The Radiation Theories of Tomonaga, Schwinger, and Feynman

F.J. Dyson

Institute for Advanced Study, Princeton, New Jersey

(Received October 6, 1948)

—— ⋄ ◇ ⋄ ——
—— ⋄ ◇ ⋄ ——

Abstract

A unified development of the subject of quantum electrodynamics is outlined, embodying the main features both of the Tomonaga-Schwinger and of the Feynman radiation theory. The theory is carried to a point further than that reached by these authors, in the discussion of higher order radiative reactions and vacuum polarization phenomena. However, the theory of these higher order processes is a program rather than a definitive theory, since no general proof of the convergence of these effects is attempted.

The chief results obtained are (a) a demonstration of the equivalence of the Feynman and Schwinger theories, and (b) a considerable simplification of the procedure involved in applying the Schwinger theory Co particular problems, the simplification being the greater the more complicated the problem.

1   Introduction

As a result of the recent and independent discoveries of Tomonaga,[1] Schwinger,[2] and Feynman[3] the subject of quantum electrodynamics has made two very notable advances. On the one hand, both the foundations and the applications of the theory have been simplified by being presented in a completely relativistic way; on the other, the divergence difficulties have been at least partially overcome. In the reports so far published, emphasis has naturally been placed on the second of these advances; the magnitude of the first has been somewhat obscured by the fact that the new methods have been applied to problems which were beyond the range of the older theories, so that the simplicity of the methods was hidden by the complexity of the problems. Furthermore, the theory of Feynman differs so profoundly in its formulation from that of Tomonaga and Schwinger, and so little of it has been published, that its particular advantages have not hitherto been available to users of the other formulations. The advantages of the Feynman theory are simplicity and ease of application, while those of Tomonaga-Schwinger are generally and theoretical completeness.

The present paper aims to show how the Schwinger theory can be applied to specific problems in such a way as to incorporate the ideas of Feynman. To make the paper reasonably self-contained it is necessary to outline the foundations of the theory, following the method of Tomonaga; but this paper is not intended as a substitute for the complete account of the theory shortly to be published by Schwinger. Here the emphasis will be on the application of the theory, and the major theoretical problems of gauge-invariance and of the divergencies will not be considered in detail. The main results of the paper will be general formulas from which the radiative reactions on the motions of electrons can be calculated, treating the radiation interaction as a small perturbation, to any desired order of approximation. These formulas will be expressed in Schwinger's notation, but are in substance identical with results given previously by Feynman. The contribution of the present paper is thus intended to be twofold: first, to simplify the Schwinger theory for the benefit of those using it for calculations, and second, to demonstrate

---

[1]Sin-itiro Tomonaga, Prog. Theoret. Phys. **1**, 27 (1946); Koba. Tati,and Tomonaga, Prog. Theoret. Phys. **2**, 101 198 (1947); S. Kanesawa and S. Tomonaga, Prog. Theoret. Phys. **3**, **1**, 101 (1948); S. Tomonaga, Phys. Rev. **74** 224 (1948).

[2]Julian Schwinger. Phys. Rev. **73**, 416 (1948); Phys. Rev, **74,** 1439 (1948). Several papers, giving a complete exposition of the theory, are in course of publication.

[3]R. P. Feynman, Rev. Mod. Phys. **20,** 367 (1948); Phys. Rev. **74**, 939, 1430 (1948); J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. **17**, 157 (1945). These articles describe early stages in the development of Feynman's theory, little of which is yet published.

the equivalence of the various theories within their common domain of applicability.[4]

2 Outline of Theoretical Foundations

Relativistic quantum mechanics is a special case on non-relativistic quantum mechanics, and it is convenient to use the usual non-relativistic terminology in order to make clear the relation between the mathematical theory and the results of physical measurements. In quantum electrodynamics the dynamical variables are the electromagnetic potentials $A_\mu(\mathbf{r})$ and the spinor electron-positron field $\psi_\alpha(\mathbf{r})$ each component of each field at each point $\mathbf{r}$ of space is a separate variable. Each dynamical variable is, in the Schrödinger representation of quantum mechanics, a time-independent operator operating on the state vectorΦ of the system. The nature of Φ (wave function or abstract vector) need not be specified; its essential property is that, given the Φ of a system at a particular time, the results of all measurements made on the system at that time are statistically determined. The variation of Φ with time is given by the Schrödinger equation

$$ih[\partial/\partial t]\Phi = \left\{ \int H(\mathbf{r})dr \right\} \Phi, \tag{1}$$

where $H(\mathbf{r})$ is the operator representing the total energy-density of the system at the point $\mathbf{r}$. The general solution of (1) is

$$\Phi(t) = \exp\left\{ [-it/h] \int H(\mathbf{r})dr \right\} \Phi_0, \tag{2}$$

with Φ_0 any constant state vector.

Now in a relativistic system, the most general kind of measurement is not the simultaneous measurement of field quantities at different points of space. It is also possible to measure independently field quantities at different points of space at different times, provided that the points of space-time at which the measurements are made lie outside each other's light cones,

---

[4]After this paper was written, the author was shown a letter, published in Progress of Theoretical Physics **3**, 205 (1948) by Z. Koba and G. Takeda. The letter is dated May 22, 1948, and briefly describes a method of treatment of radiative problems, similar to the method of this paper. Results of the application of the method to a calculation of the second-order radiative correction to the Klein-Nishina formula are stated. All the papers of Professor Tomonaga and his associates which have yet been published were completed before the end of 1946. The isolation of these Japanese workers has undoubtedly constituted a serious loss to theoretical physics.

so that the measurements do not interfere with each other. Thus the most comprehensive general type of measurement is a measurement of field quantities at each point $\mathbf{r}$ of space at a time $t(\mathbf{r})$, the locus of the points $(\mathbf{r}, t(\mathbf{r}))$ in space-time forming a 3-dimensional surface σ which is space-like (i.e., every pair of points on it is separated by a space-like interval). Such a measurement will be called "an observation of the system on σ." It is easy to see what the result of the measurement will be. At each point $\mathbf{r}'$ the field quantities will be measured for a state of the system with state vector $\Phi(t(\mathbf{r}'))$ given by (2). But all observable quantities at $\mathbf{r}'$ are operators which commute with the energy-density operator $H(\mathbf{r})$ at every point $\mathbf{r}$ different from $\mathbf{r}'$, and it is a general principle of quantum mechanics that if B is a unitary operator commuting with A, then for any state Φ the results of measurements of A are the same in the state Φ as in the state $B\Phi$. Therefore, the results of measurement of the field quantities at $\mathbf{r}'$ in the state $\Phi(t(\mathbf{r}'))$ are the same as if the state of the system were

$$\Phi(\sigma) = \exp\left\{-[i/h]\int t(\mathbf{r})H(\mathbf{r})dr\right\}\Phi_0, \tag{3}$$

which differs from $\Phi(t(\mathbf{r}'))$ only by a unitary factor commuting with these field quantities. The important fact is that the state vector $\Phi(\sigma)$ depends only on σ and not on $\mathbf{r}'$. The conclusion reached is that observations of a system on a give results which are completely determined by attributing to the system the state vector $\Phi(\sigma)$ given by (3).

The Tomonaga-Schwinger form of the Schrödinger equation-is a differential form of (3). Suppose the surface σ to be deformed slightly near the point $\mathbf{r}$ into the surface σ' the volume of space-time separating the two surfaces being V. Then the quotient

$$[\Phi(\sigma') - \Phi(\sigma)]/V$$

tends to a limit as $V \to 0$, which we denote by $\partial\Phi/\partial\sigma(\mathbf{r})$ and call the functional derivative of Φ with respect to σ at the point $\mathbf{r}$. From (3) it follows that

$$ihc[\partial\Phi/\partial\sigma(\mathbf{r})] = H(\mathbf{r})\Phi, \tag{4}$$

and (3) is, in fact, the general solution of (4).

The whole meaning of an equation such as (4) depends on the physical meaning which is attached to the statement "a system has a constant state vector Φ_0" In the present context, this statement means "results of measurements of field quantities at any given point of space are independent of

4

time." This statement is plainly non-relativistic, and so (4) is, in spite of appearances, a non-relativistic equation.

The simplest way to introduce a new state vector Ψ which shall be a relativistic invariant is to require that the statement "a system has a constant state vector Ψ" shall mean "a system consists of photons, electrons, and positrons, traveling freely through space without interaction or external disturbance." For this purpose, let

$$H(\mathbf{r}) = H_0(\mathbf{r}) + H_1(\mathbf{r}), \tag{5}$$

where H_0 is the energy-density of the free electromagnetic and electron fields, and H_1 is that of their interaction with each other and with any external disturbing forces that may be present. A system with constant Ψ is, then, one whose H_1 is identically zero; by (3) such a system corresponds to Φ of the form

$$\begin{aligned} \Phi(\sigma) =\ & T(\sigma)\Phi_0, \\ & T(\sigma) = \exp\left\{-[i/h]\int t(\mathbf{r})H_0(\mathbf{r})dr\right\}. \end{aligned} \tag{6}$$

It is therefore consistent to write generally

$$\Phi(\sigma) = T(\sigma)\Psi(\sigma), \tag{7}$$

thus defining the new state vector Ψ of any system in terms of the old Φ. The differential equation satisfied by Ψ is obtained from (4), (5), (6), and (7) in the form

$$ihc[\partial\Psi/\partial\sigma(\mathbf{r})] = (T(\sigma))^{-1}H_1(\mathbf{r})T(\sigma)\Psi. \tag{8}$$

Now if $q(\mathbf{r})$ is any time-independent field operator, the operator

$$q(x_0) = (T(\sigma))^{-1}q(\mathbf{r})T(\sigma)$$

is just the corresponding time-dependent operator as usually defined in quantum electrodynamics.[5] It is a function of the point x_0 of space-time whose coordinates are $(\mathbf{r}, ct(\mathbf{r}))$, but is the same for all surfaces σ passing through this point, by virtue of the commutation of $H_1(\mathbf{r})$ with $H_0(\mathbf{r}')$ for $\mathbf{r}' \neq \mathbf{r}$. Thus (8) may be written

$$ihc[\partial\Psi/\partial\sigma(x_0)] = H_1(x_0)\Psi, \tag{9}$$

where $H_1(x_0)$ is the time-dependent form of the energy-density of interaction of the two fields with each other and with external forces. The left side

---

[5]See for example, Gregor Wentzel, *Einführung in die Quantentheorie der Wellenfelder* (Franz Deuticke, Wien, (1943), pp. 18-26.

of (9) represents the degree of departure of the system from a system of freely traveling particles and is a relativistic invariant; $H_1(x_0)$ is also an invariant, and thus is avoided one of the most unsatisfactory features of the old theories, in which the invariant H_1 was added to the non-invariant H_0. Equation (9) is the starting point of the Tomonaga-Schwinger theory.

3   Introduction of Perturbation Theory

Equation (9) can be solved explicitly. For this purpose it is convenient to introduce a one-parameter family of space-like surfaces filling the whole of space-time, so that one and only one member $\sigma(x)$ of the family passes through any given point x. Let $\sigma_0, \sigma_1, \sigma_2, \ldots$ be a sequence of surfaces of the family, starting with σ_0 and proceeding in small steps steadily into the past. By

$$\int_{\sigma_1}^{\sigma_0} H_1(x)dx$$

is denoted the integral of $H_1(x)$ over the 4-dimensional volume between the surfaces σ_1 and σ_0 similarly, by

$$\int_{-\infty}^{\sigma_0} H_1(x)dx, \quad \int_{\sigma_0}^{\infty} H_1(x)dx$$

are denoted integrals over the whole volume to the past of σ_0 and to the future of σ_0 respectively. Consider the operator

$$
\begin{aligned}
U = U(\sigma_0) \quad &= \left(1 - [i/hc]\int_{\sigma_1}^{\sigma_0} H_1(x)dx\right) \\
&\times \left(1 - [i/hc]\int_{\sigma_2}^{\sigma_1} H_1(x)dx\right)\ldots,
\end{aligned}
\tag{10}
$$

the product continuing to infinity and the surfaces $\sigma_0, \sigma_1, \ldots$ being taken in the limit infinitely close together. U satisfies the differential equation

$$ihc[\partial U/\partial \sigma(x_0)] = H_1(x_0)U, \tag{11}$$

general solution of (9) is

$$\Psi(\sigma) = U(\sigma)\Psi_0, \tag{12}$$

6

with Ψ_0 any constant vector.

Expanding the product (10) in ascending powers of H_1 gives a series

$$
\begin{aligned}
U = 1 \quad &+ (-i/hc) \int_{-\infty}^{\sigma_0} H_1(x_1)dx_1 + (-i/hc)^2 \\
&\times \int_{-\infty}^{\sigma_0} dx_1 \int_{-\infty}^{\sigma(x_1)} H_1(x_1)H_1(x_2)dx_2 + \ldots..
\end{aligned}
\tag{13}
$$

Further, U is by (10) obviously unitary, and

$$
\begin{aligned}
U^{-1} = \tilde{U} \quad &= 1 + (i/hc) \int_{-\infty}^{\sigma_0} H_1(x_1)dx_1 + (i/hc)^2 \\
&\times \int_{-\infty}^{\sigma_0} dx_1 \int_{-\infty}^{\sigma(x_1)} H_1(x_2)H_1(x_1)dx_2 + \ldots.
\end{aligned}
\tag{14}
$$

It is not difficult to verify that U is a function of σ_0 alone and is independent of the family of surfaces of which σ_0 is one member. The use of a finite number of terms of the series (13) and (14), neglecting the higher terms, is the equivalent in the new theory of the use of perturbation theory in the older electrodynamics.

The operator $U(\infty)$ obtained from (10) by taking σ_0 in the infinite future, is a transformation operator transforming a state of the system in the infinite past (representing, say, converging streams of particles) into the same state in the infinite future (after the particles have interacted or been scattered into their final outgoing distribution). This operator has matrix elements corresponding only to real transitions of the system, i.e., transitions which conserve energy and momentum. It is identical with the Heisenberg S matrix.[6]

4 Elimination of the Radiation Interaction

In most of the problem of electrodynamics, the energy-density $H_1(x_0)$ divides into two parts —

$$
H_1(x_0) = H^i(x_0) + H^{e(x_0)},
\tag{15}
$$

$$
H^i(x_0) = -[1/c]J_\mu(x_0)A_\mu(x_0),
\tag{16}
$$

[6]Werner Heisenberg, Zeits. f. Physik **120**, 513 (1943), **120**, 673 (1943), and Zeits. f. Naturforschung **1**, 608 (1946).

the first part being the energy of interaction of the two fields with each other, and the second part the energy produced by external forces. It is usually not permissible to treat H^e as a small perturbation as was done in the last section. Instead, H^i alone is treated as a perturbation, the aim being to eliminate H^i but to leave H^e in its original place in the equation of motion of the system.

Operators $S(\sigma)$ and $5S(\infty)$ are defined by replacing H_1 by H^i in the definitions of $U(\sigma)$ and $U(\infty)$. Thus $S(\sigma)$ satisfies the equation

$$ihc[\partial S/\partial \sigma(x_0)] = H^i(x_0)S. \tag{17}$$

Suppose now a new type of state vector $\Omega(\sigma)$ to be introduced by the substitution

$$\Psi(\sigma) = S(\sigma)\Omega(\sigma). \tag{18}$$

By (9), (15), (17), and (18) the equation of motion for $\Omega(\sigma)$ is

$$ihc[\partial \Omega/\partial \sigma(x_0)] = (S(\sigma))^{-1}H^e(x_0)S(\sigma)\Omega. \tag{19}$$

The elimination of the radiation interaction is hereby achieved; only the question, "How is the new state vector $\Omega(\sigma)$ to be interpreted?" remains.

It is clear from (19) that a system with a constant Ω is a system of electrons, positrons, and photons, moving under the influence of their mutual interactions, but in the absence of external fields. In a system where two or more particles are actually present, their interactions alone will, in general, cause real transitions and scattering processes to occur. For such a system, it is rather "unphysical" to represent a state of motion including the effects of the interactions by a constant state vector; hence, for such a system the new representation has no simple interpretation. However, the most important systems are those in which only one particle is actually present, and its interaction with the vacuum fields gives rise only to virtual processes. In this case the particle, including the effects of all its interactions with the vacuum, appears to move as a free particle in the absence of external fields, and it is eminently reasonable to represent such a state of motion by a constant state vector. Therefore, it may be said that the operator,

$$H_T(x_0) = (S(\sigma))^{-1}H^e(x_))S(\sigma), \tag{20}$$

on the right of (19) represents the interaction of a physical particle with an external field, including radiative corrections. Equation (19) describes the extent to which the motion of a single physical particle deviates, in the

external field, from the motion represented by a constant state-vector, i.e., from the motion of an observed "free" particle.

If the system whose state vector is constantly Ω undergoes no real transitions with the passage of time, then the state vector Ω is called "steady." More precisely, Ω is steady if, and only if, it satisfies the equation

$$S(\infty)\Omega = \Omega. \tag{21}$$

As a general rule, one-particle states are steady and many-particle states unsteady. There are, however, two important qualifications to this rule.

First, the interaction (20) itself will almost always cause transitions from steady to unsteady states. For example, if the initial state consists of one electron in the field of a proton, H_T will have matrix elements for transitions of the electron to a new state with emission of a photon, and such transitions are important in practice. Therefore, although the interpretation of the theory is simpler for steady states, it is not possible to exclude unsteady states from consideration.

Second, if a one-particle state as hitherto denned is to be steady, the definition of $S(\sigma)$ must be modified. This is because $S(\infty)$ includes the effects of the electromagnetic self-energy of the electron, and this self-energy gives an expectation value to $S(\infty)$ which is different from unity (and indeed infinite) in a one-electron state, so that Eq. (21) cannot be satisfied. The mistake that has been made occurred in trying to represent the observed electron with its electromagnetic self-energy by a wave field with the same characteristic rest-mass as that of the "bare" electron. To correct the mistake, let δm denote the electromagnetic mass of the electron, i.e., the difference in rest-mass between an observed and a "bare" electron. Instead of (5), the division of the energy-density $H(\mathbf{r})$ should have taken the form

$$H(\mathbf{r}) = (H_0(\mathbf{r}) + \delta mc^2\psi^*(\mathbf{r})\beta\psi(\mathbf{r})) + (H_1(\mathbf{r}) - \delta mc^2\psi^*(\mathbf{r})\beta\psi(\mathbf{r})).$$

The first bracket on the right here represents the energy-density of the free electromagnetic and electron fields with the observed electron rest-mass, and should have been used instead of $H_0(\mathbf{r})$ in the definition (6) of $T(\sigma)$. Consequently, the second bracket should have been used instead of $H_1(\mathbf{r})$ in Eq. (8).

The definition of $S(\sigma)$ has therefore to be altered by replacing $H^i(x_0)$ by[7]

$$H^I(x_0) = H^i(x_0) + H^S(x_0) = H^i(x_0) - \delta mc^2\tilde{\psi}(x_0)\psi(x_0). \tag{22}$$

---

[7]Here Schwinger's notation $\tilde{\psi} = \psi^*\beta$ is used.

The value of δm can be adjusted so as to cancel out the self-energy effects in $S(\infty)$ (this is only a formal adjustment since the value is actually infinite), and then Eq. (21) will be valid for one-electron states. For the photon self-energy no such adjustment is needed since, as proved by Schwinger, the photon self-energy turns out to be identically zero.

The foregoing discussion of the self-energy problem is intentionally only a sketch, but it will be found to be sufficient for practical applications of the theory. A fuller discussion of the theoretical assumptions underlying this treatment of the problem will be given by Schwinger in his forthcoming papers. Moreover, it must be realized that the theory as a whole cannot be put into a finally satisfactory form so long as divergencies occur in it, however skilfully these divergencies arc circumvented; therefore, the present treatment should be regarded as justified by its success in applications rather than by its theoretical derivation.

The important results of the present paper up to this point are Eq. (19) and the interpretation of the state vector Ω. The state vector Ψ of a system can be interpreted as a wave function giving the probability amplitude of finding any particular set of occupation numbers for the various possible states of free electrons, positrons, and photons. The state vector Ω of a system with a given Ψ on a given surface σ is, crudely speaking, the Ψ which the system would have had in the infinite past if it had arrived at the given Ψ on a under the influence of the interaction $H^I(x_0)$ alone.

The definition of Ω being unsymmetrical between past and future, a new type of state vector Ω' can be defined by reversing the direction of time in the definition of Ω. Thus the Ω' of a system with a given Ψ on a given σ is the Ψ which the system would reach in the infinite future if it continued to move under the influence of $H^I(x_0)$ alone. More simply, Ω' can be defined by the equation

$$\Omega'(\sigma) = S(\infty)\Omega(\sigma). \tag{23}$$

Since $S(\infty)$ is a unitary operator independent of σ, the state vectors Ω and Ω' are really only the same vector in two different representations or coordinate systems. Moreover, for any steady state the two are identical by (21).

5 Fundamental Formulas of the Schwinger and Feynman Theories

The Schwinger theory works directly from Eqs. (19) and (20), the aim being to calculate the matrix elements of the "effective external potential

energy" H_T between states specified by their state vectors Ω. The states considered in practice always have Ω of some very simple kind, for example, Ω representing systems in which one or two free-particle states have occupation number one and the remaining free-particle states have occupation number zero. By analogy with (13), $S(\sigma_0)$ is given by

$$
\begin{aligned}
S(\sigma_0) = & \ 1 + (-i/hc) \int_{-\infty}^{\sigma_0} H^I(x_1)dx_1 + (-i/hc)^2 \\
& \times \int_{-\infty}^{\sigma_0} dx_1 \int_{-\infty}^{\sigma(x_1)} H^I(x_1)H^I(x_2)dx_2 + \ldots,
\end{aligned}
\tag{24}
$$

and $(S(\sigma_0))^{-1}$ by a corresponding expression analogous to (14). Substitution of these series into (20) gives at once

$$
\begin{aligned}
H_T(x_0) \ = & \ \sum_{n=0}^{\infty} (i/hc)^n \int_{-\infty}^{\sigma(x_0)} dx_1 \int_{-\infty}^{\sigma(x_1)} dx_2 \ldots \\
& \times \int_{-\infty}^{\sigma(x_{n-1})} dx_n \times [H^I(x_n), [\ldots, [H^I(x_2), [H^I(x_1), H^e(x_0)]] \ldots]].
\end{aligned}
\tag{25}
$$

The repeated commutators in this formula are characteristic of the Schwinger theory, and their evaluation gives rise to long and rather difficult analysis. Using the first three terms of the series, Schwinger was able to calculate the second-order radiative corrections to the equations of motion of an electron in an external field, and obtained satisfactory agreement with experimental results. In this paper the development of the Schwinger theory will be carried no further; in principle the radiative corrections to the equations of motion of electrons could be calculated to any desired order of approximation from formula (25).

In the Feynman theory the basic principle is to preserve symmetry between past and future. Therefore, the matrix elements of the operator H_T are evaluated in a "mixed representation;" the matrix elements are calculated between an initial state specified by its state vector Ω_1 and a final state specified by its state vector Ω_2'. The matrix element of H_T between two such states in the Schwinger representation is

$$
\Omega_2^* H_T \Omega_1 = \Omega_2'^* S(\infty) H_T \Omega_1,
\tag{26}
$$

and therefore the operator which replaces H_T in the mixed representation is

$$
H_F(x_0) = S(\infty) H_T(x_0) = S(\infty)(S(\sigma))^{-1} H^e(x_0) S(\sigma).
\tag{27}
$$

11

Going back to the original product definition of $S(\sigma)$ analogous to (10), it is clear that $S(\infty) \times (S(\sigma))^{-1}$ is simply the operator obtained from $S(\sigma)$ by interchanging past and future. Thus,

$$
\begin{aligned}
R(\sigma) \quad &= S(\infty)(S(\sigma))^{-1} = 1 + (-i/hc) \\
&\times \int_{\sigma}^{\infty} H^I(x_1)dx_1 + (-i/hc)^2 \int_{\sigma}^{\infty} dx_1 \\
&\times \int_{\sigma(x_1)}^{\infty} H^I(x_2)H^I(x_1)dx_2 + \dots .
\end{aligned}
\tag{28}
$$

The physical meaning of a mixed representation of this type is not at all recondite. In fact, a mixed representation is normally used to describe such a process as bremsstrahlung of an electron in the field of a nucleus when the Born approximation is not valid; the process of bremsstrahlung is a radiative transition of the electron from a state described by a Coulomb wave function, with a plane ingoing and a spherical outgoing wave, to a state described by a Coulomb wave function with a spherical ingoing and a plane outgoing wave. The initial and final states here belong to different orthogonal systems of wave functions, and so the transition matrix elements are calculated in a mixed representation. In the Feynman theory the situation is analogous; only the roles of the radiation interaction and the external (or Coulomb) field are interchanged; the radiation interaction is used instead of the Coulomb field to modify the state vectors (wave functions) of the initial and final states, and the external field instead of the radiation interaction causes transitions between these state vectors.

In the Feynman theory there is an additional simplification. For if matrix elements are being calculated between two states, either of which is steady (and this includes all cases so far considered), the mixed representation reduces to an ordinary representation. This occurs, for example, in treating a one-particle problem such as the radiative correction to the equations of motion of an electron in an external field; the operator $H_F(x_0)$ although in general it is not even Hermitian, can in this case be considered as an effective external potential energy acting on the particle, in the ordinary sense of the words.

This section will be concluded with the derivation of the fundamental formula (31) of the Feynman theory, which is the analog of formula (25) of the Schwinger theory. If

$$
F_1(x_1), \ \dots, \ F_n(x_n)
$$

are any operators defined, respectively, at the points $x_1, \dots x_n$ of space-time,

then
$$P(F_1(x_1), \ \ldots, \ F_n(x_n)) \tag{29}$$

will denote the product of these operators, taken in the order, reading from right to left, in which the surfaces $\sigma(x_1),\ldots,\sigma(x_n)$ occur in time. In most applications of this notation $F_i(x_i)$ will commute with $F_j(x_j)$ so long as x_i and x_j are outside each other's light cones; when this is the case, it is easy to see that (29) is a function of the points $x_1,\ldots,x_n$ only and is independent of the surfaces $\sigma(x_i)$. Consider now the Integral

$$I_n = \int\limits_{-\infty}^{\infty} dx_1 \ \ldots \int\limits_{-\infty}^{\infty} dx_n P(H^e(x_0),$$
$$H^I(x_1),\ldots,H^I(x_n)).$$

Since the integrand is a symmetrical function of the points $x_1,\ldots,x_n$, the value of the integral is just $n!$ times the integral obtained by restricting the integration to sets of points $x_1,\ldots,x_n$ for which $\sigma(x_i)$ occurs after $\sigma(x_{i+1})$ for each i.

The restricted integral can then be further divided into $(n+1)$ parts, the j'th part being the integral over those sets of points with the property that $\sigma(x_0)$ lies between $\sigma(x_{j-1})$ and $\sigma(x_j)$ (with obvious modifications for $j=1$ and $j=n+1$). Therefore,

$$I_n = n! \sum_{j=1}^{n+1} \int\limits_{-\infty}^{\sigma(x_0)} dx_j \ldots \int\limits_{-\infty}^{\sigma(x_{n-1})} dx_n$$

$$\times \int\limits_{\sigma(x_0)}^{\infty} dx_{j-1} \ldots \int\limits_{\sigma(x_2)}^{\infty} dx_1 \times H^I(x_1) \ldots \tag{30}$$

$$H^I(x_{j-1})H^e(x_0)H^I(x_j)\ldots H^I(x_n).$$

Now if the series (24) and (28) are substituted into (27), sums of integrals appear which are precisely of the form (30). Hence finally

$$H_P(x_0) \ = \sum_{n=0}^{\infty} (-i/hc)^n [1/n!] I_n$$

$$= \sum_{n=0}^{\infty} (-i/hc)^n [1/n!] \int\limits_{-\infty}^{\infty} dx_1 \ldots \int\limits_{-\infty}^{\infty} dx_n \tag{31}$$

$$\times P(H^e(x_0), H^I(x_1),\ldots,H^I(x_n)).$$

13

By this formula the notation $H_F(x_0)$ is justified, for this operator now appears as a function of the point x_0 alone and not of the surface σ. The further development of the Feynman theory is mainly concerned with the calculation of matrix elements of (31) between various initial and final states.

As a special case of (31) obtained by replacing H^e by the unit matrix in (27),

$$S(\infty) = \sum_{n=0}^{\infty} (i/hc)^n [1/n!] \int_{-\infty}^{\infty} dx_1 \ldots \int_{-\infty}^{\infty} dx_n \\ \times P(H^I(x_1), \ldots, H^I(x_n)). \tag{32}$$

6 Calculation of Matrix Elements

In this section the application of the foregoing theory to a general class of problems will be explained. The ultimate aim is to obtain a set of rules by which the matrix element of the operator (31) between two given states may be written down in a form suitable for numerical evaluation, immediately and automatically. The fact that such a set of rules exists is the basis of the Feynman radiation theory; the derivation in this section of the same rules from what is fundamentally the Tomonaga-Schwinger theory constitutes the proof of equivalence of the two theories.

To avoid excessive complication, the type of matrix element considered will be restricted in two ways. First, it will be assumed that the external potential energy is

$$H^e(x_0) = -[1/c]j_\mu(x_0)A_\mu^e(x_0), \tag{33}$$

that is to say, the interaction energy of the electron-positron field with electromagnetic potentials $A_\mu^e(x_0)$ which are given numerical functions of space and time. Second, matrix elements will be considered only for transitions from a state A, in which just one electron and no positron or photon is present, to another state B of the same character. These restrictions are not essential to the theory, and are introduced only for convenience, in order to illustrate clearly the principles involved.

The electron-positron field operator may be written

$$\psi_\alpha(x) = \sum_u \phi_{u\alpha}(x)a_u, \tag{34}$$

where the $\phi_{u\alpha}(x)$ are spinor wave functions of free electrons and positrons, and the a_u are annihilation operators of electrons and creation operators of

14

positrons. Similarly, the adjoint operator

$$\tilde{\psi}_\alpha(x) = \sum_u \tilde{\phi}_{u\alpha}(x)\tilde{a}_u, \tag{35}$$

where $\tilde{a}_u$ are annihilation operators of positrons and creation operators of electrons. The electromagnetic field operator is

$$A_\mu(x) = \sum_v (A_{v\mu}(x)b_v + A_{v\mu}^*(x)\tilde{b}_v), \tag{36}$$

where b_v and $\tilde{b}_v$ are photon annihilation and creation operators, respectively. The charge-current 4-vector of the electron field is

$$j_\mu(x) = iec\tilde{\psi}(x)\gamma_\mu\psi(x); \tag{37}$$

strictly speaking, this expression ought to be antisymmetrized to the form[8]

$$j_\mu(x) = \frac{1}{2}iec\left\{\tilde{\psi}_\alpha(x)\psi_\beta(x) - \psi_\beta(x)\tilde{\psi}_\alpha(x)\right\}(\gamma_\mu)_{\alpha\beta}, \tag{38}$$

but it will be seen later that this is not necessary in the present theory.

Consider the product P occurring in the n'th integral of (31); let it be denoted by P_n. From (16), (22), (33), and (37) it is seen that P_n is a sum of products of $(n+1)$ operators ψ_α, $(n+1)$ operators $\tilde{\psi}_\alpha$ and not more than n operators A_μ, multiplied by various numerical factors. By Q_n may be denoted a typical product of factors $\psi_\alpha, \tilde{\psi}_\alpha$, and A_μ, not summed over the indices such as α and μ, so that P_n is a sum of terms such as Q_n. Then Q_n will be of the form (indices omitted)

$$\begin{aligned}Q_n = {}&\tilde{\psi}(x_{i_0})\psi(x_{i_0})\tilde{\psi}(x_{i_1})\psi(x_{i_1})\ldots\tilde{\psi}(x_{i_n})\psi(x_{i_n})\\ &\times A(x_{j_1})\ldots A(x_{j_m}),\end{aligned} \tag{39}$$

where $i_0, i_1, \ldots, i_n$ is some permutation of the integers $0, 1, \ldots, n$, and $j_1, \ldots j_m$ are some, but not necessarily all, of the integers $1, \ldots, n$ in some order. Since none of the operators $\tilde{\psi}$ and ψ commute with each other, it is especially important to preserve the order of these factors. Each factor of Q_n is a sum of creation and annihilation operators by virtue of (34), (35), and (36), and so Q_n itself is a sum of products of creation and annihilation operators.

Now consider under what conditions a product of creation and annihilation operators can give a non-zero matrix element for the transition $A \to B$. Clearly, one of the annihilation operators must annihilate the electron in

---

[8]See Wolfgang Paul;, Rev. Mod. Phys. **13,** 203 (1941), Eq. (96), p. 224.

state A, one of the creation operators must create the electron in state B, and the remaining operators must be divisible into pairs, the members of each pair respectively creating and annihilating the same particle. Creation and annihilation operators referring to different particles always commute or anticommute (the former if at least one is a photon operator, the latter if both are electron-positron operators). Therefore, if the two single operators and the various pairs of operators in the product all refer to different particles, the order of factors in the product can be altered so as to bring together the two single operators and the two members of each pair, without changing the value of the product except for a change of sign if the permutation made in the order of the electron and positron operators is odd. In the case when some of the single operators and pairs of operators refer to the same particle, it is not hard to verify that the same change in order of factors can be made, provided it is remembered that the division of the operators into pairs is no longer unique, and the change of order is to be made for each possible division into pairs and the results added together.

It follows from the above considerations that the matrix element of Q_n for the transition $A \to B$ is a sum of contributions, each contribution arising from a specific way of dividing the factors of Q_n into two single factors and pairs. A typical contribution of this kind will be denoted by M. The two factors of a pair must involve a creation and an annihilation operator for the same particle, and so must be either one $\tilde{\psi}$ and one ψ or two A; the two single factors must be one $\tilde{\psi}$ and one ψ. The term M is thus specified by fixing an integer k, and a permutation $r_0, r_1, \ldots, r_n$ of the integers $0, 1, \ldots, n$, and a division $(s_1, t_1), (s_2, t_2)), \ldots, (s_h, t_h)$ of the integers $j_1, \ldots, j_m$ into pairs; clearly $m = 2h$ has to be an even number; the term M is obtained by choosing for single factors $\tilde{\psi}(x_k)$ and $\psi(x_{r_k})$, and for associated pairs of factors $(\tilde{\psi}(x_i), \psi(x_{r_i}))$ for $i = 0, 1, \ldots, k-1, k+1, \ldots, n$ and $(A((x_{s_i}), A(x_{t_i}))$ for $i = 1, \ldots, h$. In evaluating the term M, the order of factors in Q_n is first to be permuted so as to bring together the two single factors and the two members of each pair, but without altering the order of factors within each pair; the result of this process is easily seen to be

$$
\begin{aligned}
Q'_n \quad &= \epsilon P(\tilde{\psi}(x_0), \psi(x_{r_0})) \ldots P(\tilde{\psi}(x_n), \psi(x_{r_n})) \\
&\times P(A(x_{s_1}), A(x_{t_1})) \ldots P(A(x_{s_h}), A(x_{t_h})),
\end{aligned} \tag{40}
$$

a factor ϵ being inserted which takes the value ± 1 according to whether the permutation of $\tilde{\psi}$ and ψ factors between (39) and (40) is even or odd. Then in (40) each product of two associated factors (but not the two single factors) is to be independently replaced by the sum of its matrix elements

16

for processes involving the successive creation and annihilation of the same particle.

Given a bilinear operator such as $A_\mu(x)A_\nu(y)$, the sum of its matrix elements for processes involving the successive creation and annihilation of the same particle is just what is usually called the "vacuum expectation value" of the operator, and has been calculated by Schwinger. This quantity is, in fact (note that Heaviside units are being used)

$$\langle A_\mu(x)A_\nu(y)\rangle_0 = \frac{1}{2}hc\delta_{\mu\nu}\left\{D^{(1)} + iD\right\}(x-y),$$

where $D^{(1)}$ and D are Schwinger's invariant D functions. The definitions of these functions will not be given here, because it turns out that the vacuum expectation value of $P(A_\mu(x), A_\nu(y))$ takes an even simpler form. Namely,

$$\langle P(A_\mu(x), A_\nu(y))\rangle_0 = \frac{1}{2}hc\delta_{\mu\nu}D_F(x-y), \tag{41}$$

where D_F is the type of D function introduced by Feynman. $D_F(x)$ is an even function of x, with the integral expansion

$$D_F(x) = -[i/2\pi^2]\int_0^\infty \exp[i\alpha x^2]d\alpha, \tag{42}$$

where x^2 denotes the square of the invariant length of the 4-vector x. In a similar way it follows from Schwinger's results that

$$\langle P(\tilde{\psi}_\alpha(x), \psi_\beta(y))\rangle_0 = \frac{1}{2}\eta(x,y)S_{F\beta\alpha}(x-y), \tag{43}$$

where

$$S_{F\beta\alpha}(x) = -(\gamma_\mu(\partial/\partial x_\mu) + \kappa_0)_{\beta\alpha}\Delta_F(x), \tag{44}$$

κ_0 is the reciprocal Compton wave-length of the electron, $\eta(x,y)$ is -1 or $+1$ according as $\sigma(x)$ is earlier or later than $\sigma(y)$ in time, and Δ_F is a function with the integral expansion

$$\Delta_F(x) = -[i/2\pi^2]\int_0^\infty \exp[i\alpha x^2 - i\kappa_0^2/4\alpha]d\alpha. \tag{45}$$

Substituting from (41) and (44) into (40), the matrix element M takes the form (still omitting the indices of the factors $\tilde{\psi}, \psi$ and A of Q_n)

$$\begin{aligned} M = \epsilon \quad &\Pi_{i\neq k}\left(\frac{1}{2}\eta(x_i, x_{r_i})S_F(x_i - x_{r_i})\right) \\ &\times\Pi_j\left(\frac{1}{2}hcD_F(x_{s_j} - x_{t_j})\right)P(\tilde{\psi}(x_k), \psi(x_{r_k})). \end{aligned} \tag{46}$$

17

The single factors $\tilde{\psi}(x_k)$ and $\psi(x_{r_k})$ are conveniently left in the form of operators, since the matrix elements of these operators for effecting the transition $A \to B$ depend on the wave functions of the electron in the states A and B. Moreover, the order of the factors $\tilde{\psi}(x_k)$ and $\psi(x_{r_k})$ is immaterial since they anticommute with each other; hence it is permissible to write

$$P(\tilde{\psi}(x_k), \psi(x_{r_k})) = \eta(x_k, x_{r_k})\tilde{\psi}(x_k)\psi(x_{r_k}).$$

Therefore (46) may be rewritten

$$M = \epsilon'\Pi_{i \neq k}\left(\tfrac{1}{2}S_F(x_i - x_{r_i})\right)\Pi_j\left(\tfrac{1}{2}hcD_F(x_{s_j} - x_{t_j})\right) \times \tilde{\psi}(x_k)\psi(x_{r_k}), \tag{47}$$

with

$$\epsilon' = \epsilon\Pi_i\eta(x_i, x_{r_i}). \tag{48}$$

Now the product in (48) is $(-1)^p$, where p is the number of occasions in the expression (40) on which the ψ of a P bracket occurs to the left of the $\tilde{\psi}$. Referring back to the definition of ϵ after Eq. (40), it follows that ϵ' takes the value $+1$ or -1 according to whether the permutation of $\tilde{\psi}$ and ψ factors between (39) and the expression

$$\tilde{\psi}(x_0)\psi(x_{r_0})\ldots\tilde{\psi}(x_n)\psi(x_{r_n}) \tag{49}$$

is even or odd. But (39) can be derived by an even permutation from the expression

$$\tilde{\psi}(x_0)\psi(x_0)\ldots\tilde{\psi}(x_n)\psi(x_n), \tag{50}$$

and the permutation of factors between (49) and (50) is even or odd according to whether the permutation $r_0, \ldots, r_n$ of the integers $0, \ldots, n$ is even or odd. Hence, finally, ϵ' in (47) is $+1$ or -1 according to whether the permutation $r_0, \ldots, r_n$ is even or odd. It is important that ϵ' depends only on the type of matrix element M considered, and not on the points $x_0, \ldots, x_n$; therefore, it can be taken outside the integrals in (31).

One result of the foregoing analysis is to justify the use of (37), instead of the more correct (38), for the charge-current operator occurring in H^e and H^i. For it has been shown that in each matrix element such as M the factors $\tilde{\psi}$ and ψ in (38) can be freely permuted, so that (38) can be replaced by (37), except in the case when the two factors form an associated pair. In the exceptional case, M contains as a factor the vacuum expectation value of the operator $J_\mu(x_i)$ at some point x_i; this expectation value is zero according to the correct formula (38), though it would be infinite according to (37); thus the matrix elements in the exceptional case are always zero.

The conclusion is that only those matrix elements are to be calculated for which the integer r_i differs from i for every $i \neq k$, and in these elements the use of formula (37) is correct.

To write down the matrix elements of (31) for the transition $A \rightarrow B$, it is only necessary to take all the products Q_n replace each by the sum of the corresponding matrix elements M given by (47), reassemble the terms into the form of the P_n from which they were derived, and finally substitute back into the series (31). The problem of calculating the matrix elements of (31) is thus in principle solved. However, in the following section it will be shown how this solution-in-principle can be reduced to a much simpler and more practical procedure.

7 GRAPHICAL REPRESENTATION OF MATRIX ELEMENTS

Let an integer n and a product P_n occurring in (31) be temporarily fixed. The points $x_0, x_1, \ldots, x_n$ may be represented by $(n+1)$ points drawn on a piece of paper. A type of matrix element M as described in the last section will then be represented graphically as follows. For each associated pair of factors $(\tilde{\psi}(x_i), \psi(x_{r_i})$ with $i \neq k$, draw a line with a direction marked in it from the point x_i to the point x_{r_i}. For the single factors $\tilde{\psi}(x_k), \psi(x_{r_k})$, draw directed lines leading out from x_k to the edge of the diagram, and in from the edge of the diagram to x_{r_k}. For each pair of factors $(A(x_{s_i}), A(x_{t_i}))$, draw an undirected line joining the points x_{s_i} and x_{t_i}. The complete set of points and lines will be called the "graph" of M; clearly there is a one-to-one correspondence between types of matrix element and graphs, and the exclusion of matrix elements with $r_i = i$ for $i \neq k$ corresponds to the exclusion of graphs with lines joining a point to itself. The directed lines in a graph will be called "electron lines," the undirected lines "photon lines."

Through each point of a graph pass two electron lines, and therefore the electron lines together form one open polygon containing the vertices x_k and x_{r_k} and possibly a number of closed polygons as well. The closed polygons will be called "closed loops," and their number denoted by l. Now the permutation $r_0, \ldots, r_n$ of the integers $0, \ldots, n$ is clearly composed of $(l+1)$ separate cyclic permutations. A cyclic permutation is even or odd according to whether the number of elements in it is odd or even. Hence the parity of the permutation $r_0, \ldots, r_n$ is the parity of the number of even-number cycles contained in it. But the parity of the number of odd-number cycles in it is obviously the same as the parity of the total number $(n+1)$ of

elements. The total number of cycles being $(l+1)$, the parity of the number of even-number cycles is $(l-n)$. Since it was seen earlier that the ϵ' of Eq. (47) is determined just by the parity of the permutation $r_0, \ldots, r_n$, the above argument yields the simple formula

$$\epsilon' = (-1)^{l-n}. \tag{51}$$

This formula is one result of the present theory which can be much more easily obtained by intuitive considerations of the sort used by Feynman.

In Feynman's theory the graph corresponding to a particular matrix element is regarded, not merely as an aid to calculation, but as a picture of the physical process which gives rise to that matrix element. For example, an electron line joining x_1 to x_2 represents the possible creation of an electron at x_1 and its annihilation at x_2, together with the possible creation of a positron at x_2 and its annihilation at x_1. This interpretation of a graph is obviously consistent with the methods, and in Feynman's hands has been used as the basis for the derivation of most of the results, of the present paper. For reasons of space, these ideas of Feynman will hot be discussed in further detail here.

To the product P_n correspond a finite number of graphs, one of which may be denoted by G all possible G can be enumerated without difficulty for moderate values of n. To each G corresponds a contribution $C(G)$ to the matrix element of (31) which is being evaluated.

It may happen that the graph G is disconnected, so that it can be divided into subgraphs, each of which is connected, with no line joining a point of one subgraph to a point of another. In such a case it is clear from (47) that $C(G)$ is the product of factors derived from each subgraph separately. The subgraph G_1 containing the point x_0 is called the "essential part" of G, the remainder G_2 the "inessential part." There are now two cases to be considered, according to whether the points x_k and x_{r_k} lie in G_2 or in G_1 (they must clearly both lie in the same subgraph). In the first case, the factor $C(G_2)$ of $C(G)$ can be seen by a comparison of (31) and (32) to be a contribution to the matrix element of the operator $S(\infty)$ for the transition $A \to B$. Now letting G vary over all possible graphs with the same G_1 and different G_2, the sum of the contributions of all such G is a constant $C(G_1)$ multiplied by the total matrix element of $S(\infty)$ for the transition $A \to B$. But for one-particle states the operator $S(\infty)$ is by (21) equivalent to the identity operator and gives, accordingly, a zero matrix element for the transition $A \to B$. Consequently, the disconnected G for which x_k and x_{r_k} lie in G_2 give zero contribution to the matrix element of (31), and can be

omitted from further consideration. When x_k and x_{r_k}, lie in G_1, again the $C(G)$ may be summed over all G consisting of the given G_1 and all possible G_2; but this time the connected graph G_1 itself is to be included in the sum. The sum of all the $C(G)$ in this case turns out to be just $C(G_1)$ multiplied by the expectation value in the vacuum of the operator $S(\infty)$. But the vacuum state, being a steady state, satisfies (21), and so the expectation value in question is equal to unity. Therefore the sum of the $C(G)$ reduces to the single term $C(G_1)$, and again the disconnected graphs may be omitted from consideration.

The elimination of disconnected graphs is, from a physical point of view, somewhat trivial, since these graphs arise merely from the fact that meaningful physical processes proceed simultaneously with totally irrelevant fluctuations of fields in the vacuum. However, similar arguments will now be used to eliminate a much more important class of graphs, namely, those involving self-energy effects. A "self-energy part" of a graph G is defined as follows; it is a set of one or more vertices not including x_0, together with the lines joining them, which is connected with the remainder of G (or with the edge of the diagram) only by two electron lines or by one or two photon lines. For definiteness it may be supposed that G has a self-energy part F, which is connected with its surroundings only by one electron line entering F at x_1, and another leaving F at x_2; the case of photon lines can be treated in an entirely analogous way. The points x_1 and x_2 may or may not be identical. From G a "reduced graph" G_0 can be obtained by omitting F completely and joining the incoming line at x_1 with the outgoing line at x_2 to form a single electron line in G_0, the newly formed line being denoted by λ. Given G_0 and λ, there is conversely a well determined set Γ of graphs G which are associated with G_0 and λ in this way; G_0 itself-is considered also to belong to Γ. It will now be shown that the sum $C(\Gamma)$ of the contributions $C(G)$ to the matrix element of (31) from all the graphs G of Γ reduces to a single term $C'(G_0)$.

Suppose, for example, that the line λ in G_0 leads from a point x_3 to the edge of the diagram. Then $C(G_0)$ is an integral containing in the integrand the matrix element of

$$\tilde{\psi}(x_3) \tag{52}$$

for creation of an electron into the state B. Let the momentum-energy 4-vector of the created electron be p; the matrix element of (52) is of the form

$$Y_\alpha(x_3) = a_\alpha \exp[-i(p \cdot x_3)/h] \tag{53}$$

with a_α independent of x_3. Now consider the sum $C(\Gamma)$. It follows from an

analysis of (31) that $C(\Gamma)$ is obtained from $C(G_0)$ by replacing the operator (52) by

$$\sum_{n=0}^{\infty} (-i/hc)^n [1/n!] \int_{-\infty}^{\infty} dy_1 \ldots \int_{-\infty}^{\infty} dy_n \tag{54}$$
$$\times P(\tilde{\psi}_\alpha(x_3), H^I(y_1), \ldots, H^I(y_n)).$$

(This is, of course, a consequence of the special character of the graphs of Γ.) It is required to calculate the matrix element of (54) for a transition from the vacuum state O to the state B, i.e., for the emission of an electron into state B. This matrix element will be denoted by $Z_\alpha; C(\Gamma)$ involves $\dot{Z}_\alpha$, in the same way that $C(G_0)$ involves (53). Now Z_α can be evaluated as a sum of terms of the same general character as (47); it will be of the form

$$Z_\alpha = \sum_i \int_{-\infty}^{\infty} K_i^{\alpha\beta}(y_i - x_3) Y_\beta(y_i) dy_i,$$

where the important fact is that K_i is a function only of the coordinate differences between y_i and x_3. By (53), this implies that

$$Z_\alpha = R_{\alpha\beta}(p) Y_\beta(x_3), \tag{55}$$

with R independent of x_3. From considerations of relativistic invariance, R must be of the form

$$\delta_{\beta\alpha} R_1(p^2) + (p_\mu \gamma_\mu)_{\beta\alpha} R_2(p^2),$$

where p^2 is the square of the invariant length of the 4-vector p. But since the matrix element (53) is a solution of the Dirac equation,

$$p^2 = -h^2 \kappa_0^2, \quad (p_\mu \gamma_\mu)_{\beta\alpha} Y_\beta = ih\kappa_0 Y_\alpha,$$

and so (55) reduces to

$$Z_\alpha = R_1 Y_\alpha(x_3),$$

with R_1 an absolute constant. Therefore the sum $C(\Gamma)$ is in this case just $C'(G_0)$, where $C'(G_0)$ is obtained from $C(G_0)$ by the replacement

$$\tilde{\psi}(x_3) \longrightarrow R_1 \tilde{\psi}(x_3). \tag{56}$$

In the case when the line λ leads into the graph G_0 from the edge of the diagram to the point x_3, it is clear that $C(\Gamma)$ will be similarly obtained from $C(G_0)$ by the replacement

$$\psi(x_3) \longrightarrow R_1^* \psi(x_3). \tag{57}$$

There remains the case in which λ leads from one vertex x_3 to another x_4 of G_0. In this case $C(G_0)$ contains in its integrand the function

$$\frac{1}{2}\eta(x_3, x_4)S_{F\beta\alpha}(x_3 - x_4), \tag{58}$$

which is the vacuum expectation value of the operator

$$P(\tilde{\psi}_\alpha(x_3), \psi_\beta(x_4)) \tag{59}$$

according to (43). Now in analogy with (54), $C(\Gamma)$ is obtained from $C(G_0)$ by replacing (59) by

$$\sum_{n=0}^{\infty}(-i/hc)^n[1/n!]\int_{-\infty}^{\infty}dy_1\ldots\int_{-\infty}^{\infty}dy_n \\ \times P(\tilde{\psi}_\alpha(x_3), \psi_\beta(x_4), H^I(y_1),\ldots,H^I(y_n)), \tag{60}$$

and the vacuum expectation value of this operator will be denoted by

$$\frac{1}{2}\eta(x_3, x_4)S'_{F\beta\alpha}(x_3 - x_4). \tag{61}$$

By the methods of Section VI, (61) can be expanded as a series of terms of the same character as (47); this expansion will not be discussed in detail here, but it is easy to see that it leads to an expression of the form (61), with $S'_F(x)$ certain universal function of the 4-vector x. It will not be possible to reduce (61) to a numerical multiple of (58), as Z_α was in the previous case reduced to a multiple of Y_α. Instead, there may be expected to be a series expansion of the form

$$S_{F\beta\alpha}(x) = (R_2 + a_1(\Box^2 - \kappa_0^2) + a_2(\Box^2 - \kappa_0^2)^2 \\ + \ldots)S_{F\beta\alpha}(x) + (b_1 + b_2(\Box^2 - \kappa_0^2) + \ldots) \\ \times(\gamma_\mu[\partial/\partial x_\mu] - \kappa_0)_{\beta\gamma}S_{F\gamma\alpha}(x), \tag{62}$$

where $\Box^2$ is the Dalembertian operator and the a, b are numerical coefficients. In this case $C(\Gamma)$ will be equal to the $C'(G_0)$ obtained from $C(G_0)$ by the replacement

$$S_F(x_3 - x_4) \to SF'(x_3 - x_4). \tag{63}$$

Applying the same methods to a graph G with a self-energy part connected to its surroundings by two photon lines, the sum $C(\Gamma)$ will be obtained as a single contribution $C'(G_0)$ from the reduced graph G_0, $C'(G_0)$ being formed from $C(G_0)$ by the replacement

$$D_F(x_3 - x_4) \to D'_F(x_3 - x_4). \tag{64}$$

The function D'_F is defined by the condition that

$$\frac{1}{2}hc\delta_{\mu\nu}D'_F(x_3 - x_4) \tag{65}$$

is the vacuum expectation value of the operator

$$\sum_{n=0}^{\infty}(-i/hc)^n[1/n!]\int_{-\infty}^{\infty}dy_1\ldots\int_{-\infty}^{\infty}dy_n$$
$$\times P(A_\mu(x_3), A_\nu(x_4), H^I(y_1), \ldots, H^I(y_n)), \tag{66}$$

and may be expanded in a series

$$D'_F(x) = (R_3 + c_1\square^2 + c_2(\square^2)^2 + \ldots)D_f(x). \tag{67}$$

Finally, it is not difficult to see that for graphs G with self-energy parts connected to their surroundings by a single photon line, the sum $C(\Gamma)$ will be identically zero, and so such graphs may be omitted from consideration entirely.

As a result of the foregoing arguments, the contributions $C(G)$ of graphs with self-energy parts can always be replaced by modified contributions $C'(G_0)$ from a reduced graph G_0. A given G may be reducible in more than one way to give various G_0, but if the process of reduction is repeated a finite number of times a G_0 will be obtained which is "totally reduced," contains no self-energy part, and is uniquely determined by G. The contribution $C'(G_0)$ of a totally reduced graph to the matrix element of (31) is now to be calculated as a sum of integrals of expressions like (47), but with a replacement (56), (57), (63), or (64) made corresponding to every line in G_0. This having been done, the matrix element of (31) is correctly calculated by taking into consideration each totally reduced graph once and once only.

The elimination of graphs with self-energy parts is a most important simplification of the theory. For according to (22), H^I contains the subtracted part H^S which will give rise to many additional terms in the expansion of (31). But if any such term is taken, say, containing the factor $H^S(x_i)$ in the integrand, every graph corresponding to that term will contain the point x_i joined to the rest of the graph only by two electron lines, and this point by itself constitutes a self-energy part of the graph. Therefore, all terms involving H^S are to be omitted from (31) in the calculation of matrix elements. The intuitive argument for omitting these terms is that they were only introduced in order to cancel out higher order self-energy terms arising from H^i, which are also to be omitted; the analysis of the foregoing paragraphs

is a more precise form of this argument. In physical language, the argument can be stated still more simply; since δm is an unobservable quantity, it cannot appear in the final description of observable phenomena.

8 VACUUM POLARIZATION AND CHARGE RENORMALIZATION

The question now arises: What is the physical meaning of the new functions D'_F and S'_F, and of the constant R_1? In general terms, the answer is clear. The physical processes represented by the self-energy parts of graphs have been pushed out of the calculations, but these processes do not consist entirely of unobservable interactions of single particles with their self-fields, and so cannot entirely be written off as "self-energy processes." In addition, these processes include the phenomenon of vacuum polarization, i.e., the modification of the field surrounding a charged particle by the charges which the particle induces in the vacuum. Therefore, the appearance of D'_F, S'_F, and R_1 in the calculations may be regarded as an explicit representation of the vacuum polarization phenomena which were implicitly contained in the processes now ignored.

In the present theory there are two kinds of vacuum polarization, one induced by the external field and the other by the quantized electron and photon fields themselves; these will be called "external" and "internal," respectively. It is only the internal polarization which is represented yet in explicit fashion by the substitutions (56), (57), (63), (64); the external will be included later.

To form a concrete picture of the function D'_F it may be observed that the function $D_F(y - z)$ represents in classical electrodynamics the retarded potential of a point charge at y, acting upon a point charge at z, together with the retarded potential of the charge at z acting on the charge at y. Therefore, D_F may be spoken of loosely as "the electromagnetic interaction between two point charges." In this semiclassical picture, D'_F is then the electromagnetic interaction between two point charges, including the effects of the charge-distribution which each charge induces in the vacuum.

The complete phenomenon of vacuum polarization, as hitherto understood, is included in the above picture of the function D'_F. There is nothing left for S'_F to represent. Thus, one of the important conclusions of the present theory is that there is a second phenomenon occurring in nature, included in the term vacuum polarization as used in this paper, but additional to vacuum polarization in the usual sense of the word. The nature of

the second phenomenon can best be explained by an example.

The scattering of one electron by another may be represented as caused by a potential energy (the Miller interaction) acting between them. If one electron is at and the other at z, then, as explained above, the effect of vacuum polarization of the usual kind is to replace a factor D_F in this potential energy by D_F'. Now consider an analogous, but unorthodox, representation of the Compton effect, or the scattering of an electron by a photon. If the electron is at y and the photon at z, the scattering may be again represented by a potential energy, containing now the operator $S_F(y-z)$ as a factor; the potential is an exchange potential, because after the interaction the electron must be considered to be at z and the photon at y, but this does not detract from its usefulness. By analogy with the 4-vector charge-current density j_μ which interacts with the potential D_F, a spinor Compton-effect density u_α may be defined by the equation

$$u_\alpha(x) = A_\mu(x)(\gamma_\mu)_{\alpha\beta}\psi_\beta(x),$$

and an adjoint spinor by

$$\tilde{u}_\alpha(x) = \tilde{\psi}_\beta(x)(\gamma_\mu)_{\beta\alpha}A_\mu(x).$$

These spinors are not directly observable quantities, but the Compton effect can be adequately described as an exchange potential, of magnitude proportional to $S_F(y-z)$, acting between the Compton-effect density at any point y and the adjoint density at z. The second vacuum polarization phenomenon is described by a charge in the form of this potential from S_F to S_F'. Therefore, the phenomenon may be pictured in physical terms as the including, by a given element of Compton-effect density at a given point, of additional Compton-effect density in the vacuum around it.

In both sorts of internal vacuum polarization, the functions D_F and S_F, in addition to being altered in shape, become multiplied by numerical (and actually divergent) factors, R_3 and R_2; also the matrix elements of (31) become multiplied by numerical factors such as $R_1 R_1^*$. However, it is believed (this has been verified only for second-order terms) that n'th-order matrix elements of (31) will involve these factors only in the form of a multiplier

$$(eR_2 R_3^{1/2})^n;$$

this statement includes the contributions from the higher terms of the series (62) and (67). Here e is defined as the constant occurring in (37). Now the only possible experimental determination of e is by means of measurements

of the effects described by various matrix elements of (31), and so the directly measured quantity is not e but $eR_2R_3^{1/2}$. Therefore, in practice the letter e is used to denote this measured quantity, and the multipliers R no longer appear explicitly in the matrix elements of (31); the charge in the meaning of the letter e is called "charge-renormalization," and is essential if e is to be identified with the observed electronic charge. As a result of the renormalization, the divergent coefficients R_1, R_2, and R_3 in (56), (57), (62), and (67) are to be replaced by unity, and the higher coefficients a, b, and c by expressions involving only the renormalized charge e.

The external vacuum polarization induced by the potential A_μ^e is, physically speaking, only a special case of the first sort of internal polarization; it can be treated in a precisely similar manner. Graphs describing external polarization effects are those with an "external polarization part," namely, a part including the point x_0 and connected with the rest of the graph by only a single photon line. Such a graph is to be "reduced" by omitting the polarization part entirely and renaming with the label x_0 the point at the further end of the single photon line. A discussion similar to those of Section VII leads to the conclusion that only reduced graphs need be considered in the calculation of the matrix element of (31), and that the effect of external polarization is explicitly represented if in the contributions from these graphs a replacement

$$A_\mu^e(x) \to A_\mu^{e'}(x) \tag{68}$$

is made. After a renormalization of the unit of potential, similar to the renormalization of charge, the modified potential $A_\mu^{e'}$ takes the form

$$A_\mu^{e'}(x) = (1 + c_1\Box^2 + c_2(\Box^2)^2 + \ldots)A_\mu^e(x), \tag{69}$$

where the coefficients are the same as in (67).

It is necessary, in order to determine the functions D_F', S_F', and $A_\mu^{e'}$, to go back to formulas (60) and (66). The determination of the vacuum expectation values of the operators (60) and (66) is a problem of the same kind as the original problem of the calculation of matrix elements of (31), and the various terms in the operators (60) and (66) must again be split up, represented by graphs, and analyzed in detail. However, since D_F' and S_F' are universal functions, this further analysis has.only to be carried out once to be applicable to all problems.

It is one of the major triumphs of the Schwinger theory that it enables an unambiguous interpretation to be given to the phenomenon of vacuum polarization (at least of the first kind), and to the vacuum expectation value

27

of an operator such as (66). In making this interpretation, profound theoretical problems arise, particularly concerned with the gauge invariance of the theory, about which nothing will be said here. For Schwinger's solution of these problems, the reader must refer to his forthcoming papers. Schwinger's argument can be transferred without essential change into the framework of the present paper.

Having overcome the difficulties of principle, Schwinger proceeded to evaluate the function D'_F explicitly as far as terms of order $\alpha = (e^2/4\pi hc)$ (heaviside units). In particular, he found for the coefficient c_1 in (67) and (69) the value $(-\alpha/15\pi\kappa_0^2)$ to this order.[9] in a sequel to the present paper a similar evaluation of the function S'_F the analysis involved is too complicated to be summarized here.

9   SUMMARY OF RESULTS

In this section the results of the preceding pages will be summarized, so far as they relate to the performance of practical calculations. In effect, this summary will consist of a set of rules for the application of the Feynman radiation theory to a certain class of problems.

Suppose an electron to be moving in an external field with interaction energy given by (33). Then the interaction energy to be used in calculating the motion of the electron, including radiative corrections of all orders, is

$$
\begin{aligned}
H_E(x_0) \quad &= \sum_{n=0}^{\infty} (-i/hc)^n [1/n!] J_n \\
&= \sum_{n=0}^{\infty} (-i/hc)^n [1/n!] \int_{-\infty}^{\infty} dx_1 \ldots \int_{-\infty}^{\infty} dx_n \\
&\times P(H^e(x_0), H^i(x_1), \ldots, H^i(x_n)),
\end{aligned} \tag{70}
$$

with H^i given by (16), and the P notation as defined in (29).

To find the effective n'th-order radiative correction to the potential acting on the electron, it is necessary to calculate the matrix elements of J_n for transitions from one one-electron state to another. These matrix elements can be written down most conveniently in the form of an operator K_n bilinear in $\tilde{\psi}$ and ψ, whose matrix elements for one-electron transitions are the same as those to be determined. In fact, the operator K_n itself is

---

[9]Schwinger's results agree with those of the earlier, theoretically unsatisfactory treatment of vacuum polarization. The best account of the earlier work is V. F. Weisskopf. Kgl. Danske Sels. Math.-Fys. Medd. **14,** No. 6 (1936).

already the matrix element to be determined if the $\tilde{\psi}$ and ψ contained in it are regarded as one-electron wave functions.

To write down K_n the integrand P_n in J_n is first expressed in terms of its factors $\tilde{\psi}, \psi$, and A, all suffixes being indicated explicitly, and the expression (37) used for j_μ. All possible graphs G with $(n+1)$ vertices are now drawn as described in Section VIII omitting disconnected graphs, graphs with self-energy parts, and graphs with external vacuum polarization parts as defined in Section VIII. It will be found that in each graph there are at each vertex two electron lines and one photon line, with the exception of x_0 at which there are two electron lines only; further, such graphs can exist only for even n. K_n is the sum of a contribution $K(G)$ from each G.

Given $G, K(G)$ is obtained from J_n by the following transformations. First, for each photon line joining x and y in G, replace two factors $A_\mu(x)A_\nu(y)$ in P_n (regardless of their positions) by

$$\frac{1}{2}hc\delta_{\mu\nu}D'_F(x-y),\tag{71}$$

with D'_F given by (67) with $R_3 = 1$, the function D_F being denned by (42). Second, for each electron line joining x to y in G, replace two factors $\tilde{\psi}_\alpha(x)\psi_\beta(y)$ in P_n (regardless of positions) by

$$\frac{1}{2}S'_{F\beta\alpha}(x-y)\tag{72}$$

with S_F given by (62) with $R_2 = 1$, the function S_F being denned by (44) and (45). Third, replace the remaining two factors $P(\tilde{\psi}_\gamma(z), \psi_\delta(\omega))$ in P_n by $\tilde{\psi}_\gamma(z)\psi_\delta(\omega)$ in this order. Fourth, replace $A_\mu^e(x_0)$ by $A_\mu^{e'}(x_0)$ given by

$$A_\mu^{e'}(x) = A_\mu^e(x) - [\alpha/15\pi\kappa_0^2]\square^2 A_\mu^e(x)\tag{73}$$

or, more generally, by (69). Fifth, multiply the whole by $(-1)^l$, where l is the number of closed loops in G as denned in Section VII.

The above rules enable K_n to be written down very rapidly for small values of n. It should be observed that if K_n is being calculated, and if it is not desired to include effects of higher order than the n'th, then D'_F, S'_F and $A_\mu^{e'}$ in (71), (72), and (73) reduce to the simple functions D_F, S_F, and A_μ^e. Also, the integrand in J_n is a symmetrical function of $x_1, \ldots, x_n$; therefore, graphs which differ only by a relabeling of the vertices $x_1, \ldots, x_n$ give identical contributions to K_n and need not be considered separately.

The extension of these rules to cover the calculation of matrix elements of (70) of a more general character than the one-electron transitions hitherto

considered presents no essential difficulty. All that is necessary is to consider graphs with more than two "loose ends," representing processes in which more than one particle is involved. This extension is not treated in the present paper, chiefly because it would lead to unpleasantly cumbersome formulas.

10 EXAMPLE-SECOND-ORDER RADIATIVE CORRECTIONS

As an illustration of the rules of procedure of the previous section, these rules will be used for writing down the terms giving second-order radiative
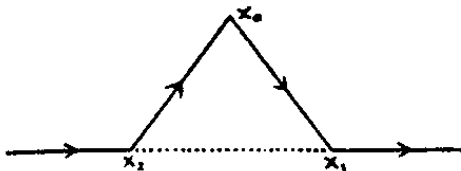

Figure 1:

corrections to the motion of an electron in an external field. Let the energy of the external field be

$$-[1/c]j_\mu(x_0)A_\mu^e(x_0). \tag{74}$$

Then there will be one second-order correction term

$$U = [\alpha/15\pi\kappa_0^2][1/c]j_\mu(x_0)\Box^2 A_\mu^e(x_0)$$

arising from the substitution (73) in the zero-order term (74). This is the well-known vacuum polarization or Uehling term.[10]

The remaining second-order term arises from the second-order part J_2 of (70). Written in expanded form, J_2 is

$$\begin{aligned}
J_2 = ie^3 \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 P(&\tilde{\psi}_\alpha(x_0)(\gamma_\lambda)_{\alpha\beta}\psi_\beta(x_0)A_\lambda^e(x_0), \\
&\tilde{\psi}_\gamma(x_1)(\gamma_\mu)_{\gamma\delta}\psi_\delta(x_1)A_\mu(x_1), \\
&\tilde{\psi}_\epsilon(x_2)(\gamma_\nu)_{\epsilon\zeta}\psi_\zeta(x_2)A_\nu(x_2)).
\end{aligned} \tag{75}$$

---

[10]Robert Serber, Phys. Rev. 48, 49 (1935); E. A. Uehling, Phys. Rev. **48,** 55 (1935).

Next, all admissable graphs with the three vertices x_0, x_1, x_2 are to be drawn. It is easy to see that there are only two such graphs, that G shown in Fig. 1, and the identical graph with x_1 and x_2 interchanged. The full lines are electron lines, the dotted line a photon line. The contribution $K(G)$ is obtained from J_2 by substituting according to the rules of Section IX; in this case $l = 0$, and the primes can be omitted from (71), (72), (73) since only second-order terms are required. The integrand in $K(G)$ can be reassembled into the form of a matrix product, suppressing the suffixes $\alpha, \ldots, \zeta$. Then, multiplying by a factor 2 to allow for the second graph, the complete second-order correction to (74) arising from J_2 becomes

$$
\begin{aligned}
L = -i[e^3/8hc] \int\limits_{-\infty}^{\infty} dx_1 \int\limits_{-\infty}^{\infty} dx_2 D_F(x_1 - x_2) A_\mu^e(x_0) \\
\times \tilde{\psi}(x_1)_{\gamma_\nu} S_F(x_0 - x_1)_{\gamma_\mu} S_F(x_2 - x_0)_{\gamma_\nu} \psi(x_2).
\end{aligned}
\tag{76}
$$

This is the term which gives rise to the main part of the Lamb-Retherford line shift,[11] the anomalous magnetic moment of the electron,[12] and the anomalous hyperfine splitting of the ground state of hydrogen.[13]

The above expression L is formally simpler than the corresponding expression obtained by Schwinger, but the two are easily seen to be equivalent. In particular, the above expression does not lead to any great reduction in the labor involved in a numerical calculation of the Lamb shift. Its advantage lies rather in the ease with which it can be written down.

In conclusion, the author would like to express his thanks to the Commonwealth Fund of New York for financial support, and to Professors Schwinger and Feynman for the stimulating lectures in which they presented their respective theories.

Notes added in proof (To Section II). The argument of Section II is an oversimplification of the method of Tomonaga,[14] and is unsound. There is an error in the derivation of (3); derivatives occurring in $H(\mathbf{r})$ give rise to non-commutativity between $H(\mathbf{r})$ and field quantities at $\mathbf{r}'$ when $\mathbf{r}'$ is a point on $\mathbf{r}$ infinitesimally distant from $\mathbf{r}'$. The argument should be amended as follows. Φ is defined only for flat surfaces $t(\mathbf{r}) = t$, and for such surfaces (3) and (6) are correct. Ψ is defined for general surfaces by (12) and (10), and is verified to satisfy (9). For a flat surface, Φ and Ψ are then shown to be related by (7). Finally, since H_1 does not involve the derivatives in H, the argument leading to (3) can be correctly applied to prove that

[11] W.E. Lamb and R.C. Retherford, Phys. Rev. **72**, 241 (1947).

[12] P. Kusch and H.M. Foley, Phys. Rev. **74**, 250 (1948).

[13] J.E. Nafe and E.B. Nelson, Phys. Rev. **73**, 718 (1948); Aage Bohr, Phys. Rev. **73**, 1109 (1948).

[14] P. Kusch and H. M. Foley, Phys. Rev. **74,** 250 (1948). J. E. Nafe and E. B. Nelson, Phys. Rev. **73,** 718 (1948); Aage Bohr, Phys. Rev. **73,** 1109 (1948).

for general σ the state-vector $\Psi(\sigma)$ will completely describe results of observations of the system on σ.

(To Section III). A covariant perturbation theory similar to that of Section III has previously been developed by E. C. G. Stueckelberg, Ann. d. Phys. **21,** 367 (1934); Nature, **153,** 143 (1944).

(To Section V). Schwinger's "effective potential" is not H_T given by (25), but is $H_T' = QH_TQ^{-1}$. Here Q is a "square-root" of $S(\infty)$ obtained by expanding $(S(\infty))^{1/2}$ by the binomial theorem. The physical meaning of this is that Schwinger specifies states neither by Ω nor by Ω', but by an intermediate state-vector $\Omega'' = Q\Omega = Q^{-1}\Omega'$, whose definition is symmetrical between past and future. H_T' is also symmetrical between past and future. For one-particle states, H_T and H_T' are identical.

Equation (32) can most simply be obtained directly from the product expansion of $S(\infty)$.

(To Section VII). Equation (62) is incorrect. The function S_F' is well-behaved, but its fourier transform has a logarithmic dependence on frequency, which makes an expansion precisely of the form (62) impossible.

(To Section X). The term L still contains two divergent parts. One is an "infra-red catastrophe" removable by standard methods. The other is an "ultraviolet" divergence, and has to be interpreted as an additional charge-renormalization, or, better, cancelled by part of the charge-renormalization calculated in Section VIII.