Massive Electrodynamics*

G. FELDMAN† AND P. T. MATTHEWS
Imperial College, London, England
(Received 14 January 1963)

The relation between gauge invariance, canonical quantization, and the photon mass is examined. It is found that gauge invariance does not require the bare photon mass to be zero. In fact, quantization in an arbitrary covariant gauge is only possible for massive photons. This is a generalization of the usual Fermi procedure. A similar generalization of the conventional Coulomb gauge quantization leads to a gauge-invariant, but noncovariant, theory of massive transverse photons.

INTRODUCTION

THE invariance of Lagrangians under constant phase transformations has often been generalized to include local phase transformations.[1] To maintain the invariance it is usual[2] to introduce a vector field. For example, in electrodynamics if $\Psi(x)$ is the charged field, which undergoes

$$\Psi(x) \rightarrow e^{ie\Lambda(x)}\Psi(x), \tag{1.1}$$

invariance is maintained if one introduces the photon field $A_\mu(x)$, such that

$$A_\mu(x) \rightarrow A_\mu(x)+\partial_\mu\Lambda(x). \tag{1.2}$$

The interaction of the photon field is determined by the replacement

$$\partial_\mu \rightarrow \partial_\mu - ieA_\mu, \tag{1.3}$$

in the Lagrangian of the charged particles. The interaction Lagrangian, L_{int}, generated in this way, defines a current j_μ, where

$$j_\mu=\delta L_{\text{int}}/\delta A_\mu. \tag{1.4}$$

This current is conserved,

$$\partial_\mu j_\mu=0. \tag{1.5}$$

The transformations (1.1) and (1.2) are referred to as gauge transformations. It is generally assumed that in addition to inducing a conserved current, gauge invariance also implies that the mass of the photon is zero.

The extension of these ideas to isotopic spin and unitary symmetry[3] has led to the prediction of an octet of vector mesons with precisely the spin and isotopic properties of ρ, ω, and K^*, which, however, have nonzero mass. We are, thus, led to reconsider the implications of gauge invariance for the mass of the induced vector particles.

The simplest instance of such a particle is the photon, since it does not itself carry the conserved quantity, charge. (In contrast, ρ carries isotopic spin.) Our discussion in this paper is restricted to this simple case.

We consider the restrictions imposed on the theory of the free photon part of the Lagrangian by the requirement that the observable properties of the system are invariant under gauge transformations.[4] We are, thus, led to consider first the theory of free photons.

We show below that there are two independent ways in which the free photon Lagrangian can be modified, in a gauge invariant manner, to include mass terms and which admit of a conventional canonical quantization. The two methods are based on the introduction of two gauge-invariant vector fields $A_\mu{}^\tau(x)$ (the Landau field) and $A_\mu{}^t(x)$ (the transverse field). The first formulation is a generalization of the usual quantization in the Fermi gauge, and leads naturally to a covariant theory of massive photons. The second formulation generalizes the usual quantization in the Coulomb gauge. The procedure is noncovariant and leads, in general, to a noncovariant theory of massive transverse photons. Only in the special case of zero-mass photons is the resulting theory covariant, since only for zero-mass particles is the notion of transverse polarization a covariant one.

2. CLASSICAL THEORY OF THE FREE ELECTROMAGNETIC FIELD WITH DISPERSION

We discuss first a covariant, gauge-invariant, generalization of the classical Maxwell theory.

Consider[5]

$$\tau_{\mu\nu}(x,y)\equiv\delta_{\mu\nu}\delta(x-y)-\partial_\mu\partial_\nu D(x-y), \tag{2.1}$$

where $D(x-y)$ is defined by the equation

$$\partial^2 D(x-y)=\delta(x-y). \tag{2.2}$$

This operator has the properties

$$\partial_\mu\tau_{\mu\nu}(x,y)=0, \tag{2.3}$$

and

$$\tau^n=\tau. \tag{2.4}$$

* The research reported in this document has been sponsored by the National Science Foundation and Air Force Office of Scientific Research, OAR, through the European Office, Aerospace Research, United States Air Force.

† John Simon Guggenheim Fellow on leave of absence from The Johns Hopkins University, Baltimore, Maryland.

[1] R. Utiyama, Phys. Rev. **101**, 1597 (1956); C. N. Yang and R. L. Mills. *ibid.* **96**, 191 (1954).

[2] See, however, V. I. Ogievetskij and I. V. Polubarinov, in *Proceedings of the 1962 Annual International Conference on High-Energy Physics at CERN*, edited by J. Prentki (CERN, Geneva, 1962).

[3] A. Salam and J. C. Ward. Nuovo Cimento **11**, 568 (1960).

[4] Our approach differs from that of J. Schwinger [Phys. Rev. **125**, 397 (1962)] who accepts the conventional theory of zero-mass bare photons and considers the possibility of mass being induced by the interaction.

[5] We use natural units $\hbar=c=1$, and the notation $\partial_\mu=\partial/\partial x_\mu$, ($\mu=1,2,3,4$) and $\partial^2\equiv\partial_\mu\partial_\mu\equiv\nabla^2+\partial_4{}^2$.

It is thus a projection operator for the "covariant transverse" part of any vector field. Given any such field, $A_\mu(x)$, we define

$$A_\mu{}^\tau(x)\equiv\tau_{\mu\nu}(x,y)A_\nu(y), \tag{2.5}$$

where summation or integration is implied over any repeated suffix or variable. $A_\mu{}^\tau$ is a gauge-invariant field, since it is invariant for quite arbitrary gauge transformations of the type (1.2).

In terms of this operator the conventional equation of motion for the free photon field may be written

$$\partial^2\tau_{\mu\nu}(x,y)A_\nu(y)=0. \tag{2.6}$$

We consider the possibility of generalizing this to

$$(\partial^2-\mu^2)\tau_{\mu\nu}(x,y)A_\nu(y)=0, \tag{2.7}$$

which is again clearly gauge invariant. The equation is nonlocal, but in the Lorentz gauge, in which

$$\partial_\nu A_\nu(x)=0, \tag{2.8}$$

the equation reduces to

$$(\partial^2-\mu^2)A_\mu(x)=0. \tag{2.9}$$

We may thus anticipate that the nonlocality is of a fairly innocuous character, and that after quantization the equations refer to photons of mass μ.

The classical plane wave solutions of (2.7), of the form

$$\epsilon_\mu\exp(ik_\mu x_\mu), \tag{2.10}$$

have to satisfy

$$(k^2+\mu^2)\tau_{\mu\nu}(k)\epsilon_\nu(k)=0, \tag{2.11}$$

where

$$\tau_{\mu\nu}(k)\equiv(\delta_{\mu\nu}-k_\mu k_\nu/k^2). \tag{2.12}$$

Since

$$\tau_{\mu\nu}(k)k_\nu=0, \tag{2.13}$$

we may take four solutions

$$\epsilon_\nu{}^{(\mu)}(k)=\tau_{\mu\nu}(k), \tag{2.14}$$

for which

$$k^2+\mu^2=0, \tag{2.15}$$

of which only *three* are independent. A *fourth independent* solution is $\epsilon_\mu\sim k_\mu$, for which k^2 is completely arbitrary. It is convenient to introduce the parameter λ, such that

$$k^2+\lambda^2=0, \tag{2.16}$$

and to define the fourth polarization vector

$$\epsilon_\mu=k_\mu/\lambda. \tag{2.17}$$

We can now make a Fourier expansion of the field operator $A_\mu(x)$ in terms of these classical solutions [using (2.14) and (2.17)]:

$$A_\mu(x)=\int\frac{d^4k\vartheta(k_0)}{(2\pi)^3}\{\tau_{\mu\nu}(k)[a_\nu(k)e^{ikx}+a_\nu{}^\dagger(k)e^{-ikx}]\times\delta(k^2+\mu^2)+(k_\mu/\lambda)\times[a(k)e^{ikx}+a^\dagger(k)e^{-ikx}]\delta(k^2+\lambda^2)\}, \tag{2.18}$$

where

$$d^4k=dk_0d\mathbf{k}\quad(k_0=-ik_4), \tag{2.19}$$

and

$$\vartheta(k_0)=1,\quad k_0>0$$
$$=0,\quad k_0<0. \tag{2.20}$$

Since it follows immediately from (2.19) that[6]

$$(\partial^2-\lambda^2)\partial_\nu A_\nu=0, \tag{2.21}$$

it is clear that the parameter λ serves to specify an arbitrary gauge.[7] In the particular, but arbitrary, gauge specified by λ, the equation of motion may be written in the *local* form[8]

$$\partial_\nu[\partial_\nu A_\mu(x)-\partial_\mu A_\nu(x)]-\mu^2(\delta_{\mu\nu}-\partial_\mu\partial_\nu/\lambda^2)A_\nu(x)=0. \tag{2.22}$$

The gauge transformation which takes the system from the gauge specified by λ to that specified by λ' is induced by $\Lambda(x)$, where

$$(\partial^2-\lambda'^2)\Lambda(x)=(\lambda'^2/\lambda^2-1)\partial_\nu A_\nu. \tag{2.23}$$

From (2.21), this implies that

$$(\partial^2-\lambda^2)(\partial^2-\lambda'^2)\Lambda(x)=0. \tag{2.24}$$

3. QUANTUM THEORY OF THE FREE FIELD

The equation of motion (2.22) may be derived from the Lagrangian density[9]

$$L=-\tfrac{1}{2}[\partial_\mu A_\nu(x)\partial_\mu A_\nu(x)-\partial_\mu A_\nu(x)\partial_\nu A_\mu(x)]-(\mu^2/2)\{A_\mu{}^2(x)+[\partial_\mu A_\mu(x)]^2/\lambda^2\}. \tag{3.1}$$

Since this Lagrangian is local, we may proceed to develop a quantized theory in the conventional way.

We define $\Pi_\mu(x)$, related to the canonical momentum,[10] by

$$\Pi_\mu(x)=\delta L/\delta[\partial_4A_\mu(x)]=-[\partial_4A_\mu(x)-\partial_\mu A_4(x)+(\mu^2/\lambda^2)\delta_{\mu4}\partial_\nu A_\nu(x)]. \tag{3.2}$$

Note that the mass term plays the role of the usual Fermi term in preventing Π_4 from vanishing identically.

The canonical commutation relations are

$$[A_\mu(\mathbf{x},t),\,\Pi_\nu(\mathbf{y},t)]=-\delta_{\mu\nu}\delta^3(\mathbf{x}-\mathbf{y}). \tag{3.3}$$

[6] Note that for the solution (2.17), k^2 is quite arbitrary. Instead of introducing the parameter λ, (2.16) could have been replaced by $f(k^2)=0$, which would have replaced the gauge condition (2.21) by $f(\partial^2)\partial_\nu A_\nu=0$. To carry through the subsequent quantization we would have to restrict $f(k^2)$ to those functions [e.g., $f(k^2)=\Pi(k^2+\lambda_i{}^2)$] considered by A. Pais and G. E. Uhlenbeck, Phys. Rev. **79**, 145 (1950).

[7] The particular case $\lambda=\mu$ has been discussed by Ogievetskij and Polubarinov. See reference 2.

[8] Note that the equation analogous to (2.11) is now

$$[(k^2+\mu^2)\delta_{\mu\nu}-(1-\mu^2/\lambda^2)k_\mu k_\nu]\epsilon_\nu=0.$$

This equation has the solutions found for (2.11) provided $\mu\neq0$. The equation has four independent solutions in the limit $\mu\to0$, *only* if $\mu^2/\lambda^2\to1$, which is the Fermi gauge.

[9] This Lagrangian is formally identical to the free Lagrangian used by T. D. Lee and C. N. Yang [Phys. Rev. **128**, 885 (1962)] to describe *charged* vector mesons, with their ξ equal to μ^2/λ^2 in our notation.

[10] The canonical momentum is $\Pi_\mu{}^{(0)}=-i\Pi_\mu=\delta L/\delta\dot{A}_\mu$.

It follows from (2.18) and (3.2) that

$$\Pi_\mu(x)=i\int\frac{d^4k}{(2\pi)^3}\vartheta(k_0)\{(k_\mu\delta_{4\nu}-\delta_{\mu\nu}k_4)\delta(k^2+\mu^2)\times[a_\nu(k)e^{ikx}-a_\nu{}^\dagger(k)e^{-ikx}]+\delta(k^2+\lambda^2)\delta_{\mu4}\frac{\mu^2}{\lambda}[a(k)e^{ikx}-a^\dagger(k)e^{-ikx}]\}. \quad (3.4)$$

The commutation relations for the operators $a_\nu(k)$, $a(k)$, $a_\nu{}^\dagger(k)$, and $a^\dagger(k)$ follow immediately from (3.3). They are

$$2\pi\vartheta(k_0)\delta(k^2+\mu^2)[a_\mu(k),\, a_\nu{}^\dagger(p)]=(2\pi)^4\tau_{\mu\nu}(k)\delta^4(k-p), \quad (3.5)$$

$$2\pi\vartheta(k_0)\delta(k^2+\lambda^2)[a(k),\, a^\dagger(p)]=-(2\pi)^4\delta^4(k-p)\lambda^2/\mu^2. \quad (3.6)$$

All other commutators vanish.

The Hamiltonian of the system is

$$P_0=\int(\Pi_\mu\partial_4A_\mu-L)d^3x, \quad (3.7)$$

where this is interpreted, as usual, as the normal product to remove the zero-point energy. Substituting the expressions (2.18) and (3.4) into (3.7), one obtains after some calculation

$$P_0=\int\frac{d^4k}{(2\pi)^3}k_0\vartheta(k_0)\left[a_\mu{}^\dagger(k)\tau_{\mu\nu}(k)a_\nu(k)\delta(k^2+\mu^2)-\frac{\mu^2}{\lambda^2}a^\dagger(k)a(k)\delta(k^2+\lambda^2)\right]. \quad (3.8)$$

The negative signs appearing in (3.6) and (3.8) in connection with the time-like photons give rise to the familiar difficulties, which can be eliminated by introducing an indefinite metric. Thus, if we define

$$a^*\equiv-a^\dagger, \quad (3.9)$$

we can take the vacuum state to satisfy

$$a_\mu\rangle_0=0,\quad a\rangle_0=0. \quad (3.10)$$

The completeness relation is

$$1=\sum|n\rangle(-1)^{n_\lambda}\langle n|,$$

where the summation is over all states and n_λ is the number of time-like photons in the state $|n\rangle$.[11] The normalization condition is

$$\langle n|n'\rangle=(-1)^{n_\lambda}\delta_{nn'}. \quad (3.12)$$

[11] The states $|n\rangle$ can be defined in terms of the creation operators on the vacuum state. Thus, for example, the one-particle states $a_\mu{}^\dagger(k)\rangle_0=|k,\mu\rangle$ are normalized so that $\langle k,\mu|p,\nu\rangle 2\pi\vartheta(k_0)\delta(k^2+\mu^2)=(2\pi)^4\delta^4(k-p)\tau_{\mu\nu}$.

The system consists of vector mesons of mass μ, and time-like mesons of mass λ. The gauge field $\Lambda(x)$, defined by (2.24) and (2.25), is associated with particles of mass λ and λ' with opposite metric (see Pais and Uhlenbeck, reference 6). The corresponding gauge transformation removes the former and substitutes the latter.

4. THE PROPAGATOR

Using (3.10), the vacuum expectation value of the product of two field operators at arbitrary space-time points can be calculated directly from the expansion (2.18) and the commutation relations (3.5) and (3.6). The result is

$$\langle A_\mu(x)A_\nu(y)\rangle_0=\frac{1}{(2\pi)^3}\int d^4k\, e^{ik(x-y)}\vartheta(k_0)\times[\tau_{\mu\nu}(k)\delta(k^2+\mu^2)-(k_\mu k_\nu/\mu^2)\delta(k^2+\lambda^2)]. \quad (4.1)$$

From this the vacuum expectation value of any other type of product of two operators may be obtained directly. Thus, the anticommutator is similar to (4.1), but with $\vartheta(k_0)$ replaced by 1; in the commutator $\vartheta(k_0)$ is replaced by

$$\epsilon(k_0)=\vartheta(k_0)-\vartheta(-k_0). \quad (4.2)$$

Finally the time-ordered product is (4.1) with the replacement

$$\vartheta(k_0)\to\tfrac{1}{2}+\tfrac{1}{2}\epsilon(x_0)\epsilon(k_0). \quad (4.3)$$

The evaluation of the second term gives rise to non-covariant contributions[12] arising from the presence of $k_\mu k_\nu$ which cancel. The result is consequently covariant.[13]

$$\langle T(A_\mu(x),\, A_\nu(y))\rangle=\frac{-i}{(2\pi)^4}\int d^4k\, e^{ik(x-y)}D_{F\mu\nu}(k), \quad (4.4)$$

where

$$D_{F\mu\nu}(k)=\left(\delta_{\mu\nu}+\frac{k_\mu k_\nu}{\mu^2}\right)\frac{1}{k^2+\mu^2-i\epsilon}-\frac{k_\mu k_\nu}{\mu^2}\frac{1}{k^2+\lambda^2-i\epsilon}=\tau_{\mu\nu}(k)\frac{1}{k^2+\mu^2-i\epsilon}+\frac{\lambda^2}{\mu^2}\frac{k_\mu k_\nu}{k^2}\frac{1}{k^2+\lambda^2-i\epsilon}. \quad (4.5)$$

The propagator thus contains a gauge-invariant term representing the propagation of heavy photons of mass μ, and a gauge-dependent term arising from the timelike photons of mass λ. The choice of particular values of the parameter λ reproduces the generalization of particular gauges which have been used in the past. Thus, for example, $\lambda=0$ is the true Landau gauge for heavy photons.

[12] See, for example, L. Evans, G. Feldman, and P. T. Matthews, Ann. Phys. (N. Y.) **13**, 268 (1961).

[13] This result is consistent with the general conclusion of reference 12 that a covariant T product is a consequence of the four fields $A_\mu(x)$ being treated as independent canonical variables. Gauges in which this condition is satisfied were termed "true."

One can see from this propagator why it has been difficult to quantize conventional electrodynamics in an arbitrary covariant gauge. The limit $\mu\rightarrow 0$ must be taken with some care. To obtain a well-defined propagator we must simultaneously take the limits

$$\lambda\rightarrow 0, \tag{4.6}$$

$$\lambda^2/\mu^2\rightarrow a, \tag{4.7}$$

where a is an arbitrary constant. However, if μ is put equal to zero in the Lagrangian, the quantization procedure can only be carried out for $a=1$, which is the Fermi gauge. (See footnote 8.)

Another interesting limit is

$$\lambda\rightarrow\infty, \tag{4.8}$$

for which the propagator reduces to

$$D_{F\mu\nu}(k)=\left(\delta_{\mu\nu}+\frac{k_\mu k_\nu}{\mu^2}\right)\frac{1}{k^2+\mu^2-i\epsilon}. \tag{4.9}$$

It is clear from (2.21) that in this limit we must have

$$\partial_\nu A_\nu(x)=0 \tag{4.10}$$

satisfied as an operator condition. We, thus, arrive at the standard results for neutral vector mesons.

5. INTERACTION AND ELIMINATION OF TIME-LIKE PHOTONS

If we now introduce the interaction, the equation of motion, according to (1.3), is replaced by

$$\partial_\nu[\partial_\nu A_\mu(x)-\partial_\mu A_\nu(x)]$$
$$-\mu^2\left(A_\mu(x)-\int D(x-y)\partial_\mu\partial_\nu A_\nu(y)d^4y\right)=-j_\mu(x). \tag{5.1}$$

By operating with ∂_μ on this equation it follows that

$$\partial_\mu j_\mu=0. \tag{5.2}$$

Using the suitably modified Lagrangian[14] of the form (3.1), the interaction representation can be set up in an arbitrary gauge parameterized by λ, in which the photon propagator will be (4.5), derived in the previous section.

Consider any matrix element involving a real time-like photon of four-momentum k_μ. This can be written[15]

$$M=\int\epsilon_\mu e^{ikx}\langle i|j_\mu(x)|f\rangle d^4x, \tag{5.3}$$

where $|i\rangle$ and $|f\rangle$ are the residual initial and final states after the timelike photon has been explicitly extracted. But, by (2.18), this implies

$$\mathrm{M}\sim\int k_\mu e^{ikx}\langle i|j_\mu(x)|f\rangle d^4x,$$

$$\sim\int e^{ikx}\langle i|\partial_\mu j_\mu(x)|f\rangle d^4x\equiv 0. \tag{5.4}$$

Thus, all matrix elements involving real time-like photons as external particles vanish identically owing to the conservation of current and our covariant definition of time-like polarization. This simple argument replaces the complicated cancellation between longitudinal and time-like photons in the Gupta-Bleuler formalism and avoids the awkward subsidiary condition of the usual Fermi formulation.[16]

Furthermore, the conservation of current also ensures that there is no contribution to any matrix element from the terms involving $k_\mu k_\nu$ in the propagator.[17] Thus, there is no dependence of any S-matrix element on λ, and the physical consequences of the theory are independent of the gauge, as they should be.

In order to obtain conventional theory we must take the limit

$$\mu\rightarrow 0.$$

For the propagator this limit has already been discussed in Sec. 4. We now consider its implication for the longitudinal photons, which have normalized polarization vector[18]

$$\epsilon_\mu{}^{(l)}=\frac{[n_\mu\mu+(n\cdot k)k_\mu/\mu]}{[k^2+(n\cdot k)^2]^{1/2}}, \tag{5.5}$$

where n_μ is a unit timelike vector. Just as in the case of time-like photons, there will be no contribution to matrix elements involving longitudinal photons from the term in $\epsilon_\mu{}^{(l)}$ proportional to k_μ. In the limit of $\mu\rightarrow 0$, the n_μ term also vanishes. Thus, only the two transverse photons survive in this limit in the physical matrix elements.

6. NONCOVARIANT THEORY

In the previous sections we have developed a guage-invariant theory based on the projection operator $\tau_{\mu\nu}$. This operator is covariant, but nonlocal. There exists

[14] Note that, as in conventional (massless) theory, current conservation is a consequence of the gauge invariant equation of motion (5.1). However, in a particular gauge, specified by λ, current conservation has to be imposed as an extra condition. This is again analogous to conventional thoery where, in the Fermi gauge, the equation of motion is $\partial^2A_\mu=-j_\mu$. In the present theory it follows from the modified Lagrangian and current conservation that $(\partial^2-\lambda^2)\partial_\mu A_\mu=0$, as in the noninteracting case [see Eq. (2.21)]. Thus, the time-like photons do not interact.

[15] H. Lehmann, K. Symanzik, and W. Zimmermann, Nuovo Cimento 1, 205 (1955).

[16] J. M. Jauch and F. Rohrlich, *Theory of Photons and Electrons* (Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, 1955).

[17] See, for example, N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantised Fields* (Interscience Publishers, Inc., New York, 1959).

[18] The projection operator for transverse photons is $t_{\mu\nu}$, defined in (6.1). The longitudinal photons have polarization $\epsilon_\mu{}^{(l)}$ which is orthogonal to this and to k_μ. The operator $a^{(l)}$, which annihilates a longidutinal photon is given by $a^{(l)}=\epsilon_\mu{}^{(l)}\tau_{\mu\nu}a_\nu$, since in the expansion for $A_\mu(x)$ one can write $\tau_{\mu\nu}a_\nu=t_{\mu\nu}a_\nu{}^{(t)}+\epsilon_\mu{}^{(l)}a^{(l)}$. It is easily checked that $[a^{(l)}(k),a^{(l)\dagger}(p)]\vartheta(k_0)\delta(k^2+\mu^2)=(2\pi)^3\times\delta^4(k-p)$.

an independent gauge-invariant projection operator, $t_{\mu\nu}$, which is local (in time) but not covariant.

If we lift the restriction that the theory has to be covariant, we may develop an alternative theory based on this operator. Thus, $t_{\mu\nu}$ is given by

$$t_{\mu\nu}=\delta_{\mu\nu}-\frac{1}{\partial^2+(n\cdot\partial)^2}[\partial_\mu\partial_\nu+(n_\mu\partial_\nu+n_\nu\partial_\mu)(n\cdot\partial)-n_\mu n_\nu\partial^2]. \quad (6.1)$$

This has the properties

$$n_\mu t_{\mu\nu}=0, \quad (6.2)$$

$$\partial_\mu t_{\mu\nu}=0, \quad (6.3)$$

and

$$t^n=t. \quad (6.4)$$

In the "natural" frame of reference in which the time axis is in the direction of n_μ, we have

$$n_\mu=(0,0,0,i), \quad (6.5)$$

and

$$t_{\mu\nu}=\left(\begin{array}{c|c} \delta_{ij}-\partial_i\partial_j/|\nabla|^2 & \begin{matrix}0\\0\\0\end{matrix} \\ \hline \begin{matrix}0 & 0 & 0\end{matrix} & 0\end{array}\right). \quad (6.6)$$

It is, thus, a projection operator, which selects the transverse part of a vector. We define

$$A_\mu{}^t(x)=t_{\mu\nu}A_\nu(x), \quad (6.7)$$

A gauge-invariant Lagrangian in terms of A^t is

$$L_0=-\tfrac{1}{2}\partial_\lambda A_\mu t_{\mu\nu}\partial_\lambda A_\nu-\tfrac{1}{2}\mu^2 A_\mu t_{\mu\nu}A_\nu. \quad (6.8)$$

Since in the natural frame this does not depend on A_4, it is clear that this is not an independent variable and should be eliminated from the interaction in a gauge-invariant manner. This requires

$$\partial_i(\partial_i A_4-\partial_4 A_i)=\nabla^2 F,$$

or

$$A_4=\frac{\partial_4\partial_i A_i}{\nabla^2}+F(j,tA), \quad (6.9)$$

where F is any gauge-invariant function.[19] Ignoring F for the moment, the interaction induced by (1.3) for spin-half particles is now

$$L_{\rm int}=A_i j_i+\frac{(\partial_4\partial_i A_i)}{\nabla^2}j_4. \quad (6.10)$$

The equation of motion derived from

$$L=L_0+L_{\rm int} \quad (6.11)$$

is

$$(\partial^2-\mu^2)t_{ij}A_j=j_i+\partial_i\partial_4 j_4/\nabla^2=t_{ij}j_j. \quad (6.12)$$

[19] For example, to include the Coulomb interaction for $\mu=0$, we must take $F=j_4/\nabla^2$; however, this is not required by gauge invariance.

To obtain the last equation we have used the conservation of current. These equations of motion can alternatively, be obtained from the Lagrangian

$$L=-\tfrac{1}{2}\partial_\lambda A_\mu t_{\mu\nu}\partial_\lambda A_\nu-\tfrac{1}{2}\mu^2 A_\mu t_{\mu\nu}A_\nu+j_\mu t_{\mu\nu}A_\nu, \quad (6.13)$$

which leads to more convenient canonical variables. Since $t_{\mu\nu}$ is a local operator as far as time derivatives are concerned, this can be made the basis of a conventional quantization. However, there are now only two independent fields, $A_\mu{}^t$.

The canonical momenta are

$$\Pi_\mu=-\partial_4 t_{\mu\nu}A_\nu, \quad (6.14)$$

and the commutation relations

$$[t_{\mu\nu}A_\nu(\mathbf{x},t),\,\Pi_\rho(\mathbf{y},t)]=-t_{\mu\rho}\delta(\mathbf{x}-\mathbf{y}). \quad (6.15)$$

The expansion of the field is

$$A_\mu(x)=\int\frac{d^4k}{(2\pi)^3}\vartheta(k_0)[t_{\mu\nu}(k)(c_\nu(k)e^{ikx}+c_\nu{}^\dagger e^{-ikx})\delta(k^2+\mu^2)+k_\mu(\alpha e^{ikx}+\alpha^\dagger e^{-ikx})+n_\mu(\beta e^{ikx}+\beta^\dagger e^{-ikx})], \quad (6.16)$$

where $t_{\mu\nu}(k)$ is given by

$$t_{\mu\nu}(k)e^{ikx}\equiv t_{\mu\nu}e^{ikx}, \quad (6.17)$$

and α and β are completely arbitrary, since they do not occur in $A_\mu{}^t$. The commutation relations for the operators c_μ and $c_\mu{}^\dagger$, which follow from (6.15), are

$$2\pi\vartheta(k_0)\delta(k^2+\mu^2)[c_\mu(k),\,c_\nu{}^\dagger(p)]=(2\pi)^4 t_{\mu\nu}(k)\delta^4(k-p). \quad (6.18)$$

The Hamiltonian of the system is

$$P_0=\int\frac{d^4k}{(2\pi)^3}\vartheta(k_0)k_0 c_\mu{}^\dagger(k)t_{\mu\nu}(k)c_\nu(k)\delta(k^2+\mu^2). \quad (6.19)$$

Note that only the transverse particles occur, so that the Hamiltonian is gauge invariant.

The vacuum expectation values of pairs of field operators at arbitrary space-time points can be calculated explicitly as before. In particular, the time-ordered product is (4.4), with D_F replaced by D_c where

$$D_c(k)=\frac{t_{\mu\nu}(k)}{k^2+\mu^2-i\epsilon}+Ak_\mu k_\nu+Bn_\mu n_\nu, \quad (6.20)$$

where A and B are the gauge-dependent terms which arise from the terms involving α and β in (6.16).

If we now consider calculations in the interaction representation of (6.13), the S matrix is given by

$$S=T\exp\left[-i\int j_\mu t_{\mu\nu}A_\nu d^4x\right]. \quad (6.21)$$

The photon propagator which arises in the calculation of matrix elements is

$$\langle T(t_{\mu\rho}A_\rho(x),t_{\nu\pi}A_\pi(y))\rangle_0=t_{\mu\rho}t_{\nu\pi}\langle T(A_\rho(x),A_\pi(y))\rangle_0. \quad (6.22)$$

(The t operators commute with T, since in the natural frame they do not involve time derivatives.) Using (6.20) and the properties of the t operator, Eqs. (6.2)–(6.4), the effective propagator is

$$D^t{}_{c\mu\nu}(k)=\frac{t_{\mu\nu}}{k^2+\mu^2-i\epsilon}$$
$$\simeq\left(\delta_{\mu\nu}-\frac{n_\mu n_\nu k^2}{k^2+(n\cdot k)^2}\right)\frac{1}{k^2+\mu^2-i\epsilon}. \quad (6.23)$$

In the final expression we have dropped from $t_{\mu\nu}$ the terms proportional to k_μ or k_ν, which give no contribution to S-matrix elements, again due to current conservation.

We have, thus, been lead to a rather surprising theory of "photons" of mass μ and spin one, of which only the two transverse polarization states occur either in the Hamiltonian or the S matrix. This theory is completely gauge invariant and is also invariant for spatial rotations. It is, of course, not covariant, since the exclusion of the longitudinal component is not a covariant concept (except for $\mu=0$).

The theory can be made covariant by adding to the interaction those terms, which in the S matrix will give rise to extra contributions of the form

$$(j_\mu n_\mu)(j_\nu n_\nu)\frac{k^2}{k^2+(nk)^2}\frac{1}{k^2+\mu^2-i\epsilon}, \quad (6.24)$$

so that the effective photon propagator is

$$D_{F\mu\nu}=\delta_{\mu\nu}/(k^2+\mu^2-i\epsilon). \quad (6.25)$$

In the special case of $\mu=0$, (6.24) is precisely the static Coulomb interaction, which in the natural frame is

$$j_4{}^2/|\mathbf{k}|^2. \quad (6.26)$$

[It corresponds to taking

$$F=j_4/\nabla^2 \quad (6.27)$$

in (6.9).] Since the new terms do not involve any time derivatives, they do not introduce any new particles into the theory. One, thus, arrives at the conventional zero-mass theory of two transverse photons in the Coulomb gauge, which is highly noncovariant in appearance, but of course, covariant in its conclusions.

If the mass is not taken to be zero, the fact that the additional terms are time dependent has the effect of introducing the longitudinal particles into the theory. The details of the rederivation of the covariant theory of the previous sections from this noncovariant starting point are complicated and of little interest.

7. CONCLUSION

We conclude that gauge invariance does not require that the bare photon mass be zero. The important feature of interactions generated by local gauge transformations, as outlined in Sec. 1, is that they lead to conserved currents. All other consequences of gauge invarinace, such as the restrictions on the form of matrix elements, follow from current conservation.

The results of the present paper cannot be generalized directly to particles arising from local isotopic spin and unitary symmetry transformations. Their significance seems to us to lie mainly in the light which they throw on the quantization of the electromagnetic field. So, far from gauge invariance requiring that the photon mass be zero, it appears that covariant canonical quantization in an arbitrary gauge is only possible for massive photons. The introduction of a photon mass is a familiar computational device for dealing with the infrared divergence problem. We have found that it also removes the necessity for the Fermi supplementary condition to define physical states, and avoids the complication of the Gupta-Bleuler formalism to eliminate the longitudinal and time-like photons.

In order to obtain a local covariant theory one is forced to use a Lagrangian and Hamiltonian which are not gauge invariant, but can be written in an arbitrary gauge specified by the parameter λ. A local gauge-invariant, but noncovariant Lagrangian and Hamiltonian can be written down by using a generalization of the Coulomb gauge. With this formulation we are led to a noncovariant theory of transverse massive photons. The theory is, however, covariant in the limit of zero photon mass. Thus, that the photon mass is zero is not a consequence of gauge invariance, but follows only from the very artificial condition that the noncovariant Coulomb gauge formulation, should lead to a covariant S matrix.

ACKNOWLEDGMENTS

The authors would like to thank Professor Abdus Salam, Dr. L. Brown, and Dr. T. W. B. Kibble for stimulating discussions.

Note added in proof.—The authors are indebted to Professor G. Källen who has been kind enough to suggest the inclusion of the following additional references: F. Coester, Phys. Rev. **83**, 798 (1951); R. J. Glauber, Progr. Theoret. Phys. (Kyoto) **9**, 295 (1953); W. Pauli (private discussion with G. Källen); F. J. Belinfante, Phys. Rev. **75**, 1321 (1949); G. Källen, in *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1960), Vol. 196.