Gauge Problem in Quantum Field Theory

F. Strocchi*
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 16 March 1967)

The difficulties connected with the quantization of the electromagnetic field are analyzed in the framework of axiomatic field theory. Under the assumptions: (1) existence of the vacuum, invariant under the Poincaré group, (2) existence of a representation of the Poincaré group such that the fields have tensor transformation properties, and (3) analyticity of the two-point function in the forward tube, it is proved that the Maxwell equations $\partial^\mu F_{\mu\nu}=0$, $\epsilon^{\lambda\mu\nu\rho}\partial_\mu F_{\nu\rho}=0$ do not admit the classical solution $F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$, where the potential transforms as a four-vector. The result is very general and can be obtained without assuming either local commutativity or the spectral condition; it is also independent of whether the metric is positive or not. Thus the difficulties connected with the gauge problem have very little to do with the Lorentz condition, indefinite metric, etc., but rather they arise at the very beginning with the introduction of the four-vector potential. As a corollary of the above result, the representations of the Poincaré group for m ssless spin-one particles in quantum field theory are shown to be essentially different from the corresponding ones of the classical case. The Fock representation in the Gupta-Bleuler approach is discussed in connection with the Lehmann-Symanzik-Zimmermann formulation of quantum electrodynamics. Some care must be used in the definition of the asymptotic Hilbert space since, by the previous result, unphysical photon states must be introduced as asymptotic states in order to have a local theory. The requirement that these states should not affect any physical result imposes definite restrictions on the S matrix. These conditions guarantee that the definition of physical photon states in terms of equivalence classes still has significance in the presence of interactions.

I. INTRODUCTION

Quantum electrodynamics has probably been the greatest success of quantum field theory, the agreement between theoretical and experimental results being much better than one could expect. Surprisingly enough, quantum electrodynamics has never been a very clean theory. It has suffered many serious diseases (like unphysical photons and infinities) since its very beginning and it is rather puzzling that it has survived all of them. According to a widespread opinion, present quantum electrodynamics is only a rough approximation of the "true" theory, and it is rather hopeless to try to understand and cure its diseases. As a consequence of this attitude, the difficulties of quantum electrodynamics have been regarded as due to our approximate theory of nature, rather than being connected with our improper formulation of the interactions between photons and electrons. In fact, it is not very clear whether these troubles are characteristic features of Lagrangian quantum field theory, or whether they are connected with some specific points of the present formulation of quantum electrodynamics. We will try to analyze some of the difficulties of quantum electrodynamics and their connection with the basic assumptions of quantum field theory. We hope to clarify at least what the crucial points are and what must be expected even under very general assumptions.

The first and probably the most annoying difficulty in quantum electrodynamics arose at the very begining: the quantization of the electromagnetic field.[1] It was soon realized that in the quantization procedure there was no room for the Lorentz condition

$$\partial^\mu A_\mu=0, \tag{1}$$

for the vector potential. This difficulty has very strong consequences. It implies that the four components of the electromagnetic field must be treated as independent, in contrast with the classical theory, and therefore one must expect other particles besides photons, because of the additional degrees of freedom of the vector potential.

An alternative approach is to quantize only the transverse part of the vector potential and to leave the Coulomb field unquantized.[2] This implies, however, the splitting of the vector potential into a transverse part and a Coulomb part, and this splitting is not Lorentz-covariant. The manifest covariance of the theory and, together with it, the elegance of the tensor formalism are lost in this approach.

The difficulties connected with the impossibility of imposing the Lorentz condition have strong implications as far as the physical interpretation of the theory is concerned, as may be seen by using the theory of the representations of the Poincaré group.[3] The part of the field corresponding to $\partial^\mu A_\mu\neq0$ is in fact associated with scalar photons which have never been seen and must be regarded as unphysical. Because of the occurrence of such unphysical particles in the theory, the metric of the Hilbert space cannot be positive. One is then faced with

---

* Present address: Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Pisa, Italy.

[1] W. Heisenberg and W. Pauli, Z. Physik **56**, 1 (1929); **59**, 169 (1930).

[2] E. Fermi, Rend. Accad. Nazl. Lincei **2**, 881 (1929); Rev. Mod. Phys. **4**, 87 (1932); P. A. M. Dirac, V. A. Fock, and B. Podolsky, Z. Physik. Sowjetunion. **2**, 468 (1932); P. A. M. Dirac, Proc. Roy. Soc. (London) **A114**, 243, 710 (1927); W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, New York, 1954), 3rd ed.; L. E. Evans and T. Fulton, Nucl. Phys. **21**, 492 (1960).

[3] R. Shaw, Nuovo Cimento **37**, 1086 (1965).

the problem of giving a meaning to states of negative norm, to negative probabilities, etc.

A partial answer to these difficulties is to require

$$\partial^\mu A_\mu{}^+ \Psi = 0\,, \tag{2}$$

instead of Eq. (1); here $\partial^\mu A_\mu{}^+$ is the positive-frequency part of the field operator $\partial^\mu A_\mu$. This amounts to restricting the physical states to the subspace $\mathcal{H}' \subset \mathcal{H}$ defined by Eq. (2). In addition, an indefinite metric must be used because there are states in $\mathcal{H}'$ with zero norm. This approach is known as the Gupta-Bleuler formalism.[4] Apart from some unpleasant features such as indefinite metric, nonlocal condition,[2] etc., one may wonder what the full implications are, in perturbation theory, of admitting unphysical particles in the theory.

The widespread belief that these particles do not contribute to anything is not correct. As a matter of fact, those unphysical photons enter into the theory with a gradient-type coupling to the electromagnetic current

$$\mathcal{L}_{\text{int}}{}^{\text{scal}} \sim j_\mu \partial^\mu \varphi\,,$$

and it has been proved that the above interaction affects the renormalization constants by infinite amounts.[5] The gauge dependence of the renormalization constants suggests that the gauge problem might not be entirely disconnected from the ultraviolet problem.[6]

Another difficulty connected with the zero mass of the photon and the gauge is the infrared problem whose implications are not very clear.

The above problems are not confined to Lagrangian field theory. They arise also in the S-matrix approach to field theory. Apart from difficulties connected with the infrared problem, doubts have been raised[7] about a possible Lehmann-Symanzik–Zimmermann (LSZ) formulation of quantum electrodynamics on the basis that the photon field operator A_μ creates states containing unphysical scalar and longitudinal photons, and therefore the physical state vectors (those containing no unphysical photons) do not form a complete set of states. As a consequence, a set of integral equations for the interpolating field cannot be derived in a straightforward way.

The treatment of Evans and Fulton[8] succeeds in avoiding unphysical photons at every stage of the theory but at the price of using a definite and nonmanifestly covariant gauge (Coulomb gauge) and an unquantized Coulomb field. (In this case, of course, an LSZ formulation can be easily carried out as the above problem does not arise).

All these difficulties seem to indicate that the problem has been formulated in an improper way. At least, it is not clear what price must be paid and for what. Actually the definition of the photon field operator is a very delicate one because of the difficulties of combining Maxwell equations and commutation relations. The analysis of this definition may be carried out in many different ways and often the choice of one rather than another is merely a matter of philosophy.[9] As a matter of fact, in the present situation it is not clear what the arbitrary parameters of the problem are which may or may not be accepted, and what intrinsic features are strongly connected with the general principles of quantum field theory.

In the following we will analyze the problem of quantizing the electromagnetic field in the framework of Wightman theory.[10] We shall use the smallest number of hypotheses in order to see the roots of the difficulties connected with the gauge problem in quantum field theory.

Using only the assumptions of (1) the existence of the vacuum, and (2) the existence of a unitary representation of the Poincaré group $\{a,\Lambda\} \to U(a,\Lambda)$, such that

$$U(a,\Lambda)F_{\mu\nu}(x)U(a,\Lambda)^{-1} = \Lambda^{-1}{}_\mu{}^\rho \Lambda^{-1}{}_\nu{}^\sigma F_{\rho\sigma}(\Lambda x + a)\,,$$

we will show that the second set of Maxwell equations,

$$\epsilon^{\lambda\mu\nu\rho}\partial_\mu F_{\nu\rho} = 0\,,$$

does not admit the classical solution

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,,$$
$$U(a,\Lambda)A_\mu(x)U(a,\Lambda)^{-1} = \Lambda^{-1}{}_\mu{}^\nu A_\nu(\Lambda x + a)\,,$$

apart from the trivial case $F_{\mu\nu} = 0$.

This result is very general and can be obtained without using the positive or negative metric of the Hilbert space, the local commutativity of the fields, or the spectral condition.

Thus, the gauge problem has very deep roots and has very little to do with the Lorentz condition, indefinite metric, etc., in contrast to what is generally stated in the literature. The difficulties connected with the introduction of the vector potential are associated with its definition as a solution of the second set of Maxwell equations.

The conclusion which can be drawn from this result is: Either A_μ cannot transform as a four-vector, or the

[4] S. N. Gupta, Proc. Phys. Soc. (London) **63**, 681 (1950); K. T. Bleuler, Helv. Phys. Acta **23**, 567 (1950).

[5] S. Okubo, Nuovo Cimento **19**, 574 (1961); A. S. Wightman, lecture notes, Cargèse, 1964 (to be published); B. Klaiber, Nuovo Cimento **36**, 165 (1965).

[6] K. Johnson and B. Zumino, Phys. Rev. Letters **3**, 351 (1959).

[7] R. E. Pugh, Ann. Phys. **30**, 422 (1964). For earlier papers on asymptotic quantum electrodynamics see K. Nishijima, Phys. Rev. **119**, 485 (1960); H. Rollnik, B. Stech, and E. Nunnemann Z. Physik **159**, 482 (1960).

[8] L. E. Evans and T. Fulton, Nucl. Phys. **21**, 492 (1960).

[9] S. N. Gupta, Proc. Phys. Soc. (London) **63**, 681 (1950); S. Mandelstam, Ann. Phys. (N. Y.) **19**, **1** (1962); E. P. Wigner, *Theoretical Physics* (International Atomic Energy Agency, Vienna, 1963); R. Shaw, Nuovo Cimento **37**, 1086 (1965); F. Rohrlich and F. Strocchi, Phys. Rev. **139**, B476 (1965).

[10] A. S. Wightman, Phys. Rev. **101**, 860 (1956); A. S. Wightman, lectures given at the Faculté des Sciences, Université de Paris, 1958 (unpublished); and *Les Problèmes Mathématiques de la Théorie Quantique des Champs* (Colloques Internationaux du Centre Nationale de la Recherche Scientifique, Paris, 1959).

Maxwell equations cannot be satisfied as operator equations if A_μ is introduced as a local operator.

If the latter solution is adopted, one cannot have a local theory involving only physical photons, and states corresponding to longitudinal and scalar photons must be present in the Hilbert space. Within this framework, one can still use an LSZ formulation of the S matrix. However, suitable auxiliary conditions must be imposed in order to single out the physical states and, by the above result, they cannot be written as local equations. The scalar photons may be separated out by requiring the subsidiary condition

$$\partial^\mu A_\mu{}^+\Psi=0.$$

(See below for detailed discussion.) However, the longitudinal photons cannot be eliminated in a Lorentz-invariant way. Therefore the physical photon states cannot be sharply defined. The longitudinal photons will always be present and the only way to get rid of them is to define the subspace of the physical photon states as a quotient space $\mathcal{H}'/\mathcal{H}''$, where $\mathcal{H}''\subset\mathcal{H}'$ is the subspace of vectors with zero norm (i.e., containing at least one longitudinal photon). Within this framework,[11] an asymptotic Hilbert space may be constructed and a complete set of states may be introduced. As stressed before, in order to have a local theory (as required for LSZ), unphysical photon states must be introduced as asymptotic states, even if at the end they should not affect the physical results. This may be guaranteed by requiring that $\mathcal{H}'$ and $\mathcal{H}''$ be left invariant by the S matrix. As we shall see, these conditions are necessary if the definition of physical photons given above has a physical meaning. The classification of the representations of the Poincaré group, in terms of physical and unphysical photons, must in fact be independent of the interactions, i.e., must be left invariant by the S matrix, if we want any connection with reality.

II. QUANTIZATION OF THE ELECTROMAGNETIC FIELD IN AXIOMATIC FIELD THEORY

In this section we will analyze the problem of describing a massless spin-one particle in axiomatic field theory.[12]

According to the classification of the representations of the Poincaré group, a massless spin-one particle is described by a completely antisymmetric tensor of rank two $F_{\mu\nu}$ which obeys the following equations:

$$\partial^\mu F_{\mu\nu}=0, \tag{3}$$

$$\epsilon^{\lambda\mu\nu\rho}\partial_\mu F_{\nu\rho}=0. \tag{4}$$

This description does not suffer from the ambiguities connected with the gauge problem, and in fact it can be obtained by simply requiring a manifestly covariant description of a free spin-one massless particle without involving unphysical particles.

In order to formulate the problem of quantizing the above equations in the framework of axiomatic field theory some of the Wightman axioms[10] will be assumed. In particular we will assume:

(1) $F_{\mu\nu}(x)$ may be defined as an operator-valued distribution in a Hilbert space $\mathcal{H}$. It is not necessary to specify which type of distribution $F_{\mu\nu}$ is supposed to be. It is sufficient that the two-point Wightman function (see below)

$$W_{\mu\nu\rho\sigma}{}^F(x-y)=(\Psi^0,F_{\mu\nu}(x)F_{\rho\sigma}(y)\Psi^0)$$

can be regarded as the boundary value of an analytic function $W_{\mu\nu\rho\sigma}{}^F(z)$, analytic in the forward tube

$$-\infty<\mathrm{Re}z<+\infty,\quad \mathrm{Im}z\in V_+,$$

where

$$V_+=\{x\,|\,x \text{ is a four-vector, } x^2>0,\ x^0>0\}.$$

This condition is clearly satisfied if $F_{\mu\nu}(x)$ is assumed to be an operator-valued tempered distribution. However, the validity of the condition is more general and in fact it can be proved for a large class of operator-valued distributions.[13]

(2) There exists a unitary representation of the Poincaré group: $\{a,\Lambda\}\rightarrow U(a,\Lambda)$ such that

$$U(a,\Lambda)F_{\mu\nu}(x)U(a,\Lambda)^{-1}=\Lambda^{-1}{}_\mu{}^\rho\Lambda^{-1}{}_\nu{}^\sigma F_{\rho\sigma}(\Lambda x+a). \tag{5}$$

(3) There exists an invariant state Ψ^0 (vacuum state) such that

$$U(a,\Lambda)\Psi^0=\Psi^0. \tag{6}$$

Local commutativity and the spectral condition are not needed for what follows.[14] It must be stressed that no assumption is made about the positive or negative metric in the Hilbert space. It may very well be that the "product" of two vectors Ψ_1, Ψ_2 is defined as a sesquilinear form

$$(\Psi_1,\Psi_2)=\langle\eta\Psi_1,\Psi_2\rangle,$$

where $\langle\ ,\ \rangle$ is the scalar product in $\mathcal{H}$, and η is the metric operator. (See Sec. III for details.) In this case the vacuum expectation values would be defined as

$$(\Psi^0,F_{\mu\nu}F_{\rho\sigma}\Psi^0)=\langle\eta\Psi^0,F_{\mu\nu}F_{\rho\sigma}\Psi^0\rangle.$$

The conclusions of the following theorems are, however, independent of η being $=1$ or $\neq1$.

The problems connected with the gauge arise when Eq. (4) is solved by putting

$$F_{\mu\nu}(x)=\partial_\mu A_\nu(x)-\partial_\nu A_\mu(x), \tag{7}$$

[11] A. S. Wightman and L. Gårding, Arch. Fysik **28**, 129 (1964).

[12] The analogous problem for gravitons will be discussed in a subsequent paper.

[13] A. M. Jaffe, Phys. Rev. Letters **17**, 661 (1966); in Proceedings of the MIT Conference on Scattering, 1966 (unpublished).

[14] A detailed exposition of Wightman axioms may be found in R. Streater and A. S. Wightman, *PCT, Spin, and Statistics, and All That* (W. A. Benjamin, Inc., New York, 1964). See also Ref. (11).

where A_μ is a four-vector satisfying

$$U(a,\Lambda)A_\mu(x)U(a,\Lambda)^{-1}=\Lambda^{-1}{}_\mu{}^\nu A_\nu(\Lambda x+a)\,, \qquad (7')$$

and $A_\mu(x)$ may be considered as an operator-valued distribution. Again it is not necessary to specify which kind of distribution A_μ is assumed to be, provided that the two point function may be regarded as the boundary value of an analytic function, analytic in the forward tube.

Under these assumptions, it can be proved that the solution given by Eqs. (7), and (7′), leads to a trivial physical theory, and thus it cannot be accepted. In order to prove this we need some preliminary results.

We consider the two-point function

$$(\Psi^0,A_\mu(x)A_\nu(y)\Psi^0)=D_{\mu\nu}(x,y)\,. \qquad (8)$$

The invariance of Ψ^0 under space time translations gives

$$D_{\mu\nu}(x,y)=D_{\mu\nu}(x-y)\,,$$

and the analyticity property is

$$D_{\mu\nu}(\xi)=\text{boundary value of } D_{\mu\nu}(z)\,, \qquad (9)$$

where $D_{\mu\nu}(z)$ is analytic in the forward tube

$$-\infty<\text{Re}z<+\infty\,,\quad \text{Im}z\in V_+\,.$$

Finally, by using Eq. (7′) we get

$$D_{\mu\nu}(\xi)=(\Psi^0,U(0,\Lambda)A_\mu(x)A_\nu(y)U(0,\Lambda)^{-1}\Psi^0)$$
$$=\Lambda^{-1}{}_\mu{}^\rho\Lambda^{-1}{}_\nu{}^\sigma D_{\rho\sigma}(\Lambda\xi)\,.$$

This equation may be analytically continued to complex Λ in such a way that $D_{\mu\nu}(z)$ yields a representation of the proper homogeneous complex Lorentz group $L_+(C)$:

$$D_{\mu\nu}(z)=\Lambda^{-1}{}_\mu{}^\rho\Lambda^{-1}{}_\nu{}^\sigma D_{\rho\sigma}(\Lambda z)\,. \qquad (10)$$

Lemma 1. $D_{\mu\nu}(\xi)$ can be written in the following form:

$$D_{\mu\nu}(\xi)=g_{\mu\nu}D_1(\xi)+\partial_\mu\partial_\nu D_2(\xi)\,, \qquad (11)$$

where

$$D_1(\xi)=D_1(\Lambda\xi)\,,\quad D_2(\xi)=D_2(\Lambda\xi)\,. \qquad (12)$$

Proof: Property (10) enables us to use a theorem by Araki and Hepp[15] on the classification of the representations of $L_+(C)$ in terms of analytic functions. According to that theorem, $D_{\mu\nu}(z)$ may be written in the following form:

$$D_{\mu\nu}(z)=g_{\mu\nu}\tilde{D}_1(z)+z_\mu z_\nu\tilde{D}_2(z)\,. \qquad (13)$$

By the analyticity properties of $D_{\mu\nu}(z)$, the functions $\tilde{D}_1(z)$ and $\tilde{D}_2(z)$ are analytic in the forward tube; moreover they are invariant under the homogeneous Lorentz group

$$\tilde{D}_1(z)=\tilde{D}_1(\Lambda z)\,,\quad \tilde{D}_2(z)=\tilde{D}_2(\Lambda z)\,, \qquad (14)$$

as a trivial consequence of Eq. (10). Then by the Hall-Wightman theorem,[16] $\tilde{D}_1(z)$ and $\tilde{D}_2(z)$ may be written as functions of z^2:

$$\tilde{D}_1(z)=\tilde{D}_1(z^2)\,,\quad \tilde{D}_2(z)=\tilde{D}_2(z^2)\,. \qquad (15)$$

Now, for an arbitrary function of z^2,

$$\frac{\partial^2}{\partial z^\mu\partial z^\nu}G(z^2)=\tfrac{1}{2}g_{\mu\nu}\frac{d}{dz^2}G(z^2)+\tfrac{1}{4}z_\mu z_\nu\left(\frac{d}{dz^2}\right)^2G(z^2)\,.$$

We now define a function $G(z^2)$ as a solution of the following equation:

$$\left(\frac{d}{dz^2}\right)^2G(z^2)=4\tilde{D}_2(z^2)\,. \qquad (16)$$

Equation (16) may be easily solved by integration in the analyticity domain of $\tilde{D}_2(z^2)$:

$$G(z^2)=4\int_{c_1}^{z^2}ds^2\int_{c_0}^{s^2}\tilde{D}_2(z'^2)dz'^2\,.$$

Thus, we have

$$z_\mu z_\nu\tilde{D}_2(z^2)=\frac{\partial^2}{\partial z^\mu\partial z^\nu}G(z^2)-\tfrac{1}{2}g_{\mu\nu}\frac{d}{dz^2}G(z^2)\,,$$

and Eq. (13) takes the following form:

$$D_{\mu\nu}(z)=g_{\mu\nu}D_1(z)+\frac{\partial^2}{\partial z^\mu\partial z^\nu}D_2(z)\,, \qquad (17)$$

where D_1 and D_2 are suitable combinations of D_1 and $G(z^2)$.

Taking the boundary value of Eq. (17) gives Eq. (11) and proves Lemma 1.[17]

Lemma 2. The differential equation

$$(\Box g_{\mu\nu}+\alpha\partial_\mu\partial_\nu)F(x)=0\,,\quad (\alpha\neq 0,\ \alpha\neq 4)\,, \qquad (18)$$

where $F(x)$ is a generalized function, invariant under Lorentz transformations

$$F(x)=F(\Lambda x)\,, \qquad (19)$$

does not admit any solution apart from the trivial one

$$F(x)=\text{const}\,.$$

Proof: On multiplying Eq. (18) by $g^{\lambda\mu}$, we obtain

$$(\Box\delta^\lambda{}_\nu+\alpha\partial^\lambda\partial_\nu)F(x)=0$$

and by taking the trace we get

$$(4+\alpha)\Box F(x)=0\,,\quad \text{i.e., } \Box F(x)=0\,.$$

[15] K. Hepp, Helv. Phys. Acta **36**, 355 (1963).

[16] D. Hall and A. S. Wightman, Kgl. Danske Videnskab. Selskab. Mat.-Fys. Medd. **31**, No. 5 (1957).

[17] I am indebted to Dr. H. Epstein for useful suggestions and for drawing my attention to the paper by Hepp.

Substituting this result in Eq. (18), we have

$$\partial_\mu \partial_\nu F(x)=0, \quad \forall \mu,\nu, \tag{20}$$

i.e.,

$$\partial_\nu F(x)=a_\nu=\text{const}.$$

(The symbol $\forall$ means: for any.) Hence, $F(x)$ must have the following form:

$$F(x)=a^\nu x_\nu+\text{const}.$$

But this is not invariant under Lorentz transformations, i.e., it does not satisfy condition (19). Therefore the only acceptable solution is

$$F(x)=\text{const.},$$

which proves Lemma 2.

We may now prove the following theorem:

Theorem. The two-point function $D_{\mu\nu}(\xi)$ has the following form:

$$D_{\mu\nu}(\xi)=g_{\mu\nu}C+\partial_\mu\partial_\nu D_2(\xi), \tag{21}$$

where C is a constant.

Proof: In terms of A_μ, Eq. (3) reads

$$(\Box\delta^\mu{}_\nu-\partial^\mu\partial_\nu)A_\mu(x)=0. \tag{22}$$

Thus, the two-point function must satisfy the following equation:

$$(\Box\delta^\mu{}_\nu-\partial^\mu\partial_\nu)D_{\mu\rho}(\xi)=0. \tag{23}$$

On the other hand, when the explicit form, Eq. (17), of $D_{\mu\nu}$ is substituted in the above equation, we obtain

$$(\Box\delta^\mu{}_\nu-\partial^\mu\partial_\nu)g_{\mu\rho}D_1(\xi)=(\Box g_{\nu\rho}-\partial_\rho\partial_\nu)D_1(\xi)=0.$$

Hence, Lemma 2, ($\alpha=-1$), yields

$$D_1(\xi)=C=\text{const.}.$$

The above theorem has very strong implications. As a matter of fact, Eq. (21) gives

$$(\Psi^0,F_{\mu\nu}(x)F_{\rho\sigma}(y)\Psi^0)=0,$$

and it is clear that one cannot get anything but a trivial physical theory. The argument can be made stronger, for which purpose it is convenient to add the following assumptions:

(4) $F_{\mu\nu}(x)$ is a local operator, i.e., it satisfies local commutativity:

$$[F_{\mu\nu}(f),F_{\rho\sigma}(g)]=0,$$

if the support of f is spacelike with respect to the support of g.

(5) Let D_0 denote the set of vectors which are obtained from the vacuum state by applying polynomials in the smeared fields $F_{\mu\nu}(f)$. Then we assume that the metric operator η is positive-definite on D_0. Under these conditions one can prove the following corollary:

Corollary 1. If conditions (1)–(5) are satisfied, the following form of the two-point function

$$(\Psi^0,A_\mu(x)A_\nu(y)\Psi^0)=g_{\mu\nu}C+\partial_\mu\partial_\nu D_2(\xi) \tag{24}$$

implies that

$$A_\mu=\partial_\mu\varphi, \tag{25}$$

where φ is a scalar field.

Proof: Let us consider the state $F_{\mu\nu}(f)\Psi^0$, where f is a test function for which $F_{\mu\nu}$ is defined, and compute

$$\begin{aligned}\|F_{\mu\nu}(f)\Psi^0\|^2&=(\Psi^0,F_{\mu\nu}(f)F_{\mu\nu}(f)\Psi^0)\\ &=-\int d^3x d^3y\ f(x)f(y)[\partial_\mu\partial_\mu D_{\nu\nu}(\xi)\\ &\quad-\partial_\nu\partial_\mu D_{\mu\nu}(\xi)-\partial_\mu\partial_\nu D_{\nu\mu}(\xi)+\partial_\nu\partial_\nu D_{\mu\mu}(\xi)],\end{aligned}$$

where $\xi=x-y$. Substituting the expression (24) for $D_{\mu\nu}$, we obtain

$$\|F_{\mu\nu}(f)\Psi^0\|^2=0, \quad \text{i.e.,} \quad F_{\mu\nu}(f)\Psi^0=0.$$

Hence, a well-known theorem[18] on local operators gives

$$F_{\mu\nu}(x)=0.$$

This equation is equivalent to

$$A_\mu=\partial_\mu\varphi,$$

where φ is a scalar field.

Clearly, this is a trivial solution and cannot describe spin-one particles. This shows that the difficulties connected with the gauge problem arise *at the very beginning*, when one tries to introduce a vector potential A_μ. The use of a four-vector A_μ is in fact inconsistent with very general principles. The complications like nonpositive metric, Lorentz condition, etc., are actually *consequences* of choosing a wrong solution of Eq. (4).

As a further remark on the above results we note that the representations of the Poincaré group in the classical and in the quantum-field-theory case are essentially different. In the former case, in fact, one may require the mass-shell condition

$$\Box A_\mu=0 \tag{26}$$

together with the auxiliary condition

$$\partial^\mu A_\mu=0,$$

which eliminates the scalar photons. In the quantum-field-theory case, one has almost the opposite situation: The Lorentz condition [together with Eq. (26)] implies that A_μ is a gradientlike field.

Corollary 2. If $A_\mu(x)$ can be defined as an operator-valued distribution which transforms under the Poincaré group according to Eq. (7′), and if $F_{\mu\nu}\equiv\partial_\mu A_\nu-\partial_\nu A_\mu$ can be defined as a local field satisfying the conditions (3) and (4), then the two equations

$$\Box A_\mu=0, \quad \partial^\mu A_\mu=0,$$

[18] B. Schroer, Diplomarbeit, Hamburg 1958, (unpublished); R. Jost, Lectures on Field Theory, Naples 1959, (unpublished); P. G. Federbush and K. A. Johnson, Phys. Rev. **120**, 1926 (1960).

imply that

$$A_\mu = \partial_\mu \varphi ,$$

where φ is a scalar field.

Proof: In fact, under the above assumptions A_μ would obey Eq. (22), and the previous analysis applies.

This result should make clear which are the crucial points in the quantization of the electromagnetic field. For example, as a consequence of Corollary 2, the Fermi solution[19] of the difficulties connected with the Lorentz condition, is automatically ruled out. One cannot hope to select the physical states as those on which the Lorentz condition is satisfied

$$\partial^\mu A_\mu(f)\Psi = 0.$$

In this case, one would have a subspace $\mathcal{H}' \subset \mathcal{H}$ on which the following two equations are simultaneously satisfied:

$$\Box A_\mu = 0, \quad \partial^\mu A_\mu = 0 ,$$

and then, by Corollary 2, A_μ would reduce to the gradient of a scalar field on $\mathcal{H}'$. As a matter of fact, the Fermi condition leads to such monstrosities as states of infinite norm,[20] etc. In general, there is no way of characterizing a subspace of physical states by means of local conditions, because the Maxwell equations cannot be satisfied, as local equations in a suitable subspace, without implying a trivial theory. This is why the Lorentz condition can at most be imposed[21] in the form of a *nonlocal* equation

$$\partial^\mu A_\mu^+(f)\Psi = 0 , \tag{27}$$

where $\partial^\mu A_\mu^+$ is the positive-frequency part of the field operator $\partial^\mu A_\mu$. The extraction of the positive frequency part, is in fact a nonlocal operation. All this indicates that a Hilbert space formulation of the quantized photon field must *involve* states corresponding to unphysical particles. A local covariant field A_μ can be defined only in a Hilbert space $\mathcal{H}$ in which the Maxwell equations are not satisfied as local operator equations. We will discuss this solution later in more detail.

As a possible alternative solution one could try to formulate the theory in terms of $F_{\mu\nu}$ only. Clearly, in this case, the above difficulties would disappear, but other problems would arise. The free-field operators $F_{ij}(p)$ behave as $p_i\epsilon_j - p_j\epsilon_i$ ($\boldsymbol{\varepsilon}$ being the polarization vector) and therefore they cannot account for the production and absorption of soft photons ($p_i \to 0$) which is instead a characteristic feature of electromagnetic interactions. A theory involving only the fields $F_{\mu\nu}$ should necessarily be of a nonlocal character, and this would complicate matters because very little is known about nonlocal theories. As a matter of fact, we are not able to write down a local interaction Lagrangian in terms of $F_{\mu\nu}$. Clearly, in this case, the conventional proof of the *TCP* theorem, based on local Lagrangian field theory, does not apply. On the other hand, one does not know whether local commutativity can be required for fields, which are solutions of a nonlocal Lagrangian field theory.

Another solution could be to give up condition (7′) and to use a field A_μ which does not transform as a four-vector,

$$U(a,\Lambda)A_\mu(x)U(a,\Lambda)^{-1} = \Lambda^{-1}{}_\mu{}^\nu A_\nu(\Lambda x + a) + \partial_\mu \mathcal{F}(\Lambda,x) , \tag{28}$$

$\mathcal{F}(\Lambda,x)$ being an arbitrary function. $\mathcal{F}(\Lambda,x)$ can be determined in the free-field case from the analysis of the representations of the Poincaré group for massless spin-one particles. In the interacting case, however, $\mathcal{F}(\Lambda,x)$ is to a large extent unknown. This gives rise to serious troubles if one tries to use a Wightman-type formulation of these "noncovariant" fields. The extension of the analyticity domain of the Wightman functions from the tube $\mathcal{T}_n$ to the extended tube $\mathcal{T}_n'$ cannot be done in the standard way and many (almost all) of the interesting results of the Wightman theory (in particular *TCP*) have to be proven anew.[22]

III. ASYMPTOTIC IN AND OUT STATES

In the following sections we will discuss the solution of the gauge problem which uses a local covariant field A_μ. The problem is how to eliminate the unphysical photons. In particular one is faced with this problem if one wants to give an LSZ formulation of quantum electrodynamics.

The definition of the Hilbert space corresponding to asymptotic states is essentially a problem of free fields and no serious difficulty arises in general. The case of electromagnetic potential A_μ has, however, some peculiar features because of the impossibility of eliminating the unphysical photons completely. Therefore, some care has to be used in the definition of the in and out states corresponding to the photon field. We shall start by defining[11] the asymptotic Hilbert space $\mathcal{H}$ corresponding to the vector field A_μ as the direct sum of Hilbert spaces

$$\mathcal{H} = \sum_{n=0}^{\infty} \mathcal{H}_n ,$$

where $\mathcal{H}_n$ consists of all tensors $\Phi_{\mu_1\cdots\mu_n}(k^1,\cdots k^n)$ symmetric under simultaneous permutations of the μ's and k's, defined for the values of the arguments satisfying the following conditions:

$$k^2 \equiv k^\mu k_\mu = 0, \quad k_0 \geqslant 0.$$

The scalar product between two states $\Phi^{(n)}$ and $\Psi^{(n)}$ is

[19] E. Fermi, Rev. Mod. Phys. 4, 87 (1932).
[20] G. Källén, in *Handbuch der Physik,* edited by S. Flügge (Springer-Verlag, Berlin, 1958), Band V, Vol. 1.
[21] See Ref. 4.
[22] The analysis of this problem will be done in a subsequent paper.

defined as

$$\langle\Phi^{(n)},\Psi^{(n)}\rangle=\int\frac{d^3k^1}{k_0{}^{(1)}}\cdots\frac{d^3k^n}{k_0{}^{(n)}}\sum_{\mu_i=0}^{3}\Phi^{(n)}{}_{\mu_1\cdots\mu_n}(k^1,\cdots k^n)^*\times\Psi^{(n)}{}_{\mu_1\cdots\mu_n}(k^1,\cdots k^n),$$

$$\langle\Phi^{(0)},\Psi^{(0)}\rangle=\Phi^{(0)*}\Psi^{(0)}.$$

In general a state $\Phi\in\mathcal{H}$ has "components" on each $\mathcal{H}_n$ such that

$$\sum_n\langle\Phi^{(n)},\Psi^{(n)}\rangle<\infty$$

and the scalar product between Φ and Ψ is defined as

$$\langle\Phi,\Psi\rangle=\sum_n\langle\Phi^{(n)},\Psi^{(n)}\rangle.$$

Already at this point, difficulties arise if one wants to give a meaning to the state $\Phi^{(n)}{}_{\mu_1\cdots\mu_n}$ in terms of one-particle states. For this purpose one should introduce a vector-field operator A_μ and decompose it into creation and destruction operators, having definite meaning with respect to the states $\Phi^{(n)}$. By the results of the previous sections, a well-defined covariant operator A_μ can be introduced only if we give up Maxwell's equations. The more natural way to do so is to use a vector field A_μ and to require that the following equation be satisfied:

$$\Box A_\mu=0.$$

However, not all the difficulties are eliminated in this way. As a matter of fact, by using local commutativity for A_μ, one can show[11] that the metric cannot be positive-definite. Thus a sesquilinear form must be introduced in the Hilbert space

$$(\Phi,\Psi)=\sum_n(\Phi^{(n)},\Psi^{(n)}),$$

where

$$(\Phi^{(n)},\Psi^{(n)})=\langle\eta\Phi^{(n)},\Psi^{(n)}\rangle,$$

η being the metric operator.[4]

The vector field operator A_μ, smeared with a test function $f_\mu(x)$, will be denoted by $A(f)$. Its explicit definition in terms of the vectors $\Phi^{(n)}$ is the following[11]:

$$(A(f)\Phi)^{(n)}{}_{\mu_1\cdots\mu_n}(k^1,\cdots k^n)=\pi^{1/2}(n+1)^{1/2}\times\int\frac{d^3k}{k_0}\tilde{f}^\mu(k)\Phi^{(n+1)}{}_{\mu\mu_1\cdots\mu_n}(k,k^1,\cdots k^n)+\frac{1}{n}\sum_{j=1}^{n}\tilde{f}_{\mu_j}(-k^j)\times\Phi^{(n-1)}{}_{\mu_1\cdots\mu_{j-1}\mu_{j+1}\cdots\mu_n}(k^1,\cdots k^{j-1},k^{j+1},\cdots k^n),$$

where $\tilde{f}_\mu(k)$ is the Fourier transform of $f_\mu(x)$.

In order to have a consistent theory one has to introduce unphysical quantities and work in a Hilbert space in which longitudinal and scalar photons are allowed. (Maxwell equations involving only physical photons are satisfied as operator equations.) One may, however, try to characterize the physical states by suitable conditions.

Actually, the scalar photons may be eliminated from the state amplitude by imposing the auxiliary condition

$$A(f)^+\Psi=0, \qquad (29)$$

for all f_μ of the form $\partial_\mu g$, $g\in\mathfrak{D}$. Here $A(f)^+$ is defined by

$$(A(f)^+\Phi)^{(n)}{}_{\mu_1\cdots\mu_n}(k^1,\cdots k^n)=\pi^{1/2}(n+1)^{1/2}\int\frac{d^3k}{k_0}\tilde{f}^\mu(k)\times\Phi^{(n+1)}{}_{\mu\mu_1\cdots\mu_n}(k,k^1,\cdots k^n). \qquad (30)$$

Condition (29) determines a subspace $\mathcal{H}'$ of $\mathcal{H}$, consisting of those vectors $\Phi^{(n)}{}_{\mu_1\cdots\mu_n}$, $n=1,\cdots\infty$, for which

$$k^\mu\Phi^{(n)}{}_{\mu\mu_1\cdots\mu_{n-1}}(k,k^1,\cdots k^{n-1})=0. \qquad (31)$$

The subspace $\mathcal{H}'$ does not contain only physical states. There is in fact a subspace $\mathcal{H}''\subset\mathcal{H}'$ consisting of those vectors $\Phi_{\mu_1\cdots\mu_n}$ of the form

$$\Phi_{\mu_1\cdots\mu_n}(k^1,\cdots k^n)=k^j{}_{\mu_j}\Phi'{}_{\mu_1\cdots\mu_{j-1}\mu_{j+1}\cdots\mu_n}(k^1,\cdots k^n),$$

for at least one index μ_j. These vectors are associated with states in which at least one longitudinal photon is present. In contrast with the previous case (scalar photons), the separation of longitudinal photons cannot be done in a Lorentz invariant way.[23] One can show in fact that the spaces $\mathcal{H}'$ and $\mathcal{H}''$ are both left invariant under a Lorentz transformation, but that is *not* the case for $\mathcal{H}'-\mathcal{H}''$. This means that a state belonging to $\mathcal{H}'-\mathcal{H}''$ in a certain reference frame, will acquire a nonvanishing component in the space $\mathcal{H}''$ after a Lorentz transformation. Thus, the number of longitudinal photons is not a Lorentz invariant quantity and the only way to get rid of the longitudinal photons in a Lorentz invariant way is to define the physical states as equivalence classes: Two states are regarded as equivalent if they differ only by the presence of longitudinal photons. It is easy to prove that the quotient space $\mathcal{H}'/\mathcal{H}''$ is in fact invariant under Lorentz transformations. Thus a physical photon state will be associated with an equivalence class of $\mathcal{H}'/\mathcal{H}''$ rather than with a ray in the Hilbert space $\mathcal{H}'$. This correspondence will have a physical meaning if it is not destroyed by switching on the interaction. Equivalence classes may be associated with photon states only if the S matrix, defined on $\mathcal{H}'$, maps equivalence classes into equivalence classes. This imposes definite conditions on the S matrix. They play the same role as the gauge invariance conditions in the conventional approach. (They are, however, weaker.) In this way we will also guarantee that unphysical photons do not contribute to any physical process, so

[23] R. Shaw, Nuovo Cimento 37, 1086 (1965).

that they may be really regarded as "unobservable" particles.

In order to formulate the above problem in a clear way we need some mathematical tools.

IV. MATHEMATICAL TOOLS

For the reader's convenience we will list some definitions and theorems. The theorems, which can be found in the literature,[24] will be stated without proof.

Definition 1. A subset M of a linear space L is a subspace if

$$(1)\quad x+y\in M,\quad \forall x,y\in M;$$
$$(2)\quad \lambda x\in M,\quad \forall x\in M,\quad \forall\lambda\in\mathbb{C}.$$

Definition 2. If M is a subspace of L, one may define equivalence classes in L, with respect to M: Two elements x_1, x_2 are equivalent if $x_1-x_2\in M$. Clearly a class is completely determined by any one of its representatives. The set of all vectors equivalent to x will be denoted by ξ_x. The classes can be considered as vectors in a new linear space L/M, with the definitions

$$(\alpha)\ \xi_x+\xi_y=\xi_{x+y},\quad (\beta)\ \lambda\xi_x=\xi_{\lambda x}.$$

The zero vector is the class ξ_0 which contains M. The space L/M is called a factor space.

Definition 3. A functional $p(x)$ defined on all of L is said to be convex if

$$(1)\ p(x)\geqslant 0,\ \forall x\in L;$$
$$(2)\quad p\left(\frac{x+y}{2}\right)\leqslant\tfrac{1}{2}[p(x)+p(y)],\ \forall x,y\in L;$$
$$(3)\quad p(\lambda x)=\lambda p(x),\ \forall\lambda\geqslant 0,\ \forall x\in L.$$

A convex functional is said to be symmetric (or a seminorm) if

$$p(\lambda x)=|\lambda|p(x),\quad \forall\lambda\in\mathbb{C},\quad \forall x\in L.$$

Theorem 1. If $p(x)$ is a symmetric convex functional, then

$$M=\{x\,|\,p(x)=0\}\text{ is a subspace in }L.$$

Theorem 2. If $p(x)$ is an arbitrary symmetric convex functional in L, and M is the subspace of all vectors $x\in L$ on which $p(x)=0$, then the equation

$$|\xi|=p(x),\quad x\in\xi$$

defines a norm on L/M.

Definition 4. A Hermitian bilinear form in L is a function of the two variables x, $y\in L$, such that

$$(1)\quad (x,y)=(y,x)^*,\quad (2)\quad (x,\lambda y)=\lambda(x,y),$$
$$(3)\quad (x,y_1+y_2)=(x,y_1)+(x,y_2).$$

[24] See e.g., M. A. Naimark, *Normed Rings* (P. Noordhoff Ltd., Groningen, The Netherlands, 1964).

A Hermitian bilinear form is positive if $(x,x)\geq 0$. In this case $p(x)=(x,x)$ is a convex functional.

Theorem 3. If (x,y) is a positive definite bilinear Hermitian form in L and $M=\{x\,|\,(x,x)=0\}$, then

$$(\xi,\eta)=(x,y),\quad\text{where}\quad x\in\xi,\quad y\in\eta,$$

defines an inner product in L/M, i.e., (ξ,η) satisfies $(\xi,\xi)=0$ if and only if $\xi=0$, in addition to conditions (1)–(3) of Definition 4.

This is the first step in the standard procedure to get a Hilbert space out of a pre-Hilbert space. One encounters this problem in quantum electrodynamics where the longitudinal photons have zero norm. A Hilbert space of physical photon states is obtained by completing $\mathcal{H}'/\mathcal{H}''$. The compatibility of this definition with the S-matrix theory imposes definite conditions. For that we will need the following theorem.

Definition 5. We will say that a linear operator S, defined in $\mathcal{H}'$, *induces* a linear operator $\tilde{S}$ in $\mathcal{H}'/\mathcal{H}''$, where $\mathcal{H}''$ is a subspace of $\mathcal{H}'$, if

$$y'=Sy\quad\text{implies}\quad \xi_{y'}=\tilde{S}\xi_y,$$

$y\in\mathcal{H}'$, $\xi_y\in\mathcal{H}'/\mathcal{H}''$.

A linear operator S defined in $\mathcal{H}'$ will be said to be *compatible with* a linear operator $\tilde{S}$ defined in $\mathcal{H}'/\mathcal{H}''$ if $\xi'=\tilde{S}\xi$ implies that for any $x\in\xi$ there exists a $x'\in\xi'$ such that $x'=Sx$.

Theorem 4. A linear operator S defined in $\mathcal{H}'$ induces (is compatible with) a linear operator $\tilde{S}$ defined in $\mathcal{H}'/\mathcal{H}''$, if and only if $\mathcal{H}''$ is invariant under S, i.e.,

$$S\mathcal{H}''\subseteq\mathcal{H}''.$$

Proof:

Clearly if $\mathcal{H}''$ is invariant under S, we may define an operator $\tilde{S}$ in $\mathcal{H}'/\mathcal{H}''$ in the following way:

$$\tilde{S}\xi_x=\xi_y\quad\text{if}\quad Sx=y.$$

This definition is independent of the choice of x and y in ξ_x and ξ_y, respectively, so that the above equation defines an operator in $\mathcal{H}'/\mathcal{H}''$. As a matter of fact, if $x_1-x_2\in\mathcal{H}''$ and $Sx_1=y_1$, $Sx_2=y_2$, then by the invariance of $\mathcal{H}''$ we have $S(x_1-x_2)=y_1-y_2\in\mathcal{H}''$ and $S\xi_{x_1}=\xi_{y_1}=\xi_{y_2}=S\xi_{x_2}$.

Conversely, let S induce an operator $\tilde{S}$ in $\mathcal{H}'/\mathcal{H}''$. Then

$$x'=Sx\quad\text{implies}\quad \xi_{x'}=\tilde{S}\xi_x.$$

Moreover, if $y\in\xi_x$ and $Sy=y'$, we have

$$\xi_{y'}=\tilde{S}\xi_y=\tilde{S}\xi_x.$$

Hence

$$\xi_{y'}-\xi_{x'}=\tilde{S}(\xi_y-\xi_x)=\xi_0,$$

i.e., $y'-x'\in\mathcal{H}''$ or $S(x-y)=x'-y'\in\mathcal{H}''$ whenever $x-y\in\mathcal{H}''$.

Similarly, let $\tilde{S}$ be defined on $\mathcal{H}'/\mathcal{H}''$. We require that there exist an operator S in $\mathcal{H}'$, such that S is compatible with $\tilde{S}$. Then S must leave $\mathcal{H}''$ invariant. In fact $\tilde{S}\xi_0=\xi_0$

implies that for any $x \in \xi_0$, there exists a x', $x' \in S\xi_0 = \xi_0$ such that $x' = Sx$, i.e.,

$$S\mathcal{H}'' \subseteq \mathcal{H}''.$$

V. CONDITIONS ON THE S MATRIX

We will now work in the whole Hilbert space of the system, in which other physical particles may be present, besides photons. For simplicity, only the indices and the variables referring to photon states will be spelled out explicitly. In the Hilbert space $\mathcal{H}$ of the asymptotic states, we will denote by

$$|e^{\text{in}}{}_{\nu_1\cdots\nu_n}(k_1,\cdots k_n)\rangle$$

the *basic* vectors with n "photons" of momenta $k_1\cdots k_n$, and "polarization" $\nu_1\cdots\nu_n$. These properties refer to the "in observables". Similarly

$$|e^{\text{out}}{}_{\mu_1\cdots\mu_m}(q_1,\cdots q_m)\rangle$$

will denote the basic vectors with m "out-photons" of momenta $q_1\cdots q_m$ and "polarizations" $\mu_1\cdots\mu_m$.

By the assumption that both the set of in and out basic vectors are complete, every vector $|\Phi\rangle \in \mathcal{H}$ may be expanded in terms of any of them. Shortly, one may write

$$|\Phi\rangle = \sum_{n,q,\mu} \int \Phi^{(n)}{}_{\mu_1\cdots\mu_n}(q)\,|e_{\mu_1\cdots\mu_n}(q)\rangle. \qquad (32)$$

A change of basis is given by the following formula:

$$\begin{aligned}|\Phi\rangle &= \sum_{n,k,\nu} \int \Phi^{\text{in}}{}_{\nu_1\cdots\nu_n}(k)\,|e^{\text{in}}{}_{\nu_1\cdots\nu_n}(k)\rangle \\ &= \sum_{k,q,n,m,\mu,\nu} \int \Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n}(k) \\ &\qquad\qquad \times S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}(q,k)\,|e^{\text{out}}{}_{\mu_1\cdots\mu_m}(q)\rangle \\ &= \sum_{m,q,\mu} \int \Phi^{\text{out}(m)}{}_{\mu_1\cdots\mu_m}(q)\,|e^{\text{out}}{}_{\mu_1\cdots\mu_m}(q)\rangle,\end{aligned}$$

where

$$\Phi^{\text{out}(m)}{}_{\mu_1\cdots\mu_m}(q) = \sum_{n,k,\nu} \int S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}(q,k)\,\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n}(k). \qquad (33)$$

The transformation

$$\langle e^{\text{out}}{}_{\mu_1\cdots\mu_m}(q)\,|\,e^{\text{in}}{}_{\nu_1\cdots\nu_n}(k)\rangle \equiv S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}(q,k) \qquad (34)$$

is called the S matrix.

By Eq. (32), a vector $|\Phi\rangle$ is completely determined by the sequence of symmetric tensors $\{\Phi^{(n)}{}_{\mu_1\cdots\mu_n}(q), n=0, 1, \cdots\}$ so that instead of working in the abstract Hilbert space of vectors $|\Phi\rangle$, one may work in one of its realizations by the sequence of tensors $\{\Phi^{(n)}{}_{\mu_1\cdots\mu_n}, n=0, 1, \cdots\}$. The former notation will, however, keep clear the distinction between a vector $|\Phi\rangle$ and its representations $\{\Phi^{(n)}{}_{\mu_1\cdots\mu_n}, n=0, 1, \cdots\}^{\text{in/out}}$ in terms of the basic in or out states.

By applying the previous analysis of the free photon states, we may define the subspace $\mathcal{H}'_{\text{in}}$ as the set of vectors $|\Phi\rangle$ such that

$$\sum_{\nu,\nu_j=0}^{3} g^{\nu\nu_j} k_\nu \Phi^{\text{in}}{}_{\nu_1\cdots\nu_j\cdots\nu_n}(k) = 0, \quad \forall j. \qquad (35)$$

Similarly, $\mathcal{H}'_{\text{out}}$ is the set of vectors $|\Phi\rangle$, such that

$$q^{\mu_i}\Phi^{\text{out}}{}_{\mu_1\cdots\mu_i\cdots\mu_n}(q) = 0, \quad \forall i. \qquad (36)$$

The completeness of the in or out basic vectors, implies $\mathcal{H}^{\text{in}} = \mathcal{H}^{\text{out}} = \mathcal{H}$.

Finally, we will denote by $\mathcal{H}_{\text{in}}''$, the subspace consisting of all those vectors $|\Phi\rangle$ such that

$$\Phi^{\text{in}}{}_{\nu_1\cdots\nu_i\cdots\nu_n}(k) = k_{\nu_i}\Phi^{\text{in}}{}_{\nu_1\cdots\hat{\nu}_i\cdots\nu_n}(k) \qquad (37)$$

for at least one n and one j. $\mathcal{H}''_{\text{out}}$ is defined in a similar way. They describe states containing at least one longitudinal photon, in or out, respectively.

We shall restrict ourselves to the subspaces $\mathcal{H}'_{\text{in/out}}$. In this way we will eliminate in a Lorentz invariant way the scalar photons, which do not satisfy the Lorentz condition. It can be shown that in $\mathcal{H}'_{\text{in/out}}$, η is ≥ 0. By the standard procedure outlined in the previous section we can generate two Hilbert spaces $\mathcal{H}'_{\text{in}}/\mathcal{H}''_{\text{in}}$ ($\mathcal{H}'_{\text{out}}/\mathcal{H}''_{\text{out}}$), such that η defines an inner product in $\mathcal{H}'_{\text{in}}/\mathcal{H}''_{\text{in}}$ ($\mathcal{H}'_{\text{out}}/\mathcal{H}''_{\text{out}}$). Physical in or out photon states will correspond to equivalence classes in $\mathcal{H}'_{\text{in}}$ or $\mathcal{H}'_{\text{out}}$, respectively. As the scalar photons, which do not satisfy the Lorentz condition, should not contribute to any physical process, the S matrix must map $\mathcal{H}'_{\text{in}}$ into itself. This implies

$$\mathcal{H}'_{\text{in}} = S\mathcal{H}'_{\text{in}} = \mathcal{H}'_{\text{out}}. \qquad (38)$$

This equation is essentially a transcription of an analogous equation derived by requiring gauge invariance.[25]

The equality $\mathcal{H}'_{\text{in}} = \mathcal{H}'_{\text{out}}$ could also have been anticipated from the beginning. It means that the characterization of the space of transversal photons is independent of the basis, in or out. The same result would be obtained by describing the electromagnetic field by means of $F_{\mu\nu}$ instead of A_μ. In this case, one may choose $\mathcal{H}'_{\text{in}}$ as the whole Hilbert space ($\mathcal{H}'_{\text{in}}$ is left invariant by $F^{\text{in}}{}_{\mu\nu}$) and therefore $\mathcal{H}'_{\text{in}} = \mathcal{H}'_{\text{out}}$. From now on we may talk of $\mathcal{H}'$ without specifying in or out.

Condition (38) may be transcribed into explicit conditions on the elements of the S matrix. In fact, the invariance of $\mathcal{H}'$ under the transformations induced by the S matrix means that the equation

$$k^{\nu_i}\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_i\cdots\nu_n}(k) = 0, \quad \forall n, \quad \forall i \qquad (39)$$

must imply

$$q^{\mu_i}\Phi^{\text{out}(m)}{}_{\mu_1\cdots\mu_i\cdots\mu_m}(q) = 0, \quad \forall m, \forall i \qquad (40)$$

[25] See: J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1959), p. 85.

and vice versa. Here $\Phi^{\text{out}}{}_{\mu_1\cdots\mu_m}$ is obtained from $\Phi^{\text{in}}{}_{\nu_1\cdots\nu_n}$ by means of the S transformation, Eq. (34). Thus Eq. (39) must imply

$$q^\mu \sum \int S_{\mu\cdots\mu_m;\nu_1\cdots\nu_n}(q,k)\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n}(k)=0.$$

This must be true for any vector whose components satisfy Eq. (39). In particular it must be true if the vector has only one component, say $\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n}(k)$. Moreover the k dependence of $\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n}(k)$ is to a large extent arbitrary, since we may multiply $\Phi^{\text{in}(n)}$ by an arbitrary function of k. Thus, Eqs. (39) and (40) imply that

$$k^{\nu_i}\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_i\cdots\nu_n}=0, \quad \forall i$$

must yield

$$q^{\mu_1}S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}(q,k)\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n}(k)=0.$$

This gives

$$q^{\mu_1}S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}(q,k)=\sum_{\text{perm.}} S'_{\mu_2\cdots\mu_m;\nu_2\cdots\nu_n}(q,k)k_{\nu_1}. \quad (41)$$

Conversely, by requiring that Eq. (40) must imply Eq. (39), one obtains in a similar way

$$\bar{S}_{\nu_1\cdots\nu_n;\mu_1\cdots\mu_m}(k,q)q_{\mu_1}=\sum_{\text{perm.}} k_{\nu_1}\bar{S}'_{\nu_2\cdots\nu_n;\mu_2\cdots\mu_m}(k,q),$$

where

$$\bar{S}_{\nu_1\cdots\nu_n;\mu_1\cdots\mu_m}(k,q)\equiv\langle e^{\text{in}}{}_{\nu_1\cdots\nu_n}(k)\,|\,e^{\text{out}}{}_{\mu_1\cdots\mu_m}(q)\rangle =(S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_m}(q,k))^{*T}.$$

The problem of eliminating any contribution by the longitudinal photons is slightly more complicated.

The amplitudes for physical processes are given as quantities of direct physical meaning. They will be described in terms of scalar products of physical states, i.e., elements of $\mathcal{H}'/\mathcal{H}''$. The probability amplitude for the transition $i\to f$ will be described by

$$(\xi^{\text{out}}{}_f,\xi^{\text{in}}{}_i)=(\tilde{S}\xi^{\text{in}}{}_f,\xi^{\text{in}}{}_i), \quad (42)$$

where ξ_f, ξ_i are the classes corresponding to the initial and final states, respectively. Equation (42) defines an operator $\tilde{S}$ in $\mathcal{H}'_{\text{in}}/\mathcal{H}''_{\text{in}}$. We want to know under which conditions $\tilde{S}$ may be regarded as induced by a linear operator S defined on $\mathcal{H}'_{\text{in}}$; and conversely which are the conditions on S in order that S should be "compatible" with $\tilde{S}$. (See Sec. IV, Definition 5). The answer is provided by Theorem 4 of the previous section:

$$S\mathcal{H}''_{\text{in}}\subseteq\mathcal{H}''_{\text{in}}. \quad (43)$$

In an explicit form Eq. (43) implies that if there is a longitudinal photon in the in-vector, at least one longitudinal photon must also appear in the out-vector, i.e., if

$$\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n}(k)=k_{\nu_1}\Phi^{\text{in}}{}_{\nu_2\cdots\nu_n},$$

it must be

$$\Phi^{\text{out}}{}_{\mu_1\cdots\mu_m}=\sum\int S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}\Phi^{\text{in}(n)}{}_{\nu_1\cdots\nu_n} =\sum_{\text{perm.}} q_{\mu_1}\Phi^{\text{out}}{}_{\mu_2\cdots\mu_m}.$$

This yields

$$S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}(q,k)k^{\nu_1}=\sum_{\text{perm.}} q_{\mu_1}S'_{\mu_2\cdots\mu_m;\nu_2\cdots\nu_n}(q,k).$$

This condition guarantees that longitudinal photons are not created out of physical states and that the two subspaces $\mathcal{H}''_{\text{in}}$ and $\mathcal{H}''_{\text{out}}$ coincide. By combining this result with the previous one, one may say that the characterization of the subspace of physical photons in the larger Hilbert space $\mathcal{H}$ does not depend on the basis in or out. The same conclusion would, of course, be reached by working with only $F_{\mu\nu}$, without introducing A_μ.

The invariance of $\mathcal{H}''$ can also be derived by requiring the invariance of the S-matrix under the gauge transformation[26]

$$\Phi^{\text{in}}{}_{\nu_1\cdots\nu_n}\to\Phi^{\text{in}}{}_{\nu_1\cdots\nu_n}+k_{\nu_1}\tilde{\Phi}^{\text{in}}{}_{\nu_2\cdots\nu_n}.$$

The usual requirement is, however, stronger than ours, because by invariance of the S operator we do not mean that

$$\delta S\equiv S_{\mu_1\cdots\mu_m;\nu_1\cdots\nu_n}k^{\nu_1}\tilde{\Phi}^{\text{in}}{}_{\nu_2\cdots\nu_n}=0, \quad (44)$$

but only that

$$\delta S=\sum_{\text{perm.}} q_{\mu_1}S'_{\mu_2\cdots\mu_m;\nu_1\cdots\nu_n}k^{\nu_1}\tilde{\Phi}^{\text{in}}{}_{\nu_2\cdots\nu_n},$$

so that δS does not contribute to physical processes. This means that S may very well have a non-gauge-invariant part in the sense of Eq. (44); the only important thing is that S and $S+\delta S$ coincide in the physical subspace $\mathcal{H}'/\mathcal{H}''$.

ACKNOWLEDGMENT

I wish to thank Professor A. S. Wightman for his kind hospitality at Palmer Physical Laboratory, for his interest in this work, and for useful remarks.

[26] See, e.g., J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1959), pp. 189, 195.