Signal velocity in quantum electrodynamics

M. I. Shirokov

Joint Institute for Nuclear Research, Dubna (Moscow Region)
Usp. Fiz. Nauk **124**, 697–715 (April 1978)

The fulfillment in quantum electrodynamics of the principle of relativistic causality—the signal velocity does not exceed the velocity of light—is discussed. In Sec. 1 it is argued that this principle is not guaranteed automatically merely by the local commutativity of the theory. Section 2 is a critical review of the signal transmission problems which have been formulated and solved in the literature. In Sec. 3, this problem is considered in the framework of a simple but fairly realistic model of quantum electrodynamics. The signal source is an external current localized in some region S. The arrival of a signal in a region D at a distance R is established by a change in the coordinate, momentum, or energy of a charged particle. It is shown that proof of relativistic causality of the theory requires one to take into account appropriately quantum-mechanical and, in particular, quantum-electrodynamical features of the problem. In the final fourth section, a general formulation and exact solution of the signal transmission problem are given. The treatment is of sufficient generality for one to be able to assert that in the framework of quantum electrodynamics none of the possible methods of signal transmission violate relativistic causality.

PACS numbers: 11.10. – z

CONTENTS

1. INTRODUCTION. PRINCIPLE OF RELATIVISTIC CAUSALITY

The prinicple that the signal velocity should not exceed the velocity of light is not guaranteed merely by relativistic invariance of a theory. This principle is a synthesis of the ordinary causality condition (the effect should not precede the cause) and the special theory of relativity. If a signal could be transmitted with superluminal velocity, then there would exist a Lorentz frame in which the cause (switching on of the signal source) would be later than the effect (arrival of the signal). And one could then have a purely logical contradiction of the type in which the effect could countermand its cause (see §2.7 in the book[1]). In this review, the principle we have just formulated is called the principle of relativistic causality.

The opinion is widely held that local commutativity[1)] of quantum field theory guarantees the fulfillment of relativistic causality. Two field quantities defined at points x and y are locally commutative if their commutator (or anticommutator) vanishes when x and y are separated by a spacelike interval (for a mathematically more precise formulation, see, for example, §1 in Ch. 3 of the book[2]). Of this kind are the commutation relations calculated in all textbooks for the Heisenberg operators **E** (electric field), **H** (magnetic field), and ψ (electron–positron field) in the Lorentz gauge in quantum electrodynamics.

Relativistic causality is deduced from local commutativity, for example, in the following manner. Locality of the commutation relations for **E** and **H** means that independent exact measurements of fields in two neighboring spacelike regions are possible: the one measurement does not interfere with the other. But if fields were to propagate with superluminal velocity, this would be impossible (see the end of §48 in the book[3]).

Nevertheless, there are theories in which local commutativity holds but phenomena occur which indicate that local commutativity alone does not guarantee relativistic causality. Let us give some examples.

The phenomenon of instantaneous spreading of the wave packet of a relativistic particle is well known. For example, in the framework of the theory of a quantized free scalar field (in which the condition of local commutativity is of course satisfied) one can pose and solve the problem of the motion of a packet for a corresponding particle. Initially, at $t = t_0$, the particle is localized in a finite volume V, so that the probability of finding the particle outside V is zero (finite packet). It

[1)]Local commutativity follows from Bogolyubov's causality condition; see Ch. 4, §3 of the book.[2]

 0038-5670/78/040345-14$01.10

can be shown that at the time $t_0+\tau$ there is a nonvanishing probability of finding the particle in a region separated from V by a distance R greater than $c\tau$ (see, for example, Ref. 4). Hegerfeldt's proposal[4] for avoiding this contradiction with relativistic causality is to suggest that finite packets are impossible. But this is to overcome difficulties by introducing into the theory a new postulate, so that relativistic causality is then guaranteed by not just local commutativity alone but also by this postulate.

Let us consider now the propagation of, not a particle, but a free quantized field (scalar, electromagnetic, etc). It is known from classical electrodynamics that if an electromagnetic field at time $t=t_0$ is localized in some manner in a region V_S, it then propagates with velocity c. Therefore, a detector which measures, for example, the electric field (test charge at rest) in a region V_D at distance R from V_S, gives an indication only after a time R/c. In the quantum case, the corresponding problem must be formulated somewhat differently. We cannot assume that outside V_S at $t=t_0$ both E and H are zero: in virtue of the equal-time commutation relation

$$[E_x(\mathbf{x}, t), H_y(\mathbf{x}', t)]=i\frac{\partial}{\partial z}\delta(\mathbf{x}-\mathbf{x}'), \tag{1.1}$$

the fields $\mathbf{E}(\mathbf{x}, t)$ and $\mathbf{H}(\mathbf{x}, t)$ at any point $\mathbf{x}$ cannot simultaneously assume exact values, in particular, zero values. We shall assume that outside V_S at $t=t_0$ the electric field $\mathbf{E}$ is zero. If in addition at points $x_S\in V_S$ it takes on certain definite values, then the initial state is completely specified. This state is nonstationary since E does not commute with the Hamiltonian $\int d^3x(\mathbf{E}^2+\mathbf{H}^2)$. Therefore, the distribution with respect to $\mathbf{E}$ in the region V_D must change with the time in some manner. This change can be found by solving the Heisenberg equations (free Maxwell equations) for the fields $\mathbf{E}$ and $\mathbf{H}$; see, for example, Sec. 4 later. Instead, we here adduce some qualitative arguments which lead to the same result as the exact solution. If $\mathbf{H}$ takes nonzero values outside V_S at the time $t=t_0$, as we have said above, then by virtue of the Maxwell equations this nonzero $\mathbf{H}$ generates an electric field everywhere outside V_S immediately after the time t_0, this including the region V_D as well. Note that we have used a particular (equal-time) value of one of the local commutation relations in the derivation of this paradoxical result.

However it does not mean that in the given case $\mathbf{E}$ propagates instantaneously. The point is that if $\mathbf{E}$ at $t=t_0$ were zero everywhere, then also in this case we should have a nonzero $\mathbf{E}$ everywhere immediately after the time t_0 (on the basis of the same qualitative arguments). If this "background" of quantum fluctuations of $\mathbf{E}$ is subtracted, the propagation velocity of $\mathbf{E}$ is finite (see later in Sec. 4). However, it does not follow from local commutativity that this "background" must be subtracted. One of the reasons for the existence of the "background" are the equal-time commutation relations, which hold in theories in which there is local commutativity and also in theories in which there is no local commutativity (for example, in nonlocal theories).

Finally, we give an example of a theory with interaction in which the condition of local commutativity is satisfied but there are difficulties with relativistic causality. This is Lee and Wick's theory of "finite quantum electrodynamics" (see Ref. 5, in particular, Sec. 8 of Lee and Wick's paper). A feature of this theory is that it makes essential use of an indefinite metric. We emphasize in connection with this example that the most popular variant of ordinary quantum electrodynamics—the Gupta–Bleuler formulation—also uses an indefinite metric (the other well-known formulation—the Coulomb gauge or radiation gauge—contains the instantaneous Coulomb interaction, and is not a local theory).

In addition to local commutativity, other general criteria can serve to guarantee relativistic causality. We mention, for example, the principle that the group velocity should not exceed c. The problem with this is that it does not work in the case of a medium with anomalous dispersion (see Refs. 6, 7). Another example which is instructive in this connection is discussed in Sec. 4 of the collection of articles of Ref. 8. Another criterion is proposed for quantum field theory in Ref. 9: relativistic causality of a theory is guaranteed by a "causal" form of the equations for the Heisenberg operators in the representation of coherent states. However, it will be shown in Sec. 2 that the solution of various signal transmission problems leads to results that do not agree with relativistic causality despite the fact that the equations of quantum electrodynamics have a "causal form." The choice of one or other representation of the theory cannot change these results. We shall see that they are determined, not by the method of solution, but above all by the formulation of these problems.

On the basis of what we have said above, we assume that neither local commutativity nor the nature of the Heisenberg equations guarantees relativistic causality. We shall show that many other ingredients of the theory are important when one is considering relativistic causality of a theory: the manner in which the signal source is described, the choice of the observables measured by the signal detector, and so forth. In this review, rather than attempting to give other general criteria guaranteeing relativistic causality, we discuss theoretical descriptions of definite signal transmission experiments. All the important aspects of the theory are then automatically taken into account. The treatment will be sufficiently general for us to be able to assert that no quantum electrodynamic methods of signal transmission lead to a violation of relativistic causality.

On the basis of Refs. 78, 10 and of other investigations we can formulate two general requirements which must be satisfied by a properly posed signal transmission problem:

1) it is necessary to consider the transmission of information from one finite region of space V_S (the source) to another finite region, V_D, where the detector is localized;

2) the person transmitting the signal must be free to send the signal (or decide not to) at any arbitrary time t_0 (see §7 of Ref. 10). In particular, a periodic process

cannot be used for signal transmission.

Thus, we are concerned with an essentially nonstationary problem, and this must be solved in quantum field theory.

In Sec. 2, we give a critical review of the formulations of the signal velocity problems in quantum electrodynamics known in the literature. The main aim of this section is to introduce the problem and not to be exhaustive. Section 3 solves a problem close to the ones considered in Sec. 2; namely, a rather simple but fairly realistic model of quantum electrodynamics. The signal source is an external current localized in S. The arrival of a signal in D is determined by a change of familiar observables: the coordinate, momentum, and energy of a charged particle localized in D. It is shown that to prove relativistic causality one must take into account appropriately quantum-mechanical and, in particular, quantum-electrodynamical features of the signal transmission problem. In Sec. 4, we give a general formulation and exact solution of the signal transmission problem in quantum electrodynamics, the charged particles also being described by a second-quantized field.

2. HISTORY OF THE SIGNAL VELOCITY PROBLEM IN QUANTUM ELECTRODYNAMICS

a) Soon after the creation of quantum electrodynamics, Kikuchi[11] calculated in 1930 the following problem at the suggestion of Heisenberg. An excited atom S serves as the source of the signal. More precisely, at time $t=0$ the following initial state is specified: The atom S is excited and there are no photons. The energy density of the electromagnetic field at times $t>0$ at a distance R from the atom is then calculated. Instead of this, Fermi[12] in 1932 calculated the probability of excitation at time $t>0$ of a second atom D at a distance R from the first (Fig. 1). In both cases, a causal result was obtained—the energy density or the excitation probability of D is equal to zero until the time $t=R/c$.

The calculations were made in the first nonvanishing order of perturbation theory. Besides this approximation, two others were made. Let us consider what they were. Fermi calculated the probability amplitude $a_{SD*}(t)$ that the initial state S^*D (S excited, D not excited, no photons) goes over at time t into the state SD^* (S not excited, D excited, no photons). The main mechanism of excitation transfer is as follows: S emits a (virtual) photon, and D absorbs it (for details, see Appendix A, in which the solution of Fermi's problem by "covariant" perturbation theory[13] is presented). The first simplifying approximation can be described as follows: integration with respect to the modulus of the momentum k of the photon which carries the excitation from S to D can be made, not within limits from 0 to ∞, but between $-\infty$ and $+\infty$. Later, in Ref. 14 it was shown that the causal result is basically a consequence of this assumption: a more accurate calculation of the integrals with respect to k does not give zero at times $t<R/c$.

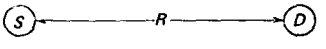

FIG. 1.

The second approximation is usually called the dipole approximation. It is assumed that the electron of the atom emits or absorbs a photon, not at its own position, but at the center of the potential which binds the electron. This is a much weaker assumption than the first; see Appendix A. If it is not made, the amplitude $a_{SD*}(t)$ acquires a trivial noncausal contribution of the following origin. In the case of real potentials that bind an electron, there is a nonvanishing probability of finding the electron of the atom S far from the center of the potential, in particular, near the atom D. As a result, the atom may be excited immediately after the time $t=0$. However, the probability of this is exponentially small if R is much greater than the characteristic atomic length l (for the hydrogen atom, it is proportional to $e^{-R/l}$). The dipole approximation is indeed admissable since it rids the result of the trivial noncausal contribution. Note that the noncausal effect discovered in Ref. 14 is proportional to some inverse power of the ratio R/λ, where λ is the mean wavelength of the exchange photon, $\lambda \gg l$, i.e., it is incomparably larger than the trivial contribution just discussed.

b) Fermi's problem was solved once more in 1949 with allowance for damping.[15,16] On the basis of these calculations, Heitler asserts at the end of §20 of his book[17] in connection with the probability $w=|a_{SD*}(t)|^2$ that "the probability w is exactly zero for all $t<R/c$." But in fact this causal result is a consequence of the same replacement of $\int_0^\infty dk$ by $\int_{-\infty}^{+\infty} dk$. In reality, it was shown in Refs. 15, 16 merely that in the theory of damping the situation as regards relativistic causality is no worse than in ordinary perturbation theory.

c) After the creation of covariant perturbation theory, the discussion of relativistic causality in quantum electrodynamics was continued in the framework of the same problem of Fermi, but in somewhat different terms. In Appendix A it is shown that the amplitude $a_{SD*}(t)$ of Fermi's problem is proportional to the expression

$$\int d^4x \int d^4y \sum_{mn} j^m_{D*D}(y)\, D^c_{mn}(y-x)\, j^n_{SS*}(x), \tag{2.1}$$

where j_{D*D} is the matrix element of the electron current of atom D. The aim of the investigations that we shall now discuss was to prove the following assertion: although the propagator D^c does not vanish outside the light cone, the expression (2.1) nevertheless does not lead to a violation of relativistic causality. The idea behind the arguments is presented in its most general form in Ref. 18. We take the expression (2.1), but we assume that the integration with respect to x_0 and y_0 is made between $-\infty$ and $+\infty$ (and not between 0 and t, as in Fermi's problem; see the passage in Appendix A after Eq. (A.3)). We take the Lorentz gauge, when m, $n=1$, 2, 3, 4 and $D^c_{mn}=\delta_{mn}D^c$. We obtain the expression

$$M=\int_{-\infty}^{\infty} dx_0 \int d^3x \int_{-\infty}^{\infty} dy_0 \int d^3y\, j^\mu_{D*D}(y)\, D^c(y-x)\, j^\mu_{SS*}(x), \tag{2.2}$$

which was considered in Ref. 18. One can show that $D^c=D^{(-)}_{\text{adv}}+D^{(+)}_{\text{ret}}$, where $(-)$ and $(+)$ denote, respectively, the negative-frequency and positive-frequency parts of

the corresponding function[2] (i.e., for example, the function $\theta(-k_0)$ is added to the momentum transform of D_{adv}). For all functions Δ and j expanded in a Fourier integral, we have

$$\int_{-\infty}^{+\infty} \Delta^{(\pm)}(x-y)\, j^{(\mp)}(y)\, dy_0 \sim \int_{-\infty}^{+\infty} \tilde{\Delta}(k_0)\,\theta(\pm k_0)\,\tilde{j}(k_0)\,\theta(\mp k_0)\,dk_0 = 0. \quad (2.3)$$

Therefore

$$M = \int\int d^4x\, d^4y\, j_\mu(y)\,[D^{(-)}_{adv} + D^{(+)}_{ret}]\,[j^{(+)}_\mu(x) + j^{(-)}_\mu(x)]$$
$$= \int\int d^4x\, d^4y\, j_\mu(y)\,[D_{adv} j^{(-)}_\mu(x) + D_{ret} j^{(+)}_\mu(x)]. \quad (2.4)$$

If the current $j_\mu(x)$ is such that $j_\mu(x) = j^{(+)}_\mu(x)$, then

$$M = \int d^4x \int d^4y\, j_\mu(y)\, D_{ret}(y-x)\, j_\mu(x), \quad (2.5)$$

and then the interaction between the currents is purely retarded and causal. The matrix element of a current for which $j_\mu = j^{(+)}_\mu$ contains only positive-energy components, and the energy of the initial state is greater than that of the final state and a photon (with positive energy) can be emitted. For the special case

$$j_{SS*}(x) \exp[\,i(E_S - E_{S*})\,x_0\,], \quad E_{S*} > E_S,$$

the result (2.5) is proved in §37 of the book of Ref. 19. For a generalization of Akhiezer and Berestetskiĭ's treatment see the end of Ch. 12 in the book of Ref. 20. Similar considerations can be found in other investigations. We shall show that they do not bear on the problem of the signal velocity. The first assumption in the derivation of (2.5) is that of the infinite (and not finite) limits of integration with respect to the time. The second is the assumption that $j_\mu(x) = j^{(+)}_\mu(x)$. It rules out the possibility of currents localized in time (equal, for example, to zero until the time $t = 0$)[3]). This means that one is essentially considering the stationary problem of the interaction of two currents of the type of electron scattering. In such a problem, there is no signal propagation. To transmit a signal, it is necessary to change the current $j_\mu(x)$ at some time, without changing it at earlier times. A signal source of this kind cannot be described by a positive frequency current.

d) Let us also consider Fierz's paper.[21] The opinion is frequently advanced that in it the following is proven: even if the expression (2.1) exhibits a noncausal behavior, the time–energy uncertainty relation does not permit one to say that it is actually observable.

To elucidate this assertion, let us suppose that the probability of excitation of atom D in Fermi's problem before the time R/c is appreciable only if t is less than R/c over a time of order λ/c, where λ is the wavelength of the photon emitted by atom S, $1/\lambda \sim E_{S*} - E_S$. Note further than in Fermi's problem the process of preparing the excited state of the atom S is not considered (in what follows, we shall consider investigations in which this shortcoming is rectified). The photon which excites D can, however, appear near S during the time of excitation of S before the time $t = 0$. The excitation has a duration of order λ/c. This can be seen, for example, from the well-known expression for the excitation probability (see Sec. 29 in Ref. 3)

$$|b(t)|^2 = \left[4\,|H_{BS}|^2 \sin^2(E_B - E_0)\,\frac{t}{2}\right](E_B - E_0)^{-2}.$$

It is seen that the noncausal advance of the signal is comparable with the uncertainty in the time of the actual emission of the signal.

However, exact calculation of (2.1) shows that for $t < R/c$ the amplitude $a_{SD*}(t)$ depends on t through the quantity $R - t$ as a power: $a_{SD*}(t) \sim \lambda/|R-t|$ if $|R-t|/\lambda \gg 1$; see Ref. 14 and the end of Appendix A. Such a decrease of $a_{SD*}(t)$ with increasing $R - t$ is too slow for one to say that the noncausal advance is of order λ/c (we could say this to be so if, for example $a_{SD*}(t) \sim \exp(-|R-t|/\lambda)$). Note that Fierz in his calculations used, in particular, the inadmissible replacement discussed above of D^c by D_{ret}. Therefore, it cannot be assumed that Fierz established an illusory nature of the noncausal advance.

We may mention that Fermi's problem can be modified in such a way that the excitation time and the magnitude of the noncausal advance are completely independent quantities. Suppose, for example, that the atom S is a three-level system with levels $E_0 < E_S < E_{S*}$ such that $E_{S*} - E_S \ll E_{S*} - E_0$.[14] Such a system can be excited in a time $\sim 1/(E_{S*} - E_0)$. The level E_{S*} may be metastable and de-excited first preferentially to the level E_S, emitting a photon of energy $\Delta = E_{S*} - E_S$, which is then absorbed by the atom D.

e) The significance of the already mentioned paper[14] of 1964 is in showing that the solution of Fermi's problem leads to a noncausal result that must be regarded as an observable effect in the framework of quantum mechanics. (The same noncausal result was obtained in Ref. 22 in the framework of a somewhat different but, in principle, similarly posed problem.)

Holding the same opinion, Ferretti in 1968 pointed out that the problem solved by Fermi does not completely correspond to a signal transmission problem.[23] The condition that there be no photons in the final state SD^* in fact presupposes that everywhere in space there are instruments that detect photons and cases are selected when these instruments detect nothing. But the signal detector must be localized in the finite volume V_D (see Sec. 1). For example, it is necessary to measure only the state of atom D irrespective of the state of other parts of the system. For the signal transmission problem, a more appropriate quantity is

$$S_{n\gamma}\,|\langle SD^* n\gamma \,|\, U(t,0) \,|\, S^* D\rangle|^2, \quad (2.6)$$

where $U(t, 0)$ is the evolution operator, $SD^* n\gamma$ is the

---

[2]From the relations

$$D_{ret}(x) = \theta(x_0)\,D(x) = \theta(x_0)\,D^{(-)} + \theta(x_0)\,D^{(+)},\ D_{adv}(x) = -\theta)-x_0)\,D(x),$$
$$D^c(x) = \theta(x_0)\,D^{(+)}(x) - \theta(-x_0)\,D^{(-)}(x)$$

there follow two equations: $D_{ret} = D^c + D^{(-)}$ and $D_{adv} = D^c - D^{(+)}$. The positive-frequency part of the first equation has the form $D^{(+)}_{ret} = D^{c(+)}$, the negative-frequency part of the second equation has the form $D^{(-)}_{adv} = D^{c(-)}$, from which it follows that $D^c = D^{(-)}_{adv} + D^{(+)}_{ret}$.

[3]The Fourier transform of the function $j(k_0)\theta(k_0)$ is an analytic function of the time and cannot vanish in any time interval (for example, $(-\infty, 0)$). Conversely, a current localized in a finite interval cannot be a purely positive-frequency current.

state with n photons, and $S_{n\gamma}$ stands for not only summation over n but also for integration with respect to the momenta of the photons and summation over their polarizations. The expression (2.6) is not equal to

$$|a_{SD^*}(t)|^2 = |\langle SD^*|U(t,0)|S^*D\rangle|^2 \tag{2.7}$$

even in the first nonvanishing order of perturbation theory. Indeed, if, as usual, we describe the initial and final states by the eigenfunctions of the free part H_0 of the total Hamiltonian (i.e., by "bare" states), then the state S^*D can go over into the state $SD^*\gamma\gamma$: S has emitted a photon, D has also emitted a "different" photon, but it has also been excited. This occurs in the same e^2 order of perturbation theory. Although the corresponding amplitude is much smaller than (2.7), the noncausal effects under consideration are also small.

However, having started to modify the object of our calculations in this manner, we must then also take into account another consequence of the use of the "bare" formalism: at a time $t>0$ the atom D may be found to be excited even if S was not initially excited (D can be excited with the emission of a photon). Therefore, Ferretti subtracts the "background" from (2.6). He defines this background as the quantity analogous to (2.6) but in the case when the atom S is absent,

$$S_{n\gamma}|\langle D^*n\gamma|U(t,0)|D\rangle|^2. \tag{2.8}$$

(Actually, apparently it would be sufficient to subtract a quantity of the type (2.8) with the initial state SD, i.e., the atom S is not excited.) Ferretti obtained a causal result: the difference between (2.6) and (2.8) in the first nonvanishing approximation is exactly zero for $t<R/c$. Unfortunately, the calculation was made for the case that in ordinary (3+1)-dimensional electrodynamics can be described as follows. Two infinite planar (thin) layers are taken as the "atoms" of the source and the detector. Then the problem reduces to a spatially one-dimensional case and is simplified. However, to establish the fact of excitation of such an infinite "atom" requires an infinite time. In the signal velocity problem, the source and the detector must be localized in finite volumes. One can of course accept that Ferretti proved relativistic causality for two-dimensional electrodynamics. However, the relativistic causality of electrodynamics in the case of the real (3+1) dimensionality remained open.

f) In Refs. 24, 25 an attempt was made to improve the formulation of the problem of the velocity of transfer of excitation from atom S to atom D by a different method of describing the excited states of the atoms which takes into account the presence of interaction. Instead of "bare" creation and annihilation operators (of photons and excitation quanta of the atoms), "physical" operators were introduced (whose no-particle vector coincides with the physical vacuum). In contrast to Ref. 23 the result was found to be noncausal. However, it can basically be explained by the circumstance that in terms of "physical" operators the interaction turns out to be nonlocal; see Sec. 4 of Ref. 26. One can introduce other "physical" operators, for which the magnitude of the noncausal effect is reduced. The result of Ref. 27 is explained in the same manner. In contrast to all the previously discussed investigations, in Ref. 27 the process of excitation of atom S is described theoretically (by means of a change of the potential that binds the electron of the atom). However, the signal arrival is determined by the transition of atom D to the excited state, which is described by means of "physical" operators.

g) In Ref. 28 the present author considered not only the signal source just described but also a source in the form of an external current localized in a region V_S and switched on at the time $t=0$. The "background" subtraction device was used. The transition amplitude was not calculated, but rather the changes in the distributions with respect to the coordinate q, the momentum p, and the energy of the electron of the atom D due to the signal source. All calculations were made without the use of perturbation theory in the framework of exactly soluble models of quantum electrodynamics (one or two nonrelativistic electrons in an oscillator potential interacting in a dipole manner with a quantized electromagnetic field). It was shown that the distribution with respect to q (or the square of the modulus $|u(q)|^2$ of the wave function of the D electron in the coordinate representation) begins to change only after a time R/c. But the momentum distribution (or the phase of the function $u(q)$), and with it the distribution with respect to the energy of the electron, begins to change immediately after the source is switched on. In the following section, these results will be reproduced for a different model, and it will be shown that the noncausal change in the phase of $u(q)$ is to be regarded as just as unobservable as the change of a vector potential by the gradient of a scalar function. Thus, all of the problems just described lead to signal transmission with velocity not exceeding c.

h) Of course, a signal source in the form of a localized external current or potential that is switched on satisfies conditions 1) and 2) on the signal source formulated at the end of Sec. 1. It would seem that these conditions are also satisfied (to a reasonable extent) by Fermi's source in the form of an excited state of an atom which is simply specified at the initial time. However, in Sec. 5 of Ref. 28 it was shown that the coordinate distribution of the D electron behaves noncausally in the case of the Fermi source even if the "background" is subtracted. In contrast to Ferretti's calculation, this was shown for ordinary (3+1) spacetime. One can relate this result to the instantaneous spreading of a wave packet; see the introduction. In this case, the signal is transmitted basically by a single photon. The fact that this photon is emitted by the atom S means that immediately after the emission the photon can be assumed to be fairly well localized in a region with the dimensions of the atom. The excitation of the atom D at time t shows that the probability of finding the photon in the region V_D at this time is not zero. The noncausal result, in both the case of packet spreading and the case of Fermi's problem, is due to the nonvanishing outside the light cone of the function D^c (which coincides with the function $D^{(+)}$ in the problem of packet spreading).

In contrast to Fermi's source, a localized current

generates a state of the electromagnetic field (with indeterminate number of photons) that does not spread instantaneously (see Refs. 29, 30, in which these states are said to be "strictly localized"; in Ref. 9 they are said to be coherent). It must be recognized that for the formulation of the signal velocity problem an important aspect is that the signal source be described by a current or potential that is switched on.

1) The general solution of the problem of relativistic causality in quantum electrodynamics will be set forth in Sec. 4 on the basis of Refs. 31, 32. A different possible general approach has been sketched by Bell in Sec. 7 of Ref. 33.

j) The present review does not pretend to completeness. In particular, we do not discuss Refs. 10, 34–36. We have not considered the literature devoted to the instantaneous spreading of packets of free relativistic particles. One can take the point of view that the theory of free particles is too lacking in content to describe the problem of signal transmission.[4)]

3. EXTERNAL CURRENT AS SOURCE OF A SIGNAL AND A NONRELATIVISTIC ELECTRON AS DETECTOR

We consider the problem of the signal velocity in which the conditions 1) and 2) on the signal transmission formulated at the end of Sec. 1 are satisfied. This problem is fairly similar to the one proposed by Fermi (see Sec. 2a–2e), but is solved differently. The problem can serve as an introduction to the general formulation presented in Sec. 4, which uses not entirely familiar field descriptions of observable physical quantities.

a) The source of the signal is an external current J_μ which is localized in a certain region V_S and is switched on at time $t=0$. The current emits a quantized electromagnetic field. The arrival of the signal is detected by means of a nonrelativistic spinless charged particle, which for brevity we shall call an electron. This electron is localized in the region V_D by an external potential and interacts with the quantized electromagnetic field. We assume that there is an instrument which measures a state of the electron.

The state vector $\Phi(t)$ of the system describes both the electron and the field (the photons). The state of the electron can be described by a density matrix (obtained by suitable summation and integration of $|\Phi(t)|^2$ with respect to the photon variables). Knowing it, we can find the probability that the electron coordinate is equal to q and the momentum equal to p, i.e., we can obtain the distributions with respect to q and with respect to p. We shall find these distributions directly by calculating the mean values, mean squares, and other moments of the operators $\hat{q}$ and $\hat{p}$ (we recall that the distribution with respect to q is equal to $\sum_n q^n\langle\Phi|\hat{q}^n|\Phi\rangle$). More precisely, we shall calculate the changes in these distributions due to the switching on of the external current. It is natural to assume that the signal arrives at the time when these changes appear.

We consider the expectation value of any Schrödinger operator A (A may be equal to $\hat{q}, \hat{q}^2, \ldots, \hat{p}, \hat{p}^2$, etc.) in the state $\Phi(t)$, which is related to the initial state Φ_0 by the relation $\Phi(t)=U(t,0)\Phi_0$ (where U satisfies the Schrödinger equation: $i\partial_t U = HU$):

$$\langle\Phi(t),\ A\Phi(t)\rangle = \langle U\Phi_0,\ AU\Phi_0\rangle = \langle\Phi_0,\ U^+AU\Phi_0\rangle = \langle\Phi_0,\ A_H(t)\Phi_0\rangle. \quad (3.1)$$

We see that this expectation value is equal to the expectation value of the corresponding Heisenberg operator $A_H(t)=U^+(t,0)AU(t,0)$ in the initial state Φ_0 (Φ_0 here is a constant state vector in the Heisenberg picture). Instead of calculating $U(t,0)$ and $\Phi(t)$, we shall find $A_H(t)$. This is a simpler problem, and in addition, its solution has a more perspicuous meaning, as we shall see.

We compare the moments of the operator q at the time t in the case when the current is switched on and when it is not:

$$\langle\Phi_J(t),\ q^n\Phi_J(t)\rangle - \langle\Phi(t),\ q^n\Phi(t)\rangle = \langle\Phi_0,\ [q_J^n(t)-q^n(t)]\Phi_0\rangle. \quad (3.2)$$

Here we have used the relation (3.1). When the current is switched on, the total Hamiltonian depends explicitly on the time, and we have

$$\Phi_J(t)=U_J(t,0)\Phi_0,\quad q_J(t)=U_J^+qU_J,$$
$$\frac{i\partial U_J}{\partial t}=\mathscr{H}_J(t)U_J. \quad (3.3)$$

If the current is not switched on at the time $t=0$, the Hamiltonian does not depend on the time (and is, for example, equal to $\mathscr{H}_J(t=0)$).

It can be seen from (3.2) that the moments (and the q distribution) begin to differ when the Heisenberg operator $q_J(t)$ begins to differ from the Heisenberg operator

$$q(t)=\exp[it\mathscr{H}_J(t=0)]\,q\exp[-it\mathscr{H}_J(t=0)]. \quad (3.4)$$

b) We first consider the problem of finding the Heisenberg operators in the Coulomb gauge (the Lorentz gauge is discussed later). The Hamiltonian $\mathscr{H}(t)$ has the form

$$\mathscr{H}(t)=\frac{1}{2m}[\mathbf{p}-e\mathbf{A}_\perp(\mathbf{q})]^2+W(\mathbf{q})+\frac{1}{8\pi}\int d^3x\,[E_\perp^2(\mathbf{x})+H^2(\mathbf{x})]$$
$$+\int_{V_S}d^3x\,(\mathbf{J}(\mathbf{x},t)\mathbf{A}_\perp(\mathbf{x}))+e\int_{V_S}\frac{d^3xJ_0(\mathbf{x},t)}{|\mathbf{x}-\mathbf{q}|},\ \operatorname{div}\mathbf{A}_\perp=0 \quad (3.5)$$

(see, for example, §13 of Ref. 17 or §17 of Ref. 38). We find the equations for $\mathbf{q}_J(t)$ and $\mathbf{p}_J(t)$. First, we calculate

$$\frac{\partial\mathbf{q}_J(t)}{\partial t}=-i[\mathbf{q}_J(t),\ \mathscr{H}_T(t)]; \quad (3.6)$$

here, $\mathscr{H}_H(t)=U_J^+\mathscr{H}(t)U_J$ is the Heisenberg operator of the Hamiltonian. In view of the relation of the form $U_J^+p^2U_J=U_J^+pU_JU_J^+pU_J=p_J^2$, it depends on the Heisenberg operators $q_J(t), p_J(t)$, etc, in the same way as $\mathscr{H}(t)$ (see (3.4)) depends on the Schrödinger operators $q, p, \ldots$. The equal-time commutation relations for the Heisenberg operators are also the same as for the Schrödinger operators. Therefore, (3.6) can be readily calculated:

$$\frac{\partial\mathbf{q}_J(t)}{\partial t}=\frac{1}{m}[\mathbf{p}_J(t)-e\mathbf{A}_{J\perp}^T(\mathbf{q}_J(t),t)]\equiv\mathbf{v}. \quad (3.7)$$

[4)] In particular, one cannot therefore deduce from the fact of instantaneous packet spreading that it is impossible to have finite packets (when the wave function of the particle vanishes outside a finite region).[4]

Further, we can calculate $\partial^2 \mathbf{q}_J/\partial t^2 = -i[\mathbf{v}, \mathscr{H}_H]$. The calculation of $[\mathbf{v}, \mathscr{H}_H]$ simplifies considerably if, using the commutation relations of the Coulomb gauge (see, for example, §§48–49 of the book of Ref. 3), we find the commutators

$$[v_i, v_j] = \frac{ie}{m^2}\varepsilon_{ijk}H_k \qquad (i, j, k = x, y, z), \tag{3.8}$$

$$\left[v_i, \frac{1}{8\pi}\int d^3x E_\perp^2 + e\int \frac{d^3x J_0(\mathbf{x}, t)}{|\mathbf{x}-\mathbf{q}|}\right] = ieE_i(\mathbf{q}, t), \tag{3.9}$$

$$\mathbf{E}(x) = -\frac{\partial \mathbf{A}_\perp(x)}{\partial t} - \operatorname{grad}_x \int \frac{d^3y J_0(\mathbf{y}, t)}{|\mathbf{y}-\mathbf{x}|}. \tag{3.10}$$

By means of them, we obtain

$$m\frac{\partial^2 \mathbf{q}_J(t)}{\partial t^2} = -\operatorname{grad} W(\mathbf{q}_J(t)) + e\mathbf{E}_J(\mathbf{q}_J(t), t) + \frac{e}{2}\{[\mathbf{v}\mathbf{H}_J(\mathbf{q}_J(t), t)] - [\mathbf{H}_J(\mathbf{q}_J(t), t)\mathbf{v}]\}. \tag{3.11}$$

The right-hand side of (3.11) contains the operator of the Lorentz force. A similar equation is obtained for $\mathbf{q}(t)$ without the index J. Equation (3.11) must be solved simultaneously with the equations for the operators $\mathbf{E}_J$ and $\mathbf{H}_J$. We shall not write out these (Maxwell) equations. We take the results that we need for the operators $\mathbf{E}$ and $\mathbf{H}$ from Appendix B.

To solve approximately the Heisenberg equations obtained above, we expand $\mathbf{q}_J(t)$, $\mathbf{E}_J$, $\mathbf{H}_J$ and $\mathbf{q}(t), \mathbf{E}, \mathbf{H}$ in terms of the coupling constant e: $q_J = \sum_n e^n q_J^{(n)}$, etc. We substitute these expansions in the equations and equate the coefficients of equal powers of e. In the zeroth approximation we obtain

$$m\frac{\partial^2 \mathbf{q}_J^{(0)}(t)}{\partial t^2} = -\operatorname{grad} W(q_J^{(0)}(t)),$$
$$m\frac{\partial^2 \mathbf{q}^{(0)}(t)}{\partial t^2} = -\operatorname{grad} W(q^{(0)}(t)). \tag{3.12}$$

We seek solutions of the Heisenberg equations in the form of definite expressions for the Heisenberg operators in terms of always the same Schrödinger operators, i.e., the operators with which the Heisenberg operators must coincide at $t=0$ when the current J_μ was not yet switched on. In particular,

$$q_J(t=0) = q(t=0) = q \tag{3.13}$$

(see (3.3) and (3.4)). We shall assume that the operators $q_J^{(0)}(t)$ and $q^{(0)}(t)$ coincide at $t=0$ with the Schrödinger operator q, and the remaining $q_J^{(n)}(t)$ and $q^{(n)}(t)$ with $n \geq 1$ are equal to zero at $t=0$. In this case, (3.13) will of course be satisfied.

Since Eqs. (3.12) for $q_J^{(0)}(t)$ and $q^{(0)}(t)$ are the same and the initial conditions are the same,[5] $q_J^{(0)}(t) = q^{(0)}(t)$ for all t.

The equation for $q_J^{(1)}$ has the form

$$m\frac{\partial^2 \mathbf{q}_J^{(1)}}{\partial t^2} = \mathbf{E}_J^{(1)} + \frac{1}{2}\{[\partial_t \mathbf{q}^{(0)}\mathbf{H}_J^{(1)}] - [\mathbf{H}_J^{(1)}\partial_t \mathbf{q}^{(0)}]\}. \tag{3.14}$$

Since $E(q^{(0)} + eq^{(1)}) = E(q^{(0)}) + eq^{(1)}E'(q^{(0)}) + \ldots$, we can assume that $\mathbf{E}_q^{(0)}$ and $\mathbf{H}_J^{(0)}$ in (3.14) do not depend on $q^{(0)}(t)$.

[5]Equations (3.12) are of second order, and therefore, besides (3.13), it is also necessary to remember that $(\partial q_J/\partial t)_{t=0} = (\partial q/\partial t)_{t=0}$ by virtue of $p_J(t=0) = p(t=0)$ and $A_{J\perp}(t=0) = A_\perp(t=0)$ (see (3.7)).

The equation for $q^{(1)}(t)$ has the same form except that none of the quantities have the subscript J. Suppose the current J_μ is localized in a region V_S near the origin of coordinates, and the electron is in the region V_D,[6] which is situated at distance R from V_S. It follows from the results of Appendix B that the differences $\mathbf{E}_J^{(0)}(\mathbf{x}, t) - \mathbf{E}^{(0)}(\mathbf{x}, t)$ and $\mathbf{H}_J^{(0)}(\mathbf{x}, t) - \mathbf{H}^{(0)}(\mathbf{x}, t)$ are retarded functions of the current J_μ. Therefore, they vanish for $x \in V_D$ and $t < R/c$. Because of this, the right-hand side of (3.14) is equal to the right-hand side of the analogous equation for $\mathbf{q}^{(1)}(t)$, from which it follows that $\partial^2[\mathbf{q}_J^{(1)}(t) - \mathbf{q}^{(1)}(t)]/\partial t^2 = 0$ for $t < R/c$. By virtue of the null initial conditions, $q_J^{(1)}(t) - q^{(1)}(t) = 0$ for $t < R/c$, i.e., in the first approximation

$$q_J(t) = q(t), \qquad t < \frac{R}{c}. \tag{3.15}$$

An exact solution (of the more general problem) is discussed in the next section. We now note only the result (3.15) was obtained in Ref. 28 without the use of perturbation theory (but in the dipole approximation and for the special case $W(q) \sim q^2$). Turning to (3.2), we conclude that the coordinate distribution of the electron does not change prior to the time R/c.

However, the behavior of the distribution with respect to the electron momentum p is noncausal; for it follows from (3.6) that

$$\mathbf{p}_J(t) - \mathbf{p}(t) - e[\mathbf{A}_{J\perp}(\mathbf{q}_J(t), t) - \mathbf{A}_\perp(\mathbf{q}(t), t)] = m\frac{\partial}{\partial t}[\mathbf{q}_J(t) - \mathbf{q}(t)]. \tag{3.16}$$

We show that the difference $\mathbf{A}_{J\perp} - \mathbf{A}_\perp$ in (3.16) does not vanish for $t < R/c$. The transverse vector potential $\mathbf{A}_\perp$ can be expressed as follows in terms of the vector potential $\mathbf{A}$ of the Lorentz gauge (see, for example, §80 in Ref. 39):

$$\mathbf{A}_\perp(\mathbf{x}, t) = \mathbf{A}(\mathbf{x}, t) - \nabla u(\mathbf{x}, t),$$
$$u(\mathbf{x}, t) = -\frac{1}{4\pi}\int d^3y \operatorname{div} A(\mathbf{y}, t)|\mathbf{x}-\mathbf{y}|^{-1}; \tag{3.17}$$

here $u(\mathbf{x}, t)$ is the potential of the longitudinal part $\mathbf{A}_\perp$ of the vector potential. Using

$$\Delta\frac{1}{|\mathbf{x}-\mathbf{y}|} = -4\pi\delta^{(3)}(\mathbf{x}-\mathbf{y}) \tag{3.18}$$

we can verify that the divergence of the expression $\mathbf{A} - \nabla u$ really is zero. In Appendix B, it is shown that $\mathbf{A}$ changes in V_D only after a time R/c. However, u in (3.17) is expressed nonlocally in terms of $\mathbf{A}$, and therefore $\mathbf{A}_{J\perp}(\mathbf{x}, t) - \mathbf{A}_\perp(\mathbf{x}, t)$ does not vanish for $x \in V_D$ and $t < R/c$:

$$\mathbf{A}_{J\perp}^{(0)}(\mathbf{x}, t) - \mathbf{A}_\perp^{(0)}(\mathbf{x}, t) = \nabla(u^{(0)} - u_J^{(0)}) = \nabla\frac{1}{4\pi}\int d^3y \frac{1}{|\mathbf{x}-\mathbf{y}|}\sum_k \frac{\partial}{\partial y_k}\int d^4z D_{\mathrm{ret}}(y-z) J_k(z) \equiv \nabla\lambda(\mathbf{x}, t) \tag{3.19}$$

(we have used (3.17) and (B.8).

[6]If the potential $W(q)$ in (3.5) is infinitely high on the boundary of the region V_D, then all the eigenfunctions of $(p^2/2m) + W(q)$ vanish outside V_D and the electron is strictly localized within V_D. In the case of realistic potentials, one can speak of localization of the electron in V_D only if one ignores the exponentially small probability that the electron may also be outside V_D. Then (3.15) must also be regarded as true under the condition that the same quantities are neglected.

Turning now to (3.16) and bearing in mind that the right-hand side of (3.16) vanishes for $t<R/c$ by virtue of (3.15), we establish that

$$\mathbf{p}_J(t) - \mathbf{p}(t) = e\nabla\lambda(q(t), t), \tag{3.20}$$

which is not equal to zero for $t<R/c$.

However, the expressions (3.19) and (3.20) for $t<R/c$ have the form of a gauge transformation:

$$\mathbf{A}'_\perp = \mathbf{A}_\perp + \nabla\chi, \quad \mathbf{p}' = \mathbf{p} + e\nabla\chi, \quad \Delta\chi = 0. \tag{3.21}$$

which is permitted in the framework of the Coulomb gauge since it conserves the transversality of the vector potential. Indeed, $\mathrm{div}\mathbf{A}_\perp = \mathrm{div}\mathbf{A}_\perp = 0$, because div grad $\lambda = \Delta\lambda(\mathbf{x}, t) = 0$ for $x \in V_D$ and $t<R/c$ (using (3.18), one can show that $\Delta\lambda(\mathbf{x}, t)$ is a retarded function of the current and equal to zero for $|\mathbf{x}-\mathbf{x}_S| > tc$ for all $\mathbf{x}_S \in V_S$). We see that (3.19) and (3.20) are a guage transformation only for $|\mathbf{x}-\mathbf{x}_S| > tc$. A further difference from an ordinary gauge transformation is that (3.19) and (3.20) have a dynamical origin: the function λ is not an arbitrary (harmonic) function but is determined by the external current. Finally, one can show that the equations of the theory in the Coulomb gauge are not invariant under (3.19) and (3.20) are only a special case of (3.21) from the point of view of an observer who is situated together with his instruments in V_D and is not able to observe anything outside V_D. On the basis of this, we may assert that the transformations (3.19) and (3.20) from $\mathbf{A}_{J\perp}$ and $\mathbf{p}_J$ to $\mathbf{A}_\perp$ and $\mathbf{p}$ do not lead to observable consequences in V_D as long as $t<R/c$. We shall say that these are quasigauge transformations. The transformations (3.19) and (3.20) can be written in the form

$$\mathbf{A}'_\perp = \mathbf{A}_\perp + \nabla\lambda, \quad \mathbf{p}' = \mathbf{p}, \quad \Phi' = \Phi e^{ie\lambda}, \tag{3.22}$$

i.e., instead of transforming the Heisenberg operator p one can transform the wave function of the system, multiplying it by a phase factor that depends only on the electron coordinate (in particular, a gauge transformation is expressed in precisely this way in §18 of the book of Ref. 38). One can therefore say that (3.20) is a "premature" (but unobservable) change in the phase of the electron wave function in the coordinate representation (strictly speaking, one should speak of the phases of the elements $\langle q_1|\rho|q_2\rangle$ of the electron density matrix; see subsection a) of this section).

We now demonstrate how an apparently noncausal result can be obtained by means of ordinary perturbation theory. Suppose the bound states of the electron are described by the eigenfunctions of the operator $\mathcal{H}_e = (p^2/2m) + W(q)$. The distribution with respect to $\mathcal{H}_e$ behaves noncausally, like the distribution with respect to p. This means that if the atom D up to the time $t=0$ was in the ground state, then immediately after the switching on of the current it could already be found in an excited state. How is this reconciled with what we have said above (before and after Eq. (3.22))? The point is that the change $\Phi \to \Phi\exp[ie\lambda(\mathbf{x}, t)]$ in the wave function has the readily recognized form of a gauge transformation if Φ is a function of the electron coordinate $\mathbf{x}$ (and certain photon variables). The same transformation expressed, for example, in the representation of the electron moment p no longer reduces to multiplication by a phase factor: the modulus of the wave function also changes, i.e., the p distribution changes. Perturbation theory actually establishes that in the representation of the eigenfunctions of $\mathcal{H}_e$ the wave function changes when the current is switched on in such a way that the distribution with respect to the eigenvalues of $\mathcal{H}_e$ changes immediately. We have succeeded in showing above that in the x representation this change has the form of an unobservable quasigauge transformation (3.22). This result can also be formulated as follows: the Heisenberg operators $p_J^2(t)/2m + W(q_J)$ and $(p^2(t)/2m) + W(q)$ describe the same observable as long as $t<R/c$ (just as $\mathbf{A}_{J\perp}$ and $\mathbf{A}_{J\perp}$ and $\mathbf{A}_\perp$ describe the same field strengths $\mathbf{E}$ and $\mathbf{H}$).

c) In the Lorentz gauge, the Hamiltonian $\mathcal{H}(t)$ has a different form. In (3.5), it is necessary to drop the subscript $\perp$ of $A_\perp$ and $E_\perp$, replace the last two terms by $\int d^3x J_\mu A_\mu$, and replace the energy operator of the electromagnetic field by the expression (17.7) of Ref. 38. In addition, it is necessary to take into account the Lorentz auxiliary condition: physically realizable states can correspond only to vectors Φ that satisfy the equation $\mathcal{L}\Phi = 0$. The operator $\mathcal{L}$ in the Heisenberg picture is equal either to $\partial_\mu A_\mu$ in the form of the auxiliary condition proposed by Fermi and used by Dirac in the book of Ref. 39 or to $(\partial_\mu A_\mu)^{(-)}$ (the Gupta–Bleuler form). In the Heisenberg picture, the initial vector Φ_0 can be taken as Φ. Since $\Box\partial_\mu A_\mu = 0$,[17,39] the auxiliary condition is satisfied at any time if it is satisfied at the initial time:

$$\begin{aligned} &\partial_\mu A_\mu(x, 0)\,\Phi_0 = 0, \\ &\partial_t(\partial_\mu A_\mu)\,\Phi_0 = [\mathrm{div}\,E(x, 0) - j_0(x, 0)]\,\Phi_0 = 0. \end{aligned} \tag{3.23}$$

As an example of the application of (3.23), we point out that one cannot take as Φ_0 a vector which is an eigenfunction of the operator $(p_L^2/2m) + W(q)$. The point is that the momentum operator p_L of the Lorentz gauge does not commute with the operator div $E - j_0$ because of $j_0 \sim \delta(\mathbf{x}-\mathbf{q})$. Therefore, the operator $(p_L^2/2m) + W$ does not commute with div $E - j_0$ either; its eigenfunctions cannot be eigenfunctions of div $E - j_0$ with zero eigenvalue, i.e., they cannot satisfy the auxiliary condition. On the basis of this and more general arguments, Dirac[39] (see §§77 and 80) points out that the observables in the Lorentz gauge can be described only by operators O satisfying the condition $[\mathcal{L}, O]\Phi = 0$ (in particular, by ones that commute with $\mathcal{L}$). Dirac calls such operators "physical." For example, the operators $\mathbf{p}_c$ and $\mathbf{A}_\perp$ of the Coulomb gauge, which can be expressed in terms of operators of the Lorentz gauge (see (3.17)), are "physical," and also

$$\mathbf{p}_c = \mathbf{p}_L - e\mathbf{A}_L, \quad \mathbf{A}_L = \nabla u. \tag{3.24}$$

Proceeding from the Hamiltonian written down at the beginning of this subsection, we can obtain an equation for $q_J(t)$ and the result (3.15). Instead of (3.7), we shall have $m\partial q/\partial t = p_L - eA$. Since A behaves "causally," so does p_L. However, as we have just said, p_L cannot describe an observable. If the electron momentum is associated with the operator p_c, then using (3.24) we of course obtain the same results as in the Coulomb gauge; see the preceding subsection, in which p_c was denoted simply by p.

4. GENERAL PROOF OF FINITENESS OF THE SIGNAL VELOCITY IN QUANTUM ELECTRODYNAMICS

The general formulation of the signal transmission problem will first be described schematically. In Sec. 4f we show that it is indeed general and suitable for describing a large class of real signal transmission experiments.

a) An external current J_μ localized in the region S and switched on at the time t_0 will serve as the source of the signal (more precisely, as the primary source; see later in Sec. 4f, in which there is also a discussion of the adjective "external"). The arrival of a signal is determined by an instrument localized in the region D, which is at distance R from S (see Fig. 1). It is assumed that it measures a certain physical quantity, for example, the electric field E (by means of a classical "test" charge). In the theoretical description, one can partly reflect the construction of the instrument. For example, one can introduce in the Hamiltonian the variables of a test charge (or electron, as is done in Sec. 3). But then one again requires an assumption about the existence of an instrument that measures the coordinate and velocity of the charge. The impossibility of dispensing with the assumption that there exists a certain finite part of the instrument not described in the framework of quantum mechanics is a general result of the theory of measurements.

In the regions S and D and also in the remaining part of space there may be photons and electrons (free or bound by constant external potentials). The state of this complete device for transmitting a signal at the time t_0 is described by the vector Φ_0. It is natural to assume that Φ_0 is stationary for $t<t_0$ (but this is not necessary).

We have already frequently emphasized the need to take into account the quantum mechanical features in the description of the signal transmission experiment. One cannot, for example, assume that up to the time t_0 the electric field E within D is zero. A state with a definite, for example, vanishing value of E is not stationary since the operator E does not commute even with the Hamiltonian of the free electromagnetic field. For the same reason E in stationary states (described by eigenvector of the total Hamiltonian $\mathscr{H}$ of the system for $t<t_0$) cannot have definite values, but is characterized by a certain distribution. We shall assume that the signal has arrived at the time t when the distribution with respect to the considered local observable has changed within D from the "background" distribution one would have had at time t if the external current had not been switched on.

The "distribution with respect to E at time t in the situation in which the current was switched on minus the distribution with respect to E at time t with no current switched on" can be found if one calculates the moments of this difference, i.e., the expectation value, the mean square, etc. The problem reduces to calculating expressions of the form

$$\langle U_J(t, t_0)\Phi_0, O(x)\, U_J(t, t_0)\Phi_0\rangle - \langle \exp[-i(t-t_0)\mathscr{H}(0)]\Phi, O(x)\exp[-i(t-t_0)\mathscr{H}(0)]\Phi_0\rangle, \quad (4.1)$$

where $O(\mathbf{x})$ may denote $E^m(\mathbf{x})$ $(m=1,2,3,4,\ldots)$, $\mathbf{x}\in V_D$,

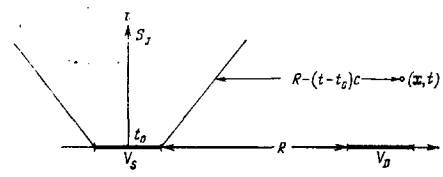

FIG. 2.

or the magnetic field, etc; $U_J(t, t_0)$ is the evolution operator of the system: $i\partial_t U_J = \mathscr{H}(t)U_J$. By $\exp[-i(t-t_0)\times\mathscr{H}(0)]$ we denote the evolution operator of the same system but when the external current is not switched on and the total Hamiltonian does not depend explicitly on the time (is equal to $\mathscr{H}(t_0)$ at all times). Since

$$\langle U_J(t, t_0)\Phi_0, OU_J(t, t_0)\Phi_0\rangle = \langle \Phi_0, U_J^+ OU_J\Phi_0\rangle, \quad (4.2)$$

and $U_J^+ OU_J$ is the Heisenberg operator $O_J(t)$ (which at $t=t_0$ coincides with the Schrödinger operator O), we can rewrite (4.1) in the form

$$\langle \Phi_0, [O_J(x, t) - O(x, t)]\Phi_0\rangle,$$
$$O(x, t) = \exp[i(t-t_0)\mathscr{H}(0)]\, O\exp[-i(t-t_0)\mathscr{H}(0)]. \quad (4.3)$$

b) The problem has been reduced to the calculation of Heisenberg operators and expectation values thereof. All Heisenberg operators of quantum electrodynamics can be expressed in terms of $A_\mu(\mathbf{x}, t)$ and $\psi(\mathbf{x}, t)$. In Appendix B it is shown that

$$\psi_J(\mathbf{x}, t) = \psi(\mathbf{x}, t), \quad A_{J\mu}(\mathbf{x}, t) = A_\mu(\mathbf{x}, t), \quad (\mathbf{x}, t)\notin S_J, \quad (4.4)$$

if the point $(\mathbf{x}, t)$ lies outside the future light cone S_J constructed on the region of localization of J_μ (Fig. 2), i.e., $(\mathbf{x}, t)$ must satisfy the condition $(\mathbf{x}-\mathbf{x}_S)^2-(t-t_0)^2>0$ for all $\mathbf{x}_S\in V_S$. Since $\mathbf{E}_J-\mathbf{E}=-\partial_t(\mathbf{A}_J-\mathbf{A})-\nabla(A_{J0}-A_0)$, then $\mathbf{E}_J(\mathbf{x}_D, t)=\mathbf{E}(\mathbf{x}_D, t)$ also for $t-t_0<R/c$. Returning to (4.3), we see that the distribution of the electric field within V_D is changed by the external current only at times R/c after the switching on of the current. The magnetic field H, the density of the Poynting vector, and the current density j_μ have the same causal behavior.

c) The quantities just listed do not exhaust all possible observables. The theory really does satisfy the principle of relativistic causality if causal behavior is shown for a sufficiently large class of observables (sufficiently large for one to be confident that all observables have this behavior).

An example of an observable that cannot be constructed from E, H, and j_μ is the number of electrons in a finite volume V. Since there exist instruments capable of measuring such a quantity (this can be done, for example, by means of a Wilson chamber), the theory must also include an operator associated with this observable.

The number of electrons in a volume V is characterized by the mean square $|u(x)|^2$ of the modulus of the wave function of the electrons in this volume. The state of an electron is also characterized by the phase of $u(x)$. In Sec. 3, we drew conclusions about the modulus and phase of u from the distributions with respect to q and p. In the second-quantized theory of relativistic electrons we must also in some manner describe the

observable characteristics of the state of the electrons.

We introduce an operator of the electron number in the macroscopic volume V. An operator of this kind was defined by Wightman and Schweber[40] by means of the operator of the particle number density (see Eqs. (53) and (54) in Ref. 40). They considered different definitions of the particle coordinate. Here, we shall use the coordinate x that arises naturally in field theory. One can show that the choice of the coordinate is unimportant if we need the integral of the density over a macroscopically large volume.[7)]

For the operator of the electron number density we take the expression $\psi^{(-)+}(x)\psi^{(-)}(x)$, where $\psi^{(-)}$ is the part of the electron–positron field ψ of the Lorentz gauge that annihilates electrons. In Appendix B it is shown that the behavior of $\int_{V_D}\psi^{(-)+}\psi^{(-)}d^3x$ does not contradict relativistic causality. However, an important feature of quantum electrodynamics is the fact that the operator $\int\psi^{(-)+}\psi^{(-)}$ cannot describe an observable number of electrons because (as one can show) it does not commute with the operator of the Lorentz auxiliary condition, i.e., it is not "physical"; see Sec. 3c.

A possible example of a "physical" electron number density is provided by the operator $\varphi^{(-)+}(x)\varphi^{(-)}(x)$ constructed by means of the "physical" operator of the electron–positron field φ of the Coulomb gauge:

$$\varphi(\mathbf{x}, t) = \psi(\mathbf{x}, t)\, e^{-ieu(\mathbf{x},t)} \tag{4.5}$$

(see §80 of the book of Ref. 39; here, e is the electron charge; the operator $u(\mathbf{x}, t)$ is written down in (3.17)).

Let us illustrate the physical meaning of the density operator $N^{(-)}(x)=\varphi^{(-)+}(x)\varphi^{(-)}(x)$. We calculate its expectation value in the single-electron state $\alpha_n^+\Omega$. The Schrödinger operator α_n is determined by the expansion of the Schrödinger operator $\varphi(\mathbf{x})$ with respect to a complete system of eigenfunctions of the Hamiltonian of the Dirac equation (with arbitrary constant external potential)

$$\varphi(\mathbf{x}) = S_n u_n(\mathbf{x})\,\alpha_n + S_p v_p(\mathbf{x})\,\beta_p^+ \equiv \varphi^{(-)}(\mathbf{x}) + \varphi^{(+)}(\mathbf{x}).$$

The symbol S_n denotes summation over the discrete values of n and integration over the continuous values. We have

$$\langle\alpha_n^+\Omega,\ \varphi^{(-)+}(\mathbf{x})\,\varphi^{(-)}(\mathbf{x})\,\alpha_n^+\Omega\rangle = \sum_{\mu=1}^{4} u_n^+(\mathbf{x}, \mu)\, u_n(\mathbf{x}, \mu). \tag{4.6}$$

We obtain the density of the probability distribution with respect to the coordinates in the considered state (in (4.6), we have written explicitly the spinor index μ). The expectation value of $N^{(-)}$ in the state $\alpha_n^+\alpha_m^+\beta_p^+\Omega$ is equal to the sum of the densities $|u_n|^2+|u_m|^2$ (the density of the positrons drops out of the expression). In the state in which there are only positrons and/or photons, the expectation value of $N^{(-)}$ is zero. In the general case, we obtain not the sum of the squares of the moduli of the wave functions of the electrons, but the diagonal elements $\langle x|\rho|x\rangle$ of the density matrix.

It can be shown that information about the phase of the electron wave function is given by the expectation value of the operator $\mathbf{P}(x)$ of the momentum density of the electron–positron field.[8)] We consider the "physical" operator $\mathbf{P}(x)$ constructed from $\varphi(x)$:

$$\mathbf{P}(x) = \frac{1}{2}:[\varphi^+(x)(-i\nabla)\varphi(x) + (-i\nabla\varphi)^+\varphi]: \tag{4.7}$$

We calculate the expectation value of (4.7) in the state $\alpha_n^+\Omega$:

$$\langle\alpha_n^+\Omega,\ \mathbf{P}(\mathbf{x})\,\alpha_n^+\Omega\rangle = \sum_{\mu}|u_n(\mathbf{x}, \mu)|^2\nabla\eta_n(\mathbf{x}, \mu), \quad u_n = |u_n|\,e^{i\eta_n} \tag{4.8}$$

The right-hand side of (4.8) differs from (4.6) by the presence of the gradient of the phase of $u_n(x)$.

On the basis of (4.6) and (4.8) we can assume that the operators $N^{(-)}$ and $\mathbf{P}(x)$ play a role analogous to the operators q and p of a single electron in Sec. 3.

d) In order to calculate the change in the number of electrons in V_D and the change in the phase of their wave function, we must in accordance with (4.1) and (4.3) calculate the expectation values of the differences $N_J^{(-)} - N^{(-)}$ and $\mathbf{P}_J(x) - \mathbf{P}(x)$ of the Heisenberg operators in the initial state Φ_0. For this, we must in turn calculate φ_J and φ. Using (4.4) and (4.5), we obtain

$$\varphi_J(x) = \psi_J e^{-ieu_J} = \psi e^{-ieu_J} = \varphi(x)\, e^{ie(u-u_J)}, \quad x \equiv (\mathbf{x}, t) \notin S_J. \tag{4.9}$$

It is shown in Ref. 32 that (4.9) implies that $N_J^{(-)}(x) - N^{(-)}(x)$ and $\mathbf{P}_J(x) - \mathbf{P}(x)$ are nonzero for $x \in V_D$ and $t - t_0 < R/c$, and they are in fact proportional to $(t-t_0)/R^3$ (see Ref. 31). However, this macroscopically noncausal behavior must be declared unobservable. The point is that (4.9) and (3.19) have the form of a quasigauge transformation;

$$\varphi_J(\ ) = \varphi(x)\ \exp ie\lambda(x),\ \mathbf{A}_{J\perp} = \mathbf{A}_\perp + \Delta\lambda,\ x \notin S_J\ x \equiv (\mathbf{x}, t),$$
$$\lambda(\mathbf{x}, t) = u(\mathbf{x}, t) - u_J(\mathbf{x}, t),\ \Delta\lambda(x, t) = 0,\ (x, t) \notin S_J, \tag{4.10}$$

as we discussed in Sec. 3b. We must therefore assume that $\int N_J^{(-)}(x)d^3x$ describes the same observable as $\int N^{(-)}(x)d^3x$ as long as $x \in S_J$. The same applies to $\mathbf{P}_J(x)$ and $\mathbf{P}(x)$.

Note that the difference $N_J^{(-)}(x) - N^{(-)}(x)$ for $x \in S_J$ contains also an exponentially small noncausal term of the same origin as the one in $\psi_J^{(-)+}\psi_J^{(-)} - \psi^{(-)+}\psi^{(-)}$ (see Sec. 3 in Ref. 31 and the end of Appendix B). However, the presence of such terms does not contradict relativistic causality.

e) In Ref. 32 it is shown that one can find an expression for the number density operator of Dirac electrons that is a "physical" operator and also has (macroscopically) causal behavior. Other expressions for the momentum density are also discussed in Ref. 32.

Since an electron has spin, for the final demonstration that quantum electrodynamics is relativistically causal

[7)] The point is that the transition to a different (say the Newton–Wigner) coordinate $\mathbf{q}$ is made by means of the function $\langle\mathbf{x}|\mathbf{q}\rangle$, which is exponentially small for $|\mathbf{x}-\mathbf{q}| \gg h/mc$.

[8)] It would seem to be difficult to give a physical interpretation of $\mathbf{P}(\mathbf{x})$ since one cannot speak of the momentum of an electron at the point $\mathbf{x}$ in quantum mechanics. But we have already shown in Sec. 3 that the momentum distribution gives information about the behavior of the phase of the wave function $\mathbf{u}(\mathbf{x})$ in the coordinate representation.

we must also show that the spin state is not changed "prematurely" in V_D. Leaving out the proof, we merely mention that this can be done by means of the expression for the density of the spin part of the angular momentum tensor (which is given, for example, in §20.5 of the book of Ref. 19).

f) The proof presented here that the signal velocity in quantum electrodynamics does not exceed the velocity of light is fairly general and applies to perfectly real signal transmission experiments. To demonstrate this assertion, we use two basic facts.

1) Mathematically, the proof reduces essentially to the derivation of the operator equations (4.4). In Appendix B this proof is given for the case which involves any arbitrary constant external potentials W_μ. It can also be carried through for various species of charged particles in addition to electrons and positrons (protons, etc).

2) We have not needed to know the actual form of the initial-state vector Φ_0.

We now show that the proof is also applicable in the case of a real signal source consisting, for example, of atoms capable of being excited and subsequently decaying. They can be excited, for example, by means of an electron beam. These atoms and electrons can be described theoretically by means of Φ_0; W_μ can describe the potentials which bind the electrons of the atoms or form the exciting beams.

We emphasize that the concept of the current J_μ which can be switched on is needed to describe a signal source of this kind too. In a signal transmission experiment we must have the possibility of beginning to excite the source atoms (by directing an exciting beam onto them) at a time t_0 which we choose arbitrarily. The current J_μ is used to describe the "first cause" of the signal.[9)] For our purposes, we do not need to know the actual dependence of J_μ on the time after t_0. Therefore, J_μ can describe much better the real current than what is called an "external current." This last, by definition, depends on the time in a prescribed manner, whereas of course the fact that the current radiates affects the time dependence of a real current. Instead of J_μ, one could also use the analogous concept of a potential which is switched on.

In the same way, one can justify the applicability of the proof for signal sources such as radar, an instrument with a Kerr shutter, etc.

The general treatment of this section (in contrast to Sec. 3) shows that before the time R/c in the region V_D there can be no change in directly observable quantities such as the current at the output of photomultipliers, counters, or other parts of a signal detector. Finally, the source S and the detector D may be separated, not by a vacuum, but by some material medium (fluid or crystal). The proof also holds for this case in so far as Φ_0 and W_μ can describe the electrons and protons of this medium.

[9)] One should perhaps regard the pressing of a switch button as the "first cause". But in virtually all cases this will have as a consequence the change of some current, and this can be reflected in the theoretical description.

APPENDICES

A. Calculation of the probability amplitude for excitation of an atom used as a signal detector

In the interaction picture, the amplitude $a_{SD^*}(t)$ of the Fermi problem has the form

$$a_{SD^*}(t)=\langle SD^*|\,U(t,0)\,|S^*D\rangle, \tag{A.1}$$

where U satisfies the equation $i\partial U/\partial t=\mathscr{H}_{\text{int}}U$, and the states $|S^*D\rangle$ and $|SD^*\rangle$ are, as usual, assumed to be two-electron eigenfunctions of the free part $\mathscr{H}_0$ of the total Hamiltonian $\mathscr{H}=\mathscr{H}_0+\mathscr{H}_{\text{int}}$. The attempt to calculate (A.1) by means of a theory expressed in the Lorentz gauge immediately encounters the difficulty described in Sec. 3c: The eigenfunctions of $\mathscr{H}_0$ do not satisfy the Lorentz auxiliary condition, and cannot therefore describe physical states. In the Coulomb gauge, this difficulty does not occur, and we therefore choose that gauge (see also the end of §13.2 in Ref. 17). We adopt a system of units in which $\hbar=1$ and $c=1$. We have

$$\mathscr{H}_{\text{int}}=\mathscr{H}_\perp+\mathscr{H}_c=-\int(\mathbf{j}\mathbf{A}_\perp)\,d^3x+\int\frac{j_0(x)\,j_0(x')d^3x\,d^3x'}{|\mathbf{x}-\mathbf{x}'|}. \tag{A.2}$$

Like Fermi,[12] we discuss in what follows only the part of the transition amplitude (A.1) due to $\mathscr{H}_\perp$, i.e., due to the exchange of a transverse photon. The contribution from $\mathscr{H}_c$ is under usual conditions smaller than that from $\mathscr{H}_\perp$ (see (A.13) below). Since there are no photons in the initial or the final state, we can separate out from $U(t,0)$ an operator U' that couples precisely such states in the e^2 approximation (as is done in §37 of the book of Ref. 19):

$$U'_\perp(t,0)=-\frac{1}{2}\int\int d^4x\,d^4y\sum_{\mu\nu}j_\mu(y)\,D^c_{\mu\nu}(y-x)\,j_\nu(x). \tag{A.3}$$

The integration with respect to the times x_0 and y_0 is from 0 to t. The propagator for the transverse photons has the form

$$D^c_{mn}=\left(\delta_{mn}+\frac{\partial_m\partial_n}{\Delta}\right)D^c_0\qquad(m,n=1,2,3),$$

$$D^c_{0n}=D^c_{m0}=D^c_{00}=0,\qquad\Delta=\sum_{i=1}^{3}\frac{\partial^2}{\partial x_i^2}, \tag{A.4}$$

$$D^c_0(\mathbf{x},t)=\frac{1}{4\pi i}\left[\delta(t^2-\mathbf{x}^2)-\frac{i}{\pi}P\frac{1}{t^2-x^2}\right]. \tag{A.5}$$

(See, for example, Eq. (14.54) in the book of Ref. 41.) The operators φ and $\bar\varphi$ of the electron–positron field in the currents j_μ and j_ν can be expanded in terms of the eigenfunctions of the electron part of $\mathscr{H}_0$; in particular, in terms of the functions $w(x)$ of bound states of electrons in the atoms. For example, the operator $\varphi(x)$ annihilates an electron in the state $w_{S*}(x)=\exp(-iE_{S*}x_0)\times u_{S*}(\mathbf{x})$ and creates it in the state w_S. Therefore, the part of the amplitude (A.1) corresponding to (A.3) takes the form

$$a_\perp(t)=\left\langle SD^*\left|\frac{e^2}{2}\int\int d^4x\,d^4y\,N[\bar\varphi(y)\,\gamma_m\varphi(y)]\,D^c_{mn}(y-x)\,N[\bar\varphi(x)\,\gamma_n\varphi(x)]\right|S^*D\right\rangle$$
$$=e^2\int\int d^4x\,d^4y\,w^+_{D^*}(y)\,\alpha_m w_D(y)\,D^c_{mn}(y-x)\,w^*_S(x)\,\alpha_n w_{S^*}(x)-\text{the exchange term}. \tag{A.6}$$

The exchange term is very small if R is much greater than the atomic dimension l (the overlapping of the electron wave functions in this case is negligible). The amplitude (A.6) can be calculated in different ways. To reduce it to the form in which Fermi made his simplifying assumptions, we use the representation for D^c given in §17.3 of the book of Ref. 19:

$$D^c_{mn}(y-x)=\theta(y_0-x_0)\,D^{(+)}_{mn}(y-x)+\theta(x_0-y_0)\,D^{(+)}_{mn}(x-y),$$
$$D^{(+)}_{mn}(x-y)=\frac{i}{2(2\pi)^3}\int\frac{d^3k}{k}\exp\{i[\mathbf{k}(\mathbf{x}-\mathbf{y})-k(x_0-y_0)]\}\left(\delta_{mn}-\frac{k_mk_n}{k^2}\right). \quad (A.7)$$

Assuming for simplicity that $E_{S*}-E_S=E_{D*}-E_D\equiv\Delta$, we obtain from (A.6) and (A.7)

$$a_\perp(t)=\frac{-e^2}{2(2\pi)^3}\sum_{mn}\int d^3y\,u^*_{D*}(\mathbf{y})\,\alpha_m u_D(\mathbf{y})$$
$$\times\int d^3x u^*_S(\mathbf{x})\,\alpha_n u_{S*}(\mathbf{x})\int\frac{d^3k}{k}\left\{\int_0^t dy_0\int_0^{y_0}dx_0\exp(i[\mathbf{k}(\mathbf{y}-\mathbf{x})-k(y_0-x_0)])\right.$$
$$\left.+\int_0^t dx_0\int_0^{x_0}dy_0\exp(i[\mathbf{k}(\mathbf{x}-\mathbf{y})-k(x_0-y_0)])\right\}[\delta_{mn}-k_mk_n/k^2]\exp[i\Delta(y_0-x_0)]. \quad (A.8)$$

Further, we can first integrate with respect to x_0 and y_0 (the reversal of the order of integration is justified in the same way as in the proof of Parseval's equation in the theory of integral Fourier transforms):

$$a_\perp(t)=\frac{-e^2}{2(2\pi)^3}\int\frac{d^3k}{k}\sum_{mn}(\delta_{mn}-k_mk_n/k^2)$$
$$\times\left\{\left[\frac{t}{i(k-\Delta)}-\frac{e^{-i(k-\Delta)t}-1}{(k-\Delta)^2}\right]\int d^3y\,u^*_{D*}(\mathbf{y})\,\alpha_m e^{i\mathbf{k}\mathbf{y}}u_D(\mathbf{y})\right.$$
$$\times\int d^3x u^*_S(\mathbf{x})e^{-i\mathbf{k}\mathbf{x}}\alpha_n u_{S*}(\mathbf{x})$$
$$\left.+\left[\frac{t}{i(k+\Delta)}-\frac{e^{-i(k+\Delta)t}}{(k+\Delta)^2}\right]\int d^3y\,u^*_{D*}(\mathbf{y})\,\alpha_m e^{-i\mathbf{k}\mathbf{y}}u_D(\mathbf{y})\right.$$
$$\left.\times\int d^3x\,u^*_S(\mathbf{x})\,e^{i\mathbf{k}\mathbf{x}}\alpha_n u_{S*}(\mathbf{x})\right\}. \quad (A.9)$$

The integral with respect to k converges rapidly for large k and therefore the main contribution is made by k values not greater in order of magnitude than Δ ($k<2\Delta$, for example). If $l\Delta\ll1$, we can replace $e^{\pm i\mathbf{k}\mathbf{x}}$ by 1 (the S electron is localized near the origin of coordinates and $e^{(\pm i\mathbf{k}\mathbf{y})}$ by $\exp(\pm i\mathbf{k}\mathbf{R})$ (the center of the potential which binds the D electron is at a point $\mathbf{R}$ on the z axis). This is called the dipole approximation. Making the inverse transformations from (A.9), in which the indicated substitution has been made, to (A.6), we obtain expression (A.6), in which $D^c(\mathbf{y}-\mathbf{x},y_0-x_0)$ is replaced by $D^c(\mathbf{R},y_0-x_0)$. Thus, the dipole approximation can be interpreted as follows: it is assumed that the photon is emitted not at the point at which the electron is situated but at the center of the potential which binds the electron. After the dipole approximation, we introduce the notation

$$j^m_D=\int d^3y\,u^*_{D*}(\mathbf{y})\,\alpha_m u_D(\mathbf{y}),\quad j^n_S=\int d^3x\,u^*_S(\mathbf{x})\,\alpha_n u_{S*}(\mathbf{x}). \quad (A.10)$$

These quantities can be expressed in terms of the matrix elements of the dipole moments: $j_S=id_S(E_{S*}-E_S)$ (see §7 in Ch. 1 of the book of Ref. 37). Further, we integrate with respect to the angles of $\mathbf{k}$ ($d^3k=k^2dkd\Omega$), obtaining

$$a_\perp(t)=\frac{i\alpha}{(2\pi)^2}\int_0^\infty k\,dk\,4\pi\left\{(\mathbf{j}_D\mathbf{j}_S)\left[\frac{\sin kR}{kR}+\frac{\cos kR}{(kR)^2}-\frac{\sin kR}{(kR)^3}\right]\right.$$
$$\left.-\frac{1}{R^2}(\mathbf{j}_S\mathbf{R})(\mathbf{j}_D\mathbf{R})\left[\frac{\sin kR}{kR}+\frac{3\cos kR}{(kR)^2}-\frac{3\sin kR}{(kR)^3}\right]\right\}^{-1}$$
$$\times\left\{\left[\frac{t}{k-\Delta}+\frac{\exp[-i(k-\Delta)t]-1}{i(k-\Delta)^2}\right]+\left[\frac{t}{k+\Delta}+\frac{\exp[-i(k+\Delta)t]-1}{i(k+\Delta)}\right]\right\}. \quad (A.11)$$

To calculate the integral with respect to k, Fermi made a much stronger assumption than the one which justifies the dipole approximation: he assumed that because of the presence of $k-\Delta$ in the denominator the principal contribution is made by k values approximately equal to Δ. Then the terms with $k+\Delta$ in the denominator must be ignored altogether and one can integrate with respect to k from $-\infty$ to $+\infty$. In addition, because of the condition $(kR)\approx(\Delta R)\gg1$ one can ignore terms of order $1/(kR)^2$ and $1/(kR)^3$ in the first figure brackets in (A.11). After this, the integral can be calculated by means of residues

$$a^F_\perp(t)=\frac{i\alpha}{\pi lR}\left[(\mathbf{j}_D\mathbf{j}_S)-\frac{(\mathbf{j}_D\mathbf{R})(\mathbf{j}_S\mathbf{R})}{R^2}\right]\int_{-\infty}^{\infty}dk\left[t-\frac{e^{-i(k-\Delta)t}-1}{i(k-\Delta)}\right]\frac{\sin kR}{k-\Delta}$$
$$=\frac{i\alpha}{\pi lR\Delta}\left[(\mathbf{j}_D\mathbf{j}_S)-\frac{(\mathbf{j}_D\mathbf{R})(\mathbf{j}_S\mathbf{R})}{R^2}\right]\pi e^{-iR\Delta}[(t-R)\Delta]\,\theta(t-R)]. \quad (A.12)$$

The function $\theta(t-R)$ is equal to zero for $t<R$, and therefore $a^F_\perp(t)=0$ for $t<R$.

However, the integral (A.11) can also be calculated without Fermi's assumption: it reduces to integral sines and cosines. For $R\Delta\gg1$ and $(R-t)\Delta\gg1$ the term in the figure brackets in (A.12) is then augmented by terms of order $1/R\Delta$ and $[1/(R-t)\Delta+1/(R+t)\Delta]$ (see Ref. 14), so that the amplitude $a_{SD*}(t)$ for $t<R$ is no longer zero. The ratio of $a_\perp(t_-)$ at time $R/c-\tau=t_-$ (less than R/c by some amount $\tau\ll R/c$) to $a_\perp(t_+)$, where $t_+=R/c+\tau$, is equal in order of magnitude $1/(\tau\Delta)^2$ if $1\ll\tau\Delta\ll R\Delta$. One can show that the cause of this result resides in the nonvanishing of D^c outside the light cone; see (A.5).

We give also the part of the amplitude $a_{SD*}(t)$ due to the Coulomb interaction $\mathscr{H}_c$:

$$a_c(t)\approx ie^2\frac{t}{R}\frac{1}{(R\Delta)^2}[(\mathbf{j}_S\mathbf{j}_D)-3(\mathbf{j}_S\mathbf{R})(\mathbf{j}_D\mathbf{R})/R^2]. \quad (A.13)$$

For $R\Delta\gg1$ it is much smaller then the part (A.11) calculated without the Fermi approximation.

B. Exact calculation of the differences $A_J - A$ and $\psi_J - \psi$

We adopt the notation of the book of Ref. 19. We write down the equations for A_μ and ψ in the presence of a constant external potential $W_\mu(\mathbf{x})$ and a nonstationary external current $J_\mu(\mathbf{x},t)$:

$$(\gamma_\mu\partial_\mu+m)\,\psi_J(x)=ie[A_{J\mu}(x)+W_\mu(x)]\,\gamma_\mu\psi_J(x), \quad (B.1)$$
$$\Box A_{J\mu}(x)=-[j_{J\mu}(x)+J_\mu(x)]; \quad (B.2)$$

where x denotes $(\mathbf{x},x_0)$. In the equations for the operators A_μ and ψ without subscript J (the case when there is no external current) it is necessary to set $J_\mu=0$. By j_μ in (B.2) we understand the antisymmetrized expression $ie\bar{\psi}\gamma_\mu\psi$.

The solution of these equations amounts to establishing how the Heisenberg operators are expressed in terms of the initial, i.e., Schrödinger, operators. The operators ψ_J, A_J and ψ, A must in accordance with their definition (see Sec. 2) coincide at $t=t_0$ with the same Schrödinger operator:

$$\psi_J(\mathbf{x},t_0)=U^+(t_0,t_0)\,\psi(\mathbf{x})\,U(t_0,t_0)=\psi(\mathbf{x})=\psi(\mathbf{x},t_0),$$
$$A_{J\mu}(\mathbf{x},t_0)=A_\mu(\mathbf{x})=A_\mu(\mathbf{x},t_0). \quad (B.3)$$

To solve the problem, we expand ψ_J, A_J and ψ, A in series in the coupling constant e (Källén; see §23 of the book of Ref. 19):

$$\psi_J(\mathbf{x}) = \sum_n e^n \psi_J^{(n)}(x), \qquad A_{J\mu} = \sum_n e^n A_{J\mu}^{(n)}, \tag{B.4}$$

and similarly for ψ and A_μ. We shall satisfy the initial conditions (B.3) if

$$\psi_J^{(0)}(\mathbf{x}, t_0) = \psi^{(0)}(\mathbf{x}, t_0) = \psi(\mathbf{x}), \qquad A_{J\mu}^{(0)}(\mathbf{x}, t_0) = A_\mu^{(0)}(\mathbf{x}, t_0) = A_\mu(\mathbf{x}), \tag{B.5}$$

and the operators $\psi_J^{(n)}(\mathbf{x}, t_0)$ and $A_{J\mu}^{(n)}(\mathbf{x}, t_0)$ with $n \geq 1$ are zero. We do not assume that the external current is small. The solutions of Eqs. (B.1) and (B.2) in the zeroth approximation:

$$(\gamma_\mu \partial_\mu + m)\psi_J^{(0)} = 0, \qquad \Box A_{J\mu}^{(0)} = -J_\mu \tag{B.6}$$

are well known. The first equation is the free equation, and therefore for all times $\psi_J^{(0)}(x) = \psi^{(0)}(x)$, and both coincide with the free operator $\psi_0(x)$:

$$\psi_0(x) = -i\int d^3x' S(x - x')\gamma_0 \psi(x'), \qquad x_0' = t_0 \tag{B.7}$$

(see (8.67) in the book of Ref. 13). The solution of the second equation has the form

$$A_{J\mu}^{(0)}(x) = A_{0\mu}(x) + \int d^4x' D_{\mathrm{ret}}(x - x') J_\mu(x'), \tag{B.8}$$

$$A_{0\mu}(x) = \int d^3y \left[\frac{\partial}{\partial y_0} D(x-y) A_\mu(y) - D(x-y)\frac{\partial}{\partial y_0} A_\mu(y)\right], \qquad y_0 = t_0. \tag{B.9}$$

Indeed, (B.8) satisfies the equation $\Box A_{J\mu} = -J_\mu$, it satisfies the ordinary equal-time commutation relations (because $A_{0\mu}$ satisfies them), and, finally, $A_{J\mu}^{(0)}(x)$ for $x_0 = t_0$ is transformed into the Schrödinger operator $A_\mu(\mathbf{x}, t_J)$ (the second term in (B.9) for $x_0 = t_0$ is equal to zero). We obtain

$$A_{J\mu}^{(0)}(x) - A_\mu^{(0)}(x) = \int D_{\mathrm{ret}}(x - x') J_\mu(x') d^4x' = \int d^3x' \frac{J_\mu(x', x_0 - |\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|}, \tag{B.10}$$

which is equal to zero if $\mathbf{x} \in V_D$ and $x_0 - t_0 < R/c$ (we recall that the current $J_\mu(x')$ is concentrated in V_S and equal to zero for $x_0' < t_0$).

Further, we write down the equations (B.1) and (B.2) for operators of the first approximation, form their differences, and write the result in the form

$$(\gamma_\mu \partial_\mu + m)(\psi_J^{(1)} - \psi^{(1)}) = i[(A_{J\mu}^{(0)} + W_\mu)\gamma_\mu(\psi_J^{(0)} - \psi^{(0)}) + (A_{J\mu}^{(0)} - A_\mu^{(0)})\gamma_\mu \psi^{(0)}], \tag{B.11}$$

$$\Box(A_{J\mu}^{(1)} - A_\mu^{(1)}) = -i[\bar{\psi}_J^{(0)}\gamma_\mu(\psi_J^{(0)} - \psi^{(0)}) + (\bar{\psi}_J^{(0)} - \bar{\psi}^{(0)})\gamma_\mu \psi^{(0)})]; \tag{B.12}$$

here $\psi_J^{(0)} - \psi^{(0)} \equiv 0$, and $A_{J\mu}^{(0)} - A_\mu^{(0)}$ is equal to zero for $x_0 - t_0 < R$. The right-hand sides of (B.11) and (B.12) are therefore zero for $x_0 - t_0 < R$. Since the initial conditions for $\psi_J^{(1)} - \psi^{(1)}$ and $A_J^{(1)} - A^{(1)}$ are null, the solution for these differences is also null for $x_0 - t_0 < R$, $\mathbf{x} \in V_D$. With regard to the expressions for the solutions of Eqs. (B.11) and (B.12) in terms of given right-hand side and in terms of the initial conditions see, respectively, §23 of the book of Ref. 19 and Eqs. (B.7) and (B.9).

The right-hand sides of equations of the form (B.11) and (B.12) for the next approximation can be expressed in terms of the operators of the zeroth and first approximations. For $x_0 - t_0 < R$, these right-hand sides vanish since for $x \in S_J$ the differences $\psi_J - \psi$ and $A_J - A$ on the right-hand sides are zero in the zeroth and the first approximation. Therefore, $\psi_J^{(2)} - \psi^{(2)}$ and $A_J^{(2)} - A^{(2)}$ are zero for $x \in S_J$. By induction we find that for $x \in S_J$ (see Fig. 2) in all orders in e the differences of the operators $\psi_J(x) - \psi(x)$ and $A_{J\mu}(x) - A_\mu(x)$ are zero (and in contrast to the operators occurring in them do not contain infinities).

Let us consider how $\int_{V_D} \psi^{(-)+}(x)\psi^{(-)}(x)d^3x$ (see Sec. 4c) behaves when the current J_μ is switched on in V_S. We write $\psi^{(-)}$ in the form $\psi^{(-)} = \Pi^{(-)}\psi$, where $\Pi^{(-)}$ separates out from ψ the part that annihilates electrons:

$$\int_{V_D} \psi_J^{(-)+}\psi_J^{(-)} d^3x - \int_{V_D} \psi^{(-)+}\psi^{(-)} d^3x = \int d^3y' \int d^3y'' \Pi^{(-)}(\mathbf{y}', \mathbf{x})\Pi^{(-)}(\mathbf{x}, \mathbf{y}'') \times [\psi_J^+(y')\psi_J(y'') - \psi^+(y')\psi(y'')], \qquad y_0' = y_0'' = x_0. \tag{B.13}$$

The expression in square brackets does not vanish for $x_0 < R/c$ because the integrand contains points y', y'' situated in the future cone S_J of the current J; see Fig. 2 (for such points, $\psi_J \neq \psi$). Their distance to points $\mathbf{x} \in V_D$ is equal to or greater than $|R - (t - t_0)c|$; see Fig. 2. The projection operator $\Pi^{(-)}(\mathbf{x}, \mathbf{y}) = -iS^{(-)}(\mathbf{x}, x_0; \mathbf{y}, x_0)$ for $|\mathbf{x} - \mathbf{y}| \gg \lambda_e$ decreases faster than $\exp(-|\mathbf{x} - \mathbf{y}|/\lambda_e)$ (here, λ_e is the electron Compton wavelength $\lambda_e = \hbar/mc$). Therefore, points y', y'' for which $\psi_J \neq \psi$ are represented in the integrand with weight $\sim\exp[-|R - (t - t_0)c|/\lambda_e]$, and the right-hand side of (B.13) is of order $\exp[-R(t - t_0)c|/\lambda_e]$. Such a noncausal effect does not mean that relativistic causality is violated since one cannot localize a Dirac particle in a region measuring $\lesssim \lambda_e$.

[1] C. Møller, The Theory of Relativity, Oxford University Press, London (1972) (Russian translation, Atomizdat, Moscow, 1975).
[2] N. N. Bogolyubov, A. A. Logunov, and I. T. Todorov, Osnovy aksiomaticheskogo podkhoda v kvantovoĭ teorii polya, Nauka, Moscow (1969) (Translation: Introduction to Axiomatic Quantum Field Theory, Benjamin, New York, 1975).
[3] L. I. Schiff, Quantum Mechanics, New York (1955) (Russian translation IL, Moscow, 1957).
[4] G. C. Hegerfeldt, Phys. Rev. D 10, 3320 (1974).
[5] T. D. Lee and G. C. Wick, Phys. Rev. D 2, 1033 (1970). T. L. Trueman, Nucl. Phys. B 26, 83 (1971).
[6] L. Brillouin, Wave Propagation and Group Velocity, Academic Press, New York (1960).
[7] V. L. Ginzburg, Teoreticheskaya fizika i astrofizika (Theoretical Physics and Astrophysics), Nauka, Moscow (1975), Ch. 8.
[8] D. A. Kirzhnits and V. N. Sazonov, in: Einshteĭnovskiĭ sbornik (Einstein Collection). 1973, Nauka, Moscow (1974), p. 84.
[9] I. V. Polubarinov, Soobshcheniya (Communications) JINR R2-7896, R2-8362 (in Russian), Dubna (1974).
[10] D. Gromes, Z. Phys. 236, 276 (1970).
[11] S. Kikuchi, Z. Phys. 66, 558 (1930).
[12] E. Fermi, Rev. Mod. Phys. 4, 87 (1932).
[13] S. S. Schweber, An Introduction to Relativistic Quantum Field Theory, Row, Peterson, and Co., Evanston, Ill. (1961) (Russian translation, IL, Moscow, 1963).
[14] M. I. Shirokov, Preprint JINR R-1719 (in Russian), Dubna (1964); Yad. Fiz. 4, 1077 (1966). [Sov. J. Nucl. Phys. 4,
[15] W. Heitler and S. T. Ma, Proc. R. Ir. Acad. 52, 123 (1949).
[16] J. Hamilton, Proc. Phys. Soc. A 62, 12 (1949).
[17] W. Heitler, The Quantum Theory of Radiation, Oxford (1964) (Russian translation of earlier edition, IL, Moscow, 1956).
[18] S. T. Ma, Nucl. Phys. 7, 163 (1958).
[19] A. I. Akhiezer and V. B. Berestetskiĭ, Kvantovaya Elektrodinamika, Fizmatgiz, Moscow (1959) (Translation:

Quantum Electrodynamics, New York, 1965).
[20]W. E. Thirring, Principles of Quantum Electrodynamics, New York (1958) (Russian translation: Vysshaya shkola, Moscow, 1964).
[21]M. Fierz, Helv. Phys. Acta 23, 731 (1950) (Russian translation in: Noveĭshee razvitie kvantovoĭ elektrodinamiki, IL, Moscow, 1954).
[22]B. A. Zon and M. I. Shirokov, Yad. Fiz. 7, 1316 (1968) [Sov. J. Nucl. Phys. 7, 780 (1968).
[23]B. Ferretti, in: Old and New Problems in Elementary Particles, Acad. Press, New York (1968), p. 108.
[24]I. A. Eganova and M. I. Shirokov, Preprint JINR R2-4645 (in Russian), Dubna (1969).
[25]I. A. Eganova, Preprint JINR R2-4936 (in Russian), Dubna (1970); Izv. Akad. Nauk Azerb. SSR, Ser. Fiz. Matem. 1, 94 (1971).
[26]M. I. Shirokov, Yad. Fiz. 21, 674 (1975) [Sov. J. Nucl. Phys. 21, 347 (1975)].
[27]M. I. Shirokov, Preprint JINR R2-5338 (in Russian), Dubna (1970).
[28]M. I. Shirokov, Soobshcheniya (Communications) JINR R2-8022 (in Russian), Dubna (1974).
[29]J. M. Knight, J. Math. Phys. 2, 459 (1961).
[30]A. L. Licht, J. Math. Phys. 4, 1443 (1963).
[31]M. I. Shirokov, Yad. Fiz. 23, 1131 (1976) [Sov. J. Nucl. Phys. 23, 599 (1976).
[32]M. I. Shirokov, Preprint JINR R2-9723 (in Russian), Dubna (1976).
[33]J. S. Bell, The Theory of Local Beables. Preprint TH 2053, CERN (1975).
[34]B. Gerlach, D. Gromes, J. Petzold, and P. Rosenthal, Z. Phys. 208, 381 (1968).
[35]B. Gerlach, D. Gromes, and J. Petzold, Z. Phys. 221, 141 (1969).
[36]J. Petzold, Ann. Phys. (Leipzig) 31, 361, 372, 381 (1974).
[37]A. J. W. Sommerfeld, Atomic Structure and Spectral Lines, London (1934) (Russian translation, Vol. 2, Gostekhizdat, Moscow, 1956).
[38]T. Venttsel', Vvedenie v kvantovuyu teoriyu volnovykh poleĭ (Introduction to the Quantum Theory of Wave Fields), Gostekhizdat, Moscow (1947).
[39]P. A. M. Dirac, The Principles of Quantum Mechanics, Oxford (1958) (Russian translation, Fizmatgiz, Moscow, 1960).
[40]A. S. Wightman and S. S. Schweber, Phys. Rev. 98, 812 (1955).
[41]J. D. Bjorken and S. D. Drell, Relativistic Fields, McGraw-Hill, New York (1965).

Translated by Julian B. Barbour