PHYSICAL REVIEW D VOLUME 13, NUMBER 8 15 APRIL 1976

Canonical quantization of gauge theories

P. Hasenfratz and P. Hraskó

Central Research Institute for Physics, Budapest, Hungary

(Received 3 March 1975; revised manuscript received 14 July 1975)

A canonical quantization scheme in a Hilbert space with positive-definite metric is proposed for local gauge theories, based on Dirac's general theory of singular dynamical systems. A theorem on the subsidiary conditions, permitting a perturbation treatment, is stated and proved. Unitarity in the subspace of allowed states is demonstrated. The method is applied to the case of free electrodynamics in a nonlinear gauge and Yang-Mills theory in the covariant Feynman gauge. The description does not require introduction of ghost particles. The rules for calculating graphs are shown to be equivalent to those in a Lagrangian approach with a ghost.

I. INTRODUCTION

The quantization of field theories, invariant with respect to a group of non-Abelian local transformations, leads to certain difficulties. If one calculates scattering amplitudes using graphical rules analogous to those of quantum electrodynamics in a covariant gauge, then, as pointed out first by Feynman,[1] one comes into conflict with unitarity. The origin of this difficulty lies in the fact that in a covariant gauge the polarization states of the particles on the internal lines of a diagram are more numerous than on the cut lines appearing in the unitarity relations. Feynman[1] suggested that unitarity could be restored by inclusion of fictitious "ghost" particles into the theory. Following the suggestion of Feynman, several authors[2-5] have succeeded in constructing a manifestly covariant unitary quantum theory of fields possessing local gauge invariance. All these schemes are based on the Lagrangian rather than the Hamiltonian and we will call them the Lagrangian approach.

In classical theories with local gauge groups, constraints of the type $\phi=0$ always occurs. In order to guarantee the correspondence between the quantum and classical theories the physical state vectors $|\psi\rangle$ must satisfy the subsidiary condition $\phi|\psi\rangle=0$.

It is well known that covariant quantization requires the introduction of an indefinite metric. In this case the equation $\phi|\psi\rangle=0$ has to be replaced by the weakened form $\phi^{(+)}|\psi\rangle=0$, suggested by Gupta[6] and Bleuler.[7] This reduction is possible when ϕ satisfies the equation $\Box\phi=0$. In addition, the modified subsidiary condition $\phi^{(+)}|\psi\rangle=0$ ensures correspondence with the classical theory only if the equations of motion do not contain terms in which ϕ is multiplied with fields. Let us assume, for example, that the equations of motion contain a term $A\phi$, where A is some field. In the classical case this term has to be omitted because of the subsidiary condition $\phi=0$. Accordingly, in the quantum theory the expectation value $\langle\psi|A\phi|\psi\rangle$ must be equal to zero for any physical state. The original subsidiary condition $\phi|\psi\rangle=0$ ensures that this expectation value vanishes, but when it is reduced to the form $\phi^{(+)}|\psi\rangle=0$, the part $\langle\psi|A\phi^{(-)}|\psi\rangle$ does not vanish for an arbitrary A.

Since, in the cases to be considered below, the equations of motion generally do contain terms like $A\phi$, it seems impossible to reconcile covariant quantization with the validity of the canonical equations of motion for the expectation values. Even the fulfillment of the first condition $\Box\phi=0$ alone requires, as shown in Ref. 5, a nontrivial modification of the Lagrange-multiplier method developed in Refs. 8 and 9 for the quantization of the electromagnetic field in different covariant gauges.

In the present work we propose a canonical quantization scheme based on Dirac's general treatment[10] of singular dynamical systems. For the reasons explained above we follow Dirac in keeping subsidiary conditions in their original $\phi|\psi\rangle=0$ form. In this case we can work in a positive-definite-metric Hilbert space, since a noncovariant quantization procedure has to be employed. As a result, the propagator of the particles will contain noncovariant terms. It will be shown that the S-matrix elements between the mathematically complicated physical states are equal to the matrix elements between simple states not satisfying the subsidiary conditions, provided the propagator is supplemented by *additional* noncovariant terms. In this form of the canonically quantized theory we have the usual vector-meson vertices, the usual vector-meson propagators except for their $i\epsilon$ prescription, but we need not introduce ghosts. In the last section we prove the equivalence of this scheme with the Lagrangian theory involving ghosts. In other words, we shall show that the contribution of the ghost loops can be absorbed into vector-meson propagators with a special type of singularity. The proof of this statement in Sec. VI is independent of the quantization procedure which has led to this particular form of the vector-meson propagator.

We will consider free electrodynamics in a nonlinear gauge and Yang-Mills theory in the covariant Feynman gauge. In the last case the expressions ϕ contain the coupling constant g explicitly in both the Heisenberg and the interaction pictures. It would be an extremely difficult task to satisfy these subsidiary conditions exactly. A similar problem occurs in quantum electrodynamics as it was discussed and solved by Haller *et al.*[11] Following a similar line of reasoning it will be shown in Sec. III that the scattering states, computed to a given order in g, automatically satisfy the subsidiary conditions up to this order, provided the asymptotic states satisfy the subsidiary conditions with g set equal to zero. The limitations of this theorem are also discussed. The proof is based on the fact that in singular theories the Hamiltonian is a first-class quantity.[10]

Throughout the discussion we assume that infinities of the perturbation series can be appropriately handled by some gauge-invariant regularization procedure. Problems of renormalization will not be touched here.

II. DIRAC'S THEORY OF SINGULAR DYNAMICAL SYSTEMS

A dynamical system is called singular if the expression of the momenta $p_i = \partial L(q,\dot q)/\partial \dot q_i$ do not determine unambiguously the velocities $\dot q_i$ as functions of the momenta p_i. Dirac[10] worked out the general framework for the canonical quantization of singular systems. We give here a very brief summary of the method. For a detailed treatment we refer the reader to Dirac's book.

In a singular theory the expressions $\partial L(q,\dot q)/\partial \dot q_i$ and the coordinates satisfy a number of identities and, therefore, the momenta are subjected to the so-called primary constraints $\phi_1^l(q,p)=0$. The consistency of these primary constraints with the equations of motion usually leads to a number of secondary constraints $\phi_2^l(q,p)=0$.

The functions $F(q,p)$ of momenta and coordinates fall into two classes: The quantities in the first class are those whose Poisson bracket with any of the constraints is zero or equal to the linear combination of the constraints; if this is not the case, the quantity is called second class. In particular, the constraints themselves can be divided into constraints of the first class and those of the second class.

The Hamiltonian is determined up to a linear combination of the first-class constraints with coefficients which are arbitrary functions of the dynamical variables. Therefore, the time evolution of a dynamical quantity is determined only up to arbitrary functions. For those quantities, however, whose Poisson bracket with the first-class constraints is a linear combination of the constraints, these arbitrary functions are multiplied by constraints in the equation of motion and give no contribution, if the constraints are satisfied. These quantities are, therefore, the physical quantities of the theory. In particular, the Hamiltonian can be shown to be of first class, viz.,

$$[\phi^a, H] = r^{ab}\phi^b \tag{1}$$

and it is a physical quantity.

The indeterminateness, connected with the arbitrary functions, reflects the gauge freedom of the theory. Different choices of these functions correspond to different gauge conditions imposed on the generalized coordinates.

The change of the nonphysical quantities under the influence of the first-class constraints, as generators of canonical transformations, does not correspond to any change in the dynamical state of the system. Therefore, the first-class constraints generate symmetry transformations of the system in the sense, e.g., of the gauge transformations in electrodynamics.[12]

In quantizing the theory the first-class constraints are imposed on the state vectors as subsidiary conditions $\phi^a|\psi\rangle = 0$, while the second-class constraints are satisfied as operator identities through a suitable redefinition of the Poisson bracket. In the cases to be considered below no second-class constraints arise.

The method outlined above will be applied to special cases in the subsequent sections, and most of the statements of the general theory will be illustrated explicitly.

III. THE SUBSIDIARY CONDITION THEOREM

When the constraints are expressed through the canonical variables their form is the same in any picture. This is because the constraints never contain time derivatives of the canonical variables, as independent arguments, and therefore, the dependence on the time of the unitary transformation connecting the different pictures does not require special attention. Hence the subsidiary conditions $\phi|\psi\rangle = 0$ will depend explicitly on the coupling constant g in any picture, whenever the classical expression of ϕ contains explicit g dependence.

It would be extremely difficult to satisfy g-dependent subsidiary conditions even in the interaction picture, where the constraints depend on g through this explicit dependence alone. Fortunately it can be shown that the scattering state, calculated up to the order n, automatically satisfies the subsidiary conditions up to the same order, if the unperturbed state satisfies the subsidiary condition with $g=0$. This is the consequence of the relation (1)—valid in any picture—which expresses the fact that the constraints are related to the gauge symmetry of the theory. The precise statement can be formulated as follows.

Let the first-class constraints ϕ^a, the Hamiltonian H, and the coefficients r^{ab} in (1) be of the form

$$\begin{aligned}\phi^a &= \chi^a + g\,\xi^a\,,\\ H &= H_0 + gH_1, \\ r^{ab} &= r_0^{ab} + g r_1^{ab},\end{aligned} \tag{2}$$

where $\chi^a, \xi^a, H_0, H_1, r_0^{ab}, r_1^{ab}$ are independent of g.[13] Then the solution of the Lippmann-Schwinger equation

$$|\psi^+\rangle = |\varphi\rangle + g\,\frac{1}{E-H+i\epsilon} H_1|\varphi\rangle \tag{3}$$

satisfies the equation

$$\underline{\phi}|\psi^+\rangle = i\epsilon\,\frac{1}{E-\hat{H}-\hat{r}+i\epsilon}\, g\underline{\xi}|\varphi\rangle \tag{4}$$

provided the unperturbed state $|\varphi\rangle$ satisfies the lowest-order subsidiary conditions[14]

$$\chi^a|\varphi\rangle = 0. \tag{5}$$

Here we introduced the vector $\underline{\phi}$, whose components are the constraints ϕ^a, and the matrices $\hat{r}$ and $\hat{H}$ with the matrix elements r^{ab} and $H\delta^{ab}$, respectively. Equation (4) shows that $|\psi^+\rangle$ is a physical state in the limit $i\epsilon\to 0$, provided the real part of the energy denominator does not vanish. The states $|\varphi\rangle$, satisfying (5), will be called allowed states. They are not physical since $\phi^a|\varphi\rangle \neq 0$. The states orthogonal to the allowed states will be referred to as forbidden.

In the proof we follow the method of Haller *et al.*,[11] who investigated the problems of the subsidiary condition in QED in detail. The basis for the proof is Eq. (1), which can be written in the form

$$[\underline{\phi}, \hat{H}] = \hat{r}\underline{\phi}.$$

Comparing the coefficients of the different powers of g we obtain

$$[\underline{\chi}, \hat{H}_0] = \hat{r}_0\underline{\chi}, \tag{6}$$

$$[\underline{\chi}, \hat{H}_1] + [\underline{\xi}, \hat{H}_0] = \hat{r}_0\underline{\xi} + \hat{r}_1\underline{\chi}, \tag{7}$$

$$[\underline{\xi}, \hat{H}_1] = \hat{r}_1\underline{\xi}. \tag{8}$$

Using Eq. (3) we have

$$\underline{\phi}|\psi^+\rangle = g\underline{\xi}|\varphi\rangle + \underline{\phi}\,\frac{1}{E-\hat{H}+i\epsilon}\, g\hat{H}_1|\varphi\rangle.$$

Let us write (1) in the form

$$\underline{\phi}\hat{H} = (\hat{H}+\hat{r})\underline{\phi} \tag{9}$$

and we get

$$\begin{aligned}\underline{\phi}\,\frac{1}{E-\hat{H}+i\epsilon}\, g\hat{H}_1|\varphi\rangle &= \frac{1}{E-\hat{H}-\hat{r}+i\epsilon}\,\underline{\phi} g\hat{H}_1|\varphi\rangle \\ &= \frac{g}{E-\hat{H}-\hat{r}+i\epsilon}([\underline{\chi}, \hat{H}_1] + g\underline{\xi}\hat{H}_1)|\varphi\rangle.\end{aligned} \tag{10}$$

Using the relations (5), (7), and (8) we can write

$$\begin{aligned}\underline{\phi}|\psi^+\rangle &= g\underline{\xi}|\varphi\rangle \\ &\quad + \frac{g}{E-\hat{H}-\hat{r}+i\epsilon}([\hat{H}_0, \underline{\xi}] + \hat{r}_0\underline{\xi} + \hat{r}_1\underline{\chi} + g\underline{\xi}\hat{H}_1)|\varphi\rangle \\ &= g\underline{\xi}|\varphi\rangle \\ &\quad + \frac{g}{E-\hat{H}-\hat{r}+i\epsilon}(\hat{H}_0 - E + \hat{r}_0 + g\hat{H}_1 + g\hat{r}_1)\underline{\xi}|\varphi\rangle \\ &= i\epsilon\,\frac{1}{E-\hat{H}-\hat{r}+i\epsilon}\, g\underline{\xi}|\varphi\rangle,\end{aligned}$$

which proves our statement.

We have therefore shown that in order to have $\underline{\phi}|\psi^+\rangle = 0$ it is sufficient to satisfy the condition $\underline{\chi}|\varphi\rangle = 0$. This fails, however, if $|\psi^+\rangle$ goes into a matrix element, which is singular in the limit $i\epsilon\to 0$, as, e.g., in the case of the wave-function renormalization constant Z_2. When matrix elements of this type are computed starting from the allowed states instead of the physical ones, one gets results, depending on the gauge chosen. In quantum electrodynamics, for example, the allowed states satisfy the equation $\operatorname{div}\vec{E}|\varphi\rangle = 0$ while the physical states obey the subsidiary condition $(\operatorname{div}\vec{E} - e\rho)|\psi\rangle = 0$. When allowed states are taken for the unperturbed states, Z_2 turns out to be gauge dependent. Of course, the renormalized S-matrix elements are independent of the gauge chosen. In quantum electrodynamics one can build up the physical states $|\psi\rangle$ from the allowed states $|\varphi\rangle$ by an appropriate similarity transformation, and it can be shown that if the asymptotic states are physical, then the renormalization constants are gauge independent.[11] In the Yang-Mills theory such a transformation does not exist, because the commutation relations of the constraints ϕ^a differ from those of χ^a. The replacement of the physical states by the allowed ones leads to the gauge dependence of the renormalization constants also in this theory.

$\underline{\chi}|\varphi\rangle = 0$ is also a necessary condition; i.e., if $|\psi^+\rangle$ is the scattering state developed from the eigenstate $|\varphi\rangle$ of H_0, then from $\underline{\phi}|\psi^+\rangle = 0$ the relation $\underline{\chi}|\varphi\rangle = 0$ follows. To see this it is sufficient to expand $\underline{\phi}|\psi^+\rangle$ in powers of g. Since $\underline{\phi}|\psi^+\rangle = 0$, the coefficients of this power series vanish, and we get $\underline{\chi}|\varphi\rangle = 0$. Evidently the theorem also holds for the scattering state $|\psi^-\rangle$.

An important straightforward consequence of the theorem discussed above is that when $|\varphi\rangle$ is an allowed state then $S|\varphi\rangle$ is also allowed, i.e., the S-matrix elements between allowed and forbidden states vanish. In order to prove this we show that the two sets of states $\Omega_+|\varphi\rangle \equiv \mathcal{H}_+$ and $\Omega_-|\varphi\rangle \equiv \mathcal{H}_-$, which can be obtained from the allowed states, coincide. Suppose that this is not true. Then one can find a vector $|\overline{\psi}\rangle$ in $\mathcal{H}_+$ which is outside $\mathcal{H}_-$. But according to the subsidiary condition theorem for the states $|\psi\rangle$ in $\mathcal{H}_+$ or $\mathcal{H}_-$ we have $\underline{\phi}|\psi\rangle = 0$, while for those outside $\mathcal{H}_+$ or $\mathcal{H}_-$, $\underline{\phi}|\psi\rangle \neq 0$. Hence no vector $|\overline{\psi}\rangle$ can exist, and $\mathcal{H}_+ = \mathcal{H}_- \equiv \mathcal{H}$.

Let $\mathcal{H}'_\pm$ be the sets $\Omega_\pm|\varphi'\rangle$, where $|\varphi'\rangle$ are the forbidden states. If there are no bound states, $\Omega_\pm$ are unitary operators and the sets $\mathcal{H}'_\pm$ are orthogonal to $\mathcal{H}$. Hence $\mathcal{H}'_+ = \mathcal{H}'_- \equiv \mathcal{H}'$, and we have $(\mathcal{H}', \mathcal{H}) = 0$.

Consider the S-matrix element $S_{\beta\alpha} = \langle \psi_\beta^- | \psi_\alpha^+ \rangle$, where $|\psi_\alpha^+\rangle = \Omega_+|\varphi_\alpha\rangle$ and $|\psi_\beta^-\rangle = \Omega_-|\varphi'_\beta\rangle$. Since $|\psi_\alpha^+\rangle \in \mathcal{H}$ and $|\psi_\beta^-\rangle \in \mathcal{H}'$ we have $S_{\beta\alpha} = 0$, which is the statement we wanted to prove.

From the Hermiticity of the Hamiltonian it follows that S is unitary in the Hilbert space of all of the asymptotic states. As a consequence of the above considerations the S matrix is unitary also in the Hilbert space of the allowed states alone.

IV. CANONICAL QUANTIZATION OF THE FREE ELECTROMAGNETIC FIELD IN NONLINEAR GAUGE

We start from the gauge-invariant Lagrangian

$$L = -\frac{1}{4}\int F_{\mu\nu}F^{\mu\nu}\,d^3x, \tag{11}$$

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu}$$

of the free electromagnetic field. The canonical momenta are

$$B^\mu = \frac{\partial L}{\partial A_{\mu,0}} = F^{\mu 0}. \tag{12}$$

The nonzero Poisson bracket is defined by the relation

$$\{A_\mu(\vec{x}), B^\nu(\vec{x}')\} = \delta_\mu^\nu \delta^3(\vec{x} - \vec{x}').$$

Since $\partial L/\partial A_{0,0}$ is identically zero, the primary constraints of the theory are

$$\phi_1 \equiv B^0(\vec{x}) = 0, \tag{13}$$

i.e., one primary constraint at each point of space.

The quantity $\int B^\mu A_{\mu,0} d^3x - L$ can be expressed through the momenta, potentials, and their space derivatives without solving (12) for the velocities $A_{\mu,0}$, and in this way one can obtain one of the possible Hamiltonians

$$H = \int (\tfrac{1}{4}F_{rs}F^{rs} + \tfrac{1}{2}B^r B^r - A_0 B^r{}_{,r})\,d^3x.$$

The indices r, s, t take on the values 1, 2, 3.

Computing the time derivative of the primary constraints $B^0(\vec{x})$ with this Hamiltonian one finds that it is equal to $B^r{}_{,r}(\vec{x})$. Therefore, the consistency of the constraint (13) with the equation of motion leads to the secondary constraints

$$\phi_2 \equiv B^r{}_{,r}(\vec{x}) = 0. \tag{14}$$

The consistency requirement of (14) leads to no further secondary constraints. The Poisson brackets of the constraints are zero, so they are of first class.

In order to obtain the most general Hamiltonian H_E one has to add to H the constraints ϕ_1, ϕ_2 multiplied by arbitrary functions of the dynamical variables:

$$H_E = \int (\tfrac{1}{4}F_{rs}F^{rs} + \tfrac{1}{2}B^r B^r - A_0 B^r{}_{,r} + C_1\phi_1 + C_2\phi_2)\,d^3x. \tag{15}$$

This Hamiltonian leads to the equations of motion

$$\begin{aligned}
A_{0,0} &= C_1 + \frac{\partial C_1}{\partial B^0}\phi_1 + \frac{\partial C_2}{\partial B^0}\phi_2,\\
A_{s,0} &= B^s + A_{0,s} - C_{2,s} + \frac{\partial C_1}{\partial B^s}\phi_1 + \frac{\partial C_2}{\partial B^s}\phi_2,\\
B^0{}_{,0} &= B^r{}_{,r} - \frac{\partial C_1}{\partial A_0}\phi_1 - \frac{\partial C_2}{\partial A_0}\phi_2,\\
B^s{}_{,0} &= A_{s,rr} - A_{r,rs} - \frac{\partial C_1}{\partial A_s}\phi_1 - \frac{\partial C_2}{\partial A_s}\phi_2.
\end{aligned} \tag{16}$$

In the classical theory one can set in (16) $\phi_1 = \phi_2 = 0$. The arbitrary functions C_i, need not disappear from the equations, since the potentials are nonphysical quantities and cannot be defined unambiguously. However, for the physical quantities $\vec{E} = \vec{B}$ and $\vec{H} = \nabla \times \vec{A}$ Eqs. (14) and (16) lead to the Maxwell equations, independent of the functions C_i.

The arbitrary functions serve to fix the gauge. Let us choose them, for example, in the following way:

$$C_1 = A_{r,r} - \lambda A_\nu A^\nu - \tfrac{1}{2}B^0, \quad C_2 = 0. \tag{17}$$

Then the first of the Eqs. (16) leads to the expression

$$B^0 = -(\partial_\nu A^\nu + \lambda A_\nu A^\nu)$$

and (13) is equivalent to the nonlinear gauge condition

$$\partial_\nu A^\nu + \lambda A_\nu A^\nu = 0.$$

After having fixed the gauge one can go back from H_E to the Lagrangian, which is already nonsingular:

$$L = \int [-\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - \tfrac{1}{2}(\partial_\nu A^\nu + \lambda A_\nu A^\nu)^2]\, d^3x. \tag{18}$$

The theory corresponding to this Lagrangian has been discussed in great detail in Ref. 15 on the basis of graph combinatorics, since it serves as a good introduction to the non-Abelian gauge theories. For the same reason we also discuss the canonical quantization first in this case.

The nonzero commutation relations are

$$[A_\nu(x), B^\mu(x')]_{x^0=x^{0\prime}} = i\delta_\nu^\mu \delta^3(\vec{x}-\vec{x}')$$

and we get the following Heisenberg equations of motion in the gauge chosen:

$$\begin{aligned} A_{0,0} &= A_{r,r} - \lambda A_\nu A^\nu - B^0, \\ A_{r,0} &= B^r + A_{0,r}, \\ B^0{}_{,0} &= B^r{}_{,r} + 2\lambda A_0 B^0, \\ B^s{}_{,0} &= A_{s,rr} - A_{r,rs} + B^0{}_{,s} - 2\lambda A_s B^0. \end{aligned} \tag{19}$$

Using these equations, the commutation rule can also be written as

$$[A_\nu(x), A_{\mu,0}(x')]_{x^0=x^{0\prime}} = -ig_{\mu\nu}\delta^3(\vec{x}-\vec{x}'),$$

i.e., in the same form as in the Feynman gauge ($\lambda = 0$).

The subsidiary conditions, corresponding to the constraints, are

$$B^0|\Omega\rangle = 0; \quad B^r{}_{,r}|\Omega\rangle = 0.$$

These conditions do not contain the "coupling constant" λ explicitly; therefore, in the case under consideration the physical states and the allowed states are the same. The subsidiary conditions cannot be weakened to the form $B^{0(+)}|\Omega\rangle = B^r{}_{,r}^{(+)}|\Omega\rangle = 0$, since for $\lambda \neq 0$ B^0 does not satisfy the equation $\Box B^0 = 0$ and so covariant quantization with indefinite metric cannot be employed.

The operators a_μ, b^μ in the interaction picture satisfy (19) with λ set equal to zero:

$$\Box a_\mu(x) = 0;$$
$$\Box b^\mu(x) = 0.$$

The subsidiary conditions in the interaction picture are

$$b^0(x)|\omega(t)\rangle = 0; \quad b^r{}_{,r}(x)|\omega(t)\rangle = 0. \tag{20}$$

The equations of motion permit us to write these equations in the form

$$\partial^\mu a_\mu(x)|\omega(t)\rangle = 0; \quad [\partial^\mu a_\mu(x)]_{,0}|\omega(t)\rangle = 0. \tag{21}$$

The "interaction" Hamiltonian is the following:

$$H_1(x^0) = -\lambda \int a_\nu(x) a^\nu(x) b^0(x)\, d^3x. \tag{22}$$

Using the Fourier expansions

$$a_0(x) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \frac{1}{(2\omega)^{1/2}} [e^{-ikx} a_0^\dagger(\vec{k}) + e^{ikx} a_0(\vec{k})],$$

$$a_r(x) = \frac{1}{\sqrt{V}} \sum_{\vec{k}} \frac{1}{(2\omega)^{1/2}} [e^{-ikx} a_r(\vec{k}) + e^{ikx} a_r^\dagger(\vec{k})]$$

(where $k^0 = \omega = |\vec{k}|$), one gets for the commutation relations of the Fourier amplitudes the oscillator-type rules

$$[a_\mu(\vec{k}), a_\nu^\dagger(\vec{k}')] = \delta_{\mu\nu}\delta_{\vec{k}\vec{k}'},$$

which lead to positive metric for all four components.

In momentum space the subsidiary conditions (21) can be written as

$$[a_0^\dagger(\vec{k}) + c(\vec{k})]|\omega(t)\rangle = 0, \quad [a_0(\vec{k}) + c^\dagger(\vec{k})]|\omega(t)\rangle = 0, \tag{23}$$

where $c(\vec{k}) = (1/\omega)k^r a_r(\vec{k})$.

Let $|0)$ be the mathematical vacuum, i.e., the Fock state, defined by the equations $a_\mu(\vec{k})|0) = 0$. It is not the physical vacuum, since it does not satisfy (23). The physical vacuum is given by the expression[16]

$$\prod_{\vec{k}} e^{-a_0^\dagger(\vec{k})c^\dagger(\vec{k})}|0) \equiv |0\rangle.$$

The unperturbed Hamiltonian H_0 in the interaction picture can be obtained from (15) and (17) replacing A_ν, B^μ by a_ν, b^μ and taking the limit $\lambda = 0$:

$$H_0 = \int [\tfrac{1}{4} f_{rs} f^{rs} + \tfrac{1}{2} b^r b^r - a_0 b^r{}_{,r} + (a_{r,r} - \tfrac{1}{2} b^0) b^0] d^3x$$

$$= \sum_{\vec{k}} \omega [a_r^\dagger(\vec{k}) a_r(\vec{k}) - a_0^\dagger(\vec{k}) a_0(\vec{k})] + (c \text{ number})$$

$$= \sum_{\vec{k}} \omega \{ a_{T_1}^\dagger(\vec{k}) a_{T_1}(\vec{k}) + a_{T_2}^\dagger(\vec{k}) a_{T_2}(\vec{k}) + [c^\dagger(\vec{k}) c(\vec{k}) - a_0^\dagger(\vec{k}) a_0(\vec{k})] \} + (c \text{ number}).$$

The term

$$c^\dagger(\vec{k}) c(\vec{k}) - a_0^\dagger(\vec{k}) a_0(\vec{k}) = \tfrac{1}{2} \{ [c^\dagger(\vec{k}) - a_0(\vec{k})][c(\vec{k}) + a_0^\dagger(\vec{k})] + [c(\vec{k}) - a_0^\dagger(\vec{k})][c^\dagger(\vec{k}) + a_0(\vec{k})] \}$$

acting on physical states gives zero; therefore, the physical states obtained from $|0\rangle$ by transverse creation operators are those which contain a definite number of photons. However, the state $|0\rangle$ is not normalizable, since the operators $a_0^\dagger + c$ and $a_0 + c^\dagger$ have a continuous spectrum. Let us confine ourselves to a single momentum component and, following Ref. 17, define the state

$$|0_\gamma\rangle = \hat{O}_\gamma |0) = e^{-\gamma a_0^\dagger(\vec{k}) c^\dagger(\vec{k})} |0), \qquad (24)$$

which for $|\gamma| < 1$ has the norm

$$\langle 0_\gamma | 0_\gamma \rangle = \frac{1}{1-\gamma^2}$$

and for $\gamma \to 1$ satisfies (23). It is well known that matrix elements between states belonging to the same eigenvalue in a continuous spectrum are not meaningful quantities, and using them one may meet contradictions. For example, the matrix element of the commutator $[\partial^\mu a_\mu(x), a_\nu(y)]$ between physical states, satisfying (23), can be proved to be zero in spite of the fact that the commutator itself is a nonzero c number. Contradictions of this type can be circumvented, if the physical bra states are generated from the vacuum bra $\langle 0_\gamma|$ with $\gamma \neq 1$, and the limit $\gamma = 1$ is taken only in the final step. This limit, when it exists, may be considered as the correct matrix element between physical states.

The S matrix is given by the formal expression

$$S = T_D \exp[-i \textstyle\int dx_0 H_1(x^0)],$$

where T_D is the Dyson chronological ordering operator.[18] Applying the method described in Ref. 19, we pass from the T_D Dyson product to Wick's T_W product. After a straightforward calculation we get

$$S = T_D \exp[-i \textstyle\int dx_0 H_1(x^0)] = T_W \exp\left[-i \int \left(\lambda \partial_\nu a^\nu a_\mu a^\mu + \frac{\lambda^2}{2} a_\nu a^\nu a_\mu a^\mu\right) dx_0\right]. \qquad (25)$$

It is easy to write down the matrix elements of this S matrix between states $|0, \mathrm{tr})$, obtained from the mathematical (nonphysical) vacuum $|0)$ with the transverse creation operators. We have the usual propagator[20] and vector-meson vertices with three and four legs. There is no ghost contribution here and it is easy to show that these rules are in contradiction with unitarity.[15]

This is because the calculation has been based incorrectly on nonphysical states. Modifications equivalent to the ghost contribution arise when the matrix elements of (25) are taken between states, satisfying the subsidiary condition. These states are complicated, since they contain longitudinal and timelike photons. They can be written as

$$|0_\gamma, \mathrm{tr}\rangle = \hat{O}_\gamma |0, \mathrm{tr})$$

and the S-matrix elements are

$$\frac{1}{\langle 0_\gamma | 0_1 \rangle} \langle 0_\gamma, \mathrm{tr}' | S | 0_1, \mathrm{tr} \rangle = \frac{1}{\langle 0_\gamma | 0_1 \rangle} (0, \mathrm{tr}' | (\hat{O}_\gamma^\dagger \hat{O}_1) \bar{S} | 0, \mathrm{tr}), \qquad (26)$$

where $\bar{S} = \hat{O}_1^{-1} S \hat{O}_1$. Since $\hat{O}_1$ is independent of the space-time coordinates, this transformation can be accounted for by the replacement

$$a_\mu(x) \to \bar{a}_\mu(x) = \hat{O}_1^{-1} a_\mu(x) \hat{O}_1. \qquad (27)$$

In Sec. V it will be shown that the operator $(1/\langle 0_\gamma | 0_1 \rangle) \hat{O}_\gamma^\dagger \hat{O}_1$ can, in fact, be omitted from (26). *Therefore, it is allowable to use transverse Fock states built up on the mathematical vacuum* $|0)$, *provided the propagator is identified with the expectation value of the chronological product of the operators* $\bar{a}_\mu(x)$. Using (24) and (27) it is a straightforward task to verify that

$$\bar{a}_0(x) = a_0(x) - \frac{1}{\sqrt{V}} \sum_{\vec{k}} \frac{1}{(2\omega)^{1/2}} e^{ikx} c^\dagger(\vec{k}),$$

$$\bar{a}_r(x) = a_r(x) - \frac{1}{\sqrt{V}} \sum_{\vec{k}} \frac{1}{(2\omega)^{1/2}} e^{-ikx} \frac{k^r}{\omega} a_0^\dagger(\vec{k}).$$

From these expressions one obtains

$$(0| T_D[\bar{a}_\mu(x) \bar{a}_\nu(y)] |0) = \int d^4k\, e^{ik(x-y)} \frac{1}{(2\pi)^4 i} \left[\frac{g_{\mu\nu}}{k^2 + i\epsilon} + \frac{2\pi i \delta(k^2) \theta(k^0) \delta_{\mu 0} k_\nu}{k^0} + \frac{2\pi i \delta(k^2) \theta(-k^0) \delta_{\nu 0} k_\mu}{k^0}\right],$$
$$(0| T_D[\partial^\mu \bar{a}_\mu(x) \bar{a}_\nu(y)] |0) = \int d^4k\, e^{ik(x-y)} i k_\nu \frac{1}{(2\pi)^4 i} \frac{1}{k^2 - i\epsilon k^0}. \qquad (28)$$

The integrand in the last line contains the Fourier component of the retarded propagator, and the integral vanishes, unless $x^0 > y^0$. This is the consequence of the fact that from the subsidiary condition $\partial^\mu a_\mu | 0\rangle = 0$ the equation $\partial^\mu \bar{a}_\mu | 0) = 0$ follows.

Therefore, it is allowable to use simple transverse Fock states, provided we replace S by $\bar{S}$, where $\bar{S}$ is given by

$$\bar{S} = T_W \exp\left[-i \int \left(\lambda \partial^\nu \bar{a}_\nu \bar{a}_\mu \bar{a}^\mu + \frac{\lambda^2}{2} \bar{a}_\nu \bar{a}^\nu \bar{a}_\mu \bar{a}^\mu\right) dx_0\right]$$

We have the following Feynman rules for the photon propagators, the three-vertex, and the four-vertex, respectively:

$$\frac{1}{(2\pi)^4 i}\left[\frac{g_{\mu\nu}}{k^2 + i\epsilon} + \frac{2\pi i \delta(k^2)\theta(k^0)\delta_{\mu 0} k_\nu}{k^0} + \frac{2\pi i \delta(k^2)\theta(-k^0)\delta_{\nu 0} k_\mu}{k^0}\right],$$

$$2(2\pi)^4 \lambda (g_{\mu\sigma} q_\nu + g_{\mu\nu} p_\sigma + g_{\nu\sigma} k_\mu),$$

$$-4(2\pi)^4 i \lambda^2 (g_{\mu\nu} g_{\rho\sigma} + g_{\mu\rho} g_{\nu\sigma} + g_{\mu\sigma} g_{\nu\rho}),$$

and one should use the usual combinatorial factors. There is no ghost here. It will be shown, however, in Sec. VI that the theory is equivalent to the theory with casual propagators in Feynman gauge and usual ghost contribution. The unitarity follows from this equivalence or from the theorem of Sec. III.

V. CANONICAL QUANTIZATION OF THE YANG-MILLS FIELD IN COVARIANT FEYNMAN GAUGE

The Lagrangian of the Yang-Mills theory[21]

$$L = -\frac{1}{4} \int F^a_{\mu\nu} F^{a\mu\nu} d^3x,$$

$$F^a_{\mu\nu} = A^a_{\nu,\mu} - A^a_{\mu,\nu} + g f^{abc} A^b_\mu A^c_\nu$$

leads to the particular Hamiltonian

$$H = \int \left(\tfrac{1}{4} F^a_{rs} F^{ars} + \tfrac{1}{2} B^{ar} B^{ar} - A^a_0 \nabla^{ab}_r B^{br}\right) d^3x,$$

where

$$\nabla^{ab}_r = \delta^{ab} \partial_r + g f^{abc} A^c_r$$

means covariant differentiation, and the indices a, b, c label the generators of the underlying compact group.

The constraints of the theory are

$$\phi^a_1 \equiv B^{a0}(\vec{x}) = 0, \tag{29}$$

$$\phi^a_2 \equiv \nabla^{ab}_r B^{br}(\vec{x}) = 0. \tag{30}$$

The consistency requirement for ϕ^a_2 is fulfilled as a consequence of (30), since

$$\dot{\phi}^a_2 = i[H, \phi^a_2] = -g f^{abc} A^b_0 \phi^c_2$$

and no further constraints arise. The constraints are of first class:

$$\begin{aligned} &[\phi^a_1(x), \phi^b_1(y)]_{x^0 = y^0} = [\phi^a_1(x), \phi^b_2(y)]_{x^0 = y^0} = 0, \\ &[\phi^a_2(x), \phi^b_2(y)]_{x^0 = y^0} = g f^{abc} \phi^c_2(x) \delta^3(\vec{x} - \vec{y}). \end{aligned} \tag{31}$$

The most general Hamiltonian is therefore

$$H_E = \int d^3x \left(\tfrac{1}{4} F^a_{rs} F^{ars} + \tfrac{1}{2} B^{ar} B^{ar} - A^a_0 \nabla^{ab}_r B^{br} + C^a_1 \phi^a_1 + C^a_2 \phi^a_2\right).$$

The subsidiary conditions in the Heisenberg and interaction pictures are

$$B^{a0} | \Omega\rangle = 0, \tag{32}$$

$$\nabla^{ab}_r B^{br} | \Omega\rangle = 0, \tag{33}$$

$$b^{a0}(x) | \omega(t)\rangle = 0, \tag{34}$$

$$[b^{ar}{}_{,r}(x) + g f^{acb} a^c_r(x) b^{br}(x)] | \omega(t)\rangle = 0. \tag{35}$$

The subsidiary conditions contain the coupling constant g explicitly. In this case, according to the theorem of Sec. III the asymptotic states need not be physical. Instead, they have to be allowed, i.e., they must satisfy (34) and (35) with g set equal to zero. Let us choose, for example,

$$C^a_1 = A^a_{r,r} - \tfrac{1}{2} B^{a0}, \quad C^a_2 = 0.$$

From the equation of motion for A^a_0 it follows that in this case $B^{a0} = -\partial_\mu A^{a\mu}$ and (32) corresponds to the Feynman gauge.

The procedure of the previous section can now be immediately applied to the Yang-Mills theory. The only step we have to add is to prove that the operator $(1/\langle 0_\gamma | 0_1\rangle)\hat{O}^\dagger_0 \hat{O}_1$ can indeed be omitted from the right-hand side of (26).

Using (24) and computing the norm $\langle 0_\gamma | 0_1\rangle$ we have

$$\frac{1}{\langle 0_\gamma | 0_1 \rangle} (0, \mathrm{tr}' | \hat{O}_\gamma^\dagger \hat{O}_1$$

$$= (1-\gamma)(0, \mathrm{tr}' | e^{-\gamma a_0(\vec{k}) c(\vec{k})} e^{-a_0^\dagger(\vec{k}) c^\dagger(\vec{k})}.$$

After simple but lengthy algebraic manipulation this expression can be brought into the form

$$(0, \mathrm{tr}' | \sum_{k=0}^{\infty} a_0^k c^k \alpha_k(\gamma), \tag{36}$$

where

$$\alpha_0(\gamma) = 1,$$

$$\alpha_k(\gamma) = \frac{(-\gamma)^k}{(k!)^2}(1-\gamma)\frac{d^k}{d\gamma^k}\left(\frac{\gamma^k}{1-\gamma}\right); \quad k = 1, 2, \ldots .$$

In the limit $\gamma = 1$ all the coefficients α_k tend to infinity except α_0, which is unity. However, the matrix element (26) has to be calculated with the aid of (36) before taking the limit $\gamma = 1$, and it turns out that the terms with $k \neq 0$ give no contribution. This follows from the fact that the state $\overline{S}| 0, \mathrm{tr})$ which multiplies (36) is generated from the mathematical vacuum by means of transverse operators alone and hence it is orthogonal to all the terms except for the first one. To see this it is sufficient to notice that, according to the theorem of Sec. III $S| 0, \mathrm{tr}\rangle$ is an allowed state, which can be denoted by $| 0, \mathrm{tr}''\rangle$. But we have the formula

$$\hat{O}_1 \overline{S} | 0, \mathrm{tr}) = S | 0, \mathrm{tr}\rangle = | 0, \mathrm{tr}''\rangle$$

showing that

$$\overline{S} | 0, \mathrm{tr}) = \hat{O}_1^{-1} | 0, \mathrm{tr}''\rangle = | 0, \mathrm{tr}'')$$

as expected. Therefore, the terms $k > 0$ in (36) can be omitted and this means that $(1/\langle 0_\gamma | 0_1\rangle)\hat{O}_1^\dagger \hat{O}_1$ can also be omitted from (26).

The rules for calculating graphs can be written down by the same way as before. Our vector-meson vertices are the usual vertices of the Lagrangian approach, but we have no ghost here, while our vector-meson propagator is

$$\frac{1}{(2\pi)^4 i}\delta_{ab}\left(\frac{g_{\mu\nu}}{k^2 + i\epsilon} + \frac{2\pi i \delta(k^2)\theta(k^0)\delta_{\mu 0} k_\nu}{k^0} + \frac{2\pi i \delta(k^2)\theta(-k^0)\delta_{\nu 0} k_\mu}{k^0}\right).$$

IV. EQUIVALENCE WITH THE LAGRANGIAN APPROACH

In this section we show that S-matrix elements calculated according to the rules of the preceding section (which are the rules corresponding to the Hamiltonian approach) are, in fact, identical to those which can be obtained by using casual propagators in the Feynman gauge and ghosts (which are the rules in the usual Lagrangian approach). Given the two types of graphical rules, the proof of their equivalence is simple graph combinatorics.

This equivalence is a special case of the following more general statement.

Let us start from the Feynman rules of the usual Lagrangian theory in the Feynman gauge, where we have causal propagators and ghosts. It is possible to change slightly the $i\epsilon$ prescription of the longitudinal and timelike vector-meson propagator as well as that of the ghost propagator (it must be done simultaneously) so as to leave the S matrix unchanged. This variation in the $i\epsilon$ prescription can be done in a continuous fashion, and we have an infinite number of different sets of Feynman rules, all of which have the same S-matrix elements. One of these equivalent vector-meson propagators is the propagator (28) of the canonical theory. It has the important property that the ghost propagator associated with it is the purely retarded one. Loops built up from retarded propagators vanish, so the contribution of the ghost loops to the S-matrix ele-

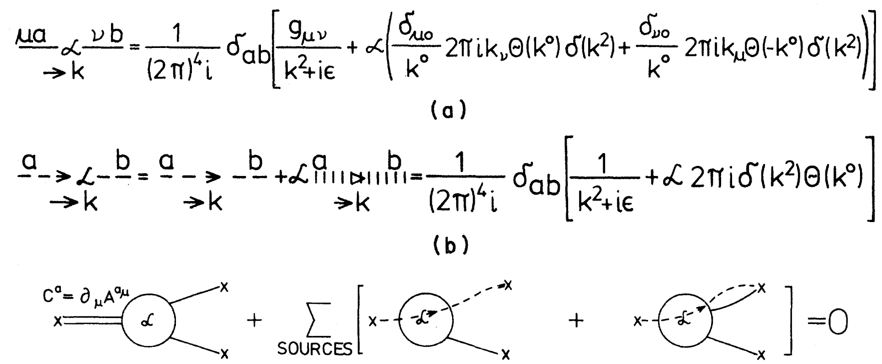

FIG. 1. (a) The $i\epsilon$ prescription of the vector-meson propagator. (b) The corresponding ghost propagator. (c) The Slavnov-Taylor identities.

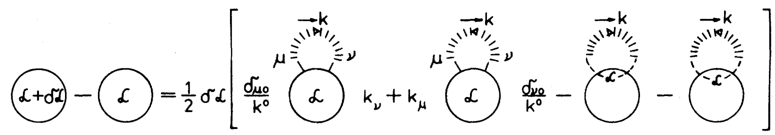

FIG. 2. The lowest-order change in the S_α-matrix element due to changing α by an infinitesimal amount.

ments vanishes when (28) is used. That is the reason that Feynman rules of the preceding section define a unitary theory in spite of the absence of ghosts. So the point is that the contribution of the ghost loops to the S-matrix elements depends on the $i\epsilon$ prescription chosen, and the canonically quantized theory automatically picks up the special type of vector-meson propagator, where the associated ghost contribution is zero.

The proof of this statement is based on the Slavnov-Taylor identities (STI) which are the generalizations of the Ward identities of electrodynamics to a theory with a non-Abelian gauge group. In order to fix the propagators in the Lagrangian approach one has to supplement the invariant Lagrangian with terms, breaking local gauge invariance, and a ghost Lagrangian L_g has to be included in order to satisfy STI. The rules for determining L_g from the symmetry-breaking part of L are given in Ref. 22. According to these rules the propagator of the ghost is that of a zero-mass scalar particle. The $i\epsilon$ prescription, which cannot be chosen arbitrarily, may be determined from STI in the lowest order.[15,22] In our gauge the latter has the form $k_\mu D^{\mu\nu}(k) = -(\hat{m}^{-1})k^\nu$ where $D_{\mu\nu}$ and $(-\hat{m}^{-1})$ are the propagators of the vector particle and the ghost, respectively. If the $i\epsilon$ prescription of the vector-meson propagator is given as in Fig. 1(a), then the corresponding ghost propagator is as shown in Fig. 1(b) and STI are of the form[23] shown in Fig. 1(c). Here α is an arbitrary continuous parameter. For $\alpha = 0$ we have the causal propagators, while for $\alpha = 1$ the vector-meson propagator is that given in (28) and the ghost propagator is purely retarded.

We now show that the S_α-matrix elements constructed from the propagators in Figs. 1(a) and 1(b) and the vertices of $(L + L_g)$ are independent of the parameter α.

Let us consider the sum of all the graphs with some given external physical particles and change α to $\alpha + \delta\alpha$ by an infinitesimal amount. We can compute the lowest-order change of the S_α-matrix element by replacing in succession each propagator by its first-order variation in every graph in all possible ways. The result can be written down in the form of the equation in Fig. 2. The blobs on the left-handed side represent matrix elements in a given order n and for the given set of the external physical particles whose lines have been omitted for simplicity. On the right-hand side the blobs are other matrix elements which have—beside the external particles already present—an additional ingoing and outgoing line, both having the same momentum. These matrix elements are multiplied by definite factors, coming from the propagators in Figs. 1(a) and 1(b) and containing $\delta(k^2)$.

Hence the two new external particles are also on the mass shell. The first and the second blobs on the right-hand side are, therefore, nth-order amplitudes to which STI may be applied. They are in a form contracted with k, and STI shows that the first and the third as well as the second and the fourth terms give zero in the sum.

There is a point which we should mention here. Our asymptotic states are allowed (and not physical) states so the renormalization constants are, in general, gauge dependent, as was discussed in the third section. So the external sources are, in fact, α (gauge) dependent,[15,22] which dependence was omitted above. On the other hand, using the STI in Fig. 2 those terms were left out, where the ghost line ended in an external source, for instance, as shown in Fig. 3(a). In general this graph does not contribute to the S matrix because the absence of a pole in P^2, except in the cases

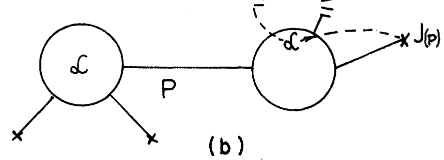

FIG. 3. (a) Class of diagrams not contributing to the S_α-matrix element. (b) Subset of graphs of (a) which is gauge dependent but just cancels the gauge dependence of the external sources.

shown in Fig. 3(b). This contribution is also gauge dependent, and it turns out[15,22] that it just cancels the gauge dependence of the external sources. As the proof of this statement is identical to that given in Refs. 15 and 22, we do not repeat it here.

This proves the independence of S_α of the parameter α and from this the equivalence of the canonical theory and the Lagrangian approach follows.

ACKNOWLEDGMENT

We are deeply indebted to Dr. A. Frenkel for valuable discussions.

[1]R. P. Feynman, Acta Phys. Pol. 24, 697 (1963).
[2]B. S. DeWitt, Phys. Rev. 162, 1195 (1967); 162, 1239 (1967).
[3]L. D. Faddeev and V. N. Popov, Phys. Lett. 25B, 30 (1967).
[4]S. Mandelstam, Phys. Rev. 175, 1580 (1968); 175, 1604 (1968).
[5]E. S. Fradkin and I. V. Tyutin, Phys. Rev. D 2, 2841 (1970).
[6]S. Gupta, Proc. R. Soc. London A63, 681 (1950).
[7]K. Bleuer, Helv. Phys. Acta 23, 567 (1950).
[8]B. Lautrup, K. Dan. Vidensk. Selsk. Mat.–Fys. Medd., 35, No. 11 (1967).
[9]N. Nakanishi, Prog. Theor. Phys. 51, 1 (1972).
[10]P. A. M. Dirac, *Lectures on Quantum Mechanics* (Academic, New York, 1965).
[11]K. Haller and L. F. Landovitz, Phys. Rev. D 2, 1498 (1970); K. Haller, *ibid.* 8, 1796 (1973); S. P. Tomczak and K. Haller, Nuovo Cimento 33, 1 (1972).
[12]The problem of finding the most general gauge transformations has recently been considered by A. Frenkel, Orsay Report No. 74-20 (unpublished).
[13]This assumption of linearity in g, which is satisfied, e.g., in fermion-photon interaction, simplifies the proof, but the theorem can be shown to hold also in the more general case.
[14]The common spectra of H_0 and H is assumed. $|\psi^+\rangle$ is the state vector at $t=0$, when the different pictures coincide.
[15]G. 't. Hooft and M. Veltman CERN Report No. 73-9 (unpublished).
[16]J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons* (Addison-Wesley, Reading, Mass., 1955).
[17]R. Utiyama, T. Imamura, S. Sunakawa, and T. Dodo, Prog. Theor. Phys. 6, 587 (1951).
[18]The Dyson chronological operator T_D is defined by the relation $T_D[A(t)B(t')] = \theta(t-t')A(t)B(t') + \theta(t'-t)B(t')A(t)$ and does not commute with differentiation with respect to time. In a Lagrangian theory Wick's chronological product is used. Acting on a product of field operators and their derivatives the operator T_W is defined to commute with differentiation.
[19]E. S. Fradkin, *Problems of Theoretical Physics* (Nauka, Moscow, 1969).
[20]In this case the propagator to be used is

$$(0|T_D[a_\mu(x)a_\nu(y)]|0) = \int d^4k\, e^{ik(x-y)}\frac{1}{(2\pi)^4 i}\left[\frac{g_{\mu\nu}}{k^2+i\epsilon}+2\pi i\,\delta(k^2)\delta_{\mu 0}\delta_{\nu 0}\right].$$

It contains a noncovariant part, which is the consequence of the noncovariant quantization procedure. We note that this is, in fact, the propagator which can be obtained via the path integral method, when the $i\epsilon$ prescription, cutting off the integrand at large values of field, is consistently handled. This shows that the path integral formulation corresponds to the quantization in a Hilbert space with positive-definite metric.
[21]C. N. Yang and R. L. Mills, Phys. Rev. 96, 191 (1954).
[22]G. 't Hooft and M. Veltman, Nucl. Phys. B50, 318 (1972).
[23]We mention that the dimensional regularization method developed by G. 't Hooft and M. Veltman [Nucl. Phys. B44, 189 (1972)] works for any arbitrary value of α.