REVIEWS OF
MODERN PHYSICS

VOLUME 15, NUMBER 3 JULY, 1943

On Dirac's New Method of Field Quantization*

W. PAULI

Institute for Advanced Study, Princeton, New Jersey

1. INTRODUCTION

IN a recent paper Dirac[1] proposed a new method of field quantization which uses an indefinite metric in the space of quantum states. Applied to particles with integral spin obeying Bose statistics, it leads particularly to a "negative probability" of states, where an odd number of particles is present. The term "negative probability" means essentially that observables with only positive eigenvalues can get negative expectation values. The field oscillators which are quantized according to this method describe particles with negative energy. In addition one uses other field oscillators quantized in the usual way with positive probabilities and positive energy. The formalism which is constructed in this way has, therefore, a greater similarity with the original interpretation of the second-order wave equation as a one-body problem without second quantization, according to which states with positive (negative) frequencies should have positive (negative) energies.

The advantage of this new method is the possibility of overcoming all well-known convergence difficulties of quantized field theories if it is coupled with a quite different and logically independent method due to Wentzel and improved by Dirac, the so-called λ-limiting process which is a purely classical procedure that can be translated into quantum theory. This positive result is, however, partly balanced by a setback in the theory of generation and annihilation of pairs of particles with opposite electric charges. While in the usual form of the theory of holes it is possible to take into account the Coulomb interaction between generated pairs, the application of the λ-process to this theory fails to make the self-energy of an electron finite (whether or not the new method of field quantization is applied at the same time). The reason is that for a consistent application of the λ-process there is the necessary condition that the charged particles do not come closer to each other than the distance λ. While this condition can be fulfilled in a theory where the number of particles present is finite and constant by choosing λ sufficiently small, difficulties occur in connection with this condition from the fact that according to the theory of holes, pairs are generated in the immediate neighborhood of a charged particle in such a way that the electrostatic (and also the electrodynamic) self-energy of an electron becomes logarithmically divergent. It is the author's opinion that this difficulty could be overcome only by using, instead of the λ-limiting process, a new and probably purely quantum theoretical method.[2] The situation in the theory of charged

* This report is an improved and amplified form of a lecture held at Purdue University, Lafayette, Indiana, in June, 1942, and at the same time a continuation of the earlier report, Rev. Mod. Phys. **13**, 203 (1941), which will be quoted as A.

[1] P. A. M. Dirac, Proc. Roy. Soc. **A180**, 1 (1942).

[2] I disagree with Dirac's statement that in the theory of holes, where one starts with the negative energy states occupied, the equations are more complicated than in the older form of the theory using a constant number of charged particles which make transitions to negative energy states.

particles with zero spin obeying Bose statistics is entirely analogous to the situation in the theory of ordinary electrons with spin $\frac{1}{2}$ obeying the exclusion principle.

Both the negative energy values and the negative probabilities of some states of a single particle according to the new formalism make rather fundamental changes indispensable in the usual physical interpretation of quantum theory. Although the new formalism has in some respects a closer similarity to the classical theory than the previous method which gave divergent results, the interpretation of the new method is not a consistent and complete system, but consists of certain preliminary rules for computing probability coefficients of radiation and collision processes. The situation also has some analogy to the old correspondence principle of Bohr in this respect, that the new theory unites different processes which actually occur under very different conditions in nature. Indeed, the theory does not give the correct dependence of the transition probabilities of these processes on the number of particles initially present in the different states. Although this circumstance shows the very preliminary character of the new formalism and seems to indicate the need of more radical changes in the fundamental assumptions of the quantum theory of fields, it may be hoped that just this situation may enable a further progress.

2. INDEFINITE METRIC IN HILBERT SPACE

In Dirac's formalism of field quantization, a generalization of the usual metric in the Hilbert space of the states of a system is used. While in the latter the normalization of a wave function $\psi(q)$ is defined by

$$\int \bar{\psi}\psi dq,$$

where $\bar{\psi}$ is the conjugate complex of ψ, and the scalar product of two complex vectors φ and ψ in the Hilbert space by

$$\int \bar{\varphi}\psi dq,$$

we consider now the more general bilinear forms

$$\int \bar{\psi}\eta\psi dq, \quad \int \bar{\varphi}\eta\psi dq, \tag{1}$$

in which the operator η is only restricted by the condition that it has to be Hermitian in order to give the form used for normalization real values. It is obvious that in a discrete coordinate system the $\int dq$ has to be replaced by a sum over the discrete index n, and η can be represented by a matrix η_{nm} so that (1) becomes

$$\sum_{n,m} \bar{\psi}_n \eta_{nm} \psi_m, \quad \sum_{n,m} \bar{\varphi}_n \eta_{nm} \psi_m. \tag{2}$$

The expectation value of an observable A represented by a linear operator has now to be defined by

$$\langle A \rangle = \int \bar{\psi}\eta A\psi dq = \sum_{n,m,l} \bar{\psi}_n \eta_{nm} A_{ml} \psi_l. \tag{3}$$

As a generalization of the Hermitian conjugate operator, we introduce the adjoint operator which we denote by A^*. This is given by

$$A^* = \eta^{-1} A^\dagger \eta^\dagger = \eta^{-1} A^\dagger \eta, \tag{4}$$

where $A^\dagger$ is the Hermitian conjugate operator, and has the property

$$\langle A^* \rangle_{Av} = \langle A \rangle_{Av}. \tag{5}$$

The physical observables have to be self-adjoint; this means that $A^* = A$ holds for them because the self-adjoint operators are those, the expectation values of which are real. The self-adjoint operators play the same role as the Hermitian operators in the usual theory. In particular the Hamiltonian operator H which determines the time dependence of the wave function ψ according to

$$\partial\psi/\partial t = -iH\psi,$$

hence

$$(\partial\bar{\psi}/\partial t)\eta = i\bar{\psi}H^\dagger\eta = i\bar{\psi}\eta H^*, \tag{6}$$

has to be self-adjoint,

$$H^* = H.$$

This has the consequence that

$$\frac{d}{dt}\int \bar{\psi}\eta\psi dq = \frac{d}{dt}\sum_{n,m} \bar{\psi}_n \eta_{nm} \psi_m = i\bar{\psi}\eta(H^* - H)\psi = 0;$$

that is, the conservation of the normalization

with time. For the time dependence of the expectation value of an observable A (not *explicitly* time dependent) one gets as usual

$$\frac{d}{dt}\langle A\rangle_{Av}=i\langle HA-AH\rangle_{Av}. \tag{7}$$

If we perform a linear transformation of the coordinate system in the Hilbert space, which has not necessarily to be a unitary transformation according to

$$\psi'=S\psi, \tag{8}$$

we have to put

$$\eta'=S^{\dagger}\eta S \tag{9}$$

in order to keep the length of a vector in the Hilbert space constant; that is, the normalization of the wave function is invariant

$$\int\bar{\psi}'\eta'\psi' dq=\int\bar{\psi}\eta\psi dq.$$

The observables such as A and their adjoints are then in accordance with (4) transformed according to

$$A'=S^{-1}AS,\quad A^{*\prime}=\eta'^{-1}A^{\dagger\prime}\eta'=S^{-1}A^{*}S$$

in order to make their expectation values invariant

$$\langle A\rangle_{Av}=\int\bar{\psi}\eta A\psi dq=\int\bar{\psi}'\eta' A'\psi' dq.$$

The quality of an observable to be self-adjoint is invariant with respect to these S-transformations, while the quality to be Hermitian is in general not invariant.

Two different forms of the matrix η connected with each other according to (9) have to be considered as equivalent. Now every Hermitian matrix η can be transformed with a suitable S according to (9) into a normal form which is diagonal and where, moreover, every diagonal element has the value 1 or -1. Only the signs of the eigenvalues of η are therefore physically relevant, and, in particular, the positive definite forms (1) are equivalent to the usual theory where η is equal to the unit matrix. We get, however, something essentially new if we take into consideration indefinite bilinear forms as defining the square of the lengths of vectors in the Hilbert space. They lead to the consequence that *operators with only positive eigenvalues can have negative expectation values*. One can express it also by saying that one introduces negative probabilities that certain positive eigenvalues of an observable are realized. While the physical meaning of this possibility is not clear, Dirac was able to show that it enables us to overcome the convergence difficulties in the quantum electrodynamics as it has been known until now.

3. APPLICATION TO THE HARMONIC OSCILLATOR

We consider a harmonic oscillator which we can describe in suitable units by two variables p, q satisfying the commutation relations

$$i[p, q]=1, \tag{10}$$

and by the Hamiltonian

$$H=\tfrac{1}{2}(p^2+q^2). \tag{11}$$

Introducing the new variables u, u^* by

$$u=(1/\sqrt{2})(p-iq),\quad u^*=(1/\sqrt{2})(p+iq), \tag{12}$$

we get

$$[u, u^*]=1, \tag{13}$$

$$H=\tfrac{1}{2}(u^*u+uu^*)=u^*u+\tfrac{1}{2}. \tag{14}$$

If we assume that p and q are Hermitian, there is $u^*=u^{\dagger}$; that is, u^* is Hermitian conjugate of u. As is well known, the quantity

$$u^*u=N \tag{15}$$

has the eigenvalues 0, 1, 2, $\cdots$ and in a representation where N is diagonal, the operators corresponding to u and u^* have the simple meaning of an absorption and emission operator applied to the wave function $\psi(N)$

$$\begin{aligned} u^*\psi(N)&=N^{\frac{1}{2}}\psi(N-1),\\ u\psi(N)&=(N+1)^{\frac{1}{2}}\psi(N+1). \end{aligned} \tag{16}$$

The wave function $\psi(N)$ has to be normalized by

$$\textstyle\sum_N\bar{\psi}(N)\psi(N)=1.$$

If $h_N(q)$ are the normalized Hermitian eigenfunctions of (11), one has

$$\psi(q)=\textstyle\sum_N h_N(q)\psi(N), \tag{17}$$

and hence

$$\int \bar{\psi}(q)\psi(q) = 1.$$

There exists, however, an alternative possibility of treating the same Hamiltonian by assuming the variables p, q which again satisfy (10) as still self-adjoint, but anti-Hermitian instead of Hermitian. In this case it will turn out to be convenient to define

$$u = (1/\sqrt{2})(p+iq), \quad u^* = (1/\sqrt{2})(p-iq). \tag{18}$$

We get

$$[u, u^*] = -1, \tag{19}$$

with u^* the negative of the Hermitian conjugate of u. We can now put

$$u^*u = -N, \tag{20}$$

where N has again the positive eigenvalues 0, 1, 2, $\cdots$, and again represent u as an absorption, u^* as an emission operator, but changing the sign of u

$$\begin{aligned} u^*\psi(N) &= N^{\frac{1}{2}}\psi(N-1), \\ u\psi(N) &= -(N+1)^{\frac{1}{2}}\psi(N+1). \end{aligned} \tag{16'}$$

The Hamiltonian becomes

$$H = \tfrac{1}{2}(u^*u + uu^*) = -(N+\tfrac{1}{2}). \tag{21}$$

We now have to find the matrix η determining the normalization of the wave function $\psi(N)$ according to

$$\sum_{N',N''} \bar{\psi}(N')(N'|\eta|N'')\psi(N'') = \text{const.}$$

In view of (4), applied to $u^* = -u^\dagger$, η has to anticommute with u^*

$$\eta u^* = -u^*\eta.$$

In the representation where N is diagonal, one sees from (19) that it is sufficient to choose η diagonal with respect to N with $(N|\eta|N) \equiv \eta_N$ fulfilling

$$\eta_{N+1} = -\eta_N$$

by

$$\eta_N = (-1)^N.$$

The wave function $\psi(N)$ has therefore to be normalized according to

$$\textstyle\sum_N (-1)^N \bar{\psi}(N)\psi(N) = \text{const.}, \tag{22}$$

which clearly shows the "negative probability" of states with odd values of N in this formalism. In view of the fact that $h_N(q)$ is an even or odd function if N is even or odd, one easily obtains, using (16), the normalization of $\psi(q)$ which corresponds to the normalization (22) of $\psi(N)$ as given by

$$\int \bar{\psi}(q)\psi(-q)dq = \text{const.}. \tag{23}$$

or, if we decompose $\psi(q)$ into an even part $\psi_{\text{ev}}(q)$ and an odd part $\psi_{\text{odd}}(q)$,

$$\int \bar{\psi}_{\text{ev}}(q)\psi_{\text{ev}}(q)dq - \int \bar{\psi}_{\text{odd}}(q)\psi_{\text{odd}}(q)dq = \text{const.} \tag{24}$$

This means η operating on $\psi(q)$ changes it into $\psi(-q)$;

$$\eta\psi(q) = \psi(-q). \tag{25}$$

If we consider p, q as operators which have to be anti-Hermitian and to satisfy (10), we have to put

$$\begin{aligned} q_{\text{op}}\psi(q) &= iq\psi(q), \\ p_{\text{op}}\psi(q) &= -\frac{d}{dq}\psi(q). \end{aligned} \tag{26}$$

One easily checks that for the operators so defined the expectation values of p and q computed according to

$$\langle q \rangle_{\text{Av}} = i\int \bar{\psi}(q)(-q)\psi(-q)dq,$$

$$\langle p \rangle_{\text{Av}} = \int \bar{\psi}(q)\frac{d}{dq}\psi(-q)dq$$

are real. We notice that from (26) and (11)

$$H_{\text{op}}\psi(q) = -\frac{1}{2}\left(-\frac{d^2}{dq^2} + q^2\right)\psi(q) \tag{27}$$

in accordance with the negative eigenvalue of H given by (21).

4. SYSTEM OF TWO OSCILLATORS WITH POSITIVE AND NEGATIVE ENERGY

We consider now a system of two oscillators of which one has positive energy (p_+, q_+ Hermitian) and the other negative energy (p_-, q_- self-adjoint but anti-Hermitian), with the commutation rules

$$i[p_+, q_+]=1, \quad i[p_-, q_-]=1, \tag{28}$$

and where both variables with $+$ commute with both variables with $-$, and with the Hamiltonian

$$H=\tfrac{1}{2}(p_+^2+p_-^2+q_+^2+q_-^2). \tag{29}$$

Introducing the variables u_+, u_- and their adjoints analogous to the Eqs. (12) and (18); that is,

$$u_+=(1/\sqrt{2})(p_+-iq_+), \quad u_+^*=(1/\sqrt{2})(p_++iq_+), \tag{30}$$

$$u_-=(1/\sqrt{2})(p_-+iq_-), \quad u_-^*=(1/\sqrt{2})(p_--iq_-), \tag{31}$$

we obtain

$$[u_+, u_+^*]=1, \quad [u_-, u_-^*]=-1; \tag{32}$$

and with

$$N_+=u_+^*u_+, \quad N_-=-u_-^*u_-, \tag{33}$$

$$H=(N_++\tfrac{1}{2})-(N_-+\tfrac{1}{2})=N_+-N_-. \tag{34}$$

The normalization of the wave function is given by

$$\sum_{N+,\,N-}(-1)^{N-}\bar{\psi}(N_+, N_-)\psi(N_+, N_-)=\text{const.}, \tag{35}$$

or

$$\int\int\bar{\psi}(q_+, q_-)\psi(q_+, -q_-)dq_+dq_-=\text{const.} \tag{36}$$

It is important for the following that for this system of oscillators, two other pairs of variables exist which fulfill the same commutation rules as u_+, u_- and their adjoints do; namely,

$$a=(1/\sqrt{2})(u_++u_-^*), \quad a^*=(1/\sqrt{2})(u_+^*+u_-), \tag{37}$$

$$b=(1/\sqrt{2})(u_+^*-u_-), \quad b^*=(1/\sqrt{2})(u_+-u_-^*), \tag{38}$$

for which we have

$$[a, a^*]=1, \quad [b, b^*]=-1; \tag{39}$$

and with

$$N_a=a^*a, \quad N_b=-b^*b, \tag{40}$$

$$N_a-N_b=N_+-N_-, \tag{41}$$

and therefore also

$$H=N_a-N_b. \tag{42}$$

We notice that the inverse formulas which express u_+, u_- and their adjoints by a, b and their adjoints are of the same form as (37), (38), namely,

$$u_+=(1/\sqrt{2})(a+b^*), \quad u_+^*=(1/\sqrt{2})(a^*+b), \tag{43}$$

$$u_-=(1/\sqrt{2})(a^*-b), \quad u_-^*=(1/\sqrt{2})(a-b^*), \tag{44}$$

and that therefore the variables a, b are entirely equivalent to the variables u_+, u_-. The transformation of the p, q variables corresponding to the new self-adjoint variables p_a, q_a, p_b, q_b is given by

$$p_a=(1/\sqrt{2})(p_++p_-), \quad q_a=(1/\sqrt{2})(q_++q_-). \tag{45}$$

$$p_b=(1/\sqrt{2})(p_+-p_-), \quad q_b=(1/\sqrt{2})(q_+-q_-), \tag{46}$$

with

$$i[p_a, q_a]=i[p_b, q_b]=1, \tag{47}$$

the a variables and the b variables commuting with each other. The inverse formulas of (45), (46) have again the same form. Obviously one has

$$\tfrac{1}{2}(p_+^2+p_-^2+q_+^2+q_-^2)=\tfrac{1}{2}(p_a^2+p_b^2+q_a^2+q_b^2)=H. \tag{48}$$

Taking into account (30), (31), (37), (38), we find

$$a=(1/\sqrt{2})(p_a-iq_a), \quad a^*=(1/\sqrt{2})(p_a+iq_a), \tag{49}$$

$$b=(1/\sqrt{2})(p_b+iq_b), \quad b^*=(1/\sqrt{2})(p_b-iq_b). \tag{50}$$

We now proceed to the investigation of the S-transformation which carries the wave function $\psi(N_+, N_-)$ to the wave function $\psi(N_a, N_b)$ according to

$$\psi(N_a, N_b)=\sum_{N+, N-}(N_a, N_b|S|N_+, N_-)\psi(N_+, N_-). \tag{51}$$

This transformation will not be unitary, but it can be chosen in such a way that the normalization is conserved; that is,

$$\sum_{N+, N-}(-1)^{N-}\bar{\psi}(N_+, N_-)\psi(N_+, N_-)=\sum_{N_a, N_b}(-1)^{N_b}\bar{\psi}(N_a, N_b)\psi(N_a, N_b), \tag{52}$$

or

$$\sum_{N+, N-}(N_a', N_b'|S(-1)^{N-}S^\dagger|N_a'', N_b'')=(-1)^{N_b'}\delta_{N_a', N_a''}\delta_{N_b', N_b''}. \tag{53}$$

The last equation can be considered as a particular case of (9) if we put

$$(N_a', N_b'|\eta'|N_a'', N_b'')=(-1)^{N_b'}\delta_{N_a', N_a''}\delta_{N_b', N_b''},$$

$$(N_+', N_-'|\eta|N_+'', N_-'')=(-1)^{N-'}\delta_{N+', N+''}\delta_{N-', N-''}.$$

Equations (51), (52) can be simplified in view of (41), and the double sums in these equations reduce to single ones. Putting

$$N_b=N_a-N\geqq 0, \quad N_-=N_+-N\geqq 0,$$

we obtain for every N

$$\psi(N_a)=\sum_{N+=N}^{\infty}(N_a|S|N_+)\psi(N_+), \tag{51'}$$

$$\sum_{N+}(N_a'|S(-1)^{N+-N}S^\dagger|N_a'')=(-1)^{N_a-N}\delta_{N_a', N_a''}. \tag{53'}$$

In order to compute the transformation function S, it is convenient to start from the equality

$$\tfrac{1}{2}(aa^*+a^*a)-\tfrac{1}{2}(bb^*+b^*b)=u_+u_-+u_+^*u_-$$

which is derived from (37), (38). Because the left side can also be written

$$N_a+N_b+1=2N_a-N+1,$$

the application of this equality to S gives

$$(2N_a-N+1)(N_a|S|N_+)=(u_+u_-+u_+^*u_-^*)(N_a|S|N_+).$$

Using (16) for u_+, u_+^* but (16′) for u_-, u_-^* we obtain

$$(2N_a-N+1)(N_a|S|N_+)=[N_+(N_+-N)]^{\frac{1}{2}}(N_a|S|N_+-1) \\ -[(N_++1)(N_++1-N)]^{\frac{1}{2}}(N_a|S|N_++1). \tag{54}$$

Putting

$$\varphi(N_+)=\frac{[N_+!]^{\frac{1}{2}}}{[(N_+-N)!]^{\frac{1}{2}}}(N_a|S|N_+), \quad (N_a|S|N_+)=\frac{[N_+!]^{\frac{1}{2}}}{[(N_+-N)!]^{\frac{1}{2}}}\psi(N_+), \tag{55}$$

we get from (54)

$$(2N_a-N+1)\varphi(N_+)=N_+\varphi(N_+-1)-(N_++1-N)\varphi(N_++1), \tag{56}$$

$$(2N_a-N+1)\psi(N_+)=(N_+-N)\psi(N_+-1)-(N_++1)\psi(N_++1). \tag{56'}$$

While (56) has a solution with $\varphi(N_+)=0$ for $N_+<N$, (56') has a solution with $\psi(N_+)=0$ for $N_+<0$. In the latter case the factor $[(N_+-N)!]^{-\frac{1}{2}}$ in the second Eq. (55) makes $(N_a|S|N_+)=0$ for $0\leqq N_+<N$ as it must be, while (56') defines $\psi(N_+)$ for $0<N_+<N$. On the other hand for $N<0$, $\varphi(N_+)$ can be defined by (56) also for negative values of N_+ greater than or equal to $N=-|N|$, while the factors in the first Eq. (55) make $(N_a|S|N_+)=0$ for $N_+<0$.

Introducing the auxiliary variable x and the power series

$$\varphi(x)=\sum_{N_+=N}^{\infty}\varphi(N_+)x^{N_+-N},\quad \psi(x)=\sum_{N_+=0}^{\infty}\psi(N_+)x^{N_+}, \tag{57}$$

one gets for $\varphi(x)$ and $\psi(x)$ from (56), (56'), the differential equations

$$(x^2-1)\varphi'-[2N_a-N+1-(N+1)x]\varphi=0 \quad \text{or} \quad \frac{\varphi'}{\varphi}=\frac{N_a-N}{x-1}-\frac{N_a+1}{x+1}, \tag{58}$$

$$(x^2-1)\psi'-[2N_a-N+1+(N-1)x]\psi=0 \quad \text{or} \quad \frac{\psi'}{\psi}=\frac{N_a}{x-1}-\frac{N_a-N+1}{x+1}. \tag{58'}$$

Their solutions are

$$\varphi=\text{const.}\ (x-1)^{N_a-N}(x+1)^{-N_a-1}, \tag{59}$$

$$\psi=\text{const.}\ (x-1)^{N_a}(x+1)^{N-N_a-1}. \tag{59'}$$

In order to compare the values of the constant factors in the last two equations we notice that from (55), it follows that

$$\varphi(N_+)=\frac{N_+!}{(N_+-N)!}\psi(N_+);$$

hence

$$\varphi(N_+)=N_+(N_+-1)\cdots(N_+-N+1)\psi(N_+),\quad \varphi(x)=\frac{d^N}{dx^N}\psi(x),\quad \text{for } N\geqq 0; \tag{60}$$

$$\psi(N_+)=(N_++|N|)(N_++|N|-1)\cdots(N_++1)\varphi(N_+),\quad \psi(x)=\frac{d^{|N|}}{dx^{|N|}}\varphi(x),\quad \text{for } N\leqq 0. \tag{60'}$$

If we compare both sides of Eq. (60) in the neighborhood of $x=1$ and bring the coefficients of $(x-1)^{N_a-N}$ in (60), and of $(x-1)^{N_a}$ in (60') into agreement, we obtain in both cases $N\geqq 0$ and $N\leqq 0$

$$\varphi(x)=K(N_a,N)\frac{N_a!}{(N_a-N)!}2^N(x-1)^{N_a-N}(x+1)^{-N_a-1}, \tag{61}$$

$$\psi(x)=K(N_a,N)(x-1)^{N_a}(x+1)^{N-N_a-1}. \tag{61'}$$

Moreover, it is easy to check (60), (60') by direct computation, which, however, is not necessary for our purpose.

We now determine the normalization factor $K(N_a, N)$ using the condition (53'). That the left side of this equation is zero for $N_a'\neq N_a''$ is already a consequence of (54). Indeed, if we insert in the conjugate of this equation $N_a=N_a''$, multiply by $(N_a'|S|N_+)(-1)^{N-N_+}$, and sum over N_+, the right side is the same as if we start with (54) for the value $N_a=N_a'$, multiply by $(-1)^{N-N_+}(N_+|S^\dagger|N_a'')$,

and sum over N_+. Subtracting the two results we obtain

$$(N_a''-N_a')(N_a'|S(-1)^{N-N_+}S^\dagger|N_a'')=0.$$

We are left now with the case $N_a'=N_a''$. In order to give the sum

$$\sum_{N_+}(N_a|S|N_+)(-1)^{N-N_+}(N_+|S^\dagger|N_a),$$

which does not converge in the proper sense, a definite meaning, we define it following Dirac as the limit

$$\lim_{r\to 1}\Sigma(r) \quad \text{with} \quad 0<r<1$$

where

$$\Sigma(r)=\sum_{N_+}(N_a|S|N_+)r^{2N_+}(-1)^{N-N_+}(N_+|S^\dagger|N_a),$$

and in view of (55) also

$$\Sigma(r)=\sum_{N_+}(-1)^{N-N_+}r^{2N_+}\varphi(N_+)\bar{\psi}(N_+).$$

As can be seen from (57) the sum $\Sigma(r)$ can be expressed with the help of the functions $\varphi(x)$, $\psi(x)$ by introducing the points

$$x=re^{i\theta}, \quad \bar{x}=re^{-i\theta}, \quad 0<r<1$$

of the complex plane as an integral over the unit circle, namely,

$$\Sigma(r)=\frac{1}{2\pi}\int_0^{2\pi}\bar{\psi}(x)x^N\varphi(-x)d\theta. \tag{62}$$

Inserting (61) and (61′), we get

$$\Sigma(r)=|K(N_a,N)|^2\frac{N_a!}{(N_a-N)!}2^N I(r) \tag{63}$$

where

$$I(r)=\frac{r^N}{2\pi}\int_0^{2\pi}e^{iN\theta}(re^{-i\theta}-1)^{N_a}(re^{-i\theta}+1)^{N-N_a-1}(-re^{i\theta}-1)^{N_a-N}(-re^{i\theta}+1)^{-N_a-1}d\theta,$$

or introducing $z=e^{-i\theta}$, $dz/iz=-d\theta$

$$I(r)=(-1)^{N_a-N}\frac{r^N}{2\pi i}\oint dz(z+r)^{N_a-N}(z-r)^{-N_a-1}(rz-1)^{N_a}(rz+1)^{N-N_a-1} \tag{64}$$

and the path of the last integral is the unit circle in the positive sense. Taking into account $N_a-N\geqq 0$, one sees that the integrand has poles for $z=r$ and $z=-r^{-1}$ from which only the first is inside the unit circle. The value of the integral is therefore the residue at the pole $z=r$, namely,

$$I(r)=(-1)^{N_a-N}r^N\frac{1}{N_a!}\frac{d^{N_a}}{dz^{N_a}}[(z+r)^{N_a-N}(rz+1)^{N-N_a-1}(rz-1)^{N_a}]_{z=r}.$$

Fortunately this expression simplifies greatly in the limit $r=1$ because the last factor in the bracket has the value $(r^2-1)^{N_a}$ for $z=r$ and vanishes for $r=1$. Therefore only the differentiation of the last factor gives a contribution which is different from zero for $r\to 1$, and we get

$$\begin{aligned}\lim_{r\to 1}I(r)&=(-1)^{N_a-N}\tfrac{1}{2},\\ \lim_{r\to 1}\Sigma(r)&=(-1)^{N_a-N}|K(N_a,N)|^2 2^{N-1}\frac{N_a!}{(N_a-N)!}.\end{aligned} \tag{65}$$

The normalization condition (53) or (53′) is therefore fulfilled if we choose

$$K(N_a, N)=\left[2^{-(N-1)}\frac{(N_a-N)!}{N_a!}\right]^{\frac{1}{2}}. \tag{66}$$

A phase factor in $K(N_a, N)$ is of course arbitrary.

Collecting our results we finally reach the conclusion

$$\begin{aligned}(N_a|S|N_+)&=\sqrt{2}\left[2^N\frac{N_a!(N_+-N)!}{(N_a-N)!N_+!}\right]^{\frac{1}{2}}\{\text{coeff. of } x^{N+-N} \text{ in } (x-1)^{N_a-N}(x+1)^{-N_a-1}\},\\ &=\sqrt{2}\left[2^N\frac{N_a!(N_+-N)!}{(N_a-N)!N_+!}\right]^{-\frac{1}{2}}\{\text{coeff. of } x^{N+} \text{ in } (x-1)^{N_a}(x+1)^{N-N_a-1}\}.\end{aligned} \tag{67}$$

Because of the complete symmetry between the variables N_a, N_b and N_+, N_- (or N_a, N and N_+, N) one obtains the inverse transformation function S^{-1} which expresses $\psi(N_+, N_-)$ by $\psi(N_a, N_b)$ according to

$$\psi(N_+, N_-)=\sum_{N_a, N_b}(N_+, N_-|S^{-1}|N_a, N_b)\psi(N_a, N_b) \tag{51′}$$

simply by interchanging N_a, N_b with N_+, N_-.

A case of particular interest is the state $N_+=N_-=0$ (hence $N=N_+-N_-=N_a-N_b=0$), for which we get from (67) putting $N_a=N_b=n$

$$(n, n|S|0, 0)=\sqrt{2}(-1)^n$$

which means that the "probability of the value $n=N_a=N_b$ in the state $N_+=N_-=0$" is given by

$$\eta_n|(n, n|S|0, 0)|^2=2(-1)^n. \tag{68}$$

The sum of the probabilities defined by

$$\lim_{r\to 1}\sum_{n=0}^{\infty}2(-1)^n r^n=\lim_{r\to 1}\frac{2}{1+r}$$

with $0<r<1$ is 1 as it has to be. Of course this result can also be derived directly from (54). It means that the ground state of our system of two oscillators is not uniquely determined when an indefinite metric in the Hilbert space is used; this circumstance plays an essential role in Dirac's theory.

5. RELATIVISTIC FIELD QUANTIZATION

(a) Uncharged Particles with Spin 0 or 1 and Bose Statistics

We consider first as the simplest case of a wave field with integral spin, a real scalar field $A(x)=A^*(x)$ in the absence of any interaction. It satisfies the wave equation

$$(\Box+\mu^2)A=0 \tag{69}$$

where

$$\Box=-\sum_{k=1}^{3}\frac{\partial^2}{\partial x_k^2}+\frac{\partial^2}{\partial x_0^2}$$

with $x_0=ct$ and μ the rest mass of the particle (the units are chosen so that $\hbar=c=1$).[3] A particularly important case is $\mu=0$, but we shall postpone this specialization. The energy E and the momentum

[3] Compare paper A, Part II, Section 1-d. The field function which we denote here by $A(x)$ is there called $V(x)$.

$\mathbf{G}$ are given by

$$E=\frac{1}{2}\int\left[(\nabla A)^2+\left(\frac{\partial A}{\partial x_0}\right)^2\right]dV, \quad \mathbf{G}=-\int\frac{\partial A}{\partial x_0}\nabla A\,dV. \tag{70}$$

The usual method of decomposing the field into Fourier components (periodic with respect to a large hole with volume V) is done according to

$$A(x)=V^{-\frac{1}{2}}\sum_k(2k_0)^{-\frac{1}{2}}[A(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}+A^*(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}] \tag{71}$$

where we define k_0 as positive and satisfying $k_0^2=\mathbf{k}^2+\mu^2$, hence

$$k_0=(\mathbf{k}^2+\mu^2)^{\frac{1}{2}}.$$

The usual quantization according to

$$[A(k), A^*(k)]=1 \tag{72}$$

and with the commutability of all variables corresponding to different values of k leads to

$$A^*(k)A(k)=N_a(k) \tag{73}$$

and to the expression

$$E=\sum_k k_0[N_a(k)+\tfrac{1}{2}], \quad \mathbf{G}=\sum_k\mathbf{k}[N_a(k)+\tfrac{1}{2}] \tag{74}$$

for energy and momentum.

Dirac proposed the procedure of decomposing the real field $A(x)$ into two complex fields $U(x)$, $U^*(x)$

$$A(x)=\frac{1}{\sqrt{2}}[U(x)+U^*(x)], \tag{75}$$

and of quantizing only the oscillators of $U(x)$ with $(-k_0x_0)$ in the phase factor in the usual way, while the other part of the oscillators of $U(x)$ with $(+k_0x_0)$ in the phase factor is quantized with the new method developed in the last two sections, which leads to negative energies of these oscillators. This means that we have to put

$$U(x)=V^{-\frac{1}{2}}\sum_k(2k_0)^{-\frac{1}{2}}[U_+(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}+U_-(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}], \tag{76}$$

$$U^*(x)=V^{-\frac{1}{2}}\sum_k(2k_0)^{-\frac{1}{2}}[U_+^*(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}+U_-^*(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}]. \tag{76'}$$

If we now define energy and momentum by

$$\begin{aligned} E&=\int\left[\nabla U^*\nabla U+\frac{\partial U^*}{\partial x_0}\frac{\partial U}{\partial x}+\mu^2U^*U\right]dV,\\ \mathbf{G}&=-\int\left[\frac{\partial U^*}{\partial x_0}\nabla U+\nabla U^*\frac{\partial U}{\partial x_0}\right]dV \end{aligned} \tag{77}$$

(with symmetrization of the order of factors in the last integrand), if we assume that $U_+^*(k)$, $U_+(k)$ commute with $U_-^*(k)$, $U_-(k)$, and if we put

$$[U_+(k), U_+^*(k)]=1, \quad [U_-(k), U_-^*(k)]=-1, \tag{78}$$

we are in accordance with the rule

$$dF/\partial x_0=i[H, F] \tag{79}$$

applied to $F=U(x)$ and $F=U^*(x)$. Moreover, this leaves the usual commutation relation for $U^*(x)$, $U(x)$ unchanged.[4] The latter equation means that $U_-^*(k)$ is the adjoint but not any longer the

[4] See paper A, Part II, Eq. (21).

Hermitian conjugate of $U_-(k)$. In accordance with (33), (34) we obtain

$$N_+(k)=U_+^*(k)U_+(k),\quad N_-(k)=-U_-^*(k)U_-(k), \tag{80}$$

and

$$\begin{aligned} E&=\textstyle\sum_k k_0[N_+(k)-N_-(k)],\\ \mathbf{G}&=\textstyle\sum_k \mathbf{k}[N_+(k)-N_-(k)], \end{aligned} \tag{81}$$

which shows the negative energy of the N_- oscillators. The normalization of the wave function is analogous to (35)

$$\sum_{N+(k),\,N-(k)}(-1)^{\Sigma_k N-(k)}\bar{\psi}(\cdots N_+(k)\cdots,\ \cdots N_-(k)\cdots)\psi(\cdots N_+(k)\cdots,\ \cdots N_-(k)\cdots)=\text{const.} \tag{82}$$

which shows the "negative probability" of states with an odd number of particles in states with negative energy.

This procedure is also equivalent to introducing besides $A(x)$ another field, $B(x)$,[5] the adjoint of which is $B^*(x)=-B(x)$,

$$B(x)=\frac{1}{\sqrt{2}}[U(x)-U^*(x)], \tag{75'}$$

or in Fourier decomposition

$$B(x)=V^{-\frac{1}{2}}\textstyle\sum_k (2k_0)^{-\frac{1}{2}}[B^*(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}-B(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}], \tag{71'}$$

and with the quantization

$$[B(k),B^*(k)]=-1 \tag{72'}$$

leading to

$$B^*(k)B(k)=-N_b(k) \tag{73'}$$

and to the energy and momentum expressions

$$\begin{aligned} E&=\frac{1}{2}\int\left[(\nabla A)^2+\left(\frac{\partial A}{\partial x_0}\right)^2-(\nabla B)^2-\left(\frac{\partial B}{\partial x_0}\right)^2\right]dV,\\ \mathbf{G}&=\int\left[-\frac{\partial A}{\partial x_0}\nabla A+\frac{\partial B}{\partial x_0}\nabla B\right]dV, \end{aligned} \tag{70'}$$

(in the latter expression one has to symmetrize the order of all non-commuting factors) from which one derives

$$\begin{aligned} E&=\textstyle\sum_k k_0[N_a(k)-N_b(k)],\\ \mathbf{G}&=\textstyle\sum_k \mathbf{k}[N_a(k)-N_b(k)]. \end{aligned} \tag{74'}$$

The connection between $U_+(k)$, $U_-(k)$ and their adjoints, on the one hand, with $A(k)$, $B(k)$ and their adjoints on the other hand is, in view of (75), (75'), given by equations in analogy with (37), (38)

$$A(k)=\frac{1}{\sqrt{2}}[U_+(k)+U_-^*(k)],\quad A^*(k)=\frac{1}{\sqrt{2}}[U_+^*(k)+U_-(k)], \tag{37'}$$

$$B(k)=\frac{1}{\sqrt{2}}[U_+^*(k)-U_-(k)],\quad B^*(k)=\frac{1}{\sqrt{2}}[U_+(k)-U_-^*(k)]. \tag{38'}$$

One obtains the inverse formulas by permutation of $A(k)$, $B(k)$ with $U_+(k)$, $U_-(k)$ and the same for their adjoints.

[5] The functions $A(x)$, $B(x)$ are the same as $V(x)$, $iW(x)$ in A, Part II. Their connection with $U(x)$, $U^*(x)$ is given there in Eq. (33).

An alternative way of decomposing the field $A(x)$ is to write all oscillators with positive energy in a field $A_+(x)=A_+^*(x)$ and all oscillators with negative energy in a field $A_-(x)=A_-^*(x)$ given by[6]

$$\begin{aligned} A_+(x) &= V^{-\frac{1}{2}} \sum_k (2k_0)^{-\frac{1}{2}}[u_+(k) \exp[i(\mathbf{k}\cdot\mathbf{x})]+u_+^*(k) \exp[-i(\mathbf{k}\cdot\mathbf{x})]], \\ A_-(x) &= V^{-\frac{1}{2}} \sum_k (2k_0)^{-\frac{1}{2}}[u_-^*(k) \exp[i(\mathbf{k}\cdot\mathbf{x})]+u_-(k) \exp[-i(\mathbf{k}\cdot\mathbf{x})]], \end{aligned} \tag{83}$$

$$A(x)=(1/\sqrt{2})[A_+(x)+A_-(x)]. \tag{84}$$

We have introduced here the time dependent quantities

$$\begin{aligned} u_+(k) &= U_+(k) \exp(-ik_0x_0), \quad & u_+^*(k) &= U_+^*(k) \exp(ik_0x_0), \\ u_-(k) &= U_-(k) \exp(ik_0x_0), & u_-^*(k) &= U_-^*(k) \exp(-ik_0x_0), \end{aligned} \tag{85}$$

besides the constant quantities $U_+(k)$, $U_-(k)$ and their adjoints in order to prepare for the treatment of the interaction of the field with charged particles. The corresponding decomposition of $B(x)$ into $B_+(x)=-B_+^*(x)$, $B_-(x)=-B_-^*(x)$ according to

$$B(x)=(1/\sqrt{2})[B_+(x)+B_-(x)] \tag{86}$$

is given by

$$\begin{aligned} B_+(x) &= V^{-\frac{1}{2}} \sum_k (2k_0)^{-\frac{1}{2}}\{u_+(k) \exp[i(\mathbf{k}\cdot\mathbf{x})]-u_+^*(k) \exp[-i(\mathbf{k}\cdot\mathbf{x})]\}, \\ B_-(x) &= V^{-\frac{1}{2}} \sum_k (2k_0)^{-\frac{1}{2}}\{-u_-^*(k) \exp[i(\mathbf{k}\cdot\mathbf{x})]+u_-(k) \exp[i(\mathbf{k}\cdot\mathbf{x})]\}. \end{aligned} \tag{87}$$

If we introduce for every value of k, the variables $p_+(k)$, $q_+(k)$, $p_-(k)$, $q_-(k)$ by (30), (31), and identify $a(k)$, $b(k)$ and their adjoints given by (37), (38) analogous to (85) with

$$\begin{aligned} a(k) &= A(k) \exp(-ik_0x_0), \quad & a^*(k) &= A^*(k) \exp(ik_0x_0), \\ b(k) &= B(k) \exp(ik_0x_0), & b^*(k) &= B^*(k) \exp(-ik_0x_0), \end{aligned} \tag{88}$$

then the quantities $A_+(x)$, $A_-(x)$, $B_+(x)$, $B_-(x)$ can be written

$$A_+(x)=V^{-\frac{1}{2}} \sum_k k_0^{-\frac{1}{2}}[p_+(k) \cos(\mathbf{k}\cdot\mathbf{x})+q_+(k) \sin(\mathbf{k}\cdot\mathbf{x})], \tag{89}$$

$$A_-(x)=V^{-\frac{1}{2}} \sum_k k_0^{-\frac{1}{2}}[p_-(k) \cos(\mathbf{k}\cdot\mathbf{x})+q_-(k) \sin(\mathbf{k}\cdot\mathbf{x})], \tag{90}$$

$$B_+(x)=V^{-\frac{1}{2}} \sum_k k_0^{-\frac{1}{2}}i[-q_+(k) \cos(\mathbf{k}\cdot\mathbf{x})+p_+(k) \sin(\mathbf{k}\cdot\mathbf{x})], \tag{91}$$

$$B_-(x)=V^{-\frac{1}{2}} \sum_k k_0^{-\frac{1}{2}}i[q_-(k) \cos(\mathbf{k}\cdot\mathbf{x})-p_-(k) \sin(\mathbf{k}\cdot\mathbf{x})]. \tag{92}$$

Further, if we define the self-adjoint quantities $p_a(k)$, $q_a(k)$, $p_b(k)$, $q_b(k)$ by (45), (46), (49), and (50), we obtain

$$A(x)=V^{-\frac{1}{2}} \sum_k k_0^{-\frac{1}{2}}[p_a(k) \cos(\mathbf{k}\cdot\mathbf{x})+q_a(k) \sin(\mathbf{k}\cdot\mathbf{x})], \tag{93}$$

$$B(x)=V^{-\frac{1}{2}} \sum_k k_0^{-\frac{1}{2}}[-q_b(k) \cos(\mathbf{k}\cdot\mathbf{x})+p_b(k) \sin(\mathbf{k}\cdot\mathbf{x})]. \tag{94}$$

The energy and momentum can also be written in the new variables as

$$E=V^{-\frac{1}{2}} \sum_k k_0\tfrac{1}{2}[p_a^2(k)+q_a^2(k)+p_b^2(k)+q_b^2(k)]=V^{-\frac{1}{2}} \sum_k k_0\tfrac{1}{2}[p_+^2(k)+q_+^2(k)+p_-^2(k)+q_-^2(k)], \tag{95}$$

$$\mathbf{G}=V^{-\frac{1}{2}} \sum_k \mathbf{k}\tfrac{1}{2}[p_a^2(k)+q_a^2(k)+p_b^2(k)+q_b^2(k)]=V^{-\frac{1}{2}} \sum_k \mathbf{k}\tfrac{1}{2}[p_+^2(k)+q_+^2(k)+p_-^2(k)+q_-^2(k)]. \tag{96}$$

One obtains the analogous quantization of Maxwell's equation by putting $\mu=0$ and by substituting for the scalar fields four-vector fields $A_\nu(x)$, $B_\nu(x)$, $U_\nu(x)$, $U_\nu^*(x)$, ($\nu=1, 2, 3$ and 4 or 0) with the same kind of connection

$$A_\nu(x)=\frac{1}{\sqrt{2}}[U_\nu(x)+U_\nu^*(x)], \quad B_\nu(x)=\frac{1}{\sqrt{2}}[U_\nu(x)-U_\nu^*(x)]. \tag{97}$$

---

[6] The quantities here called $A_+(x)$, $A_-(x)$ are called $A^1(x)$, $A^2(x)$ in Dirac's paper.

In the definition of energy and momentum one has to take into account the definition of the scalar product of four-vectors by

$$\sum_{\nu=1}^{4} A_\nu B_\nu = -A_0 B_0 + \sum_{j=1}^{3} A_j B_j,$$

and to replace in the expressions of energy and momentum of the scalar theory the squares of field function by the corresponding invariant total squares of four-vectors. Moreover, one has to add the well-known extra conditions of Fermi for all fields

$$\sum_{\nu=1}^{4} \left(\frac{\partial A_\nu}{\partial x_\nu}\right)\Psi = \sum_{\nu=1}^{4} \left(\frac{\partial B_\nu}{\partial x_\nu}\right)\Psi = \sum_{\nu=1}^{4} \left(\frac{\partial U_\nu}{\partial x_\nu}\right)\Psi = \sum_{\nu=1}^{4} \left(\frac{\partial U_\nu^*}{\partial x_\nu}\right)\Psi = 0. \tag{98}$$

Although until now the theory is symmetric with respect to the N_a, N_b on the one side and the N_+, N_- on the other side, this is no longer so if we take the interaction with charged particles into account. Dirac assumes that only the field $A_\nu(x)$ and not $B_\nu(x)$ occurs in the interaction energy with the consequence that the numbers $N_b(k)$ stay constant with time. On the other hand he assumes that at least for sufficiently large values of k the $N_+(k)$ and $N_-(k)$ are zero, entirely different from the $N_a(k)$ and $N_b(k)$. This makes it more convenient to use the first kind of variables in the computations in spite of the redundant character of the N_b.

(b) Charged Particles[7]

In view of the interaction of mesotrons with heavy particles (protons and neutrons), it is of particular interest to consider also fields describing charged particles instead of photons. We are dealing here with the simplest case of spin 0 particles, for the treatment of charged particles with spin 1 can be done in an analogous way.

In the usual theory of Pauli and Weisskopf one introduces a complex field $\psi(x)$ (instead of the real field $A(x)$ of the preceding section) which again satisfies the wave equation

$$(\Box + \mu^2)\psi = 0,$$

and the Fourier decomposition of which can be written

$$\begin{aligned} \psi(x) &= V^{-\frac{1}{2}} \textstyle\sum_k (2k_0)^{-\frac{1}{2}} [A_p(k) \exp\{i(\mathbf{k}\cdot\mathbf{x} - k_0 x_0)\} + A_n^*(k) \exp\{i(-\mathbf{k}\cdot\mathbf{x} + k_0 x_0)\}], \\ \psi^*(x) &= V^{-\frac{1}{2}} \textstyle\sum_k (2k_0)^{-\frac{1}{2}} [A_p^*(k) \exp\{i(-\mathbf{k}\cdot\mathbf{x} + k_0 x_0)\} + A_n(k) \exp\{i(\mathbf{k}\cdot\mathbf{x} - k_0 x_0)\}]. \end{aligned} \tag{99}$$

The commutation rules for the quantities $A_p(k)$, $A_n(k)$ with their conjugates are

$$[A_p(k), A_p^*(k)] = 1, \quad [A_n(k), A_n^*(k)] = 1. \tag{100}$$

We remember[8] that

$$N_{p,a}(k) = A_p^*(k) A_p(k), \quad N_{n,a}(k) = A_n^*(k) A_n(k), \tag{101}$$

represent in this theory the number of particles with positive and negative electric charge, respectively.

Analogous to Eq. (71′) we now introduce the second redundant field $\varphi(x)$ defined by

$$\begin{aligned} \varphi(x) &= V^{-\frac{1}{2}} \textstyle\sum_k (2k_0)^{-\frac{1}{2}} [B_n^*(k) \exp\{i(\mathbf{k}\cdot\mathbf{x} - k_0 x_0)\} - B_p(k) \exp\{i(-\mathbf{k}_0\cdot\mathbf{x} + k_0 x_0)\}], \\ \varphi^*(x) &= V^{-\frac{1}{2}} \textstyle\sum_k (2k_0)^{-\frac{1}{2}} [B_n(k) \exp\{i(-\mathbf{k}\cdot\mathbf{x} + k_0 x_0)\} - B_p^*(k) \exp\{i(\mathbf{k}\cdot\mathbf{x} - k_0 x_0)\}], \end{aligned} \tag{102}$$

[7] Compare A, Part II, 1, 2.
[8] A, Part II, Eqs. (11), (11*), (19), and (20).

and quantized according to

$$[B_p(k), B_p^*(k)]=-1, \qquad [B_n(k), B_n^*(k)]=-1, \tag{103}$$

$$B_p^*(k)B_p(k)=-N_{p,b}(k), \qquad B_n^*(k)B_n(k)=-N_{n,b}(k). \tag{104}$$

The energy and momentum are given by

$$\begin{aligned} E&=\int\left[\left\{\frac{\partial\psi^*}{\partial x_0}\frac{\partial\psi}{\partial x_0}+\nabla\psi^*\nabla\psi+\mu^2\psi^*\psi\right\}+\left\{\frac{\partial\varphi^*}{\partial x_0}\frac{\partial\varphi}{\partial x_0}+\nabla\varphi^*\nabla\varphi+\mu^2\varphi^*\varphi\right\}\right]dV,\\ \mathbf{G}&=-\int\left[\left\{\frac{\partial\psi^*}{\partial x_0}\nabla\psi+\frac{\partial\psi}{\partial x_0}\nabla\psi\right\}+\left\{\frac{\partial\varphi^*}{\partial x_0}\nabla\varphi+\frac{\partial\varphi}{\partial x_0}\nabla\varphi^*\right\}\right]dV. \end{aligned} \tag{105}$$

Here and in the following, one always has to symmetrize the order of all non-commuting factors. One derives from these expressions

$$\begin{aligned} E&=\textstyle\sum_k k_0[N_{p,a}(k)+N_{n,a}(k)-N_{p,b}(k)-N_{n,b}(k)],\\ \mathbf{G}&=\textstyle\sum_k \mathbf{k}[N_{p,a}(k)+N_{n,a}(k)-N_{p,b}(k)-N_{n,b}(k)]. \end{aligned} \tag{106}$$

For the current vector $s_\nu(\nu=1, 2, 3$ and 4 or 0, $s_4=is_0)$ one gets

$$s_\nu=ie\left[\left\{\frac{\partial\psi^*}{\partial x_\nu}\psi-\frac{\partial\psi}{\partial x_\nu}\psi^*\right\}+\left\{\frac{\partial\varphi^*}{\partial x_\nu}\varphi-\frac{\partial\varphi}{\partial x_\nu}\varphi^*\right\}\right]. \tag{107}$$

Hence for the total electric charge

$$Q=\int s_0 dV=e\textstyle\sum_k[N_{p,a}(k)-N_{n,a}(k)+N_{p,b}(k)-N_{n,b}(k)]. \tag{108}$$

The decomposition of the fields $\psi(x)$, $\varphi(x)$ analogous to the former decomposition of $A(x)$ in $U(x)$ and its adjoint $U^*(x)$ [see (75), (75′)] has to be done according to

$$\begin{aligned} \psi(x)&=\frac{1}{\sqrt{2}}[U_p(x)+U_n^*(x)], \qquad \varphi(x)=\frac{1}{\sqrt{2}}[U_p(x)-U_n^*(x)],\\ \psi^*(x)&=\frac{1}{\sqrt{2}}[U_p^*(x)+U_n(x)], \qquad \varphi^*(x)=\frac{1}{\sqrt{2}}[U_p^*(x)-U_n(x)], \end{aligned} \tag{109}$$

where

$$\begin{aligned} U_p(x)&=V^{-\frac{1}{2}}\textstyle\sum_k(2k_0)^{-\frac{1}{2}}[U_{p,+}(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}+U_{p,-}(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}],\\ U_p^*(x)&=V^{-\frac{1}{2}}\textstyle\sum_k(2k_0)^{-\frac{1}{2}}[U_{p,+}^*(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}+U_{p,-}^*(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}], \end{aligned} \tag{110}$$

and in the same way

$$\begin{aligned} U_n(x)&=V^{-\frac{1}{2}}\textstyle\sum_k(2k_0)^{-\frac{1}{2}}[U_{n,+}(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}+U_{n,-}(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}],\\ U_n^*(x)&=V^{-\frac{1}{2}}\textstyle\sum_k(2k_0)^{-\frac{1}{2}}[U_{n,+}^*(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}+U_{n,-}^*(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}]. \end{aligned} \tag{111}$$

The quantization is

$$[U_{p,+}(k), U_{p,+}^*(k)]=[U_{n,+}(k), U_{n,+}^*(k)]=1, \tag{112}$$

$$[U_{p,-}(k), U_{p,-}^*(k)]=[U_{n,-}(k), U_{n,-}^*(k)]=-1. \tag{113}$$

Hence the $U_{p,-}(k)$, $U_{n,-}(k)$ have an indefinite metric in Hilbert space as the $B_p(k)$, $B_n(k)$; the

$U_{p,+}(k)$, $U_{n,+}(k)$ have a definite metric in Hilbert space as the $A_p(k)$, $A_n(k)$. Moreover, we have

$$\begin{aligned} U_{p,+}^*(k)U_{p,+}(k)&=N_{p,+}(k), \qquad U_{n,+}^*(k)U_{n,+}(k)=N_{n,+}(k),\\ U_{p,-}^*(k)U_{p,-}(k)&=-N_{p,-}(k), \qquad U_{n,-}^*(k)U_{n,-}(k)=-N_{n,-}(k), \end{aligned} \tag{114}$$

and by comparison of (110), (111) with (99), (102) we find, using (109)

$$\begin{aligned} A_p(k)&=\frac{1}{\sqrt{2}}[U_{p,+}(k)+U_{n,-}^*(k)], \qquad A_p^*(k)=\frac{1}{\sqrt{2}}[U_{p,+}^*(k)+U_{n,-}(k)],\\ A_n(k)&=\frac{1}{\sqrt{2}}[U_{n,+}(k)+U_{p,-}^*(k)], \qquad A_n^*(k)=\frac{1}{\sqrt{2}}[U_{n,+}^*(k)+U_{p,-}(k)], \end{aligned} \tag{115}$$

$$\begin{aligned} B_p(k)&=\frac{1}{\sqrt{2}}[-U_{p,-}(k)+U_{n,+}^*(k)], \qquad B_p^*(k)=\frac{1}{\sqrt{2}}[-U_{p,-}^*(k)+U_{n,+}(k)],\\ B_n(k)&=\frac{1}{\sqrt{2}}[U_{p,+}^*(k)-U_{n,-}(k)], \qquad B_n^*(k)=\frac{1}{\sqrt{2}}[U_{p,+}(k)-U_{n,-}^*(k)]. \end{aligned} \tag{116}$$

Hence

$$N_{p,a}(k)-N_{n,b}(k)=N_{p,+}(k)-N_{n,-}(k), \quad N_{n,a}(k)-N_{p,b}(k)=N_{n,+}(k)-N_{p,-}(k); \tag{117}$$

$$\begin{aligned} E&=\textstyle\sum_k k_0[N_{p,+}(k)+N_{n,+}(k)-N_{p,-}(k)-N_{n,-}(k)],\\ \mathbf{G}&=\textstyle\sum_k \mathbf{k}[N_{p,+}(k)+N_{n,+}(k)-N_{p,-}(k)-N_{n,-}(k)],\\ Q&=\textstyle\sum_k [N_{p,+}(k)-N_{n,+}(k)+N_{p,-}(k)-N_{n,-}(k)]. \end{aligned} \tag{118}$$

One can see from these expressions that the field describes four kinds of particles, with positive and negative charges and with positive and negative energy, respectively. Inserting (109) into (105) and (107) we find

$$E=\int\left[\left\{\frac{\partial U_p^*}{\partial x_0}\frac{\partial U_p}{\partial x_0}+\nabla U_p^* U_p+\mu^2 U_p^* U_p\right\}+\left\{\frac{\partial U_n^*}{\partial x_0}\frac{\partial U_n}{\partial x_0}+\nabla U_n^*\nabla U_n+\mu^2 U_n^* U_n\right\}\right]dV,$$

$$\mathbf{G}=-\int\left[\left\{\frac{\partial U_p^*}{\partial x_0}\nabla U_p+\frac{\partial U_p}{\partial x_0}U_p^*\right\}+\left\{\frac{\partial U_n^*}{\partial x_0}\nabla U_n+\frac{\partial U_n}{\partial x_0}\nabla U_n^*\right\}\right]dV, \tag{119}$$

$$s_\nu=ie\left[\left\{\frac{\partial U_p^*}{\partial x_\nu}U_p-\frac{\partial U_p}{\partial x_\nu}U_p^*\right\}-\left\{\frac{\partial U_n^*}{\partial x_\nu}U_n-\frac{\partial U_n}{\partial x_\nu}U_n^*\right\}\right]. \tag{120}$$

In an external electromagnetic field with the potentials $\varphi_\nu(x)$ one has to substitute for

$$\begin{aligned} \frac{\partial U_p}{\partial x_\nu}&\to\frac{\partial U_p}{\partial x_\nu}-ie\varphi_\nu U_p, \qquad \frac{\partial U_p^*}{\partial x_\nu}\to\frac{\partial U_p^*}{\partial x_\nu}+ie\varphi_\nu U_p^*,\\ \frac{\partial U_n}{\partial x_\nu}&\to\frac{\partial U_n}{\partial x_\nu}+ie\varphi_\nu U_n, \qquad \frac{\partial U_n^*}{\partial x_\nu}\to\frac{\partial U_n^*}{\partial x_\nu}-ie\varphi_\nu U_n^*. \end{aligned} \tag{121}$$

Therefore the functions U_p, U_n^* are multiplied under the gauge transformation $\varphi_\nu\to\varphi_\nu-(i/e)(\partial\alpha/\partial x_\nu)$ by $e^{i\alpha}$, while U_p^*, U_n are multiplied by $e^{-i\alpha}$. There is a fundamental difference between this form of the theory and the older theory of Pauli-Weisskopf. According to (99) the current vector s_ν due to the field $\psi(x)$ contains terms of the form $A_p(k)A_n(k')$ and $A_p^*(k)A_n^*(k')$ which correspond to the absorption and emission of pairs of particles with positive and negative electric charge, respectively. However, the expression (120) according to (110) and (111) does not contain terms of this kind but

commutes both with the total number of p and of n particles present. Therefore in this mathematical description pair generation and annihilation do not occur, but only transitions of a particle from positive to negative energy states as was the case in Dirac's original theory of the electron with spin $\frac{1}{2}$. The physical interpretation of this description is discussed below in Section 7. We notice here that of course there is in general also a change in the number of charged particles in the new mathematical description as soon as there is an interaction of the charged particles with other charged particles (for instance of charged mesons with protons and neutrons).

For the sake of completeness we indicate here the possibility of other variables which are sometimes convenient and which comes from the separation of $\psi(x)$ into its real and imaginary parts

$$\psi_\alpha(x)=\psi_\alpha^*(x), \quad \alpha=1, 2 \tag{122}$$

according to

$$\psi(x)=\frac{1}{\sqrt{2}}[\psi_1(x)-i\psi_2(x)], \quad \psi^*(x)=\frac{1}{\sqrt{2}}[\psi_1(x)+i\psi_2(x)], \tag{123}$$

and analogously

$$\varphi_\alpha(x)=-\varphi_\alpha^*(x), \tag{122'}$$

$$\varphi(x)=\frac{1}{\sqrt{2}}[\varphi_1(x)-i\varphi_2(x)], \quad \varphi^*(x)=\frac{1}{\sqrt{2}}[\varphi_1(x)+i\varphi_2(x)]. \tag{123'}$$

Hence

$$E=\sum_{\alpha=1,2}\frac{1}{2}\int\left[\left\{\left(\frac{\partial\psi_\alpha}{\partial x_0}\right)^2+(\nabla\psi_\alpha)^2+\mu^2\psi_\alpha^2\right\}-\left\{\left(\frac{\partial\varphi_\alpha}{\partial x_0}\right)^2+(\nabla\varphi_\alpha)^2+\mu^2\varphi_\alpha^2\right\}\right]dV,$$

$$\mathbf{G}=\sum_{\alpha=1,2}\frac{1}{2}\int\left[-\frac{\partial\psi_\alpha}{\partial x_0}\nabla\psi_\alpha+\frac{\partial\varphi_\alpha}{\partial x_0}\nabla\varphi_\alpha\right]dV, \tag{124}$$

$$s_\nu=e\left[\left\{\frac{\partial\psi_1}{\partial x_\nu}\psi_2-\frac{\partial\psi_2}{\partial x_\nu}\psi_1\right\}-\left\{\frac{\partial\varphi_1}{\partial x_\nu}\varphi_2-\frac{\partial\varphi_2}{\partial x_\nu}\varphi_1\right\}\right]. \tag{125}$$

If we put in a similar way

$$U_p(x)=\frac{1}{\sqrt{2}}[U_1(x)-iU_2(x)], \quad U_p^*(x)=\frac{1}{\sqrt{2}}[U_1^*(x)+iU_2^*(x)],$$

$$U_n(x)=\frac{1}{\sqrt{2}}[U_1(x)+iU_2(x)], \quad U_n^*(x)=\frac{1}{\sqrt{2}}[U_1^*(x)-iU_2^*(x)], \tag{126}$$

we obtain, in view of (109),

$$\psi_\alpha(x)=\frac{1}{\sqrt{2}}[U_\alpha(x)+U_\alpha^*(x)], \qquad \varphi_\alpha(x)=\frac{1}{\sqrt{2}}[U_\alpha(x)-U_\alpha^*(x)], \tag{127}$$

and the corresponding expressions for energy, momentum, and current vector are

$$E=\sum_{\alpha=1,2}\int\left[\frac{\partial U_\alpha^*}{\partial x_0}\frac{\partial U_\alpha}{\partial x_0}+\nabla U_\alpha^*\nabla U_\alpha+\mu^2U_\alpha^*U_\alpha\right]dV,$$

$$\mathbf{G}=-\sum_{\alpha=1,2}\int\left[\frac{\partial U_\alpha^*}{\partial x_0}\nabla U_\alpha+\frac{\partial U_\alpha}{\partial x_0}\nabla U_\alpha^*\right]dV, \tag{128}$$

$$s_\nu=e\left[\frac{\partial U_1^*}{\partial x_\nu}U_2-\frac{\partial U_2^*}{\partial x_\nu}U_1+\frac{\partial U_1}{\partial x_\nu}U_2^*-\frac{\partial U_2}{\partial x_\nu}U_1^*\right]. \tag{129}$$

We emphasize, however, that there exists a different scheme for the description of the fields corresponding to charged particles which is more similar to the older Pauli-Weisskopf theory. We decompose the field according to

$$\psi(x)=\frac{1}{\sqrt{2}}[U_+(x)+U_-(x)] \tag{130}$$

with

$$\begin{aligned} U_+(x)&=V^{-\frac{1}{2}}\sum_k(2k_0)^{-\frac{1}{2}}[U_{p,+}(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}+U_{n,+}{}^*(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}],\\ U_-(x)&=V^{-\frac{1}{2}}\sum_k(2k_0)^{-\frac{1}{2}}[U_{p,-}(k)\exp\{i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\}+U_{n,-}{}^*(k)\exp\{i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)\}], \end{aligned} \tag{131}$$

and define the redundant field different from $\varphi(x)$ by

$$\chi(x)=\frac{1}{\sqrt{2}}[U_+(x)-U_-(x)]. \tag{132}$$

The energy, momentum, and current vector are given by

$$\begin{aligned} E&=\int\left[\left\{\frac{\partial\psi^*}{\partial x_0}\frac{\partial\psi}{\partial x_0}+\nabla\psi^*\nabla\psi+\mu^2\psi^*\psi\right\}+\left\{\frac{\partial\chi^*}{\partial x_0}\frac{\partial\chi}{\partial x_0}+\nabla\chi^*\nabla\chi+\mu^2\chi^*\chi\right\}\right]dV,\\ \mathbf{G}&=-\int\left[\left\{\frac{\partial\psi^*}{\partial x_0}\nabla\psi+\frac{\partial\psi}{\partial x_0}\nabla\psi^*\right\}+\left\{\frac{\partial\chi^*}{\partial x_0}\nabla\chi+\frac{\partial\chi}{\partial x_0}\nabla\chi^*\right\}\right]dV, \end{aligned} \tag{133}$$

$$s_\nu=ie\left[\left\{\frac{\partial\psi^*}{\partial x_\nu}\psi-\frac{\partial\psi}{\partial x_\nu}\psi^*\right\}+\left\{\frac{\partial\chi^*}{\partial x_\nu}\chi-\frac{\partial\chi}{\partial x_\nu}\chi^*\right\}\right]; \tag{134}$$

or else by

$$\begin{aligned} E&=\int\left[\left\{\frac{\partial U_+{}^*}{\partial x_0}\frac{\partial U_+}{\partial x_0}+\nabla U_+{}^*\nabla U_++\mu^2U_+{}^*U_+\right\}+\left\{\frac{\partial U_-{}^*}{\partial x_0}\frac{\partial U_-}{\partial x_0}+\nabla U_-{}^*\nabla U_-+\mu^2U_-{}^*U_-\right\}\right]dV,\\ \mathbf{G}&=-\int\left[\left\{\frac{\partial U_+{}^*}{\partial x_0}\nabla U_++\frac{\partial U_+}{\partial x_0}\nabla U_+{}^*\right\}+\left\{\frac{\partial U_-{}^*}{\partial x_0}\nabla U_-+\frac{\partial U_-}{\partial x_0}\nabla U_-{}^*\right\}\right]dV, \end{aligned} \tag{135}$$

$$s_\nu=ie\left[\left\{\frac{\partial U_+{}^*}{\partial x_\nu}U_+-\frac{\partial U_+}{\partial x_\nu}U_+{}^*\right\}+\left\{\frac{\partial U_-{}^*}{\partial x_\nu}U_--\frac{\partial U_-}{\partial x_\nu}U_-{}^*\right\}\right]. \tag{136}$$

The expressions (118) for E, $\mathbf{G}$, and Q still hold, but the current is essentially different. It contains, according to (131), no operators corresponding to transitions from positive to negative energy states of a particle, but operators corresponding to generation and annihilation of pairs with opposite charges as in the older theory. We do not need to give the details for the decomposition of the self-adjoint fields $\psi_1(x)$, $\psi_2(x)$, defined by (123), into its self-adjoint parts $U_{\alpha,+}(x)$, $U_{\alpha,-}(x)$ with positive and negative energies which has to be done in a way quite analogous to the above development. The

corresponding expression for the current is

$$s_\nu = e\left[\left\{\frac{\partial U_{1,+}}{\partial x_\nu}U_{2,+} - \frac{\partial U_{2,+}}{\partial x_\nu}U_{1,+}\right\} + \left\{\frac{\partial U_{1,-}}{\partial x_\nu}U_{2,-} - \frac{\partial U_{2,-}}{\partial x_\nu}U_{1,-}\right\}\right]. \tag{137}$$

In an external field one has to substitute both for the + and the − part

$$\frac{\partial U_\pm}{\partial x_\nu} \to \frac{\partial U_\pm}{\partial x_\nu} - ie\varphi_\nu U_\pm, \qquad \frac{\partial U_\pm{}^*}{\partial x_\nu} \to \frac{\partial U_\pm{}^*}{\partial x_\nu} + ie\varphi_\nu U_\pm{}^*, \tag{138}$$

the gauge transformation being

$$\varphi_\nu = \varphi_\nu - \frac{i}{e}\frac{\partial \alpha}{\partial x_\nu}, \quad U_\pm \to U_\pm e^{i\alpha}, \quad U_\pm{}^* \to U_\pm{}^* e^{-i\alpha}. \tag{139}$$

A more detailed consideration of the two alternative forms of the theory of charged particles has to be postponed to the following section, but we may add here some preliminary remarks. The zero-point energy of the vacuum of the positive and the negative energy part of the field cancels exactly also in an external electromagnetic field; therefore all effects due to the so-called polarizability of the vacuum disappear. On the other hand, the self-energy, both of the vacuum and of a single particle due to the electromagnetic interaction of particles, gives rise to infinities in this second form of the theory while the divergences are completely eliminated in the former alternative of a theory with a fixed number of particles in the mathematical description. Just as the former alternative has, in the case of the ordinary electron with spin $\frac{1}{2}$, its analogue in the original form of Dirac's theory of electron, the latter alternative has its analogue in a slightly modified form of the theory of holes. We have to bear in mind that the principle of defining the vacuum as the state of lowest energy in the mathematical description is already abandoned by defining the photon vacuum as given by $N_+ = N_- = 0$. Therefore (and this will become more obvious in Section 7, dealing with the physical interpretation of the theory), we introduce besides the ordinary field $u_p(x)$ (the spinor index is suppressed) which fulfills the Dirac equation, a second field $u_n(x)$ which also fulfills this equation. We assume, however, that in an external electromagnetic field one has for the first field to introduce the substitutions

$$\frac{\partial u_p}{\partial x_\nu} \to \frac{\partial u_p}{\partial x_\nu} - ie\varphi_\nu u_p, \qquad \frac{\partial u_p{}^*}{\partial x_\nu} \to \frac{\partial u_p{}^*}{\partial x_\nu} + ie\varphi_\nu u_p{}^*, \tag{140}$$

and for the second field

$$\frac{\partial u_n}{\partial x_\nu} \to \frac{\partial u_n}{\partial x_\nu} + ie\varphi_\nu u_n, \qquad \frac{\partial u_n{}^*}{\partial x_\nu} \to \frac{\partial u_n{}^*}{\partial x_\nu} - ie\varphi_\nu u_n{}^*. \tag{141}$$

The total current vector s_ν is therefore given by

$$s_\nu = e[(u_p{}^* \alpha_\nu u_p) - (u_n{}^* \alpha_\nu u_n)] \tag{142}$$

where $\alpha_4 = iI$ and $\alpha_1, \alpha_2, \alpha_3$ are the spatial part of the Dirac matrices. The "vacuum" analogous to that given by $N_+ = N_- = 0$ in the theory for Bose particles is obtained by the following assumption: *All negative energy states of the p particles (the $u_p(x)$ field) shall be occupied and all positive energy states of the n particles (the $u_n(x)$ field) shall be unoccupied.* Using the principle of charge-conjugate solutions it is easy to prove that there is no zero-point energy nor zero-point charge in this picture even if an external electromagnetic field is present. Hence the "polarizability of the vacuum" drops entirely. On the other hand infinite self-energies exist here again just as in the analogous picture for charged Bose-particles due to the electromagnetic interaction, because the p and the n particles have to be treated as different particles for the application of the exclusion principle. (See Section 7.)

6. THE INTERACTION OF ELECTRONS WITH AN ELECTROMAGNETIC FIELD

(a) The λ-Limiting Process. Electrostatic Self-Energy

We develop the interaction of electrons with an electromagnetic field in two steps: first introducing the so-called λ-limiting process, and then the negative energy photons in accordance with the previous sections. The former, which is due to Wentzel,[9] was originally a purely classical method to avoid the singularities due to the classical electromagnetic self-energy of a point source in a relativistically invariant way without altering the validity of Maxwell's equations near the source. It was based on the formalism of Dirac-Fock-Podolsky, in which separate time coordinates are used for each electric particle present and for the electromagnetic field. Wentzel showed that a translation of this method into quantum theory can easily be made as long as one does not introduce explicitly Dirac's theory of holes, but uses the original form of Dirac's theory, where the number of charged particles present is conserved and transitions to negative energy states for this particle are possible. We shall discuss later this restriction, which seems to be a matter of principle, and use here too the original form of Dirac's theory. Later it was shown by Dirac,[10] that the procedure can more easily be described by a change of the commutation relation between the potentials according to[11]

$$i[A_\mu(\mathbf{x}, x_0), A_\nu(\mathbf{x}, x_0)] = \tfrac{1}{2}\delta_{\mu\nu}[D(\mathbf{x}-\mathbf{x}'+\boldsymbol{\lambda}, x_0-x_0'+\lambda_0)+D(\mathbf{x}-\mathbf{x}'-\boldsymbol{\lambda}, x_0-x_0'-\lambda_0)], \quad (143)$$

where the rest mass of the photon in the D function has to be put equal to zero and where the new four-vector λ_ν or $\lambda_0, \boldsymbol{\lambda}$ has to be timelike,

$$\lambda_0^2 > \boldsymbol{\lambda}^2. \quad (144)$$

In the final result one goes to the limit $\lambda_0, \boldsymbol{\lambda} \to 0$ again, in such a way that (144) is always fulfilled. For finite λ the results are relativistically invariant only if λ is also transformed (like a vector) while its influence on the results drops in the limit $\lambda \to 0$. We may decompose into Fourier components according to

$$A_\mu(\mathbf{x}, x_0) = V^{-\frac{1}{2}}\textstyle\sum_k (2k_0)^{-\frac{1}{2}}[a_\mu(k, x_0)\exp[i(\mathbf{k}\cdot\mathbf{x})]+a_\mu^*(k, x_0)\exp[-i(\mathbf{k}\cdot\mathbf{x})]], \quad (145)$$

$$\frac{dA_\mu}{dx_0}(\mathbf{x}, x_0) = V^{-\frac{1}{2}}\textstyle\sum_k (2k_0)^{-\frac{1}{2}}ik_0[-a_\mu(k, x_0)\exp[i(\mathbf{k}\cdot\mathbf{x})]+a_\mu^*(k, x_0)\exp[-i(\mathbf{k}\cdot\mathbf{x})]]. \quad (146)$$

The commutation relations (143) are at a certain instant x_0, equivalent to

$$[a_\mu(k, x_0), a_\nu^*(k', x_0)] = \delta_{\mu\nu}\delta_{kk'}\cos(\lambda_0 k_0 - \boldsymbol{\lambda}\cdot\mathbf{k}). \quad (147)$$

While (143) holds also for $x_0 \neq x_0'$ in the case of interaction of the field with matter, if only the particle time is the same in both field quantities, the latter relation, which is already specialized by the assumption $x_0 = x_0'$ does not require the use of different time variables. These are very convenient for putting relativistic invariance of the method into evidence. On the other hand it may be an amendment of the existing literature, to which we refer for this proof, to give also a formulation of the method with only a single time variable and a Hamiltonian H, which is really the energy and which determines the time dependence of all observables F according to the well-known relation [compare (7)]

$$dF/dx_0 = i(HF - FH). \quad (148)$$

We shall use here the latter method. In order to find the necessary changes in the usual expression for

[9] G. Wentzel, Zeits. f. Physik **86**, 479 and 635 (1933); **87**, 726 (1934).
[10] P. A. M. Dirac, Ann. de l'Inst. Poincaré **9**, 13 (1939).
[11] We use here the notations of paper A. The indices μ, ν run from 1 to 4, where the last one corresponds to the imaginary coordinate $x_4 = ix_0$. The vector or tensor components with index zero always fulfill the relation $A\ldots_4\ldots = iA\ldots_0\ldots$. Hence the components with index zero are contravariant as x_0. Moreover we adhere to the convention, used in A, that in the case of a complex vector or tensor, $A^*\ldots_4\ldots = iA^*\ldots_0\ldots$, the latter quantity being the actual conjugate-complex, or self-adjoint.

H due to the λ-process, we notice first that the eigenvalues of $a_\mu^*(k)a_\mu(k)$ due to (147) are given by

$$a_\mu^*(k)a_\mu(k)=N_\mu(k)\cos(\lambda_0 k_0-\mathfrak{\lambda}\cdot\mathbf{k}) \tag{149}$$

where $k_0=|\mathbf{k}|$, the rest mass of the photons being zero, and the $N_\mu(k)$ are zero or positive integers for $\mu=1, 2, 3$, but negative integers for $\mu=4$. [Notice that $-a_0^*a_0=a_4^*a_4$ has negative eigenvalues, which are given by $(-1, -2, -3, \cdots)\cos(\lambda_0 k_0-\mathfrak{\lambda}\cdot\mathbf{k})$ according to (147).] This statement holds also for a negative sign of the cosine factor in (149), if we define in this case $a_u^*(k)$ not as the Hermitian conjugate of $a_u(k)$ but as the self-adjoint analogous to Eqs. (19), (20).

The Hamiltonian H consists of three parts, the energy of the free radiation field H_0, the energy of the free electrons H_m, and the interaction energy Ω,

$$H=H_0+H_m+\Omega.$$

First we have to write

$$H_0=\sum_k k_0[\sum_\mu N_\mu(k)+\tfrac{1}{2}]=\sum_k k_0[\cos(\lambda_0 k_0-\mathfrak{\lambda}\cdot\mathbf{k})]^{-1}\tfrac{1}{2}\sum_\mu[a_\mu^*(k)a_\mu(k)+a_\mu(k)a_\mu^*(k)]. \tag{150}$$

The factor $[\cos(\lambda_0 k_0-\mathfrak{\lambda}\cdot\mathbf{k})]^{-1}$ is rather striking, but unavoidable in order to stay in accordance with (146) and the rule (148) applied to $A_\mu(x)$ in view of (147). In the formulation of the λ-process with the help of several time variables this circumstance is not obvious because one then does not need H_0 explicitly. The energy of matter for a set of spin $\frac{1}{2}$ electrons characterized by a running index n is given by

$$H_m=\sum_n[(\boldsymbol{\alpha}^{(n)}\cdot\mathbf{p}^{(n)})+m\beta^{(n)}] \tag{151}$$

where the usual Dirac matrices $\boldsymbol{\alpha}$, β are introduced. The interaction energy is

$$\begin{aligned}\Omega&=(4\pi)^{\frac{1}{2}}e\sum_n[-A_0(z_n)+\boldsymbol{\alpha}^{(n)}\cdot\mathbf{A}(z_n)]\\&=\left(\frac{4\pi}{V}\right)^{\frac{1}{2}}e\sum_n\sum_k(2k_0)^{-\frac{1}{2}}[\{-a_0(k)+\boldsymbol{\alpha}^{(n)}\cdot\mathbf{a}(k)\}\exp\{i(\mathbf{k}\cdot\mathbf{z}_n)\}\\&\qquad+\{-a_0^*(k)+\boldsymbol{\alpha}^{(n)})\cdot\mathbf{a}(k)\}\exp\{-i(\mathbf{k}\cdot\mathbf{z}_n)\}].\end{aligned} \tag{152}$$

Here z_n is the place of the nth particle; the sign is chosen in such a way that for electrons with negative charge e is the positive absolute value of the charge, and the factor $(4\pi)^{\frac{1}{2}}$ is added in order to measure it in the ordinary and not in the Heaviside units, while the potentials are, according to (150), measured in Heaviside units. One sees that (146) still holds as a consequence of (148) if the interaction energy (152) is taken into account, for the reason that the latter commutes with $A_\mu(x)$. The momentum $\mathbf{p}^{(n)}$ fulfills the usual commutation rule

$$i[p_i^{(n)}, z_j^{(m)}]=\delta_{ij}\delta_{nm}, \quad (i, j=1, 2, 3). \tag{153}$$

The correct generalization of the subsidiary condition (98) if the interaction energy is taken into account, is given in the k space by

$$X(k)\Psi=0 \quad \text{and} \quad X^*(k)\Psi=0 \tag{154}$$

where

$$X(k)=\mathbf{k}\cdot\mathbf{a}(k)-k_0a_0(k)-\left(\frac{4\pi}{V}\right)^{\frac{1}{2}}e(2k_0)^{-\frac{1}{2}}\sum_n\exp[-i(\mathbf{k}\cdot\mathbf{z}_n)]\cos(\lambda_0 k_0-\mathfrak{\lambda}\cdot\mathbf{k}). \tag{155}$$

The additional term on the right side of this expression is uniquely determined by the condition that $(dX/dx_0)\Psi=i[H, X]\Psi=0$ has to be a consequence of the conditions (154) themselves. Using (147), (153), one finds indeed

$$[H, X(k)]=-k_0X(k), \quad [H, X^*(k)]=k_0X^*(k). \tag{156}$$

Moreover one has

$$[X(k), X^*(k)]=0 \tag{156'}$$

in view of $k_0^2=\mathbf{k}^2$.

A further reduction of the problem is possible if one decomposes the field in a longitudinal and a transverse part (this decomposition depends of course on the coordinate system). The corresponding decomposition of $\mathbf{a}(k)$ is given by

$$\mathbf{a}(k)=\mathbf{a}^L(k)+\mathbf{a}^T(k),\quad \mathbf{a}^L(k)=(\mathbf{k}/k_0^2)(\mathbf{k}\cdot\mathbf{a}(k)),\quad \mathbf{a}^T(k)=\mathbf{a}(k)-(\mathbf{k}/k_0^2)(\mathbf{k}\cdot\mathbf{a}(k)), \tag{157}$$

with similar formulas for $\mathbf{a}^*(k)$. $X(k)$ contains only the part $\mathbf{a}^L(k)$ and commutes with $\mathbf{a}^{*T}(k)$. Furthermore one has

$$[a_i{}^T(k),\ a_j{}^{*T}(k)]=[\delta_{ij}-(k_ik_j/k_0^2)]. \tag{158}$$

Using

$$\sum_\mu a_\mu{}^*(k)a_\mu(k)=\mathbf{a}^{*T}(k)\cdot\mathbf{a}^T(k)-a_0{}^*(k)a_0(k)+\frac{1}{k_0^2}(\mathbf{k}\cdot\mathbf{a}^*(k))(\mathbf{k}\cdot\mathbf{a}(k)),$$

one decomposes H into two parts

$$\begin{aligned}H=\sum_k \frac{1}{k_0}[\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})]^{-1}\frac{1}{2}\Big[\Big\{&\mathbf{k}\cdot\mathbf{a}^*(k)+k_0a_0{}^*(k)\\&+\Big(\frac{4\pi}{V}\Big)^{\frac{1}{2}}e(2k_0)^{-\frac{1}{2}}\sum_n\exp[i(\mathbf{k}\cdot\mathbf{z}_n)]\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})\Big\}X(k)+\Big\{\mathbf{k}\cdot\mathbf{a}(k)+k_0a_0(k)\\&+\Big(\frac{4\pi}{V}\Big)^{\frac{1}{2}}e(2k_0)^{-\frac{1}{2}}\sum_n\exp[-i(\mathbf{k}\cdot\mathbf{z}_n)]\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})\Big\}X^*(k)\Big]+H'.\end{aligned} \tag{159}$$

Introducing the abbreviation

$$\begin{aligned}\mathbf{p}'^{(n)}&=\mathbf{p}^{(n)}+\Big(\frac{4\pi}{V}\Big)^{\frac{1}{2}}e\sum_k(2k_0)^{-\frac{1}{2}}\sum_n\{\mathbf{a}^L(k)\exp[i(\mathbf{k}\cdot\mathbf{z}_n)]+\mathbf{a}^{*L}(k)\exp[-i(\mathbf{k}\cdot\mathbf{z}_n)]\}\\&=\mathbf{p}^{(n)}+\Big(\frac{4\pi}{V}\Big)^{\frac{1}{2}}e\sum_n\frac{1}{i}\frac{\partial}{\partial\mathbf{z}_n}\sum_k(2k_0)^{-\frac{1}{2}}\Big\{\frac{(\mathbf{k}\cdot\mathbf{a}(k))}{k_0^2}\exp[i(\mathbf{k}\cdot\mathbf{z}_n)]-\frac{(\mathbf{k}\cdot\mathbf{a}^*(k))}{k_0^2}\exp[-i(\mathbf{k}\cdot\mathbf{z}_n)]\Big\},\end{aligned} \tag{160}$$

one obtains for H' the expression

$$\begin{aligned}H'=\sum_k k_0[\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})]^{-1}\tfrac{1}{2}[\mathbf{a}^{*T}(k)\cdot\mathbf{a}^T(k)+\mathbf{a}^T(k)\cdot\mathbf{a}^{*T}(k)]+\sum_n[(\boldsymbol{\alpha}^{(n)}\cdot\mathbf{p}'^{(n)})+m\beta^{(n)}]\\+\Big(\frac{4\pi}{V}\Big)^{\frac{1}{2}}e\sum_k(2k_0)^{-\frac{1}{2}}\sum_n[(\boldsymbol{\alpha}^{(n)}\cdot\mathbf{a}^T(k))\exp\{i(\mathbf{k}\cdot\mathbf{z}_n)\}+(\boldsymbol{\alpha}^{(n)}\cdot\mathbf{a}^{*T}(k))\exp\{-i(\mathbf{k}\cdot\mathbf{z}_n)\}]\\+\frac{4\pi}{V}e^2\sum_k\frac{1}{k_0^2}\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})\tfrac{1}{2}\sum_n\sum_m\exp[i\mathbf{k}\cdot(\mathbf{z}_n-\mathbf{z}_m)].\end{aligned} \tag{161}$$

The operators $\mathbf{p}'^{(n)}$, $X(k)$, $X^*(k)$ have the important property of commuting with each other:

$$[p_i{}'^{(n)},\ p_j{}'^{(m)}]=0,\quad [\mathbf{p}'^{(n)},\ X(k)]=0,\quad [\mathbf{p}'^{(n)},\ X^*(k)]=0. \tag{162}$$

Only such observables are of physical significance which commute with the subsidiary conditions, that is with $X(k)$, $X^*(k)$; and it is well known that these are gauge invariant quantities. The result (162) proves, for instance, that $\mathbf{p}'^{(n)}$ belongs to these quantities, but not $\mathbf{p}^{(n)}$. So long as we are interested only in observables of this kind, we are allowed to put explicitly $X(k)=X^*(k)=0$, which reduces the Hamiltonian H to its part H', and to use a wave function Ψ' of a complete set of variables of this kind which commute with each other. One can choose for them the coordinates $\mathbf{z}_{(n)}$ of the particles and the numbers $N_\lambda(k)$ of photons in the eigenvibrations of the transverse part of the field ($\lambda=1$, 2 for every $\mathbf{k}$ corresponding to the two possible polarizations). Moreover, the relation $i[p_i{}'^{(n)},\ z_{ni}]=1$ is

sufficient to put

$$\mathbf{p}'^{(n)}\Psi' = -i\frac{\partial}{\partial \mathbf{z}_n}\Psi'. \tag{163}$$[12]

This will be done in the following, where we can now simply omit all primes.

We now discuss the last term of the Hamiltonian (161), which is due to the longitudinal part of the field and is nothing but the electrostatic energy E^{st}. Inserting

$$\frac{1}{V}\sum_k \to \frac{1}{(2\pi)^3}\int d\mathbf{k}$$

and separating the terms $n=m$ from the terms $n\neq m$, we obtain

$$E^{st} = \sum_n E_0{}^{st} + \sum_{n<m} E_{nm}^{st}$$

where

$$E_0{}^{st} = \frac{1}{2}e^2\frac{4\pi}{(2\pi)^3}\int d\mathbf{k}k_0^{-2}\cos(k_0\lambda_0 - \mathbf{k}\cdot\boldsymbol{\lambda}), \tag{164}$$

$$E_{nm}^{st} = e^2\frac{4\pi}{(2\pi)^3}\int d\mathbf{k}k_0^{-2}\cos(k_0\lambda_0 - \mathbf{k}\cdot\boldsymbol{\lambda})\exp[i\mathbf{k}\cdot(\mathbf{z}_n - \mathbf{z}_m)]. \tag{165}$$

---

[12] In order to explain this in more detail we write

$$\mathbf{k}\cdot\mathbf{a}(k) - k_0 a_0(k) = k_0[Q_1(k) + iQ_2(k)],$$
$$\mathbf{k}\cdot\mathbf{a}^*(k) - k_0 a_0{}^*(k) = k_0[Q_1(k) - iQ_2(k)],$$

where $Q_1(k)$, $Q_2(k)$ are Hermitian and commute with each other. In accordance with the commutation relation (147) we define the Hermitian operators $P_1(k)$, $P_2(k)$ which fulfill

$$[P_1(k), P_2(k)] = 0, \quad i[P_1(k), Q_1(k)] = i[P_2(k), Q_2(k)] = \cos(k_0\lambda_0 - \mathbf{k}\cdot\boldsymbol{\lambda}),$$

by putting

$$\mathbf{k}\cdot\mathbf{a}(k) + k_0 a_0(k) = ik_0[P_1(k) + iP_2(k)],$$
$$\mathbf{k}\cdot\mathbf{a}(k) + k_0 a_0(k) = -ik_0[P_1(k) - iP_2(k)].$$

Furthermore we define the Hermitian operators $X_1(k)$, $X_2(k)$ by

$$X(k) = X_1(k) + iX_2(k), \quad X^*(k) = X_1(k) - iX_2(k),$$

and get from (155)

$$X_1(k) = k_0 Q_1(k) - \left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e(2k_0)^{-\frac{1}{2}}\cos(k_0\lambda_0 - \mathbf{k}\cdot\boldsymbol{\lambda})\cos(\mathbf{k}\cdot\mathbf{z}_n),$$

$$X_2(k) = k_0 Q_2(k) + \left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e(2k_0)^{-\frac{1}{2}}\cos(k_0\lambda_0 - \mathbf{k}\cdot\boldsymbol{\lambda})\sin(\mathbf{k}\cdot\mathbf{z}_n).$$

The expression (160) for $\mathbf{p}'^{(n)}$ has now the form

$$\mathbf{p}'^{(n)} = \mathbf{p}^{(n)} + \left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e\sum_k (2k_0)^{-\frac{1}{2}}\frac{\mathbf{k}}{k_0}[(-P_2 + Q_1)\cos(\mathbf{k}\cdot\mathbf{z}_n) - (P_1 + Q_2)\sin(\mathbf{k}\cdot\mathbf{z}_n)].$$

Assuming the operators $\mathbf{p}^{(n)}$, P_1, P_2 as usual as $-i\partial/\partial\mathbf{z}_n$, $-i\partial/\partial Q_1$, $-i\partial/\partial Q_2$, respectively, we now evaluate the applicatoin of $\mathbf{p}'^{(n)}$ to an arbitrary function of the $X_1(k)$, $X_2(k)$. The contributions of $\mathbf{p}^{(n)}$, P_1, P_2 cancel, and we are left with

$$\mathbf{p}'^{(n)}F(\cdots X_1(k)\cdots, \cdots X_2(k)\cdots) = F(\cdots X_1(k)\cdots, \cdots X_2(k)\cdots)\left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e\sum_k (2k_0)^{-\frac{1}{2}}\frac{\mathbf{k}}{k_0}[Q_1\cos(\mathbf{k}\cdot\mathbf{z}_n) - Q_2\sin(\mathbf{k}\cdot\mathbf{z}_n)].$$

For

$$\Psi(Q_1(k), Q_2(k), \mathbf{z}_n, N_\lambda(k)) = F(\cdots X_1(k)\cdots, \cdots X_2(k)\cdots)\exp\left[-i\left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e\sum_k (2k_0)^{-\frac{1}{2}}k_0^{-1}\right.$$
$$\left.\times\sum_n \{Q_1\sin(\mathbf{k}\cdot\mathbf{z}_n) + Q_2\cos(\mathbf{k}\cdot\mathbf{z}_n)\}\right]\Psi'(\mathbf{z}_n, N_\lambda(k)),$$

we get therefore

$$\mathbf{p}'^{(n)}\Psi = F\exp[\cdots]\left(-i\frac{\partial}{\partial\mathbf{z}_n}\right)\Psi'.$$

To fulfill the subsidiary conditions

$$X_1(k)\Psi = X_2(k)\Psi = 0,$$

we have to choose

$$F = \Pi_k\,\delta(X_1(k))\delta(X_2(k)).$$

The phase factor does not change the expectation values of any observables which commute with the $X_1(k)$, $X_2(k)$, for these are functions of $X_1(k)$, $X_2(k)$ themselves and of $\mathbf{z}_n$, $\mathbf{p}'^{(n)}$ and the transverse part of the field only. For these observables it is therefore allowed to use the part Ψ' of the wave function alone, and to identify the $X_1(k)$, $X_2(k)$ everywhere with zero, as was indicated in the text.

The integration over the directions in (164) leads to

$$E_0^{st}=\frac{e^2}{\pi}\int_0^\infty dk_0\cos(k_0\lambda_0)\frac{\sin(k_0|\mathfrak{\lambda}|)}{k_0|\mathfrak{\lambda}|}=\frac{e^2}{2\pi}\int_0^\infty dk_0\left[\frac{\sin k_0(\lambda_0+|\mathfrak{\lambda}|)}{k_0}-\frac{\sin k_0(\lambda_0-|\mathfrak{\lambda}|)}{k_0}\right].$$

The two integrals cancel if $\lambda_0^2>\mathfrak{\lambda}^2$ as was assumed in (144), thus leading to

$$E_0^{st}=0. \tag{166}$$

For $\lambda_0^2<\mathfrak{\lambda}^2$ they would add and give a contribution proportional to $|\mathfrak{\lambda}|^{-1}$. For $n\neq m$ it is sufficient to put $\mathfrak{\lambda}=0$ which is permitted according to (144). Then performing the integrations over the directions, and putting

$$r_{nm}=|\mathbf{z}_n-\mathbf{z}_m| \tag{167}$$

we obtain

$$E_{nm}^{st}=\frac{2}{\pi}e^2\int_0^\infty dk_0\cos(k_0\lambda_0)\frac{\sin(k_0r_{nm})}{k_0r_{nm}}=\frac{1}{\pi}\frac{e^2}{r_{nm}}\int_0^\infty dk_0\left[\frac{\sin k_0(r_{nm}+\lambda_0)}{k_0}+\frac{\sin k_0(r_{nm}-\lambda_0)}{k_0}\right].$$

Hence

$$\begin{aligned}E_{nm}^{st}&=\frac{e^2}{r_{nm}}\quad\text{for}\quad r_{nm}>\lambda_0,\\ E_{nm}^{st}&=0\quad\text{for}\quad r_{nm}<\lambda_0.\end{aligned} \tag{168}$$

So long as we are dealing with a discrete set of particles, we can exclude the singular case $r_{nm}=0$, and if $r_{nm}>0$, we finally obtain, always in the limit $\lambda_0\to 0$,

$$\lim_{\lambda_0\to 0}E_{nm}^{st}=\frac{e^2}{r_{nm}}\quad\text{for}\quad r_{nm}\neq 0. \tag{168a}$$

In the case of a continuous distribution of charges, for a finite λ_0 the contribution of the particles inside a sphere with the radius λ_0 is canceled. So long as the charge density is regular, this contribution goes of course to zero for $\lambda_0\to 0$. There are cases, however, where the charge density around a particle is singular for small distances, and then it can happen that the contribution to the electrostatic energy due to the particles in a distance larger than λ_0 increases to infinity for $\lambda_0\to 0$. We shall see at the end of this section that in the theory of holes such a case does occur. It is the author's opinion that the restriction $r_{nm}\neq 0$ for the particles, which is necessary for the validity of the λ process,[13] is not quite satisfactory.

(b) Introduction of the Negative Energy Protons. Electrodynamic Part of the Self-Energy in the e^2 Approximation

We now introduce the negative energy photons by putting, according to (97), (98), (76), and (137)

$$A_\nu(x)=\frac{1}{\sqrt{2}}[U_\nu(x)+U_\nu^*(x)], \tag{169}$$

$$U_\nu(x)=V^{-\frac{1}{2}}\textstyle\sum_k(2k_0)^{-\frac{1}{2}}[u_{\nu,+}(k)\exp\{i(\mathbf{k}\cdot\mathbf{x})\}+u_{\nu,-}(k)\exp\{-i(\mathbf{k}\cdot\mathbf{x})\}], \tag{170}$$

$$\frac{dU(x)}{dx_0}=V^{-\frac{1}{2}}\textstyle\sum_k(2k_0)^{-\frac{1}{2}}[-u_{\nu,+}(k)\exp\{i(\mathbf{k}\cdot\mathbf{x})\}+u_{\nu,-}(k)\exp\{-i(\mathbf{k}\cdot\mathbf{x})\}], \tag{171}$$

$$[u_{\mu,+}(k),u_{\nu,+}^*(k)]=\delta_{\mu\nu}\cos(\lambda_0k_0-\mathfrak{\lambda}\cdot\mathbf{k}),\quad[u_{\mu,-}(k),u_{\nu,-}^*(k)]=-\delta_{\mu\nu}\cos(\lambda_0k_0-\mathfrak{\lambda}\cdot\mathbf{k}), \tag{172}$$

$$a_\mu(k)=\frac{1}{\sqrt{2}}[u_{\mu,+}(k)+u_{\mu,-}^*(k)],\quad a_\mu^*(k)=\frac{1}{\sqrt{2}}[u_{\mu,-}^*(k)+u_{\mu,-}(k)]. \tag{173}$$

[13] P. A. M. Dirac, see references 1 and 10.

One has to insert these expressions in the Hamiltonian (150), (152) (in H_0 the terms containing the cross products between $u_{\mu,+}{}^*$ and $u_{\mu,-}$ or $u_{\mu,+}$ and $u_{\mu,-}{}^*$ cancel by virtue of the contribution of the redundant field, see Section 5) and the extra condition (154) splits into

$$X_+(k)\Psi=0, \quad X_-(k)\Psi=0, \quad X_+{}^*(k)\Psi=0, \quad X_-{}^*(k)\Psi=0,$$

where

$$\begin{aligned} X_+(k)&=\mathbf{k}\cdot\mathbf{u}_+(k)-k_0u_{0,+}(k)-\left(\frac{\pi}{Vk_0}\right)^{\frac{1}{2}}e\sum_n \exp\left[-i(\mathbf{k}\cdot\mathbf{z}_n)\right]\cos\left(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k}\right), \\ X_-(k)&=\mathbf{k}\cdot\mathbf{u}_-(k)-k_0u_{0,-}(k)-\left(\frac{\pi}{Vk_0}\right)^{\frac{1}{2}}e\sum_n \exp\left[i(\mathbf{k}\cdot\mathbf{z}_n)\right]\cos\left(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k}\right), \end{aligned} \tag{155'}$$

from which their adjoints follow. For the transverse parts of the field one gets instead of (158)

$$[u^T_{i,+}(k),\ u^{*T}_{j,+}(k)]=-[u^T_{i,-}(k),\ u^{*T}_{j,-}(k)]=\left(\delta_{ij}-\frac{k_ik_j}{k_0{}^2}\right)\cos\left(k_0\lambda_0-\mathbf{k}\cdot\boldsymbol{\lambda}\right). \tag{158'}$$

For H' one has simply to insert (173).

It is convenient to eliminate the first part of H' which is due to the pure radiation field with the help of the well-known canonical substitution[14]

$$\tilde{f}=\exp\left(iH_0x_0\right)f\exp\left(-iH_0x_0\right)$$

for every observable or operator which corresponds to the transformation of the wave function

$$\tilde{\Psi}=\exp\left(iH_0x_0\right)\Psi.$$

If we put

$$H'=H_0+H_1,$$

we have now

$$\partial\tilde{\Psi}/\partial x_0=-i\tilde{H}_1\tilde{\Psi}.$$

In our particular case we have

$$\tilde{u}_{\mu,+}(k)=u_{\mu,+}(k)\exp\left(-ik_0x_0\right), \quad \tilde{u}_{\mu,-}(k)=u_{\mu,-}(k)\exp\left(ik_0x_0\right).$$

Therefore one finally obtains from (161) in view of (166), (168a), the wave equation (omitting again the tilde)

$$\begin{aligned} -\frac{1}{i}\frac{\partial\Psi}{\partial x_0}&=H_1\Psi \\ &=\Bigg\{\sum_n\left[(\boldsymbol{\alpha}^{(n)}\cdot\mathbf{p}^{(n)})+m\beta^{(n)}\right]+\left(\frac{4\pi}{V}\right)^{\frac{1}{2}}e\sum_k(2k_0)^{-\frac{1}{2}}\sum_n\boldsymbol{\alpha}^{(n)}\cdot\Bigg[\frac{1}{\sqrt{2}}\{\mathbf{U}_+{}^T(k) \\ &+\mathbf{U}_-{}^{*T}(k)\}\exp\left\{i(\mathbf{k}\cdot\mathbf{z}_n-k_0x_0)\right\}+\frac{1}{\sqrt{2}}\{\mathbf{U}_+{}^{*T}(k)+\mathbf{U}_-{}^T(k)\}\exp\left\{i(-\mathbf{k}\cdot\mathbf{z}_n+k_0x_0)\right\}\Bigg]+\sum_{n<m}\frac{e^2}{r_{nm}}\Bigg\}\Psi. \end{aligned} \tag{174}$$

The $U_\mu(k)$ are analogous to the quantities defined in (76), which are constant in the absence of H_1.

We are now particularly interested in the self-energy, that is, in the one-body problem, which we

[14] See P. Dirac, *Quantum Mechanics* (Oxford University Press), second edition, p. 287.

want to compute with the help of the perturbation theory to the order e^2. We have here with

$$\Omega=\left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e\sum_k (2k_0)^{-\frac{1}{2}}\boldsymbol{\alpha}\cdot\left[\frac{1}{\sqrt{2}}\{\mathbf{U}_+^T(k)+\mathbf{U}_-^{*T}(k)\}\exp\{i(\mathbf{k}\cdot\mathbf{z}-k_0x_0)\}\right.$$
$$\left.+\frac{1}{\sqrt{2}}\{\mathbf{U}_+^{*T}(k)+\mathbf{U}_-^T(k)\}\exp\{i(-\mathbf{k}\cdot\mathbf{z}+k_0x_0)\}\right], \quad (175)$$

$$p_0\Psi=\{(\boldsymbol{\alpha}\cdot\mathbf{p})+m\beta+\Omega\}\Psi, \quad (176)$$

with

$$p_0=-\frac{1}{i}\frac{\partial}{\partial x_0}, \quad \mathbf{p}=\frac{1}{i}\frac{\partial}{\partial \mathbf{z}}. \quad (177)$$

We write, in the sense of the perturbation theory

$$\Psi=\Psi_0+\Psi_1+\Psi_2+\cdots,$$

where the term Ψ_n is of the order e^n and obtain

$$\begin{aligned}(p_0-\boldsymbol{\alpha}\cdot\mathbf{p}-m\beta)\Psi_0&=0,\\ (p_0-\boldsymbol{\alpha}\cdot\mathbf{p}-m\beta)\Psi_1&=\Omega\Psi_0,\\ (p_0-\boldsymbol{\alpha}\cdot\mathbf{p}-m\beta)\Psi_2&=\Omega\Psi_1.\end{aligned} \quad (178)$$

We start with a state where $N_+(k)=N_-(k)=0$. Writing the corresponding photon eigenfunction ω_0, we have

$$\Psi_0=u_0\omega_0 \quad (179)$$

where u_0 is a solution of the Dirac equation for which we assume

$$u_0=a\exp[i(\mathbf{q}\cdot\mathbf{z}-q_0x_0)], \quad (180)$$

denoting energy and momentum in the initial state with q_0, $\mathbf{q}$ to distinguish them from the operators p_0, $\mathbf{p}$. Of course one has $q_0^2-\mathbf{q}^2-m^2=0$. We do not write explicitly the spin index on which both u_0 and a depend. It is important that by virtue of (16), (16′)

$$\mathbf{U}_+^T(k)\Psi_0=\mathbf{U}_-^T(k)\Psi_0=0, \quad (181)$$

both operators being absorption operators. This fact is technically useful, for every term is zero where on the right side stands a $U(k)$ and not a $U^*(k)$. Ψ_1 contains only states where one photon is present, while Ψ_2 contains states with no photons and with two photons. Writing the eigenfunctions of these states Ψ_{20} and Ψ_{22}, respectively, we can put

$$\Psi_2=\Psi_{20}+\Psi_{22}, \quad \Psi_{20}=u_{20}\omega_0.$$

Only the first summand is important for the computation of the self-energy in the second approximation.

In view of (181) we obtain from (178) first

$$(p_0-\boldsymbol{\alpha}\cdot\mathbf{p}-m\beta)\Psi_1=\left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e\sum_k (2k_0)^{-\frac{1}{2}}\frac{1}{\sqrt{2}}[(\boldsymbol{\alpha}\cdot\mathbf{U}_+^{*T}(k))\exp i\{(\mathbf{q}-\mathbf{k})\cdot\mathbf{z}-(q_0-k_0)x_0\}$$
$$+(\boldsymbol{\alpha}\cdot\mathbf{U}_-^{*T}(k))\exp i\{(\mathbf{q}+\mathbf{k})\cdot\mathbf{z}-(q_0+k_0)x_0\}]a\omega_0.$$

To evaluate Ψ_1, we use the important relation

$$(p_0-\boldsymbol{\alpha}\cdot\mathbf{p}-m\beta)^{-1}=\frac{p_0+\boldsymbol{\alpha}\cdot\mathbf{p}+m\beta}{p_0^2-\mathbf{p}^2-m^2} \quad (182)$$

which shows that the reciprocal of the operator on the left side of (178) can be expressed in a rational

way by p_0 and $\mathbf{p}$. Using this relation we now have

$$\Psi_1=\left(\frac{4\pi}{V}\right)^{\frac{1}{2}} e\sum_k (2k_0)^{-\frac{1}{2}}\frac{1}{\sqrt{2}}\Big[\frac{(q_0-k_0)+\boldsymbol{\alpha}\cdot(\mathbf{q}-\mathbf{k})+m\beta}{(q_0-k_0)^2-(\mathbf{q}-\mathbf{k})^2-m^2}(\boldsymbol{\alpha}\cdot\mathbf{U}_+^{*T}(k))\exp i\{(\mathbf{q}-\mathbf{k})\cdot\mathbf{z}-(q_0-k_0)x_0\}$$

$$+\frac{(q+k_0)+\boldsymbol{\alpha}\cdot(\mathbf{q}+\mathbf{k})+m\beta}{(q_0+k_0)^2-(\mathbf{q}+\mathbf{k})^2-m^2}(\boldsymbol{\alpha}\cdot\mathbf{U}_-^{*T}(k))\exp i\{(\mathbf{q}+\mathbf{k})\cdot\mathbf{z}-(q_0+k_0)x_0\}\Big]a\omega_0. \quad (183)$$

We must now apply the third equation (178) for Ψ_2. Here we are interested only in the part $\Psi_{20}=u_{20}\omega_0$ of Ψ_2 which comes from terms of the form $U(k)U^*(k)$ while Ψ_{22} comes from terms of the form $U^*(k)U^*(k)$. We obtain from (181) a further simplification which has by virtue of (158′) the consequence

$$U_{i,+}^T(k')U_{j,-}^{*T}(k)\Psi_0=U_{i,-}^T(k')U_{j,+}^{*T}(k)\Psi_0=0,$$

$$U_{i,+}^T(k)U_{j,+}^{*T}(k')\Psi_0=[U_{i,+}^T(k),\ U_{j,+}^{*T}(k')]_+\Psi_0=\left(\delta_{ij}-\frac{k_ik_j}{k_0^2}\right)\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})\delta_{kk'}\Psi_0,$$

and similarly

$$U_{i,-}^T(k)U_{j,-}^T(k')\Psi_0=-\left(\delta_{ij}-\frac{k_ik_j}{k_0^2}\right)\delta_{kk'}\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})\Psi_0.$$

Using this result we obtain

$$(p_0-\boldsymbol{\alpha}\cdot\mathbf{p}-m\beta)u_{20}=\frac{\pi e^2}{V}\sum_k\frac{1}{k_0}\sum_{i,j=1}^3\alpha_i\Big[\frac{q_0-k_0+\boldsymbol{\alpha}\cdot(\mathbf{q}-\mathbf{k})+m\beta}{(q_0-k_0)^2-(\mathbf{q}-\mathbf{k})^2-m^2}$$

$$-\frac{q_0+k_0+\boldsymbol{\alpha}\cdot(\mathbf{q}+\mathbf{k})+m\beta}{(q_0+k_0)-(\mathbf{q}+\mathbf{k})^2-m^2}\Big]\alpha_j\left(\delta_{ij}-\frac{k_ik_j}{k_0^2}\right)\cos(k_0\lambda_0-\mathbf{k}\cdot\boldsymbol{\lambda})u_0. \quad (184)$$

We notice that the second term can be derived from the first term simply by changing k_0, $\mathbf{k}$ into $-k_0$, $-\mathbf{k}$. We now bring α_i to the right side, using

$$\alpha_i\{\boldsymbol{\alpha}\cdot(\mathbf{q}-\mathbf{k})+m\beta\}=-\{\boldsymbol{\alpha}\cdot(\mathbf{q}-\mathbf{k})+m\beta\}\alpha_i-2(q_i-k_i).$$

Moreover we have

$$\sum_{i,j}\alpha_i\alpha_j\left(\delta_{ij}-\frac{k_ik_j}{k_0^2}\right)=2,$$

and according to the wave equation for u_0

$$(q_0-\boldsymbol{\alpha}\cdot\mathbf{q}-m\beta)u_0=0.$$

Finally the denominators can be simplified by virtue of $q_0^2-\mathbf{q}^2=m^2$ and we obtain

$$(p_0-\boldsymbol{\alpha}\cdot\mathbf{p}-m\beta)u_{20}=\frac{\pi e^2}{V}\sum_k\Big[\frac{-k_0+(\boldsymbol{\alpha}\cdot\mathbf{k})+(\boldsymbol{\alpha}\cdot\mathbf{q})-(\boldsymbol{\alpha}\cdot\mathbf{k})(\mathbf{q}\cdot\mathbf{k})/k_0^2}{k_0(-q_0k_0+\mathbf{q}\cdot\mathbf{k})}$$

$$-\frac{k_0-(\boldsymbol{\alpha}\cdot\mathbf{k})+(\boldsymbol{\alpha}\cdot\mathbf{q})-(\boldsymbol{\alpha}\cdot\mathbf{k})(\mathbf{q}\cdot\mathbf{k})/k_0^2}{k_0(q_0k_0-\mathbf{q}\cdot\mathbf{k})}\Big]\cos(k_0\lambda_0-\mathbf{k}\cdot\boldsymbol{\lambda})u_0. \quad (185)$$

It will be convenient for us to apply to this equation the operator $(p_0+\boldsymbol{\alpha}\cdot\mathbf{p}+m\beta)$ from the left. Using

$$(q_0+\boldsymbol{\alpha}\cdot\mathbf{q}+m\beta)[-k_0+(\boldsymbol{\alpha}\cdot\mathbf{k})+(\boldsymbol{\alpha}\cdot\mathbf{q})-(\boldsymbol{\alpha}\cdot\mathbf{k})(\mathbf{q}\cdot\mathbf{k})/k_0^2]$$

$$=[-k_0-(\boldsymbol{\alpha}\cdot\mathbf{k})-(\boldsymbol{\alpha}\cdot\mathbf{q})+(\boldsymbol{\alpha}\cdot\mathbf{k})(\mathbf{q}\cdot\mathbf{k})/k_0^2](-q_0+\boldsymbol{\alpha}\cdot\mathbf{q}+m\beta)+2[-q_0k_0+\mathbf{q}^2+(\mathbf{q}\cdot\mathbf{k})-(\mathbf{q}\cdot\mathbf{k})^2/k_0^2],$$

we obtain in this way

$$(p_0^2-\mathbf{p}^2-m^2)u_{20}=\frac{2\pi e^2}{V}\sum_k\left[\frac{q_0k_0-\mathbf{q}^2-(\mathbf{q}\cdot\mathbf{k})+(\mathbf{q}\cdot\mathbf{k})^2/k_0^2}{k_0(q_0k_0-\mathbf{q}\cdot\mathbf{k})}\right.$$
$$\left.-\frac{-q_0k_0-\mathbf{q}^2+(\mathbf{q}\cdot\mathbf{k})-(\mathbf{q}\cdot\mathbf{k})^2/k_0^2}{k_0(-q_0k_0+\mathbf{q}\cdot\mathbf{k})}\right]\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})u_0. \quad (186)$$

This equation is of the form

$$(p_0^2-\mathbf{p}^2-m^2)u_{20}=cu_0.$$

In the approximation of the order e^2, the term cu_{20} can be neglected and u_0+u_{20} satisfies the equation

$$(p_0^2-\mathbf{p}^2-m^2-c)(u_0+u_{20})=0.$$

Therefore c is simply the correction of the square of the rest mass of the particle due to the second-order perturbation. Inserting

$$V^{-1}\sum_k\rightarrow\frac{1}{(2\pi)^3}\int d\mathbf{k},$$

we obtain

$$\Delta m^2=2m\Delta m=\frac{e^2}{2\pi^2}\int\frac{1}{2}[f(k_0,\mathbf{k})+f(-k_0,-\mathbf{k})]\cos(\lambda_0k_0-\boldsymbol{\lambda}\cdot\mathbf{k})d\mathbf{k}$$

with

$$f(k_0,\mathbf{k})=\frac{q_0k_0-\mathbf{q}^2-(\mathbf{q}\cdot\mathbf{k})+(\mathbf{q}\cdot\mathbf{k})^2/k_0^2}{k_0(q_0k_0-\mathbf{q}\cdot\mathbf{k})}=\frac{1}{k_0}\left[1+\frac{q_0^2-q^2}{q_0k_0-\mathbf{q}\cdot\mathbf{k}}-\frac{q_0}{k_0}-\frac{(\mathbf{q}\cdot\mathbf{k})}{k_0^2}\right].$$

We put, for the sake of simplicity, $\boldsymbol{\lambda}=0$ and obtain after integration over the directions of $\mathbf{k}$ with $q=|\mathbf{q}|$

$$\frac{1}{2}\Delta m^2=m\Delta m=\frac{e^2}{\pi}\int_0^\infty\frac{1}{2}\left[\left\{k_0+\left(\frac{m^2}{2q}\log\frac{q_0+q}{q_0-q}-q_0\right)\right\}\right.$$
$$\left.+\left\{-k_0+\left(\frac{m^2}{2q}\log\frac{q_0+q}{q_0-q}-q_0\right)\right\}\right]\cos\lambda_0k_0dk_0. \quad (187)$$

We compare this result with the old result of Waller[15] for the self-energy of the one-electron problem, which we can write

$$\frac{1}{2}\Delta m^2=m\Delta m=\frac{e^2}{\pi}\int_0^{k_m}k_0dk_0+\frac{e^2}{\pi}\left(\frac{m^2}{2q}\log\frac{q_0+q}{q_0-q}-q_0\right)\int_0^{k_m}dk_0\cdots \quad (187')$$

where we took into account that for a given momentum $\mathbf{q}$ of the electron one has $q_0\Delta q_0=m\Delta m$. In the old form of the theory one had to cut off the integrals arbitrarily in order to avoid the infinity. Here we get rid of the infinities, using two different methods simultaneously, namely, the λ-limiting process and the negative energy photons. The former makes the second integral disappear because of

$$\int_0^\infty dk\cos\lambda_0k_0=0. \quad (188)$$

This is quite analogous to the cancellation of the electrostatic self-energy. Indeed the factor before the second integral disappears if the electron is initially at rest ($\mathbf{q}=0$, $q_0=m$), in which case the computation can be much simplified. Generally the second term gives the electromagnetic energy of the electron due to its motion and is in a relativistically invariant theory always present if the electrostatic

[15] I. Waller, Zeits. f. Physik **62**, 673 (1930).

self-energy exists. The first term in (187) on the other side is due to the fluctuations of the field in the vacuum and a typical quantum effect. It is directly cancelled by the contributions of the negative energy photons, as one can see from the integrand in (187). The final result is therefore

$$\Delta m = 0 \tag{189}$$

in the e^2 approximation.

We now generalize both remarks. First we note that in (188) a generalization of the ordinary concept of integral is used, which is useful for integrands of an oscillating type. One introduces a factor $g(\epsilon x)$, where $g(x)$ falls off sufficiently rapidly for large x, so that at least $g'(x) \to 0$ for $x \to \infty$ but for some integrands the decrease of some higher derivatives of $g(x)$ is required. Moreover $g(0) = 1$, and we assume that $g(x)$ has only a finite number of maxima and minima between zero and infinity. Then we define

$$\int_0^\infty f(x)dx = \lim_{\epsilon \to 0} \int_0^\infty f(x) g(\epsilon x) dx. \tag{190}$$

If the limit exists, it is independent of the particular choice of $g(x)$; for instance, $g(x) = e^{-x}$ fulfills all requirements.[16] In this way one gets

$$\int_0^\infty x^{2n} \cos x dx = 0, \quad \int_0^\infty x^{2n+1} \sin x dx = 0,$$

but

$$\int_0^\infty x^{2n} \sin x dx = (2n)!(-1)^n, \quad \int_0^\infty x^{2n+1} \cos x dx = (2n+1)!(-1)^{n+1}, \tag{188'}$$

for

$$n = 0, 1, 2, \cdots.$$

Using the λ-limiting process but not the negative energy photons one would get, instead of (187),

$$m\Delta m = (e^2/\pi m) \int_0^\infty k_0 \cos \lambda_0 k_0 dk_0 = -(e^2/\pi m \lambda_0^2)$$

which was derived by Dirac at an earlier time.

Also the cancellation of the terms in the integrand which are linear in k_0 by the negative energy photons, can be greatly generalized. If in the older theory (including the λ-limiting process) the integrand over the momenta $\mathbf{k}$ of photons in intermediate states is $f(\mathbf{k}, k_0)$, in the new theory the integrand is $\frac{1}{2}[f(\mathbf{k}, k_0) + f(-\mathbf{k}, -k_0)]$. This is equivalent to the other statement that after integration over the directions one has to take

$$\frac{1}{2} \int_{-\infty}^\infty f(k_0) dk_0 \qquad \text{instead of} \qquad \int_0^\infty f(k_0) dk_0.$$

[16] An alternative to the definition (189) is to take first

$$\int_0^x f(x)dx = F(x),$$

and to construct then, instead of $\lim_{x \to \infty} F(x)$, using the mean value

$$F_1(x) = \frac{1}{x} \int_0^\infty F(x)dx,$$

the other limit, $\lim_{x \to \infty} F_1(x)$. If $\lim_{x \to \infty} F(x)$ exists, the value of $\lim_{x \to \infty} F_1(x)$ coincides with the former, as one can easily prove. The procedure can be repeated by forming

$$F_2(x) = \frac{1}{x} \int_0^x F_1(x)dx.$$

If after a finite number of steps $\lim_{x \to \infty} F_n(x)$ exists, one can define $\int_0^\infty f(x)dx$ by its value. It can be shown that under general conditions this definition coincides with that given in (190).

This holds also in higher approximations where multiple integrals over the momenta $\mathbf{k}$, $\mathbf{k}'$, $\cdots$ of several photons in intermediate states occur and where one has, in the new theory, to take the part of the integrand which is even with respect to the k_ν-vector of every photon. The proof of it is rather elementary but becomes clearer if we introduce new variables $\varphi_+(k)$, $\varphi_-(k)$ and their adjoints instead of $U_+(k)$, $U_-(k)$ which make the square root $k_0^{-\frac{1}{2}}$ in the expression (176) for $U(x)$ disappear. We put

$$\varphi_+(k)=(k_0)^{\frac{1}{2}}U_+(k), \qquad \varphi_-(k)=-(k_0)^{\frac{1}{2}}U_-(k),$$

$$\varphi_+^*(k)=(k_0)^{\frac{1}{2}}U_+^*(k), \qquad \varphi_-^*(k)=-(k_0)^{\frac{1}{2}}U_-^*(k),$$

$$[\varphi_+(k),\ \varphi_+^*(k)]=k_0\cos(\lambda_0 k_0-\boldsymbol{\lambda}\cdot\mathbf{k}), \quad [\varphi_-(k),\ \varphi_-^*(k)]=-k_0\cos(\lambda_0 k_0-\boldsymbol{\lambda}\cdot\mathbf{k}),$$

$$U(x)=(2V)^{-\frac{1}{2}}\sum_k\left[\frac{1}{k_0}\varphi_+(k)\exp i(\mathbf{k}\cdot\mathbf{x}-k_0x_0)-\frac{1}{k_0}\varphi_-(k)\exp i(-\mathbf{k}\cdot\mathbf{x}+k_0x_0)\right].$$

Then from the beginning one obtains in every equation the contribution of the negative-energy particles by substituting for $\varphi_+(k)$, $\varphi_+^*(k)$, $\mathbf{k}$, k_0, simply $\varphi_-(k)$, $\varphi_-^*(k)$, $-\mathbf{k}$, $-k_0$ and this holds also for the final results. The notations which we have used here were chosen in order to stay in agreement with the notations of paper A.

Now it is easily seen that, as was pointed out by Dirac, every (single or multiple) integral *with a rational function as integrand* is made convergent for large $\mathbf{k}$'s by the simultaneous application of the λ-limiting process and the negative energy photons. The latter reduces the integrand to the even part in k_0 (notice that it is always permitted to replace $\mathbf{k}$ by $-\mathbf{k}$ without changing the result of the integrations over the direction of the $\mathbf{k}$'s), and this even part can be reduced to a convergent integral by using the decomposition of the rational integrand into partial fractions and using the first line of (188′). The repeated application of the operation (182), however, will certainly produce only rational functions of the vectors k_ν, k_ν', $\cdots$ of the different photons in the intermediate states. We have also to expect convergence of the new theory if applied to the one-body problem even in higher approximations. Quite the same as for the spin electron holds also for spin 0 electrons if they are treated according to this alternative of the theory where the number of particles is constant in an external electromagnetic field [see Eqs. (119), (120), and (121)]. Also for these electrons the self-energy is zero in the e^2 approximation. (Whether the higher approximations give a vanishing or only a finite self-energy is not yet known, either for spin 0 or for ordinary spin $\frac{1}{2}$ electrons.) Difficulties other than the convergence for high frequencies of intermediate photons may arise from the fact that in the new theory there does not exist a rigorous stationary solution if a particle is present, due to the spontaneous transitions which occur if starting from the "vacuum" $N_+(k)=N_-(k)=0$ as initial state. This question is closely connected with the problem of the physical interpretation of the theory which is treated in the following section.

We add some remarks about the self-energy in the theory of holes, which is treated in detail by Weisskopf.[17] For the sake of simplicity we restrict ourself to the case where the electron is initially at rest and consider only the difference of the self-energy if one electron is present, and the self-energy of the vacuum. If we apply the λ-limiting process, but not yet the negative energy photons, the electrostatic self-energy is given by

$$E^{st}=\frac{e^2m}{\pi}\int_0^\infty\frac{\cos\lambda_0k_0}{(k_0^2+m^2)^{\frac{1}{2}}}dk_0, \tag{191}$$

while the electrodynamic part is given by

$$E^{dn}=\frac{e^2}{\pi m}\int_0^\infty k_0\frac{(k_0^2+m^2)^{\frac{1}{2}}-k_0}{(k_0^2+m^2)^{\frac{1}{2}}}\cos\lambda_0k_0dk_0=\frac{e^2m}{\pi}\int_0^\infty\frac{k_0\cos\lambda_0k_0}{k_0^2+m^2+k_0(k_0^2+m^2)^{\frac{1}{2}}}dk_0. \tag{192}$$

We assume here $\boldsymbol{\lambda}=0$ which is permitted. Both integrals behave for small λ_0 as $\log\lambda_0$, and hence

[17] V. Weisskopf, Zeits. f. Physik **89**, 27 and 90 (1934); Phys. Rev. **56**, 72 (1939).

become infinite for $\lambda_0 \to 0$. Indeed, one has in this case, denoting by γ Euler's constant,

$$E^{st} \sim -\frac{e^2 m}{\pi}\left[\log\frac{\lambda_0 m}{2}+\gamma\right], \tag{191'}$$

$$E^{dn} \sim -\frac{e^2 m}{\pi}\left[\frac{1}{2}\left\{\log\frac{\lambda_0 m}{2}+\gamma\right\}+\frac{1}{4}\right]. \tag{192'}$$

This result is not improved by Dirac's negative energy photons, which have again the convention that the integrand $f(k_0)$ has to be replaced by $\frac{1}{2}[f(k_0)+f(-k_0)]$. This leaves the electrostatic self-energy unchanged, while we have now[18]

$$E^{dn} = -\frac{e^2}{\pi m}\int_0^\infty \frac{k_0^2}{(k_0^2+m^2)^{\frac{1}{2}}}\cos\lambda_0 k_0 dk_0, \tag{193}$$

which has for small λ_0 the even stronger singularity

$$E^{dn} = \frac{e^2 m}{\pi}\left[\frac{1}{(\lambda_0 m)^2}-\frac{1}{2}\left\{\log\frac{\lambda_0 m}{2}+\gamma\right\}-\frac{1}{4}\right]. \tag{193'}$$

Furthermore the self-energy of the "anti-particles" described by the second spinor field $u_n(x)$ which we introduced in the last section, Eqs. (140)–(142) is the same as that of the particles described by the $u_p(x)$ spinors.

Similar results hold for electrons without spin if they are treated according to the second alternative [Eqs. (130)–(139)]. If we include here the negative energy photons, the part E^{dn}_{fluct} of Weisskopf[19] is cancelled, and for particles at rest, we have

$$E^{dn} = 0, \tag{194}$$

but

$$E^{st} = \frac{\pi e^2}{4m}\int_0^\infty \frac{k_0^2+2m^2}{(k_0^2+m^2)^{\frac{1}{2}}}\cos\lambda_0 k_0 dk_0, \tag{195}$$

which, for small λ_0, is given by

$$E^{st} = \frac{\pi e^2 m}{4}\left[-\frac{1}{(\lambda_0 m)^2}-\frac{3}{2}\left\{\log\frac{\lambda_0 m}{2}+\gamma\right\}+\frac{1}{4}\right]. \tag{195'}$$

If one applies the λ-process in the commutation relations for the field of the electron rather than for the electromagnetic field, the singularity of the different energy expressions turns out to be exactly the same.

We see that for both spin $\frac{1}{2}$ and spin 0 electrons, the failure of the λ-limiting process in theories which allow explicitly pair generation and pair annihilation is formally due to the occurrence of a square root in the integrand in which case the theorem of convergence, which was correct for even *rational* integrands, does not hold any longer. Physically the failure of the λ-process is due to the contribution of pairs in the neighborhood of the particle considered which comes from an expectation value of the product of charge densities $\langle s_0(x)s_0(x')\rangle_{\text{Av}}$ with a singularity of the type r^{-2} for small distances $r=|\mathbf{x}-\mathbf{x}'|$ of the two points.[20] Therefore this failure is closely connected with the fact that for a given λ_0 the contribution of all particles with $r>\lambda_0$ to the electrostatic self-energy is not modified. It seems therefore likely that in a future quantum theory, which includes the description of generation and annihilation of pairs, it will be necessary to substitute for the classical method of the λ-limiting process a different new procedure.

---

[18] Weisskopf decomposes E^{dn} into two parts, one part E^{dn}_{spin} and the other part E^{dn}_{fluct}. The first part is given by (193), while the second part is cancelled by the negative energy photons.

[19] V. Weisskopf, Phys. Rev. **56**, 72 (1939) Sect. V.

[20] V. Weisskopf, reference 19, Sect. II.

7. THE PROBLEM OF THE PHYSICAL INTERPRETATION OF THE FORMALISM

If one applies the usual rules of interpretation to the new formalism, one obtains, of course, results which contradict any experience. If we start with the state where $N_+=N_-=0$, the theory leads to spontaneous emission process of negative energy photons. It can easily be seen that the transition probability A_s^- per unit of time for a process of the spontaneous emission of a single photon in an eigenstate s of the field oscillators (the single index s denotes the direction of the emitted photons, their polarization, and their frequency) is related to the probability coefficient B_s of absorption of photons of the state s according to the usual theory. The latter is defined in such a way that the transition probability of absorption per unit of time is $B_sN_s^0$ if N_s^0 is the actual number of photons of the eigenstate present in the initial state. The relation between A_s^- and B_s is

$$A_s^-=-\tfrac{1}{2}B_s \tag{196}$$

where the factor $\frac{1}{2}$ is due to the factor $1/\sqrt{2}$ in the expression of $A_\mu(x)$ by $U_\mu(x)$ and $U_\mu^*(x)$ which enters into the interaction energy. If an external field of force is present (as for instance in the hydrogen atom), there are in the new theory also spontaneous emission processes of positive energy photons with a transition probability A_s^+ which is also equal to one-half of the corresponding probability A_s of the usual theory

$$A_s^+=\tfrac{1}{2}A_s. \tag{196'}$$

This process occurs also in the case of absence of an external field if the final state of the electron has a negative energy. Dirac proposes to interpret the spontaneous emission processes of positive (negative) energy photons as actual emission (absorption) processes. The conditions under which the processes occur in nature are very different from the theoretical ones. Dirac proposes always to start with $N_+=N_-=0$ and to overtake the dependence of the probabilities of the actual processes on the number of photons actually present in the initial state (Einstein's probability laws) from experience without trying to derive them.

It is possible to improve this point a little by considering more general initial states. Two alternative ways are possible leading to the same result.

(a) Consider for an actual emission (absorption) process in the mathematical scheme (the "hypothetical world") an initial state where, for a certain eigenstate s of the field, $N_{s,+}{}^0$ positive energy photons and $N_{s,-}{}^0$ negative energy photons are present, and in the final state the number $N_{s,+}{}^0(N_{s,-}{}^0)$ has changed by $+1$ while $N_{s,-}{}^0(N_{s,+}{}^0)$ is unchanged. The transition probabilities of these processes in the hypothetical world are

$$a_s(N_{s,+}{}^0+1) \quad \text{or} \quad -a_s(N_{s,-}{}^0+1). \tag{197}$$

The probabilities of actual emission (absorption) processes, if n_s actual photons are present in the initial state, are then given by

$$2a_s(n_s+1) \quad \text{or} \quad 2a_sn_s, \text{ respectively.} \tag{197'}$$

One sees that for this given eigenstate s for which we want to consider emission and absorption processes, one has to put

$$n_s=N_{s,+}{}^0 \quad \text{or} \quad n_s=-N_{s,-}{}^0-1, \tag{198}$$

respectively.[21] For the numbers of photons in other eigenstates we have to put $N_+=N_-=0$ in both the initial and final states. We could have selected in the same way processes where one of the numbers N_+, N_- decreases by 1 while the other number keeps its value, and we could have interpreted the absorption of positive (negative) energy photons as actual absorption (emission) process, again using the connection (198). A simple possibility also is to combine these two ways as follows.

(b) Consider for the physical interpretation only those states of the hypothetical world where $N_-=0$ in both the initial and final states, but with an arbitrary number N_+ of positive energy photons in the eigenstate for which we want to consider an emission or absorption of a photon. One needs only the identification of the actual number of photons with N_+ in this case.

The latter possibility is, however, very unsymmetrical, and we prefer the interpretation (a). In any case the abundance of possible states and

[21] This has to be considered as a formal algebraic substitution in spite of the fact that in the hypothetical world $N_{s,-}{}^0$ is not negative, while in the actual world n_s is also not negative.

processes in the hypothetical world, in comparison with those of the actual world, is directly related to the introduction of redundant variables in the hypothetical world.[22] The fact that only a part of the processes in the hypothetical world can be physically interpreted at the same time, is all the more unsatisfactory as the conservation law for the total sum of probabilities in the hypothetical world does not have a simple physical interpretation in this way, because the same initial state in the hypothetical world corresponds, according to (198), in general to different initial states of the actual world if we consider different processes (the reverse being also the case). The value of the new formalism is the possibility of computing the probability coefficients of the different radiation processes, not only in the first approximation where they occur in the development in power series of the electric charge e, but rigorously because the new formalism leads to convergent results in higher approximations.

The generalization of the rule to obtain the probability coefficients for processes in the actual world for more complicated processes does not lead to new difficulties and can be formulated in the following way:

Starting with an initial state where, for certain eigenstates s which we want to consider, $N_{s,+}{}^0$ positive energy photons and $N_{s,-}{}^0$ negative energy photons are present (while for sufficiently high frequencies one has always $N_{s,+}{}^0 = N_{s,-}{}^0 = 0$), we consider a process where in the final state the numbers $N_{s,+}$, $N_{s,-}$ are

$$N_{s,+} = N_{s,+}{}^0 + p_{s,+}, \quad N_{s,-} = N_{s,-}{}^0 + p_{s,-}.$$

We interpret only those processes for which $p_{s,+}$, $p_{s,-}$ are zero or positive, and where for sufficiently high frequencies one assumes $p_{s,+} = p_{s,-} = 0$. In the mathematical expression for the transition probabilities of a process characterized by given numbers $p_{s,+}$, $p_{s,-}$ of positive and negative energy photons in its dependence on $N_{s,+}{}^0$, $N_{s,-}{}^0$ one has to introduce the formal substitution (198) for this selected group of states s and to multiply the result by

$$2^{\Sigma_s (p_{s,+} + p_{s,-})}$$

to obtain the actual probability of a process of emission of $p_{s,+}, \cdots$ photons in the states $s, \cdots$ and the absorption of $p_{s,-}, \cdots$ photons in the states $s, \cdots$ in the actual world.[23] The substitution (198) at the same time makes the result positive, while the original expression has the sign $(-1)^{\Sigma_s p_{s,-}}$.

As an example we mention the process of scattering of a photon from the state s into the state s' in the actual world, which is described in the hypothetical world as the process of the simultaneous emission of a photon with positive energy in the state s' and a photon with negative energy in the state s.

It is needless to stress the problematical and preliminary character of this rule for a translation of results derived originally for the "hypothetical world" into results for the "actual world." It is interesting, however, that these rules have a similarity to the other rule, which was already well known, for translation of results concerning negative energy states of matter into a description of the actual world. In Dirac's positron theory one has to replace the number $M_{s,-}$ of electrons (n) in a given state of negative energy (this number can have only the values 0 or 1 because of the exclusion principle) by the actual number $m_{s,p}$ of positrons (p) with positive energy which is given by

$$m_{s,p} = 1 - M_{s,-} \quad \text{while} \quad m_{s,n} = M_{s,+} \tag{199}$$

in order to obtain the actual transition probability. These relations have a close analogy to the relations (198) for Bose statistics. If one adds this rule for the interpretation to the original form of the theory (where the number of particles is conserved "in the hypothetical world") one neglects to take into account the Coulomb interaction between generated pairs, as Dirac pointed out. This neglect is intentional, in order to get rid of any convergence difficulties. In the other alternative, where the redundant second spinor field is introduced [Eqs. (140)–(142)] the Coulomb interaction between generated pairs can

[22] Dirac checks the occurrence of the factor 2 in (197′) by using the variables N_a, N_b instead of N_+ and N_-. This, however, does not change the fact that a new rule is necessary in this place to get the correct physical interpretation.

[23] Professor Dirac pointed out in a letter to me that the cases where both $p_{s,+}$, $p_{s,-}$ are different from zero for the same s, and for which $p_{s,+} - p_{s,-}$ has the same value, cannot be directly distinguished physically. Their physical interpretation, however, is expressed by a dependence of the transition probability on n_s which is different from the case where one of the $p_{s,+}$ or $p_{s,-}$ is zero.

be taken into account. One has to consider only those processes where $M_{n,-}$ and $M_{p,-}$ either decrease or stay unchanged, and $M_{n,+}$ and $M_{p,-}$ either increase or stay unchanged. The p particles in the hypothetical world are used to describe pair generation, the n particles to describe emission of radiation, the negative energy photons to describe absorption of radiation. But as we have seen, this second alternative gives rise to a divergent self-energy.

Quite similar is the situation for the physical interpretation of spinless charged particles which obey Bose statistics. For the form of the theory in which the number of particles is conserved [Eqs. (117)–(121)] it is not necessary to use at all the field $U_n(x)$, $U_n^*(x)$ (which was introduced for reasons of symmetry in order to compensate the charge of the vacuum) for the physical interpretation. We have then in the hypothetical world (eventually in an external electromagnetic field) transitions from a positive energy state s to negative energy states s' with a probability

$$-a_{ss'}N_{s,+}(1+N_{s',-}),$$

and for the reverse process

$$-a_{ss'}N_{s',-}(1+N_{s,+}).$$

For the interpretation in the actual world we have to put

$$m_{s,p}=N_{s,+}, \quad m_{s,n}=-(1+N_{s,-}) \qquad (200)$$

which makes the first process to a pair annihilation with the transition probability

$$+a_{ss'}m_{s,p}m_{s',n}$$

and the second process to a pair creation with a probability

$$+a_{ss'}(1+m_{s,p})(1+m_{s',n}).$$

If one does not use the field $U_n(x)$ for the physical interpretation, it is not necessary to include here a factor 2 as was the case with the photons. Again in this form of the theory one has to leave out the Coulomb interaction between generated pairs, and gains for it the convergence of the theory.

For the other alternative of the theory of charged particles without spin [Eqs. (135)–(139)], in which no transition from positive energy to negative energy states occurs, one can use the positive energy states for the description of pair creation and the negative energy states for the description of pair annihilation, but as we have seen, the self-energy then becomes infinite.

The arbitrariness of the rules for the translation of results concerning the hypothetical world into results concerning the actual world and the lack of uniqueness of these rules seems to indicate that new ideas and more radical changes of the present formalism will be necessary in order to get a really satisfactory quantum theory of the electromagnetic field.