_3\,dt_4\,, \tag{A2}$$

where

$$E_1^2 = (\mathbf{p}-\mathbf{k})^2 + m^2, \qquad E_1 > 0. \tag{A3}$$

To simplify Q_1 we note the relations

$$\gamma_i[\gamma_4 E_1 - \boldsymbol{\gamma}\cdot(\mathbf{p}-\mathbf{k}) + m]\gamma^i = 2[2m - \gamma_4 E_1 + \boldsymbol{\gamma}\cdot(\mathbf{p}-\mathbf{k})], \tag{A4}$$

$$(\not{p}+m)[2m - \gamma_4 E_1 + \boldsymbol{\gamma}\cdot(\mathbf{p}-\mathbf{k})]\,\mathscr{U} = (\not{p}+m)[m - \gamma_4(E_1 - E) - \boldsymbol{\gamma}\cdot\mathbf{k}]\,\mathscr{U}$$
$$= [2m^2 - 2E(E_1 - E) - 2\mathbf{k}\cdot\mathbf{p}]\,\mathscr{U}. \tag{A5}$$

Using (A4), (A5) in the integrand of (A2) we get

$$Q_1 = \xi_1\mathscr{U}e^{-iET}, \tag{A6}$$

where

$$\xi_1 = \frac{e^2}{\pi}\iiint_{t_3>t_4} Ie^{i(E-K)(t_3-t_4)}\,dK\,dt_3\,dt_4\,, \tag{A7}$$

$$I = K\int_0^\pi \frac{m^2 + E^2 - EE_1 - Kp\cos\theta}{2EE_1}\,e^{-iE_1(t_3-t_4)}\sin\theta\,d\theta, \tag{A8}$$

$$E_1^2 = E^2 + K^2 - 2pK\cos\theta, \qquad p = |\,\mathbf{p}\,|\,. \tag{A9}$$

For K large, and using E_1 as the variable in (A8),

$$I = \int \frac{2m^2 + E^2 - 2EE_1 - K^2 + E_1^2}{4pE}\,e^{-iE_1(t_3-t_4)}\,dE_1\,, \tag{A10}$$

the range with sufficient accuracy being from

$$K + \frac{m^2}{2K} - p \quad \text{to} \quad K + \frac{m^2}{2K} + p.$$

To evaluate (A10) we note that the main contribution to (A6) comes from $t = t_3 - t_4$ small, and that

$$\int e^{-iE_1 t}\, dE_1 \cong 2e^{-i[K+(m^2/2K)]t}\, \frac{\sin pt}{t}, \tag{A11}$$

$$\begin{aligned}\int E_1 e^{-iE_1 t}\, dE_1 &= i\, \frac{d}{dt} \int e^{-iE_1 t}\, dE_1 \\ &\cong 2i\, \frac{d}{dt} \left[e^{-i[K+(m^2/2K)]t}\, \frac{\sin pt}{t}\right],\end{aligned} \tag{A12}$$

$$\int E_1^{\,2} e^{-iE_1 t}\, dE_1 \cong -2\, \frac{d^2}{dt^2} \left[e^{-i[K+(m^2/2K)]t}\, \frac{\sin pt}{t}\right], \tag{A13}$$

For K large and t small it is sufficient to replace $e^{-i[K+(m^2/2K)]t} \sin pt/t$ in both (A11) and (A12) by pe^{-iKt}. Replacing $\sin pt/t$ by p also in (A13) does not lead to an error and gives

$$\begin{aligned}\int E_1^{\,2} e^{-iE_1 t}\, dE_1 &\cong 2p \left(K + \frac{m^2}{2K}\right)^2 e^{-i[K+(m^2/2K)]t} \\ &\cong 2p(K^2 + m^2)\, e^{-iKt}.\end{aligned} \tag{A14}$$

Inserting in (A10) leads to

$$I \cong \frac{3m^2 + E^2 - 2EK}{2E}\, e^{-iK(t_3 - t_4)}. \tag{A15}$$

There is, however, a point of subtlety in (A13). Writing more strictly

$$\begin{aligned}-2\, \frac{d^2}{dt^2} \left[e^{-i[K+(m^2/2K)]t}\, \frac{\sin pt}{t}\right] &\cong -2p\, \frac{d^2}{dt^2} \left[e^{-i[K+(m^2/2K)]t} \left(1 - \frac{1}{6} p^2 t^2 + \cdots\right)\right] \\ &\cong \left[2p(K^2 + m^2) - \frac{p^3}{3}(2iKt - 1)\right] e^{-iKt},\end{aligned} \tag{A16}$$

it is necessary to show that the second term of (A16) makes no contribution. This cannot be demonstrated until we come to integrate with respect to dK. We then have

$$2it \int_0^\infty Ke^{-2iKt}\, dK = -t\, \frac{d}{dt} \int_0^\infty e^{-2iKt}\, dK = -t\, \frac{d}{dt}\left(\frac{1}{2it}\right) = \int_0^\infty e^{-2iKt}\, dK. \tag{A17}$$

Omission of the second term of (A16) is justified, but only because of the subsequent integration with respect to K.

The rest of the calculation is essentially the same as that given in Section 5. We have

$$\xi_1 \sim \frac{e^2}{\pi} \iiint_{t_3 > t_4} (2M - K)\, e^{-i(E-2K)(t_3 - t_4)}\, dK\, dt_3\, dt_4\,, \tag{A18}$$

where

$$2M = \frac{3m^2 + E^2}{2E}\,. \tag{A19}$$

With

$$\xi_2 \sim \frac{e^2}{\pi} \iiint_{t_3 < t_4} (2M + K)\, e^{i(E+2K)(t_3 - t_4)}\, dK\, dt_3\, dt_4\,, \tag{A20}$$

there is no difficulty in showing that

$$\xi_1 + \xi_2 \sim \frac{ie^2 T}{\pi} \cdot \frac{3m^2}{2E} \ln(E\epsilon_4), \tag{A21}$$

which is the required result, i.e.,

$$\Delta E = -\frac{3e^2}{2\pi} \frac{m^2}{E} \ln(E\epsilon_4)\,. \tag{A22}$$

2. We also left over from Section 5 the demonstration by direct calculation that the cutoff dependent contribution from I of Fig. 4 to C_0 in (165) cancels the contributions of II and III.

It is sufficient to work with a constant potential B_i since we are concerned with terms that survive as $\not{q} \to 0$. For comparison with Section 5 we take a particle at rest at $t = 0$ with wavefunction $\mathscr{U}_0$ and compute the wavefunction at $t = T$ to first order in B_i and to order e^2 in the expansion of (166). The relevant terms are

$$\text{I. } (-ie^2)(-ie) \int \cdots \int K_+(\mathbf{x}_2\,, T; \mathbf{x}_3\,, t_3)\, \gamma^i K_+(\mathbf{x}_3\,, t_3; \mathbf{x}, t)\, \not{B} K_+(\mathbf{x}, t; \mathbf{x}_4\,, t_4)$$
$$\cdot\, \gamma_i \mathscr{U}_0 \delta_+(q_{34}^2)\, e^{-imt_4}\, d^3\mathbf{x}_3\, d^3\mathbf{x}_4\, d^3\mathbf{x}\, dt_3\, dt_4\, dt,$$

$$\text{II. } (-ie^2)(-ie) \int \cdots \int K_+(\mathbf{x}_2\,, T; \mathbf{x}_3\,, t_3)\, \gamma^i K_+(\mathbf{x}_3\,, t_3\,; \mathbf{x}_4\,, t_4)\, \gamma_i K_+(\mathbf{x}_4\,, t_4\,; \mathbf{x}, t)$$
$$\cdot\, \not{B} \mathscr{U}_0 \delta_+(q_{34}^2)\, e^{-imt}\, d^3\mathbf{x}_3\, d^3\mathbf{x}_4\, d^3\mathbf{x}\, dt_3\, dt_4\, dt,$$

$$\text{III. } (-ie^2)(-ie) \int \cdots \int K_+(\mathbf{x}_2\,, T; \mathbf{x}, t)\, \not{B} K_+(\mathbf{x}, t; \mathbf{x}_3\,, t_3)\, \gamma^i K_+(\mathbf{x}_3\,, t_3\,; \mathbf{x}_4\,, t_4)$$
$$\cdot\, \gamma_i \mathscr{U}_0 \delta_+(q_{34}^2)\, e^{-imt_4}\, d^3\mathbf{x}_3\, d^3\mathbf{x}_4\, d^3\mathbf{x}\, dt_3\, dt_4\, dt. \tag{A23}$$

Parts II and III of (A23) correspond to II and III of Fig. 4. In III the wavefunction is propagated to $(\mathbf{x}, t)$ under the influence of $\delta_+(q_{34}^2)$ only. This means that provided t is not within ϵ of o or T we have a factor

$$1 - \tfrac{1}{2}B - i\,\Delta m \tag{A24}$$

multiplying $\mathscr{U}_0 e^{-imt}$. The Δm term was identified in Section 5 as a change of mass and we shall not consider it further here. A similar term in $-\tfrac{1}{2}B$ appears in II. We now show that these $-\tfrac{1}{2}B$ terms from II and III are cancelled by a cutoff dependent term from I.

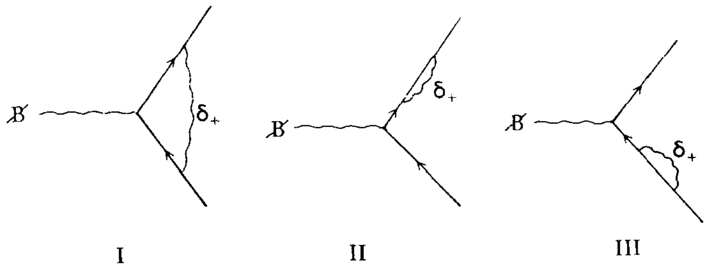

FIG. 4. The three diagrams which arise when the external potential and the response of the universe act once in the perturbation expansion. The infinite B-term in I is exactly cancelled by similar terms arising from II and III.

To begin with we write all the K_+ functions in I and the δ_+ function as three-dimensional Fourier integrals,

$$K_+(\mathbf{x}_2, T; \mathbf{x}_3, t_3) = \int \frac{\gamma_4 E'' - \boldsymbol{\gamma}\cdot\mathbf{p}'' + m}{2\,|\,E''\,|}\, e^{-i|E''(T-t_3)|}\, e^{i\mathbf{p}''\cdot(\mathbf{x}_2-\mathbf{x}_3)}\, \frac{d^3\mathbf{p}''}{(2\pi)^3},$$

$$K_+(\mathbf{x}_3, t_3; \mathbf{x}, t) = \int \frac{\gamma_4 E' - \boldsymbol{\gamma}\cdot\mathbf{p}' + m}{2\,|\,E'\,|}\, e^{-i|E'(t_3-t)|}\, e^{i\mathbf{p}'\cdot(\mathbf{x}_3-\mathbf{x})}\, \frac{d^3\mathbf{p}'}{(2\pi)^3},$$

$$K_+(\mathbf{x}, t; \mathbf{x}_4, t_4) = \int \frac{\gamma_4 E - \boldsymbol{\gamma}\cdot\mathbf{p} + m}{2\,|\,E\,|}\, e^{-i|E(t-t_4)|}\, e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}_4)}\, \frac{d^3\mathbf{p}}{(2\pi)^3},$$

$$\delta_+(q_{34}^2) = 4\pi i \int \frac{1}{2K}\, e^{-iK|t_3-t_4|+i\mathbf{k}\cdot(\mathbf{x}_3-\mathbf{x}_4)}\, \frac{d^3\mathbf{p}}{(2\pi)^3}, \tag{A25}$$

where

$$E''^2 = \mathbf{p}''^2 + m^2, \quad E'^2 = \mathbf{p}'^2 + m^2, \quad E^2 = \mathbf{p}^2 + m^2, \quad K = |\,\mathbf{k}\,|. \tag{A26}$$

Substituting (A25) in I and carrying out the integrations with respect to $d^3\mathbf{x}_3$, $d^3\mathbf{x}_4$, $d^3\mathbf{x}$ gives delta functions that require conservation of 3-momenta,

$$-\mathbf{p}'' + \mathbf{k} + \mathbf{p}' = 0, \qquad \mathbf{p}' - \mathbf{p} = 0, \qquad \mathbf{k} + \mathbf{p} = 0. \tag{A27}$$

Integrations with respect to $d^3\mathbf{p}''$, $d^3\mathbf{p}'$, $d^3\mathbf{p}$ lead to

$$(-ie)\frac{\pi e^2}{4}\int\cdots\int\frac{\gamma_4 E''+m}{|E''|}\gamma^i\frac{\gamma_4 E'+\boldsymbol{\gamma}\cdot\mathbf{k}+m}{|E'|}\not{B}\frac{\gamma_4 E+\boldsymbol{\gamma}\cdot\mathbf{k}+m}{|E|}\gamma_i\mathscr{U}_0$$
$$\cdot\exp\{-i\,|E''(T-t_3)|-i\,|E'(t_3-t)|-i\,|E(t-t_4)|-ik\,|t_3-t_4|-imt_4\}$$
$$\cdot\frac{1}{K}\frac{d^3\mathbf{k}}{(2\pi)^3}\,dt_3\,dt_4\,dt, \qquad \text{(A28)}$$

where

$$|E''|=m, \qquad |E'|=|E|=\sqrt{(K^2+m^2)}\,. \qquad \text{(A29)}$$

Significant contributions to (A28) arise only for the time orderings

$$T\geqslant t_3>t>t_4\geqslant 0, \qquad T\geqslant t_4>t>t_3\geqslant 0. \qquad \text{(A30)}$$

In both these cases t disappears from the exponential factor in (A28). This means that for a fixed relationship of t_3, t_4 to t there is a systematic contribution as t varies. Other time orderings give no such systematic summation. The contributions from (A30) will be denoted by J_1, J_2 respectively.

Consider J_1 first. We have

$$J_1=(-ie)\frac{\pi e^2}{4}\int\cdots\int(\gamma_4+1)\,\gamma^i$$
$$\times\,(\gamma_4\sqrt{K^2+m^2}+\boldsymbol{\gamma}\cdot\mathbf{k}+m)\,\not{B}(\gamma_4\sqrt{K^2+m^2}+\boldsymbol{\gamma}\cdot\mathbf{k}+m)$$
$$\times\,\gamma_i U_0\,\frac{1}{K(K^2+m^2)}\,e^{-imT+i(-\sqrt{K^2+m^2}-K+m)(t_3-t_4)}\,\frac{d^3\mathbf{k}}{(2\pi)^3}\,dt_3\,dt_4\,dt. \qquad \text{(A31)}$$

We simplify the matrix part of the integrand, retaining only those terms that can lead to a cutoff dependent factor. And we consider the integrand for $K\gg m$ using symmetrical integration to remove terms linear in the components of $\mathbf{k}$. We find

$$\gamma^i(\gamma_4\sqrt{K^2+m^2}+\boldsymbol{\gamma}\cdot\mathbf{k}+m)\,\not{B}(\gamma_4\sqrt{K^2+m^2}+\boldsymbol{\gamma}\cdot\mathbf{k}+m)\,\gamma_i$$
$$\sim -2K^2\gamma_4\not{B}\gamma_4-2(\boldsymbol{\gamma}\cdot\mathbf{k})\,\not{B}(\boldsymbol{\gamma}\cdot\mathbf{k}), \qquad \text{(A32)}$$

the remaining terms not being cutoff dependent. The average value of $k_\mu k_\nu$ ($\mu,\nu=1,2,3$) in symmetrical integration is $-\frac{1}{3}K^2\eta_{\mu\nu}$. Hence the average value of $(\boldsymbol{\gamma}\cdot\mathbf{k})\,\not{B}(\boldsymbol{\gamma}\cdot\mathbf{k})$ is

$$-\tfrac{1}{3}\gamma^\mu\not{B}\gamma^\nu\eta_{\mu\nu}K^2=\not{B}K^2-\tfrac{2}{3}B^\mu\gamma_\mu K^2. \qquad \text{(A33)}$$

The second term of (A33) gives no contribution to J_1 because

$$(\gamma_4 + 1)\,\gamma_\mu \mathscr{U}_0 = (1 + \gamma_4)\,\gamma_4\gamma_\mu \mathscr{U}_0 = -(1 + \gamma_4)\,\gamma_\mu\gamma_4 \mathscr{U}_0 = -(\gamma_4 + 1)\,\gamma_\mu \mathscr{U}_0 = 0 \tag{A34}$$

since $\gamma_4 \mathscr{U}_0 = \mathscr{U}_0$. The cutoff dependent part of J_1 is therefore

$$\begin{aligned} J_1 &\sim (-ie)\,\frac{\pi e^2}{4} \int \cdots \int (\gamma_4 + 1)(-4K^2)\,\not{B}\mathscr{U}_0\,\frac{1}{K^3}\,\frac{4\pi K^2\,dK}{(2\pi)^3} \\ &\times e^{-imT+i(m-2K)(t_3-t_4)}\,dt_3\,dt_4\,dt \\ &\sim -(-ie)\,\frac{e^2}{2\pi} \int \cdots \int (\gamma_4 + 1)\,\not{B}\mathscr{U}_0 e^{-2iK(t_3-t_4)}\,K\,dK\,e^{-imT+im(t_3-t_4)} \\ &\times dt_3\,dt_4\,dt \\ &= -\sum_{i=1,2} \mathscr{U}_0^{(i)} e^{-imT} \cdot \frac{e^2}{\pi} \int \cdots \int \bar{u}_0^{(i)}(-ie\not{B})\,\mathscr{U}_0 e^{-2iK(t_3-t_4)+im(t_3-t_4)} \\ &\times K\,dK\,dt_3\,dt_4\,dt. \end{aligned} \tag{A35}$$

In the last step we have used

$$\sum_{i=1,2} \mathscr{U}_0^{(i)} \bar{\mathscr{U}}_0^{(i)} = \tfrac{1}{2}(1 + \gamma_4), \tag{A36}$$

where $\mathscr{U}_0^{(i)}$ are the two states of zero momentum and opposite spins. We now proceed as in Section 5,

$$\int_0^\infty K e^{-2iK(t_3-t_4)}\,dK = -\frac{1}{4(t_3 - t_4)^2}\,. \tag{A37}$$

The double integral with respect to t_3, t_4 leads to a cutoff dependent term as $t_3 \to t\,+$, $t_4 \to t\,-$. We require that t_3, t_4 cannot approach closer than ϵ, and obtain integrals of the form

$$\int_t^T \frac{e^{im(t_3-t_4)}}{(t_3 - t_4)^2}\,dt_3 \sim \frac{e^{im(t_3-t_4)}}{t - t_4}\,, \tag{A38}$$

$$\int_0^{t-\epsilon} \frac{e^{im(t-t_4)}}{t - t_4}\,dt_4 \sim -\ln(m\epsilon). \tag{A39}$$

This gives

$$J_1 \sim -\sum_{i=1,2} \mathscr{U}_0^{(i)} e^{-imT}\,\frac{e^2}{4\pi}\ln(m\epsilon) \int \bar{\mathscr{U}}_0^{(i)}(-ie\not{B})\,\mathscr{U}_0\,dt. \tag{A40}$$

This is a half of what is required to cancel the term in $-B = e^2 \ln(m\epsilon)/2\pi$ from the parts II and III of (A23). J_2 gives the other half, i.e., the second time ordering in (A30).

3. Finally we have to evaluate the integral

$$J = \int \{q^2 I_+'(q^2)\}^2 \, d^4q \tag{A41}$$

encountered in Section 6. Here q is the displacement $x_2 - x_1$ between points 2 and 1. The function I_+ can be considered in the form

$$I_+(2, 1) = \int \frac{1}{p^2 - m^2} e^{-ip\cdot q} \frac{d^4p}{(2\pi)^4}, \tag{A42}$$

as well as in the form $I_+(q^2)$. Evidently

$$\nabla\!\!\!/_2 I_+(2, 1) = \int \frac{-i p\!\!\!/}{p^2 - m^2} e^{-ip\cdot q} \frac{d^4p}{(2\pi)^4}. \tag{A43}$$

The left hand side of (A43) was written in Section 6 as $2q\!\!\!/ I_+'(q^2)$, so that

$$2q_k I_+'(q^2) = -\int \frac{ip_k}{p^2 - m^2} e^{-ip\cdot q} \frac{d^4p}{(2\pi)^4}, \tag{A44}$$

and

$$\begin{aligned} 2q^2 I_+'(q^2) &= \int \frac{-ip_k q^k}{p^2 - m^2} e^{-ip\cdot q} \frac{d^4p}{(2\pi)^4} \\ &= \left[\frac{d}{d\alpha} \int \frac{e^{-i\alpha p\cdot q}}{p^2 - m^2} \frac{d^4p}{(2\pi)^4}\right]_{\alpha=1} \\ &= \left[\frac{d}{d\alpha} \left\{\frac{1}{\alpha^4} \int \frac{e^{-ip\cdot q}}{(p/\alpha)^2 - m^2} \frac{d^4p}{(2\pi)^4}\right\}\right]_{\alpha=} \\ &= 2\int \frac{2m^2 - p^2}{(p^2 - m^2)^2} e^{-ip\cdot q} \frac{d^4p}{(2\pi)^4}. \end{aligned} \tag{A45}$$

Substituting (A45) in (A41) we have

$$J = \iiint \frac{(2m^2 - p^2)}{(p^2 - m^2)^2} \frac{(2m^2 - p'^2)}{(p'^2 - m^2)^2} e^{-i(p+p')\cdot q} \frac{d^4p}{(2\pi)^4} \frac{d^4p'}{(2\pi)^4} d^4q. \tag{A46}$$

Integrations with respect to $d^3\mathbf{q}$ and $d^3\mathbf{p}'/(2\pi)^3$ give $\mathbf{p} + \mathbf{p}' = 0$. Write

$$P = |\mathbf{p}|, \qquad \Delta = (P^2 + m^2)^{1/2}. \tag{A47}$$

Then

$$J = \iiiint \frac{(m^2 - p_0^2 + \Delta^2)}{(p_0^2 - \Delta^2)^2} \frac{(m^2 - p_0'^2 + \Delta^2)}{(p_0'^2 - \Delta^2)^2} e^{-i(p_0+p_0')q_0} \frac{dp_0}{2\pi} \frac{dp_0'}{2\pi} dq_0 \frac{4\pi P^2\, dP}{(2\pi)^3}. \tag{A48}$$

There are two cases to be considered, $q_0 > 0$ and $q_0 < 0$. The minimum value of $|q_0|$ is ϵ. The contributions from these two cases will be denoted by J_+, J_- respectively.

Consider the p_0 integration for the case $q_0 > 0$. The $I_+(2, 1)$ function carries the additional definition that the pole at $p_0 = \Delta$ is to be considered as slightly below the real axis in a complex p_0 plane, and the pole at $p_0 = -\Delta$ is slightly above the real axis. Taking a contour along the real axis and over the half of the infinite circle in the lower half plane, we require $-2\pi i$ (minus because of the sense of circulation) times the residue at $p_0 = \Delta$. It is easily seen that the residue $\sim e^{-i\Delta q_0}/2\Delta$, so that the p_0 integration gives for $P \gg m$,

$$\frac{i\pi}{\Delta} e^{-i\Delta q_0}. \tag{A49}$$

The p_0' integration also gives this result. Hence we have

$$J_+ = -\frac{1}{8\pi^2} \int_\epsilon^\infty \int_0^\infty \frac{1}{\Delta^2} e^{-2i\Delta q_0} P^2\, dP\, dq_0. \tag{A50}$$

Using Δ as the variable instead of P,

$$J_+ = -\frac{1}{8\pi^2} \int_\epsilon^\infty \int_m^\infty \frac{P}{\Delta} e^{-2i\Delta q_0}\, d\Delta\, dq_0. \tag{A51}$$

We are interested in (A51) as $P \to \infty$. Then $P/\Delta \to 1$, and

$$J_+ \sim -\frac{1}{8\pi^2} \int_\epsilon^\infty \int_m^\infty e^{-2i\Delta q_0}\, d\Delta\, dq_0. \tag{A52}$$

The terms omitted here are finite, and remain finite as $\epsilon \to 0$. Then

$$\begin{aligned} J_+ &\sim -\frac{1}{8\pi^2} \int_\epsilon^\infty e^{-2imq_0} \int_0^\infty e^{-2i\Delta q_0}\, d\Delta\, dq_0 \\ &= -\frac{1}{8\pi^2} \int_\epsilon^\infty \frac{e^{-2imq_0}}{2iq_0}\, dq_0. \end{aligned} \tag{A53}$$

Since $m\epsilon \ll 1$, $e^{-2imq_0} \approx 1$ near the cutoff, and we have

$$J_+ \sim -\frac{i}{16\pi^2} \ln(m\epsilon). \tag{A54}$$

Similarly

$$J_- \sim -\frac{i}{16\pi^2} \ln(m\epsilon) \tag{A55}$$

and

$$J \sim -\frac{i}{8\pi^2} \ln(m\epsilon) \tag{A56}$$

which is the result quoted in Section 6.

Note added in proof. It is well-known that an invariant finite cut-off procedure, such as we suggest in section 5, is not consistent with a strict conservation of probabilities. Finite radiation processes involving convergent integrals are cut-off at ϵ^{-1} instead of going to infinity with respect to frequency. This leads to a very small change of order $m\epsilon$ in the calculated probabilities. The question of conservation of probabilities arises in this order. Our point of view is that this question cannot be discussed within the framework of Minkowski space-time. The gravitational effect of an electron distorts space-time to the same order as the question at issue.

While the present paper was in course of publication a similar point of view has been published by A. Salam and J. Strathdee (*Lett. Nuovo Cimento.* **IV**, 101). In their paper Salam and Strathdee discuss this same issue at greater length and in more detail than we have done.

References

1. F. Hoyle and J. V. Narlikar, *Ann. Phys. (N.Y.)* **54** (1969), 207.
2. J. A. Wheeler and R. P. Feynman, *Rev. Mod. Phys.* **17** (1945), 157.
3. R. P. Feynman, *Phys. Rev.* **80** (1950), 440.
4. R. P. Feynman, *Phys. Rev.* **76** (1949), 749.
5. J. D. Bjorken and S. D. Drell, "Relativistic Quantum Fields." McGraw-Hill, New York, 1965.