Nonlocal regularizations of gauge theories

D. Evens and J. W. Moffat
Department of Physics, University of Toronto, Toronto, Ontario, Canada M5S 1A7

G. Kleppe and R. P. Woodard
Department of Physics, University of Florida, Gainesville, Florida 32611
(Received 20 June 1990)

A procedure is given for generalizing local, gauge-invariant field theories to nonlocal ones which are finite, Poincaré invariant, and perturbatively unitary. These theories are endowed with nonlocal gauge symmetries which ensure current conservation and decoupling in the same way that their local analogs do in the parent theories. An elegant way of viewing the resulting on-shell symmetry transformations is as "quantum representations" of the local gauge group in which the representation matrices become field-dependent, nonlocal operators. By varying the scale of nonlocality one can obtain gauge-invariant regularization schemes which are manifestly Poincaré invariant, perturbatively unitary, and free of automatic subtractions. Since our method does not entail changing either the particle content or the dimension of spacetime, it may preserve global supersymmetry. As applications we work out the electron self-energy and vacuum polarization in QED at one loop. The latter gives the surprising result that no Landau ghost occurs with the regulator on and before renormalization. Another surprise is the absence of an axial-vector anomaly.

I. INTRODUCTION

Quantum field theorists have had to put up with ultraviolet divergences for a very long time now, so long, in fact, that remembrance has faded of the desperate expedients which were explored prior to the apotheosis of renormalization. In this context the putative finiteness of superstring theory came as a pleasant surprise and is still cited as one of the theory's chief virtues. It is now possible to understand this finiteness in a way which would have been very familiar to a physicist of the late 1940s and early 1950s:[1] The vertices of string field theory contain nonlocal factors of $\exp(-\alpha' p^2)$ which cause loops to converge in Euclidean space.[2]

The phenomenon is not restricted to strings; attaching such factors to the interactions of *any* otherwise local Lagrangian gives an ultraviolet-finite theory. Consider the following simple scalar model:

$$\mathcal{L}=-\tfrac{1}{2}\partial_\mu\phi\partial^\mu\phi-\tfrac{1}{2}m^2\phi^2-\tfrac{1}{6}g(\phi^\Lambda)^3 , \tag{1.1a}$$

where $\phi^\Lambda\equiv\mathcal{E}_m\phi$ and the nonlocal smearing operator $\mathcal{E}_m$ is defined, for any mass m, as follows:

$$\mathcal{E}_m\equiv\exp\left[\frac{\partial^2-m^2}{2\Lambda^2}\right] . \tag{1.1b}$$

The resulting perturbative S matrix is finite, unitary, and Lorentz invariant, just like that of string theory. The price for these benefits is the same too: off-shell noncausality at the perturbative level as well as instability and a breakdown of the initial-value problem beyond perturbation theory.[3] These problems are fatal to (1.1) as any sort of fundamental theory, but they pose no obstacle to regarding it as a perturbative *regularization* of the local action which results from taking Λ to infinity. Indeed, a simple variant of this method was successfully used by Polchinski in the context of ϕ^4 theory.[4]

This scheme, which we shall term "nonlocal regularization," has several advantages over conventional methods. First, in the simple scalar model considered above, it preserves global Poincaré invariance without changing the dimension of spacetime. This suggests the possibility of generalizations which preserve global supersymmetry, the absence of which is the only point needed for a proof of the finiteness of $N=4$ supersymmetric Yang-Mills theory.[5] Second, it lacks the notorious "automatic subtractions" present in both dimensional regularization[6] and the ζ-function method.[7] Third, it does not sacrifice perturbative unitarity as does the Pauli-Villars method.[8] Finally, this method proceeds from modifying the action, not simply changing the rules for computing certain inner products. It can therefore be used at the operator level in canonical problems.[9] As we shall see, nonlocal regularization is operationally very similar to Schwinger's proper time method,[10] but proceeds from a systematic, field-theoretic formulation.

The fact that nonlocalization could cure ultraviolet divergences was realized long ago.[1] Two things inhibited its early application, either to produce candidate fundamental theories (since the problems of Ref. 3 were unsuspected until recently) or as a regularization scheme. First, there was the problem of how to canonically formulate nonlocal actions. This affects the operator formalism in an obvious way through its dependence upon a Hamiltonian and commutation relations; the problem shows up in the functional formalism through the measure factor needed to ensure unitarity. Although procedures valid to low orders were known for certain theories, it was long believed that insuperable obstacles must appear at two loops.[11] Of course, this is false in view of the existence and perturbative unitarity of the string S matrix, and a general procedure has recently been given for the canonical formulation of any perturbatively localizable action.[3,12]

The second major obstruction was the obvious physical relevance of gauge theories and the belief that gauge invariance is inconsistent with any useful sort of nonlocalization. The problem can easily be seen in the context of QED:

$$\mathcal{L}=-\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}-\bar{\psi}(i\mathcal{D}\!\!\!/+m)\psi \ , \tag{1.2}$$

where $\mathcal{D}_\mu$ is the covariant derivative operator $\partial_\mu+ieA_\mu$ and our γ matrices obey $\{\gamma_\nu,\gamma_\nu\}=-2\eta_{\mu\nu}$. Suppose we imitate the previous scalar model by simply nonlocalizing the interaction

$$e\bar{\psi}A\!\!\!/\,\psi\rightarrow e\bar{\psi}^{\Lambda}A\!\!\!/^{\Lambda}\psi^{\Lambda} \ , \tag{1.3}$$

where $\bar{\psi}^{\Lambda}\equiv\mathcal{E}_m\bar{\psi}$, $\psi^{\Lambda}\equiv\mathcal{E}_m\psi$, and $A^{\Lambda}\equiv\mathcal{E}_0A$. Although free of ultraviolet divergences, the resulting theory is not gauge invariant because we have broken up the covariant derivative and used ordinary derivatives in smearing charged fields. It is worth emphasizing that losing the mathematical abstraction referred to as "gauge invariance" is unacceptable *physically,* because it means giving up either unitarity (if we quantize in a covariant gauge) or Lorentz invariance (if we quantize in a physical gauge). If we attempt to avoid these problems by covariantly nonlocalizing the entire covariant-derivative term,

$$-i\bar{\psi}\mathcal{D}\!\!\!/\psi\rightarrow-i\left[\exp\left[\frac{\mathcal{D}^2-m^2}{2\Lambda^2}\right]\bar{\psi}\right](\partial\!\!\!/+ieA\!\!\!/^{\Lambda})\times\left[\exp\left[\frac{\mathcal{D}^2-m^2}{2\Lambda^2}\right]\psi\right] \ , \tag{1.4}$$

then the resulting theory is invariant, but not completely finite. This is because the electron propagator carriers a factor of $1/\mathcal{E}_m^2$, which cancels the convergence factors on the vertices; hence pure fermion loop divergences are unsuppressed.

That this problem cannot apply to all nonlocal gauge theories follows from the existence of string theory. The key to understanding its success lies in extending the notion of "gauge invariance" to include nonlocal transformation laws. The raison d'être of gauge symmetry in quantum field theory is the decoupling of unphysical vector and tensor quanta while maintaining Lorentz invariance. Any symmetry which accomplishes this task is acceptable; it is not necessary that the transformation rule should be local. In fact, invariant string field theory possesses a nonlocal gauge invariance.[13,14] The transformation rule consists of a local inhomogeneous term, which preserves the (local) quadratic part of the action, and a nonlocal homogeneous term, which engenders a variation of the free action that cancels the inhomogeneous variation of the nonlocal interaction.

The burden of this paper is to show that the ability of nonlocal actions to support such symmetries is not limited to strings. In fact, *any* local gauge theory can be generalized to a finite, nonlocal theory endowed with a nonlocal gauge symmetry that reconciles unitarity and Poincaré invariance. For example, it is simple to check that the nonlocalized version of QED obtained by the replacement (1.3) is invariant at order e under the transformation

$$\delta A_\mu=-\partial_\mu\theta \ , \tag{1.5a}$$

$$\delta\psi=ie\,\mathcal{E}_m\theta^{\Lambda}\psi^{\Lambda} \ , \tag{1.5b}$$

where $\theta^{\Lambda}\equiv\mathcal{E}_0\theta$ and the explicit operator $\mathcal{E}_m$ in (1.5b) is understood to act on everything to its right, as opposed to the implicit operators in θ^{Λ} and ψ^{Λ}. Unfortunately, invariance is lost at order e^2; nor can the transformation rule be modified so as to recover it. That is an inconvenience, but nothing more; it means we must add higher-order terms, both to the action and to the symmetry as it turns out.[15] We describe a general method for doing this to QED in Sec. II. Although the modified symmetry is fixed by the higher interactions, the latter are only partially determined by the requirement of decoupling. Arbitrary choices remain to be made at each other. Section III presents a particular solution for which we compute the electron self-energy and the vacuum polarization at one loop. Section IV discusses how to extend the method beyond QED to non-Abelian gauge theories and to gravity. Our conclusions comprise Sec. V.

II. GENERALITIES FOR QED

We begin with an overview of the method. The first step is to introduce nonlocal convergence factors onto the interaction term in the manner of (1.3) so as to make Euclidean loop integrals finite. This destroys local gauge invariance, but since current conservation at order e depends only upon the (unchanged) free theory, there must be an associated symmetry at this order. One finds it by nonlocalizing the homogeneous part of the transformation law along the lines of (1.5). At order e^2 the theory's "invariance" is violated in a physically meaningful way by the breakdown of current conservation and the loss of decoupling. We postpone these disasters to order e^3 by adding a new interaction of the form $\mathcal{L}_2\sim e^2\bar{\psi}A^2\psi$. The resulting action will still not possess invariance under the old symmetry, but it must be invariant under *some* symmetry, at order e^2, since it has the physical attributes of current conservation and decoupling. We find the desired symmetry by adding a term $\delta_2\psi\sim ie^2\theta A\psi$ to the transformation law. We then restore current conservation and decoupling at order e^2 by adding a new interaction of the form $\mathcal{L}_3\sim e^3\bar{\psi}A^3\psi$, and so on.

We have found it convenient to divide this section's discussion into three parts. Section II A reviews the connection between the physically essential feature of decoupling and the progressively more off-shell abstractions of current conservation and some form of gauge invariance. Section II B explains the determination of the new interactions, which restore decoupling at each order in the classical theory. Section II C treats quantization.

A. Gauge invariance, current conservation, and decoupling

We first argue that gauge invariance implies current conservation and decoupling, in that order, and then reverse the sequence to show that current conservation implies gauge invariance. The subsection closes with a discussion of the extent to which decoupling implies the other conditions. The first two arguments are given in terms of an action with the form

$$S=\int d^4x\left[-\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}-\bar{\psi}(i\not{\partial}+m)\psi\right]-\int d^4x\, d^4y\,\bar{\psi}(x)V[eA](x,y)\psi(y)\ , \qquad (2.1)$$

where the vertex function is a spinorial matrix and can be expanded in a power series $V\sim eA+(eA)^2+\cdots$. The possibility of pure photon and multifermion interactions is ignored because they can never be used to restore decoupling. Why this is so will be explained in Sec. II B.

Suppose the action is invariant under a transformation of the form

$$\delta A_\mu(x)=-\partial_\mu\theta(x)\ , \qquad (2.2a)$$

$$\delta\psi(x)=ie\int d^4y\, d^4z\ \mathcal{T}[eA](x,y,z)\theta(y)\psi(z)\ , \qquad (2.2b)$$

where the "representation operator" $\mathcal{T}\sim 1+eA+\cdots$ is a spinorial matrix as well as a functional of the vector potential. From $\delta S=0$ we extract the condition

$$\frac{i}{e}\frac{\partial}{\partial y^\mu}\frac{\delta V[eA](x,z)}{\delta A_\mu(y)}=-(i\not{\partial}_x+m)\mathcal{T}(x,y,z)+\bar{\mathcal{T}}(z,y,x)(-i\overleftarrow{\not{\partial}}_z+m)+\int d^4u\left[V(x,u)\mathcal{T}(u,y,z)+\bar{\mathcal{T}}(u,y,x)V(u,z)\right]\ . \qquad (2.3)$$

To see current conservation first note that the equations of motion are

$$\partial_\nu F^{\nu\mu}(y)=\int d^4x\, d^4z\ \bar{\psi}(x)\frac{\delta V[eA](x,z)}{\delta A_\mu(y)}\psi(z)\ , \qquad (2.4a)$$

$$\Psi[eA,\psi](x)\equiv(i\not{\partial}+m)\psi(x)+\int d^4z\ V[eA](x,z)\psi(z)=0\ . \qquad (2.4b)$$

Taking the divergence of (2.4a) and substituting (2.3) gives

$$i\partial_\mu\partial_\nu F^{\nu\mu}=\int d^4x\, d^4z\ i\frac{\partial}{\partial y^\mu}\frac{\delta V[eA](x,z)}{\delta A_\mu(y)}\psi(z)=e\int d^4x\, d^4z[\bar{\psi}(x)\bar{\mathcal{T}}(z,y,x)\Psi(z)-\bar{\Psi}(x)\mathcal{T}(x,y,z)\psi(z)]\ , \qquad (2.5)$$

which vanishes by (2.4b) and its Dirac adjoint. Note that a failure of current conservation really means that (2.4a) cannot generally be solved for A_μ when the Fermi fields obey (2.4b). Conversely, establishing current conservation proves that a general perturbative solution exists.

"Decoupling" has two meanings. Both follow from gauge invariance and both arise in the context of covariant quantization on a large Fock space which includes massless vector particles of arbitrary polarization $\epsilon_\mu(\mathbf{p})$. The first meaning is that covariant gauge-fixing terms exist such that the resulting on-shell S matrices vanish whenever the polarization vector of even one external vector is longitudinal, $\epsilon_\mu(\mathbf{p})=-ip_\mu\theta(\mathbf{p})$. The relevant gauge-fixing terms are those which possess a residual invariance under transformation by harmonic gauge parameters. The second meaning concerns perturbative unitarity. Note that on the large Fock space this follows trivially from the Cutkosky rules, which are true for nonlocal theories just as they are for local ones, as long as the interactions are both analytic and Hermitian. One says that the negative norm, temporal polarizations have "decoupled," if Lorentz invariance and perturbative unitarity are maintained within a physical space which excludes them and any one of the spacelike polarizations.

This second usage follows simply by exploiting gauge invariance to impose a physical gauge. Since the vector transformation law is unchanged from that of the local theory, the usual local gauge conditions are attainable. The only nontrivial point is whether or not the resulting constraint equation can be solved on shell. The previous argument about current conservation guarantees that it can be, at least perturbatively. Because the substitution of this solution preserves the Hermiticity and analyticity of the action, perturbative unitarity follows on the physical space, if it held on the large one. Lorentz invariance is a consequence of having gauge fixed a manifestly invariant theory.

The second definition of decoupling also turns out to follow from the first. To show this we shall follow the treatment of Mandelstam.[16] Suppose we define the large space theory by quantizing in a gauge which has residual symmetry under harmonic gauge transformations. It then decouples in the first sense. Now let $|s\rangle$ be an arbitrary longitudinal photon with polarization $\epsilon^s_\mu(\mathbf{p})=-ip_\mu\theta(\mathbf{p})$. We define its "conjugate" state $|\bar{s}\rangle$ by reversing the temporal components of the polarization vector, $\epsilon^{\bar{s}}_\mu(\mathbf{p})=-i(-\|\mathbf{p}\|,\mathbf{p})\theta(\mathbf{p})$. Clearly, adding these two gives a state of positive norm,

$$|s_+\rangle\equiv|s\rangle+|\bar{s}\rangle\ , \qquad (2.6a)$$

while their difference gives a ghost state whose norm has the same magnitude:

$$|s_-\rangle\equiv|s\rangle-|\bar{s}\rangle\ . \qquad (2.6b)$$

In its first sense decoupling means

$$\langle f|T|s\rangle=0=\langle s|T^\dagger|g\rangle\ , \qquad (2.7)$$

where T is the transition operator, and $|f\rangle$ and $|g\rangle$ are any two states, physical or unphysical. Substitution of (2.6) into (2.7) gives

$$\langle f|T|s_+\rangle=-\langle f|T|s_-\rangle\ , \qquad (2.8a)$$

$$\langle s_+|T^\dagger|g\rangle=-\langle s_-|T^\dagger|g\rangle\ , \qquad (2.8b)$$

from which simple multiplication allows us to write

$$\langle f|T|s_+\rangle\langle s_+|T^\dagger|g\rangle=\langle f|T|s_-\rangle\langle s_-|T^\dagger|g\rangle\ . \qquad (2.9)$$

But unitarity in the large space implies

$$i\langle f|T^\dagger - T|g\rangle = \sum_{\text{phys}} \frac{1}{\langle \text{phys}|\text{phys}\rangle}\langle f|T|\text{phys}\rangle\langle \text{phys}|T^\dagger|g\rangle + \sum_{s_+}\frac{1}{\langle s_+|s_+\rangle}(\langle f|T|s_+\rangle\langle s_+|T^\dagger|g\rangle - \langle f|T|s_-\rangle\langle s_-|T^\dagger|g\rangle) \tag{2.10a}$$

$$= \sum_{\text{phys}} \frac{1}{\langle \text{phys}|\text{phys}\rangle}\langle f|T|\text{phys}\rangle\langle \text{phys}|T^\dagger|g\rangle , \tag{2.10b}$$

where $|\text{phys}\rangle$ is a general physical state. The "physical states" of this argument are of course those of the Coulomb gauge. Any other physical gauge could be reached by varying the definition of the "conjugate" polarization $\epsilon^{\bar{s}}$. For example, had we taken $\epsilon^{\bar{s}}_\mu(\mathbf{p}) = -i(-\|\mathbf{p}\|, -p_1, -p_2, p_3)\theta(\mathbf{p})$, the physical polarizations would have been those of the axial gauge.

To see that the first sense of decoupling follows from gauge invariance, we begin by noting that gauge-fixing terms obviously exist which are invariant under harmonic gauge transformations. The Feynman gauge is an example. That it can be added in the usual way follows from the fact that the vector transformation (2.2a) is the same as for the local theory. Even had there been higher-order terms, we could still have attained this gauge perturbatively.

It is now useful to make a rather extensive digression recalling the formalism of DeWitt,[17] whereby one obtains a generating functional for the tree S matrix in terms of the action evaluated for a general classical scattering solution (for a review with generalizes the method to all orders in the loop expansion, see Sec. 4 of Ref. 18). Suppose that $S[\phi]$ is the classical action of a scalar field ϕ and that the free theory is Klein-Gordon with mass m. The associated Wronskian

$$\overleftrightarrow{W}^\mu \equiv \overrightarrow{\partial}^\mu - \overleftarrow{\partial}^\mu \tag{2.11}$$

induces two useful inner products

$$[f\circ g](t) \equiv -i\int_{t=\text{const}} d^3x\, f(t,\mathbf{x})\overleftrightarrow{W}^0 g(t,\mathbf{x}) , \tag{2.12a}$$

$$f\bullet g \equiv -i\oint_{\partial R^4} d^3\sigma_\mu(x) f(x)\overleftrightarrow{W}^\mu g(x) . \tag{2.12b}$$

Assuming the field obeys the usual sort of asymptotic condition, then scattering states are annihilated in the weak sense by the limiting operators

$$a^{(\text{out}/\text{in})} \sim \lim_{t\to\pm\infty} f\circ\phi , \tag{2.13}$$

where f is any positive-energy solution to the Klein-Gordon equation. Hence the S matrix is

$$S_{12\cdots\to\cdots N} \equiv \langle\Omega|a_N^{\text{out}}\cdots(a_1^{\text{in}})^\dagger|\Omega\rangle \tag{2.14a}$$

$$= \lim_{t_1\to-\infty}\cdots\lim_{t_N\to\infty}\langle\Omega|[f_N\circ\phi](t_N)\cdots[f_1^*\circ\phi]|\Omega\rangle \tag{2.14b}$$

$$= \langle\Omega|T([f_1^*\bullet\phi]\cdots[f_n\bullet\phi])|\Omega\rangle + (\text{forward-scattering terms}) . \tag{2.14c}$$

The next step in Lehmann-Symanzik-Zimmermann (LSZ) reduction would be to rewrite the surface integrals of (2.14c) as volume ones. However, DeWitt instead expresses the result in terms of $W[J]$, the generating functional for connected Green's functions:

$$S_{12\cdots\to\cdots N} = \left[f_1^*\bullet\frac{\delta}{\delta J}\right]\cdots\left[f_n\bullet\frac{\delta}{\delta J}\right]\exp(W[J])\bigg|_{J=0} . \tag{2.15}$$

From (2.12b) it is apparent that we really only need to know W for asymptotic currents of the form

$$J_\infty[\xi,\xi^*](x) \equiv -\oint_{\partial R^4} d^3\sigma_\mu(y)\phi_1[\xi,\xi^*](y)\overleftrightarrow{W}^\mu\delta^4(y-x) , \tag{2.16a}$$

$$\phi_1[\xi,\xi^*](x) \equiv \int\frac{d^3k}{(2\pi)^3}[\xi(\mathbf{k})f(x;\mathbf{k}) + \xi(\mathbf{k})^* f^*(x;\mathbf{k})] , \tag{2.16b}$$

where the $f(x;\mathbf{k})$'s are momentum eigenfunctions with δ-function normalization under (2.12a). With this surface current, the S matrix can be written so that one immediately recognizes the exponential of $iW[J_\infty]$ as its generating functional:

$$S_{12\cdots\to\cdots N} = \frac{\delta}{\delta\xi_1^*}\cdots\frac{\delta}{\delta\xi_N}\exp(iW[J_\infty])\bigg|_{\xi=\xi^*=0} . \tag{2.17}$$

The final step is to reexpress W in terms of the classical action

$$W[J_\infty] = S[\phi_\infty] + \int d^4x\, J_\infty(x)\phi_\infty(x) , \tag{2.18a}$$

where the "scattering background" ϕ_∞ obeys the classical equation with J_∞ as its source:

$$\frac{\delta S[\phi_\infty]}{\delta\phi(x)}=-J_\infty(x)\ . \qquad (2.18b)$$

We have been a little casual about surface terms for the sake of brevity, but it is easy to make amends now. The sense of (2.18b) is that one solves the sourceless classical equation in the weak-field expansion

$$\phi_\infty=\phi_1+\phi_2+\cdots\ , \qquad (2.19)$$

where ϕ_1 is the linearized scattering field (2.16b) and all the higher terms are functionals of it obtained by inverting the Klein-Gordon operator with the Feynman propagator. Note that from the S matrix one can extract not only $W[J_\infty]$, through relation (2.17), but also the on-shell action $S[\phi_\infty]$ as a functional of $\xi(\mathbf{k})$ and $\xi^*(\mathbf{k})$. The key relation is

$$S[\phi_\infty]=W[J_\infty]-\int\frac{d^3k}{(2\pi)^3}\left[\xi(\mathbf{k})\frac{\delta W[J_\infty]}{\delta\xi(\mathbf{k})}+\xi^*(\mathbf{k})\frac{\delta W[J_\infty]}{\delta\xi^*(\mathbf{k})}\right]\ , \qquad (2.20)$$

which follows from (2.16) and (2.18).

That was rather a lot of notation, but it makes the actual argument trivial. One can clearly carry out DeWitt's construction for our nonlocalized version of QED (2.1). The general asymptotic data for a scalar particle are provided by the function $\xi(\mathbf{p})$. The asymptotic datum for the vector particle is the polarization vector $\epsilon_\mu(\mathbf{p})$; for the electron and positron it is the spinor wave functions $u_i(\mathbf{p})$ and $v_i(\mathbf{p})$. In DeWitt's language the first sense of decoupling would be equivalent to asserting the invariance of the on-shell action under the transformation

$$\epsilon_\mu(\mathbf{p})\rightarrow\epsilon_\mu(\mathbf{p})-ip_\mu\theta(\mathbf{p})\ . \qquad (2.21)$$

To show this, note that since replacement (2.21) just transforms the asymptotic data, it does not prevent the general scattering solutions A_∞ and ψ_∞ from obeying field equations (2.4). Since these equations are invariant under (2.2) and since the replacement engenders a linearized transformation of this form upon the linearized solutions, it is clear that perturbation theory just builds up the full transformation with gauge parameter θ on the full solutions. But the gauge-fixed action is invariant under such transformations, and so the on-shell action is indeed unchanged.

This completes the line of reasoning that follows from gauge invariance, and we now assume that the current in (2.4a) is conserved when the Fermi fields obey (2.4b) and its Dirac adjoint. Whenever a quantity vanishes with the field equations, we know that it must be proportional to them; hence we have the off-shell condition

$$\int d^4x\, d^4z\, \bar\psi(x)i\frac{\partial}{\partial y^\mu}\frac{\delta V[eA](x,z)}{\delta A_\mu(y)}\psi(z)=e\int d^4x\, d^4z[\bar\psi(x)\bar{\mathcal{W}}(z,y,x)\Psi(x)-\bar\Psi(x)\mathcal{W}(x,y,z)\psi(z)]\ , \qquad (2.22)$$

for some bosonic functional $\mathcal{W}[eA,\bar\psi,\psi]$. But this implies the action's invariance under the transformation

$$\delta A_\mu(x)=-\partial_\mu\theta(x)\ , \qquad (2.23a)$$

$$\delta\psi(x)=ie\int d^4y\, d^4z\, \mathcal{W}(x,y,z)\theta(y)\psi(z)\ . \qquad (2.23b)$$

So far we have allowed for the possibility that $\mathcal{W}$ might have to depend upon the Fermi fields, but once the action's invariance under (2.23) is conceded, it is apparent that any such dependence would only induce higher fermion terms. Since these can never cancel variations in the $\bar\psi V[eA]\psi$ term, the action must still be invariant if any fermion dependence in $\mathcal{W}$ is dropped. But this gives the transformation law the same form as (2.2), and so we have decoupling as well by the previous argument.

Now let us assume decoupling, in the first of the above senses, and explore the extent to which it implies current conservation and gauge invariance. One cannot hope to show that decoupling implies gauge invariance, because if an invariant action manifests decoupling, then so too will any covariantly gauge-fixed version of it. Similarly, one cannot settle upon a particular form such as (2.1) because decoupling alone is not sufficient to rule out multiphoton amplitudes or contributions to multifermion amplitudes from terms without an intermediate exchange pole. Even without such amplitudes, field redefinitions can induce multiphoton and higher fermion interactions without changing the S matrix. What we will instead show is that any S matrix which decouples is derivable from a gauge-invariant action.

First, find a manifestly Poincaré-invariant action which gives the desired S matrix and large space particle spectrum. There are many ways of doing this, and it suffices to choose any one of them. Now truncate this action by imposing the completely *ad hoc* condition $\partial^\mu A_\mu=0$. From Mandelstam's argument it is apparent that the S matrix of this truncated action is just the original S matrix restricted to the covariant subspace defined by the condition $\epsilon_0(\mathbf{p})=\mathbf{p}\cdot\boldsymbol{\epsilon}(\mathbf{p})/\|\mathbf{p}\|$. It is therefore unchanged by replacement (2.21), which carries polarizations within the truncated Fock space to other such polarizations.

Now apply the analysis of DeWitt to express the Lorentz gauge scattering solutions A_∞ and ψ_∞ as functionals of the Lorentz gauge asymptotic data. As in the preceding discussion, we see that replacement (2.21) engenders a transformation on A_∞ and ψ_∞. This transformation is not necessarily of the form (2.2), and so let us write it as

$$\Delta A_{\mu\infty}(x)=\int d^4y\, B_\mu[eA_\infty,\bar\psi_\infty,\psi_\infty](x,y)\theta_h(y)\ , \qquad (2.24a)$$

$$\Delta\psi_\infty(x)=ie\int d^4y\, C[eA_\infty,\bar\psi_\infty,\psi_\infty](x,y)\theta_h(y)\ , \qquad (2.24b)$$

where the gauge parameter $\theta_h(x)$ is understood to be harmonic. There are of course many ways to use the equations of motion and $\partial^2\theta_h=0$ to reexpress the functionals B_μ and C, and it suffices to choose any convenient form.

Decoupling means that the on-shell action is invariant under this transformation; hence any off-shell noninvariance must be proportional to the field equations

$$\Delta S = \int d^4x\, d^4y\, \theta_h(y) \left\{ \frac{\delta S}{\delta A_\mu(x)} B'_\mu[eA,\bar{\psi},\psi](x,y) + \frac{\delta S}{\delta\psi(x)} C'[eA,\bar{\psi},\psi](x,y) - \bar{C}'[eA,\bar{\psi},\psi](x,y)\frac{\delta S}{\delta\bar{\psi}(x)} \right\} . \tag{2.25}$$

Actually, the functions B'_μ and C' are a little ambiguous because the invariance of the on-shell field equations implies that ΔS vanishes *quadratically* with the field equations. Hence B'_μ and C' are themselves proportional to the field equations and are ambiguous up to mixed derivatives, and again it suffices to make any convenient choice.

It follows from (2.25) that the off-shell action is invariant under the new transformation

$$\Delta' A_\mu(x) = \int d^4y (B_\mu(x,y) - B'_\mu(x,y))\theta_h(y) , \tag{2.26a}$$

$$\Delta'\psi(x) = ie \int d^4y (C(x,y) - C'(x,y))\theta_h(y) . \tag{2.26b}$$

Note that the gauge parameter θ_h is still assumed to be harmonic even though the fields no longer obey their equations of motion. Note also that since B'_μ and C' vanish with the equations of motion, the transformation of the on-shell fields is unaffected.

We are now free to regard S as an invariant action in Lorentz gauge. The associated residual symmetry is of course (2.26). We could extend it in many ways by exploiting $\partial^2\theta_h = 0$ to rewrite (2.26) before removing the harmonic condition. The effect of such an extended gauge transformation will be to disturb the Lorentz gauge condition. Starting from an arbitrary vector potential, we can obviously solve perturbatively for the parameter $\theta[A,\bar{\psi},\psi]$, which imposes Lorentz gauge. This will of course by ambiguous up to a harmonic term, but we have just shown that the Lorentz gauge action is unaffected by transformation with a harmonic parameter. Therefore, we obtain a unique, gauge, and Lorentz-invariant action by subjecting the Lorentz gauge action to a transformation with this parameter $\theta[A,\bar{\psi},\psi]$. This is the invariant action we have been seeking. A review of this technique with worked out examples can be found in Ref. 19. Note that harmonic ambiguities in $\theta[A,\bar{\psi},\psi]$ are what prevent us from preserving both perturbative unitarity and Lorentz invariance by transforming an arbitrary Lorentz gauge action in this way.

We shall use this result over and over again in the next subsection. Because the construction can proceed from any gauge-fixed action which gives the S matrix, one can, if the amplitudes are suitable, make field redefinitions to enforce part or all of form (2.1) for the invariant action and (2.2) for the transformation. For example, if there are no pure photon amplitudes, then we can find a starting action which is free of interactions involving only the vector potential. Similarly, it might be possible to obtain a starting action which is free of interactions involving more than two Fermi fields. If both conditions can be met, then it will be possible to make the various choices called for by the construction so that the transformation has form (2.2) and the invariant action has the form (2.1).

B. Choosing the classical higher interactions

What we *want* is a perturbatively viable nonlocalization of QED. For any such theory to posses a smooth local limit, this implies some form of gauge invariance; however, we have not found it fruitful to search for the symmetry directly. Instead, we iterate higher interactions which enforce the physical attribute of a gauge-invariant theory, namely, decoupling, and then infer the symmetry which the previous discussion has shown must be present.

The method is most easily explained as we implement it. The starting point is the nonlocalized version of QED which results from replacement (1.3). As noted previously, the order e loss of gauge invariance signals no physical problem. Since current conservation continues to hold at this order, one can extend the transformation law as in (1.5) to find a symmetry under which the theory is invariant, at order e. The physical problem is the loss of decoupling which occurs at order e^2. This shows up in the Compton tree amplitude [Fig. 1(a)]:

$$\mathcal{C}[\bar{u}_0,\epsilon_1,\epsilon_2,u_3] = ie^2\bar{u}_0\not{\epsilon}_1 \frac{\exp[(s-m^2)/\Lambda^2]}{\not{p}_0+\not{p}_1+m} \times \not{\epsilon}_2 u_{-3} + ie^2\bar{u}_0\not{\epsilon}_2 \times \frac{\exp[(t-m^2)/\Lambda^2]}{\not{p}_0+\not{p}_2+m}\not{\epsilon}_1 u_{-3} , \tag{2.27}$$

where one should understand all the momenta to be positive outward and the Mandelstam parameters are $s \equiv -(p_0+p_1)^2$ and $t \equiv -(p_0+p_2)^2$. Note that any nonlocal operators residing on external legs have degenerated to unity owing to the mass shell conditions. Suppose we let the first photon be longitudinal, $\epsilon_1 = p_1$. The usual arguments give

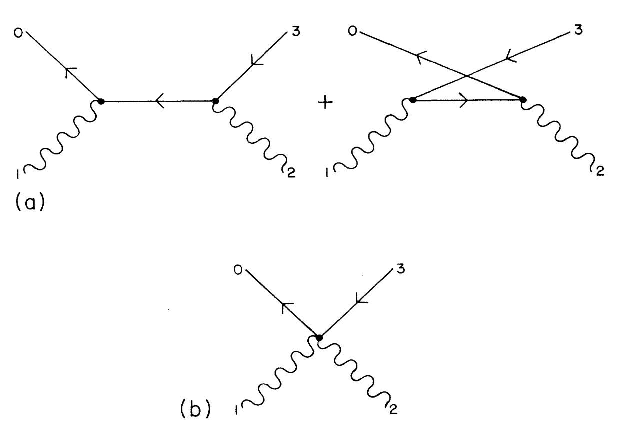

FIG. 1. (a) Compton scattering: s- and t-channel contributions. (b) Compton scattering: new interaction contribution.

$$\mathcal{C}[\bar{u}_0,p_1,\epsilon_2,u_3]=ie^2\bar{u}_0\not{\epsilon}_2u_{-3}\times\left[\exp\left[\frac{s-m^2}{\Lambda^2}\right]-\exp\left[\frac{t-m^2}{\Lambda^2}\right]\right]. \tag{2.28}$$

Hence longitudinal photons couple to physical particles.

Our procedure is to cancel this failure of decoupling with an explicit four-point coupling. Obviously, it must be Poincaré invariant, and to be a good regularization we will require that it vanishes as $\Lambda\to\infty$, and that all the fields have exponential convergence factors. The interaction should also be Hermitian and an entire function of the various momenta. The reason for this is that the absorptive parts of amplitudes must not receive contributions from anything other than internal lines going on shell.[20]

To find a suitable four-point interaction, let us work backwards from the longitudinal photon coupling we wish to cancel. If $f(s,t)$ is a symmetric function of s and t, one can write

$$-ie^2\bar{u}_0\not{\epsilon}_2u_{-3}\left[\exp\left[\frac{s-m^2}{\Lambda^2}\right]-\exp\left[\frac{t-m^2}{\Lambda^2}\right]\right]$$

$$=-ie^2\bar{u}_0\not{\epsilon}_2u_{-3}\left[\exp\left[\frac{s-m^2}{\Lambda^2}\right]-f(s,t)-\exp\left[\frac{t-m^2}{\Lambda^2}\right]+f(t,s)\right] \tag{2.29a}$$

$$=-ie^2\bar{u}_0(\not{p}_0+\not{p}_1+m)(\not{p}_0+\not{p}_1-m)\not{\epsilon}_2u_{-3}\left[\frac{\exp[(s-m^2)/\Lambda^2]-f(s,t)}{s-m^2}\right]$$
$$+ie^2\bar{u}_0\not{\epsilon}_2(\not{p}_0+\not{p}_2+m)(\not{p}_0+\not{p}_2-m)u_{-3}\left[\frac{\exp[(t-m^2)/\Lambda^2]-f(t,s)}{t-m^2}\right] \tag{2.29b}$$

$$=-ie^2\bar{u}_0\not{\epsilon}_1(\not{p}_0+\not{p}_1-m)\not{\epsilon}_2u_{-3}\left[\frac{\exp[(s-m^2)/\Lambda^2]-f(s,t)}{s-m^2}\right]$$
$$-ie^2\bar{u}_0\not{\epsilon}_2(\not{p}_0+\not{p}_2-m)\not{\epsilon}_1u_{-3}\left[\frac{\exp[(t-m^2)/\Lambda^2]-f(t,s)}{t-m^2}\right]. \tag{2.29c}$$

The last expression would come from the following (momentum-space) interaction [Fig. 1(b)]:

$$-e^2(2\pi)^4\delta^4(p_0+p_1+p_2+p_3)\hat{\bar{\psi}}_3\hat{\not{A}}_2(\not{p}_0+\not{p}_1-m)\hat{\not{A}}_1\hat{\psi}_0\left[\frac{\exp[(s-m^2)/\Lambda^2]-f(s,t)}{s-m^2}\right]. \tag{2.30}$$

We have ensured exponential convergence in Euclidean momentum space by smearing each field with the appropriate factor to $\mathcal{E}_M$, which is unity in the amplitude owing to the on-shell conditions. To meet the requirements of Hermiticity, analyticity, and vanishing in the infinite Λ limit, the function f must have the form

$$f(s,t)=1+\left[\frac{s-m^2}{\Lambda^2}\right]\left[\frac{t-m^2}{\Lambda^2}\right]e\left[\frac{s-m^2}{\Lambda^2},\frac{t-m^2}{\Lambda^2}\right], \tag{2.31}$$

where $e(x,y)$ is a symmetric function of x and y which is entire in both variables and real when they are.

Before revising the transformation law, we should comment on two points. The first is that because there is no problem with Bhabha scattering, no four-fermion interaction is required. The second point concerns why we cannot enforce decoupling by introducing a three-photon interaction. Anyone who has studied local Yang-Mills coupled to fermions will be struck by the similarity between our violation of decoupling (2.28) and the imperfect cancellation which occurs between the analogous s- and t-channel graphs [Fig. 2(a)] for the process $q\bar{q}\gamma\gamma$ when one of the gluons becomes longitudinal. The problem in that case is the ordering of the representation matrices; its resolution is the u-channel graph [Fig. 2(b)] in which the two external gluons interact via a three-gluon vertex before intersecting the quark line. Why cannot such a graph restore decoupling in our theory? The

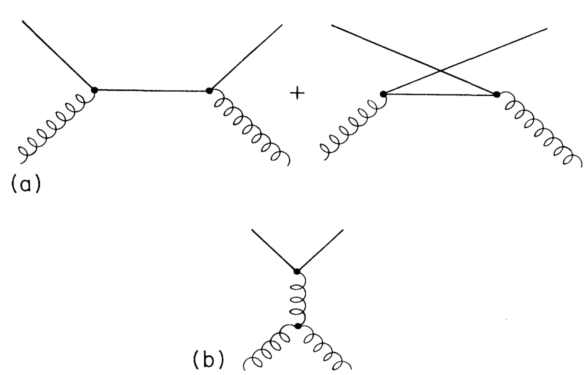

FIG. 2. (a) Non-Abelian Compton scattering: s and t channels. (b) Non-Abelian Compton scattering: u channel.

answer is that a u-channel graph could only contribute a function of u to the term in square brackets in (2.28). Even using the on-shell condition $s+t+u=2m^2$ cannot result in cancellation, and so a three-photon vertex is useless. This feature signals the important distinction between a nonlocal deformation of the Abelian *algebra* and the effect we shall really discover, which is induced field dependence in the *representations* of our nonlocal, on-shell U(1).

Because the modified theory decouples at order e^2, it must be invariant under transformation (1.5) on shell. That is, any off-shell noninvariance at order e^2 must be proportional to the field equations. In fact, since we are only interested in order e^2 effects at this stage, we know that any noninvariance resides in terms proportional to the *free* field equations. But such terms can always be absorbed into order e^2 variations of the free action under a modified transformation. Because it is really not necessary that the photons be on shell in order to prove decoupling, only the fermion transformation law requires modification. The new rule is of the form (2.2b) with representation operator $\mathcal{T}$ chosen as follows:

$$ie\mathcal{E}_m[\delta^4(x-y)]^\Lambda \left[1-e(i\partial\!\!\!/-m)\frac{\mathcal{E}_m^2-f(\partial^2,(\overleftarrow{\partial}_y+\partial_z)^2)}{+\partial^2-m^2}\boldsymbol{A}^\Lambda(x)\right]\mathcal{E}_m\delta^4(x-z)\,, \tag{2.32}$$

where f is any function of the form (2.31). In this expression all the derivatives, including those in the operators $\mathcal{E}_m$, act upon the coordinate x unless they bear subscripts y or z.

It is simple to check that these transformations form an Abelian group on shell and at order e^2. Hence our nonlocalization has not changed the group structure on shell; rather it has modified the representations. The local, field-independent representation operator of the unregulated theory,

$$\mathcal{T}[eA](x,y,z)=ie\delta^4(x-y)\delta^4(x-z)\,, \tag{2.33}$$

has been distorted into the nonlocal, field-dependent form (2.32). One might refer to these on shell as "quantum representations" of U(1). Note also that our nonlocalization results in a peculiar mixing between spinor indices, spacetime coordinates, and the local group indices—here the phase of the Fermi fields. This is very reminiscent of the concatenation of group indicates and spacetime coordinates which occurs in the gauge symmetries of invariant string field theory.[13,14]

The rest is more of the same. At each order it suffices to enforce decoupling on what might be termed the "extended Compton trees," that is tree amplitudes with two external fermions and N external photons. This is done by means of an interaction of the form $\bar{\psi}(eA)^N\psi$, which is manifestly Poincaré invariant, Hermitian, analytic, exponentially suppressed for Euclidean momenta, and zero in the infinite Λ limit. This last property is never any problem because the decoupling failures we seek to cancel also vanish with infinite Λ. The penultimate condition is similarly trivial; we can always enforce it by decorating each field with the appropriate factor of $\mathcal{E}_M$. These factors are unity for the Nth extended Compton tree, and so they cannot affect our ability to find an interaction. If the necessary interaction can be found at all, it will be ambiguous, even as was (2.30), up to terms of the form $\bar{\psi}(eF_{\mu\nu})^N\psi$, which decouple off shell, and terms which decouple when the external fermions are on shell. We have not been able to show that making an unfortunate choice at order N can never result in problems at higher orders. That there is at least one complete solution is proven in Sec. III by exhibiting it.

As with Bhabha scattering at e^2, the order e^N amplitudes with more than two external fermion lines are already free of problems by virtue of off-shell gauge invariance at lower orders. Hence no multifermion interactions ever need to be considered. Similarly, the momentum structure of the deficit term always precludes cancellation by adding any pure photon interaction. Finally, on-shell decoupling at order e^N implies off-shell invariance at the same order under a revised transformation law. Since only the fermion on-shell condition is used in proving decoupling, only the fermion transformation law ever suffers modification.

C. Quantization

A curious feature of fundamentally nonlocal field theories is that they are perturbatively acausal.[3,21] One consequence of this is that one cannot quantize in such a way as to simultaneously preserve the manifest Lorentz invariance of the functional formalism and the operator formalism.[22] This is probably an excellent reason for rejecting nonlocal actions as fundamental theories, but it poses no problem to their use as regularizations. One simply has to choose which formalism, operator or functional, to quantize covariantly. We shall make the same choice as was made for invariant string field theory, namely, to retain manifest Lorentz invariance in the functional formalism.

Once this is accepted, the problem of quantization amounts to finding an acceptable measure factor which makes the functional formalism invariant under the classical gauge transformation. The qualifier "acceptable" means that the measure factor interactions obey the same restrictions as the classical ones, namely, manifest Poincaré invariance, exponential suppression in Euclidean momentum space, reality for real momenta, and analyticity in all momentum variables throughout the complex plane. If such a functional exists, then the resulting perturbation theory has the properties we seek to all orders. Its Poincaré invariance is manifest, and finiteness is almost as obvious owing to the exponential convergence factors. As previously noted, the Cutkosky rules apply to nonlocal field theories the same as for local ones. This and the fact that the interactions are all real and analytic implies perturbative unitarity on the large Fock space, which includes unphysical photon polarizations. But since decoupling also follows, by gauge fixing inside the functional integral, we also have perturbative unitarity on the small Fock space of physical states. Similarly, current conservation and the Ward identities (for the nonlocal symmetry, of course) follow by changing variables in the usual manner.

Since the classical action was constructed to possess gauge invariance, the only problem can come from the functional measures $[dA]$, $[d\psi]$, and $[d\bar{\psi}]$. Because the ordinary photon transformation rule is unchanged, $[dA]$ is manifestly invariant and gauge fixing can be done in terms of the vector potential as in the local theory. Therefore, we need only consider the behavior of $[d\psi]$ and $[d\bar{\psi}]$ under transformation (2.2b). It is useful to introduce the "dot product" at this stage:

$$(\theta\cdot\mathcal{T}[eA])(x,z)\equiv\int d^4y\,\theta(y)\mathcal{T}[eA](x,y,z)\ . \tag{2.34}$$

The fact that Grassmann variables behave oppositely to $\mathbb{C}$ numbers under $\mathbb{C}$ number rescalings gives the following transformation rule to lowest order in θ:

$$[d\psi']=[d\psi]\det{}^{-1}(1+ie\theta\cdot\mathcal{T}[eA]) \tag{2.35a}$$

$$=[d\psi]\exp[-ie\,\mathrm{Tr}(\theta\cdot\mathcal{T}[eA])]\ . \tag{2.35b}$$

The "trace" in (2.35b) is understood to involve both summing over spinor indices and integrating over spacetime coordinates. A similar argument for $[d\bar{\psi}]$ gives the complete result

$$[d\psi'][d\bar{\psi}']=[d\psi][d\bar{\psi}]\times\exp[-ie\,\mathrm{Tr}(\theta\cdot\mathcal{T}[eA])+ie\,\mathrm{Tr}(\theta\cdot\bar{\mathcal{T}}[eA])]\ . \tag{2.36}$$

Owing to the peculiar mixing between gauge and spacetime indicates, it is generally not true that the two traces cancel; hence the fermion measures are generally not invariant. The condition for successful quantization is the existence of an acceptable measure factor $\mu[eA]$ which absorbs this noninvariance:

$$\mu[eA]\equiv\exp(iS_{\mathrm{meas}}[eA])\ , \tag{2.37a}$$

$$\partial_\mu\theta\cdot\frac{\delta S_{\mathrm{meas}}[eA]}{\delta A_\mu}=-e\,\mathrm{Tr}(\theta\cdot\mathcal{T}[eA])+e\,\mathrm{Tr}(\theta\cdot\bar{\mathcal{T}}[eA])\ . \tag{2.37b}$$

That such a term can be found is not obvious, although the expectation is that it can be for a nonanomalous theory such as QED. We prove its existence for the model of Sec. III and in fact evaluate it to lowest order. It is conceivable that poor choices for the ambiguous classical interactions give an acceptable tree theory for which no acceptable measure factor could be found.

If such a functional $S_{\mathrm{meas}}[eA]$ does exist, it will of course be ambiguous up to fully invariant terms, i.e., functionals of the field-strength tensor. To see that it can be chosen to have manifest Poincaré invariance, Euclidean suppression, and reality for real momenta, note that the transformation operator $\mathcal{T}[eA]$ is constructed to possess all of these properties. The difficult points are integrability of the variational equation (2.37b) and analyticity of the result. The latter condition requires showing that the right-hand side of (2.37b) vanishes for constant θ, and, hence, that it is really proportional to $\partial_\mu\theta=-\delta A_\mu$. Integrability follows if the variation of the right-hand side of (2.37b) under $\delta A_\mu=-\partial_\mu\phi$ is symmetric under interchange of θ and ϕ.

A point to note in passing is that the on-shell tree amplitudes of our nonlocalized action can agree with those of the unregulated theory, while the loop amplitudes disagree. This seems to be a contradiction because the loops of local field theories can be expressed as sums of integrals of trees using Feynman's tree theorem,[23] so that if two local theories agree at tree level,then their loop amplitudes must agree as well. The key word is "local;" the tree theorem does not generally apply to nonlocal theories. The theorem is proved by using the relation

$$\Delta_F=\Delta_R+\Delta^+ \tag{2.38}$$

to expand loops formed from the Feynman propagator Δ_F into a series in the on-shell projector Δ^+. This decomposes all terms with even one Δ^+ into trees. The term with no Δ^+'s is a loop formed entirely with the retarded propagator and must vanish if the interactions are local. However, when the interactions are nonlocal, this term generally survives, and there can be physical phenomena in loop amplitudes which could not have been predicted from the on-shell trees.[24]

III. EXPLICIT MODEL

We remind the reader of some useful conventions from Sec. I. First, there is the smearing operator $\mathcal{E}_M$:

$$\mathcal{E}_M\equiv\exp\left[\frac{\partial^2-M^2}{2\Lambda^2}\right]\ . \tag{3.1}$$

A field carrying the superscript Λ denotes smearing the unsuperscripted field with the appropriate mass:

$$\psi^\Lambda\equiv\mathcal{E}_m\psi\ , \tag{3.2a}$$

$$A_\mu^\Lambda\equiv\mathcal{E}_0A_\mu\ . \tag{3.2b}$$

The obvious initial nonlocalization of QED follows from (1.3):

$$\mathcal{L}_{0+1}\equiv-\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}-\bar{\psi}(i\partial\!\!\!/+m)\psi+e\bar{\psi}^\Lambda A\!\!\!/^\Lambda\psi^\Lambda\ . \tag{3.3}$$

It is invariant up to (but not including) order e^2 under transformation (1.5):

$$\delta A_\mu=\delta_0A_\mu=-\partial_\mu\theta\ , \tag{3.4a}$$

$$\delta_1\psi=ie\,\mathcal{E}_m\theta^\Lambda\psi^\Lambda\ , \tag{3.4b}$$

where the $\theta^\Lambda\equiv\mathcal{E}_0\theta$. Note that in (3.4b) the operator $\mathcal{E}_m$ acts on the product $\theta^\Lambda\psi^\Lambda$, while the $\mathcal{E}_0$ in θ^Λ acts only upon θ and the $\mathcal{E}_m$ in ψ^Λ acts only upon ψ.

From $\mathcal{E}_m$ we form the operator $\mathcal{O}$:

$$\mathcal{O} \equiv \frac{(\mathcal{E}_m)^2 - 1}{\partial^2 - m^2} = \int_0^1 \frac{d\tau}{\Lambda^2} \exp\left[\tau \frac{\partial^2 - m^2}{\Lambda^2}\right] . \tag{3.5}$$

Note that $\mathcal{O}$ is an entire function of ∂^2. Using this new operator, we can compactly express the simplest of the four-point interactions determined in the previous section:

$$\mathcal{L}_2 = -e^2 \bar{\psi}^\wedge \not{A}^\wedge (i\not{\partial} - m)\mathcal{O} \not{A}^\wedge \psi^\wedge . \tag{3.6}$$

The Compton amplitude computed with $\mathcal{L}_{0+1+2}$ is unchanged from that of QED. The way this works is that each V_2 contribution to the amplitude can be split into two terms through decomposing the operator $\mathcal{O}$ into $\mathcal{E}_m^2/(\partial^2 - m^2)$ and $-1/(\partial^2 - m^2)$. The first such term cancels the contribution from the corresponding $V_1 \cdot V_1$channel, while the second term is just the usual QED result for that channel. We can extend the process to higher photon amplitudes with interactions of the form

$$\mathcal{L}_n = -(-e)^n \bar{\psi}^\wedge \not{A}^\wedge [(i\not{\partial} - m)\mathcal{O} \not{A}^\wedge]^{(n-1)} \psi^\wedge , \tag{3.7}$$

which sum to give the total Lagrangian

$$\mathcal{L} = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} - \bar{\psi}(i\not{\partial} + m)\psi + e\bar{\psi}^\wedge \not{A}^\wedge [1 + e(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{-1} \psi^\wedge . \tag{3.8}$$

Since the extended Compton trees are the same as those of QED, decoupling is manifest. In fact, the only true amplitudes that differ from those of QED are ones containing an internal photon line. These acquire an exponential enhancement factor for each internal photon momentum.

To keep the action "invariant" with its higher interactions, the fermionic transformation must be modified at each order:

$$\delta_n \psi = -i(-e)^n \mathcal{E}_m \theta^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-1} \psi^\wedge . \tag{3.9}$$

The higher variations can also be summed:

$$\delta A_\mu = -\partial_\mu \theta , \tag{3.10a}$$

$$\delta\psi = ie\mathcal{E}_m \theta^\wedge [1 + e(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{-1} \psi^\wedge , \tag{3.10b}$$

but it is simpler to prove invariance using the expanded form. To make the argument, consider the zeroth-order variation of $\mathcal{L}_n$:

$$\delta_0 \mathcal{L}_n = (-e)^n \bar{\psi}^\wedge [\not{\partial}, \theta^\wedge][(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-1}\psi^\wedge + (-e)^n \bar{\psi}^\wedge \not{A}^\wedge (i\not{\partial} - m)\mathcal{O}[\not{\partial}, \theta^\wedge][(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-2}\psi^\wedge + \cdots + (-e)^n \bar{\psi}^\wedge [\not{A}^\wedge (i\not{\partial} - m)\mathcal{O}]^{n-1}[\not{\partial}, \theta^\wedge]\psi^\wedge . \tag{3.11a}$$

Now use the identities $[\not{\partial}, \theta^\wedge] = -i[(i\not{\partial} + m), \theta^\wedge]$ and $(i\not{\partial} + m)(i\not{\partial} - m) = \partial^2 - m^2$ to reach the form

$$\delta_0 \mathcal{L}_n = -i(-e)^n \bar{\psi}^\wedge (i\not{\partial} + m)\theta^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-1}\psi^\wedge + i(-e)^n \bar{\psi}^\wedge \theta^\wedge (\partial^2 - m^2)\mathcal{O}\not{A}^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-2}\psi^\wedge - i(-e)^n \bar{\psi}^\wedge \not{A}^\wedge (\partial^2 - m^2)\mathcal{O}\theta^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-2}\psi^\wedge + \cdots + i(-e)^n \bar{\psi}^\wedge [\not{A}^\wedge (i\not{\partial} - m)\mathcal{O}]^{n-1}\theta^\wedge (i\not{\partial} + m)\psi^\wedge . \tag{3.11b}$$

Decomposing the operator $\mathcal{O}$ and canceling some adjacent terms then gives

$$\delta_0 \mathcal{L}_n = -i(-e)^n \bar{\psi}^\wedge (i\not{\partial} + m)\theta^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-1}\psi^\wedge + i(-e)^n \bar{\psi}^\wedge \theta^\wedge \mathcal{E}_m^2 \not{A}^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-2}\psi^\wedge - i(-e)^n \bar{\psi}^\wedge \not{A}^\wedge \mathcal{E}_m^2 \theta^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-2}\psi^\wedge + \cdots + i(-e)^n \bar{\psi}^\wedge [\not{A}^\wedge (i\not{\partial} - m)\mathcal{O}]^{n-1}\theta^\wedge (i\not{\partial} + m)\psi^\wedge . \tag{3.11c}$$

Finally, we recognize pairs of terms as the variation of a lower interaction under a higher symmetry:

$$\delta_0 \mathcal{L}_n = +\bar{\psi}(i\not{\partial} + m)\delta_n \psi + (-e)^{n-1} \overline{\delta_1 \psi}^\wedge \not{A}^\wedge [(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{n-2}\psi^\wedge + \cdots + \overline{\delta_n \psi}(i\not{\partial} - m)\psi \tag{3.11d}$$

$$= -\delta_n \mathcal{L}_0 - \delta_{n-1}\mathcal{L}_1 - \cdots - \delta_1 \mathcal{L}_{n-1} , \tag{3.11e}$$

from which it follows that $\delta\mathcal{L} = 0$ at order e^n.

Though beautiful to behold, this gauge invariance does not seem to possess a conventional interpretation. The nonlocalization has resulted in a peculiar mixing between gauge "indices" (we mean, of course, the phase of the Fermi fields), spinor indices, and spacetime coordinates. For example, it is clear from expressions (3.8) and (3.10b) that the operator

$$\mathcal{T}[eA](x,y,z) \equiv \mathcal{E}_m [\delta^4(x-y)]^\wedge [1 + e(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{-1} \mathcal{E}_m \delta^4(x-z) \tag{3.12}$$

behaves in some respects like a representation matrix, but many familiar properties are lacking. Its operatorial character endows this "matrix" with nontrivial transformation properties. Further, the gauge field fails to commute with it:

$$\mathcal{E}_m \not{A}^\wedge [1 + e(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{-1} \mathcal{E}_m = \mathcal{E}_m [1 + e\not{A}^\wedge O(i\not{\partial} - m)]^{-1} \not{A}^\wedge \mathcal{E}_m . \tag{3.13}$$

And because (3.12) is not Hermitian, it follows that $\bar{\psi}\psi$ is not an invariant. One consequence is that the fermion measures of the functional formalism are not invariant; another is that the natural covariant kinetic operator

$$(i\not{\partial} + m) + e\mathcal{E}_m \not{A}^\wedge [1 + e(i\not{\partial} - m)\mathcal{O}\not{A}^\wedge]^{-1} \mathcal{E}_m \tag{3.14}$$

transforms one way on the right and another way entirely on the left. This means that products of such derivatives are not covariant; nor can we obtain a covariant field strength tensor by commuting these operators. Another eccentricity is that while the symmetry is still Abelian on shell, it does not close unless the Fermi fields obey their equations of motion

$$[\delta_{\theta_1},\delta_{\theta_2}]\psi=-e^2\mathcal{E}_m\{\theta_1^\wedge[1+eA\!\!\!/\,^\wedge\mathcal{O}(i\partial\!\!\!/-m)]^{-1}(i\partial\!\!\!/-m)\mathcal{O}\theta_2^\wedge-(1\leftrightarrow 2)\}$$
$$\times\mathcal{E}_m(i\partial\!\!\!/+m+e\mathcal{E}_m A\!\!\!/\,^\wedge[1+eA\!\!\!/\,^\wedge\mathcal{O}(i\partial\!\!\!/-m)]^{-1}\mathcal{E}_m)\psi\ . \tag{3.15}$$

Hence the transformations do not even form a group. In fact they are part of a larger group which includes some transformations that affect only the Fermi fields and vanish in the local limit.

The similarity of our trees to those of QED implies the existence of a field redefinition which reduces our Lagrangian to the form[25]

$$-\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu}-\bar\psi(i\partial\!\!\!/+m)\psi+e\bar\psi A\!\!\!/\,^\wedge\psi\ . \tag{3.16}$$

This theory has been studied previously by Efimov.[26] It does not represent a complete regularization of QED because the vertex contains no convergence factors for the Fermi fields. This means that pure fermion loops receive no regularization, and it is easy to see that the vacuum polarization diverges. On the other hand, all loops in our theory are finite. That this can be so is a manifestation of the inapplicability of Feynman's tree theorem to nonlocal field theories, as was explained in Sec. II C.

Although it is really the Euclidean loop integrals which are regulated by this method, it is simplest to keep track of the external momenta, the various traces, and the factors of i by formally working in Minkowski space. Consider, for example, the one-loop correction to the electron self-energy which derives from joining two V_1's [Fig. 3(a)]:

$$-i\Sigma_1(p)\equiv\int\frac{d^4k}{(2\pi)^4}(ie\gamma^\mu)\left[\frac{-i}{q\!\!\!/+m-i\epsilon}\right](ie\gamma^\nu)\left[\frac{-i\eta^{\mu\nu}}{k^2-i\epsilon}\right]\exp\left[-\left[\frac{p^2+m^2}{\Lambda^2}\right]-\left[\frac{q^2+m^2}{\Lambda^2}\right]-\frac{k^2}{\Lambda^2}\right], \tag{3.17a}$$

where $q\equiv p-k$. The next steps are to promote the propagators to Schwinger integrals and then perform the momentum integral:

$$-i\Sigma_1(p)=-e^2\exp\left[-\left[\frac{p^2+m^2}{\Lambda^2}\right]\right]\int_1^\infty\frac{d\tau_1}{\Lambda^2}\int_1^\infty\frac{d\tau_2}{\Lambda^2}\int\frac{d^4k}{(2\pi)^4}(2q\!\!\!/+4m)\exp\left[-\tau_1\left[\frac{q^2+m^2}{\Lambda^2}\right]-\tau_2\frac{k^2}{\Lambda^2}\right] \tag{3.17b}$$

$$=\frac{-ie^2}{8\pi^2}\exp\left[-\left[\frac{p^2+m^2}{\Lambda^2}\right]\right]\int_1^\infty d\tau_1\int_1^\infty d\tau_2\left[\frac{\tau_2}{(\tau_1+\tau_2)^3}p\!\!\!/+\frac{2}{(\tau_1+\tau_2)^2}m\right]\exp\left[-\frac{\tau_1\tau_2}{\tau_1+\tau_2}\frac{p^2}{\Lambda^2}-\tau_1\frac{m^2}{\Lambda^2}\right]. \tag{3.17c}$$

The factor of i comes from the "rotation" to Euclidean space, the sign being determined by the requirement that we avoid enclosing poles of the propagator. This trick gets the right answer, but we stress that it is entirely formal; the correct derivation proceeds from Euclidean space. Far from improving convergence of the Minkowski-space momentum integrals, our exponential factors actually worsen it.

The other correction at the same order comes from a single V_2 [Fig. 3(b)]:

$$-i\Sigma_2(p)\equiv\int\frac{d^4k}{(2\pi)^4}(-ie^2)\gamma^\mu(q\!\!\!/-m)\gamma^\nu\left[\frac{-i\eta_{\mu\nu}}{k^2-i\epsilon}\right]\int_0^1\frac{d\tau}{\Lambda^2}\exp\left[-\left[\frac{p^2+m^2}{\Lambda^2}\right]-\tau\left[\frac{q^2+m^2}{\Lambda^2}\right]-\frac{k^2}{\Lambda^2}\right]. \tag{3.18}$$

The result of performing the momentum integration is the same as expression (3.17c) except that the Schwinger parameters are integrated over the region $0\le\tau_1\le 1\le\tau_2<\infty$. Σ_1 and Σ_2 obviously add to give the simple expression

$$\Sigma(p)=\frac{e^2}{8\pi^2}\exp\left[-\left[\frac{p^2+m^2}{\Lambda^2}\right]\right]\int_0^\infty d\tau_1\int_1^\infty d\tau_2\left[\frac{\tau_2}{(\tau_1+\tau_2)^3}p\!\!\!/+\frac{2}{(\tau_1+\tau_2)^2}m\right]\exp\left[-\frac{\tau_1\tau_2}{\tau_1+\tau_2}\frac{p^2}{\Lambda^2}-\tau_1\frac{m^2}{\Lambda^2}\right] \tag{3.19a}$$

$$=\frac{e^2}{8\pi^2}\exp\left[-\left[\frac{p^2+m^2}{\Lambda^2}\right]\right]\int_0^1 dx\,(xp\!\!\!/+2m)E_1\left[(1-x)\frac{p^2}{\Lambda^2}+\left[\frac{1-x}{x}\right]\frac{m^2}{\Lambda^2}\right], \tag{3.19b}$$

where E_1 is the exponential integral:

$$E_1(z)\equiv\int_z^\infty dt\frac{\exp(-t)}{t}=-\ln(z)-\gamma-\sum_{n=1}^\infty\frac{(-z)^n}{nn!}\ . \tag{3.20}$$

Although the final parameter integral cannot be evaluated in terms of elementary functions, it is easy to develop an asymptotic expansion in Λ by expanding the exponential integral:

$$\Sigma(p)=\frac{e^2}{8\pi^2}\left[\left(\frac{1}{2}\not{p}+2m\right)\ln(\Lambda^2)-\left(\frac{1}{2}\not{p}+2m\right)\gamma+\frac{1}{2}\not{p}-\int_0^1 dx\,(x\not{p}+2m)\ln(xp^2+m^2)+O\left(\frac{\ln(\Lambda^2)}{\Lambda^2}\right)\right] . \tag{3.21}$$

Comparison with the result of dimensional regularization in D dimensions and scale μ gives

$$\Sigma(p)=e^2\frac{\Gamma(2-D/2)}{2^D\pi^{D/2}}\int_0^1 dx[(D-2)x\not{p}+Dm]\left[x(1-x)\frac{p^2}{\mu^2}+(1-x)\frac{m^2}{\mu^2}\right]^{D/2-2} \tag{3.22a}$$

$$=\frac{e^2}{8\pi^2}\left[\left(\frac{1}{2}\not{p}+2m\right)\frac{2}{4-D}+\left(\frac{1}{2}\not{p}+m\right)\left(\ln(4\pi)-\gamma+\frac{1}{2}\right)-\int_0^1 dx\,(x\not{p}+2m)\ln(xp^2+m^2)+O(4-D)\right] . \tag{3.22b}$$

This suggests the correspondence

$$\frac{2}{4-D}\sim\ln(\Lambda^2) , \tag{3.23}$$

for the coefficients of logarithmic divergences in the two methods. In fact, this relation, which we will see again with the vacuum polarization, must persist for all one-loop amplitudes in order that the two bare theories agree.

A more revealing comparison between the two methods comes from examining the nonlocally regulated electron self-energy in D dimensions:

$$\Sigma(p)=-ie^2\exp\left[-\left(\frac{p^2+m^2}{\Lambda^2}\right)\right]\int_0^\infty\frac{d\tau_1}{\Lambda^2}\int_1^\infty\frac{d\tau_2}{\Lambda^2}\int\frac{d^Dk}{(2\pi)^D}[(D-2)\not{q}+Dm]\exp\left[-\tau_1\left(\frac{q^2+m^2}{\Lambda^2}\right)-\tau_2\frac{k^2}{\Lambda^2}\right] , \tag{3.24a}$$

$$\Sigma(p)=\frac{e^2}{2^D\pi^{D/2}}\exp\left[-\left(\frac{p^2+m^2}{\Lambda^2}\right)\right]\int_0^\infty d\tau_1\int_1^\infty d\tau_2\left[\frac{(D-2)\tau_2}{(\tau_1+\tau_2)^3}\not{p}+\frac{D}{(\tau_1+\tau_2)^2}m\right]$$
$$\times\left[\frac{\Lambda^2}{(\tau+1)}\right]^{D/2-2}\exp\left[-\frac{\tau_1\tau_2}{\tau_1+\tau_2}\frac{p^2}{\Lambda^2}-\tau_1\frac{m^2}{\Lambda^2}\right] , \tag{3.24b}$$

$$=\frac{e^2}{2^D\pi^{D/2}}\exp\left[-\left(\frac{p^2+m^2}{\Lambda^2}\right)\right]\int_0^1 dx[(D-2)x\not{p}+Dm][x(1-x)p^2+(1-x)m^2]^{D/2-2}$$
$$\times\Gamma\left[2-\frac{D}{2},(1-x)\frac{p^2}{\Lambda^2}+\left(\frac{1-x}{x}\right)\frac{m^2}{\Lambda^2}\right] , \tag{3.24c}$$

where $\Gamma(n,z)$ is the incomplete γ function:

$$\Gamma(n,z)\equiv\int_z^\infty dt\,t^{n-1}e^{-t} \tag{3.25a}$$

$$=(n-1)\Gamma(n-1,z)+z^{n-1}e^{-z} . \tag{3.25b}$$

One develops an asymptotic expansion in Λ by using the recursion relation (3.25b) to reach either $\Gamma(0,z)=E_1(z)$ for D even or

$$\Gamma(\tfrac{1}{2},z)=\sqrt{\pi}\,\mathrm{erfc}(\sqrt{z}) \tag{3.26}$$

for D odd. Here erfc(x) is is the complementary error function:

$$\sqrt{\pi}\,\mathrm{erfc}(\sqrt{z})\equiv 2\int_{\sqrt{z}}^\infty dt\,\exp(-t^2)=\sqrt{\pi}-2\sqrt{z}\sum_{n=0}^\infty\frac{(-z)^n}{(2n+1)n!} . \tag{3.27}$$

As an example, the $D=5$ result is

$$\Sigma(p)=\frac{e^2}{16\pi^2}\left[(2\not{p}+10m)\frac{\Lambda}{\sqrt{\pi}}-\int_0^1 dx\,(3x\not{p}+5m)[x(1-x)p^2+(1-x)m^2]^{1/2}+O\left(\frac{1}{\Lambda}\right)\right] . \tag{3.28}$$

Note that while the finite term agrees with (3.22a), our method picks out a gauge-invariant, linear divergence that dimensional regularization misses. In general, our method gives all the divergences which would be expected by power counting and gauge invariance, while the dimensional method catches only logarithmic divergences. Another point of interest is the persistence of the infrared problem for $p^2=m^2=0$ in low dimensions. Nonlocal regularization affects only ultraviolet phenomena; any infrared problems must be treated using other methods.

It is interesting to note that (3.19a) is just $\exp[-(p^2+m^2)/\Lambda^2]$ times the result obtained from Efimov's Lagrangian (3.16). This concurrence is lost for all-fermion loops such as the vacuum polarization where Efimov's theory does not differ from the $\Lambda=\infty$ result, aside from a factor of $\mathcal{E}_0$ on each external line. The result from our theory is the sum of the measure factor contribution plus two loops built from the classical vertices. The first of these last derives from the combination of two V_1's [Fig. 4(a)]:

$$i\Pi_1^{\mu\nu}(p)\equiv-\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}\left[(ie\gamma^\mu)\left[\frac{-i}{\not q+m-i\epsilon}\right](ie\gamma^\nu)\left[\frac{-i}{-\not k+m-i\epsilon}\right]\right]\exp\left[-\left[\frac{q^2+m^2}{\Lambda^2}\right]-\left[\frac{k^2+m^2}{\Lambda^2}\right]-\frac{p^2}{\Lambda^2}\right] \tag{3.29a}$$

$$\equiv i\Pi_1^T(p^2)\left[\eta^{\mu\nu}-\frac{p^\mu p^\nu}{p^2}\right]+i\Pi_1^L(p^2)\frac{p^\mu p^\nu}{p^2}\,, \tag{3.29b}$$

where the transverse and longitudinal coefficients are

$$i\Pi_1^T(p^2)=\frac{ie^2}{4\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]\int_1^\infty d\tau_1\int_1^\infty d\tau_2\left[\frac{\Lambda^2}{(\tau_1+\tau_2)^3}-\frac{\tau_1\tau_2}{(\tau_1+\tau_2)^4}p^2+\frac{m^2}{(\tau_1+\tau_2)^2}\right]\exp\left[-\frac{\tau_1\tau_2}{\tau_1+\tau_2}\frac{p^2}{\Lambda^2}-(\tau_1+\tau_2)\frac{m^2}{\Lambda^2}\right], \tag{3.29c}$$

$$i\Pi_1^L(p^2)=\frac{ie^2}{4\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]\int_1^\infty d\tau_1\int_1^\infty d\tau_2\left[\frac{\Lambda^2}{(\tau_1+\tau_2)^3}+\frac{\tau_1\tau_2}{(\tau_1+\tau_2)^4}p^2+\frac{m^2}{(\tau_1+\tau_2)^2}\right]\exp\left[-\frac{\tau_1\tau_2}{\tau_1+\tau_2}\frac{p^2}{\Lambda^2}-(\tau_1+\tau_2)\frac{m^2}{\Lambda^2}\right]. \tag{3.29d}$$

A very similar contribution derives from a single V_2 [Fig. 4(b)]:

$$i\Pi_2^{\mu\nu}(p)\equiv-\int\frac{d^4k}{(2\pi)^4}\mathrm{Tr}\left[-2ie^2\gamma^\mu(\not q-m)\gamma^\nu\left[\frac{-i}{-\not k+m-i\epsilon}\right]\right]\int_0^1\frac{d\tau}{\Lambda^2}\exp\left[-\tau\left[\frac{q^2+m^2}{\Lambda^2}\right]-\left[\frac{k^2+m^2}{\Lambda^2}\right]-\frac{p^2}{\Lambda^2}\right]. \tag{3.30}$$

The transverse and longitudinal parts of Π_2 are the same as those of Π_1, except that the Schwinger parameters are integrated over the regions $0\le\tau_1\le1\le\tau_2<\infty$ and $0\le\tau_2\le1\le\tau_1<\infty$. The obvious conjecture is that the sum of the classical vertex contributions to the N-photon loop can be obtained from the $(V_1)^N$contribution by simply extending the range of parameter integration from $1\le\tau_i<\infty$ to $0\le\tau_i\le\infty$, excluding the unit hypercube $0\le\tau_i\le1$.

As is obvious from (3.29d), the two classical vertex contributions to $\Pi^{\mu\nu}$ do not sum to give zero longitudinal part: neither do the transverse parts sum to zero on shell. If these were the only contributions, we would have lost both decoupling and the photon's masslessness. Both features are restored by the contribution from the measure factor [Fig. 4(c)]:

$$i\Pi_3^{\mu\nu}\equiv\frac{-ie^2}{2\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]\int_0^1 d\tau\frac{\Lambda^2}{(\tau+1)^3}\exp\left[-\frac{\tau}{(\tau+1)}\frac{p^2}{\Lambda^2}-(\tau+1)\frac{m^2}{\Lambda^2}\right]\eta^{\mu\nu}\,. \tag{3.31}$$

We will shortly derive this result, but for the moment let us proceed to simply add it to the previous contributions. To see transversality, change variables in (3.29d) from τ_1 to $x=\tau_1/\tau_2$:

$$i\Pi_1^L(p^2)=\frac{ie^2}{4\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]\left[\int_1^\infty dx\int_1^\infty d\tau_2+\int_0^1 dx\int_{1/x}^\infty d\tau_2\right]$$
$$\times\frac{\partial}{\partial\tau_2}\left[-\frac{1}{\tau_2}\frac{\Lambda^2}{(x+1)^3}\exp\left[-\tau_2\frac{x}{x+1}\frac{p^2}{\Lambda^2}-\tau_2(x+1)\frac{m^2}{\Lambda^2}\right]\right] \tag{3.32a}$$

$$=\frac{ie^2}{2\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]\int_1^\infty d\tau\frac{\Lambda^2}{(\tau+1)^3}\exp\left[-\frac{\tau}{\tau+1}\frac{p^2}{\Lambda^2}-(\tau+1)\frac{m^2}{\Lambda^2}\right]. \tag{3.32b}$$

A similar set of manipulations gives

$$i\Pi_2^L(p^2)=\frac{ie^2}{2\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]\left[\int_0^1-\int_1^\infty\right]d\tau\frac{\Lambda^2}{(\tau+1)^3}\exp\left[-\frac{\tau}{\tau+1}\frac{p^2}{\Lambda^2}-(\tau+1)\frac{m^2}{\Lambda^2}\right]. \tag{3.33}$$

From (3.31) we see that $\Pi_1^L+\Pi_2^L+\Pi_3^L=0$ as desired.

If we subtract those terms whose vanishing has just been shown, the following expression results for the total transverse part:

$$\Pi^T(p^2)=-\frac{e^2p^2}{2\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]\left[\int_0^\infty\int_1^\infty+\int_1^\infty\int_0^1\right]d\tau_1 d\tau_2\frac{\tau_1\tau_2}{(\tau_1+\tau_2)^4}\exp\left[-\frac{\tau_1\tau_2}{\tau_1+\tau_2}\frac{p^2}{\Lambda^2}-(\tau_1+\tau_2)\frac{m^2}{\Lambda^2}\right] \tag{3.34a}$$

$$=-\frac{e^2p^2}{2\pi^2}\exp\left[-\frac{p^2}{\Lambda^2}\right]2\int_0^{1/2}dx\,x(1-x)E_1\left[x\frac{p^2}{\Lambda^2}+\frac{1}{1-x}\frac{m^2}{\Lambda^2}\right]. \tag{3.34b}$$

Note that gauge invariance has absorbed the naive quadratic divergence. The factor of p^2 guarantees masslessness, because $\Pi^T(p^2)\to 0$ as $p^2\to 0$.

Another interesting and possibly significant point is that the positivity of the exponential integral prevents the appearance of a Landau ghost for any real value of p^2, even when $p^2\gg\Lambda^2$. This may be of relevance to a scheme whereby local QED is embedded in a larger theory which provides an ultraviolet cutoff. On the other hand, it must be admitted that there are complex-conjugate poles on the physical sheet, at least when one ignores higher-loop effects. The consequent noncausality presumably dooms any attempt at resummation and emphasizes the fact that nonlocal QED is not viable beyond on-shell perturbation theory. The same effect has been demonstrated for the nonlocal scalar model (1.1) and doubtless occurs as well in string theory.[3]

Though the final parameter integral cannot be evaluated in terms of elementary functions, it is simple to develop an asymptotic expansion for large Λ:

$$\Pi^T(p^2)=-\frac{e^2p^2}{2\pi^2}\left[\frac{1}{6}\ln(\Lambda^2)+\frac{1}{6}\ln(2)-\frac{1}{6}\gamma-\frac{13}{72}-\int_0^1 dx\,x(1-x)\ln[x(1-x)p^2+m^2]+O\left[\frac{\ln(\Lambda^2)}{\Lambda^2}\right]\right]. \tag{3.35}$$

Our result compares with that of dimensional regularization in D dimensions with scale μ:

$$\Pi^T(p^2)=-e^2p^2 2^{D/2+1}\frac{\Gamma(2-D/2)}{2^D\pi^{D/2}}\int_0^1 dx\,x(1-x)\left[x(1-x)\frac{p^2}{\mu^2}+\frac{m^2}{\mu^2}\right]^{D/2-2} \tag{3.36a}$$

$$=-\frac{e^2p^2}{2\pi^2}\left[\frac{1}{6}\frac{2}{4-D}-\frac{1}{6}\gamma-\frac{1}{6}\ln(2\pi)-\int_0^1 dx\,x(1-x)\ln\left[x(1-x)\frac{p^2}{\mu^2}+\frac{m^2}{\mu^2}\right]+O(4-D)\right]. \tag{3.36b}$$

As with the electron self-energy, the logarithmic divergent terms are related according to (3.23).

The preceding discussion witnessed the intervention of a deus ex machina to restore transversality and masslessness. We are referring to the measure factor $\mu[eA]=\exp(iS_{\rm meas}[eA])$. Recall from Sec. II C that its purpose is to cancel the gauge variance of the fermion measures and that this requires

$$\partial_\mu\theta\cdot\frac{\delta S_{\rm meas}[eA]}{\delta A_\mu}=-e\,{\rm Tr}(\theta\cdot\mathcal{T}[eA])+e\,{\rm Tr}(\theta\cdot\overline{\mathcal{T}}[eA]). \tag{3.37}$$

Of course, this condition only determines $S_{\rm meas}[eA]$ up to a functional of the field-strength tensor, but an obvious minimal choice exists even as it did for the higher classical interactions. With a totally straightforward calculation, this choice gives the term we used above to compute Π_3:

$$S_{\rm meas}[eA]=\frac{-e^2}{4\pi^2}\int d^4x\,A_\mu^\Lambda\eta^{\mu\nu}\mathcal{O}_1A_\nu^\Lambda+O(e^4), \tag{3.38a}$$

$$\mathcal{O}_1\equiv\int_0^1 d\tau\frac{\Lambda^2}{(\tau+1)^3}\exp\left[\frac{\tau}{(\tau+1)}\frac{\partial^2}{\Lambda^2}-(\tau+1)\frac{m^2}{\Lambda^2}\right]. \tag{3.38b}$$

We have not evaluated any of the higher-order terms, though we emphasize that this is no harder than computing the associated higher-point photon loops. That they exist can be proven in three steps, each of which involves an application of the same sort of argument which is used to prove Furry's theorem. Recall that a crucial element of this is the charge-conjugation matrix C:

$$C(\gamma^\mu)^{\rm tr}C^{-1}=-\gamma^\mu. \tag{3.39}$$

Note that it is independent of spacetime and hence commutes with purely spacetime operators such as $\mathcal{E}_m$.

The first step is to show that odd powers of e cancel. To see this, consider the order e^N contribution to $e\,{\rm Tr}(\theta\cdot\overline{\mathcal{T}})$:

$$-(-e)^N\mathrm{Tr}\{\mathcal{E}_m[\boldsymbol{A}^\Lambda\mathcal{O}(i\not\partial-m)]^{N-1}\theta^\Lambda\mathcal{E}_m\}=-(-e)^N\mathrm{Tr}\{\mathcal{E}_m[C^{-1}C\boldsymbol{A}^\Lambda C^{-1}C\mathcal{O}(i\not\partial-m)C^{-1}C]^{N-1}\theta^\Lambda\mathcal{E}_m\} \quad (3.40a)$$

$$=-(-e)^N\mathrm{Tr}\{\mathcal{E}_m[-(\boldsymbol{A}^\Lambda)^{\mathrm{tr}}\mathcal{O}(i\overleftarrow{\not\partial}-m)^{\mathrm{tr}}]^{N-1}\theta^\Lambda\mathcal{E}_m\} \quad (3.40b)$$

$$=e^N\mathrm{Tr}\{\mathcal{E}_m\theta^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{N-1}\mathcal{E}_m\}\ . \quad (3.40c)$$

For N odd this just cancels the analogous contribution from $-e(\theta\cdot\mathcal{T})$, while for even N they add.

The second step is proving that the right-hand side of (3.37) vanishes for constant θ, which is of course essential if it is to equal the left-hand side. To see this, drop θ from the total contribution at order e^{2N} and commute the final factor of $(i\not\partial-m)\mathcal{O}$ through the operator $\mathcal{E}_m$ and around to the other side of the trace:

$$2e^{2N}\mathrm{Tr}\{\mathcal{E}_m[\boldsymbol{A}^\Lambda\mathcal{O}(i\not\partial-m)]^{2N-1}\mathcal{E}_m\}=2e^{2N}\mathrm{Tr}\{\mathcal{E}_m(i\not\partial-m)\mathcal{O}[\boldsymbol{A}^\Lambda\mathcal{O}(i\not\partial-m)]^{2N-2}\boldsymbol{A}^\Lambda\mathcal{E}_m\} \quad (3.41a)$$

$$=2e^{2N}\mathrm{Tr}\{\mathcal{E}_m[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-1}\mathcal{E}_m\}\ . \quad (3.41b)$$

Now introduce $C^{-1}C$ and integrate by parts as before to obtain minus the original term. It therefore vanishes, which was to be shown.

The final step is integrability. To see this we first vary the order e^{2N} contribution to the right-hand side of (3.37) with respect to gauge parameter ϕ and then show that the result is symmetric under interchange of θ and ϕ. It is convenient to drop the factor of e^{2N}:

$$\delta_\phi\mathrm{Tr}\{\mathcal{E}_m\theta^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-1}\mathcal{E}_m\}$$

$$=i\sum_{k=0}^{2N-2}\mathrm{Tr}\{\mathcal{E}_m\theta^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^k(i\not\partial-m)\mathcal{O}[(i\not\partial+m),\phi^\Lambda][(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-2-k}\mathcal{E}_m\} \quad (3.42a)$$

$$=-i\,\mathrm{Tr}\{\mathcal{E}_m^2\theta^\Lambda\phi^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-2}\}+i\sum_{k=0}^{2N-2}\mathrm{Tr}\{\mathcal{E}_m^2\theta^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^k\mathcal{E}_m^2\phi^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-2-k}\}$$

$$-i\sum_{k=0}^{2N-3}\mathrm{Tr}\{\mathcal{E}_m^2\theta^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^k(i\not\partial-m)\mathcal{O}\phi^\Lambda\mathcal{E}_m^2\boldsymbol{A}^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-3-k}\}$$

$$-i\,\mathrm{Tr}\{(i\not\partial+m)\mathcal{E}_m^2\theta^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-2}(i\not\partial-m)\mathcal{O}\phi^\Lambda\}\ . \quad (3.42b)$$

The first term of (3.42b) is manifestly symmetric under $\theta\leftrightarrow\phi$. The first of the sums is too because the interchange takes its $k=n$ term into the $k=2N-2-n$ one. (The $k=N-1$ is symmetric by itself.) To see the symmetry of the second sum, we again effect a transposition with the charge-conjugation matrix and some integrations by parts:

$$\mathrm{Tr}\{\mathcal{E}_m^2\theta^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^k(i\not\partial-m)\mathcal{O}\phi^\Lambda\mathcal{E}_m^2\boldsymbol{A}^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-3-k}\}$$

$$=\mathrm{Tr}\{\mathcal{E}_m^2\theta^\Lambda[(i\overleftarrow{\not\partial}-m)^{\mathrm{tr}}\mathcal{O}(\boldsymbol{A}^\Lambda)^{\mathrm{tr}}]^k(i\overleftarrow{\not\partial}-m)^{\mathrm{tr}}\mathcal{O}\phi^\Lambda\mathcal{E}_m^2(\boldsymbol{A}^\Lambda)^{\mathrm{tr}}[(i\overleftarrow{\not\partial}-m)^{\mathrm{tr}}\mathcal{O}(\boldsymbol{A}^\Lambda)^{\mathrm{tr}}]^{2N-3-k}\}$$

$$=\mathrm{Tr}\{[\boldsymbol{A}^\Lambda\mathcal{O}(i\not\partial-m)]^{2N-3-k}\boldsymbol{A}^\Lambda\mathcal{E}_m^2\phi^\Lambda\mathcal{O}(i\not\partial-m)[\boldsymbol{A}^\Lambda\mathcal{O}(i\not\partial-m)]^k\theta^\Lambda\mathcal{E}_m^2\}$$

$$=\mathrm{Tr}\{\mathcal{E}_m^2\phi^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^k(i\not\partial-m)\mathcal{O}\theta^\Lambda\mathcal{E}_m^2\boldsymbol{A}^\Lambda[(i\not\partial-m)\mathcal{O}\boldsymbol{A}^\Lambda]^{2N-3-k}\}\ , \quad (3.43)$$

from which it follows that the interchange takes the $k=n$ term to the $k=2N-3-n$ one. The symmetry of the final term follows through a similar transposition.

This completes the proof. Three significant final points are that the higher terms are all finite in the limit $\Lambda\to\infty$ (in fact, all but the $N=2$ term vanish in this limit), that

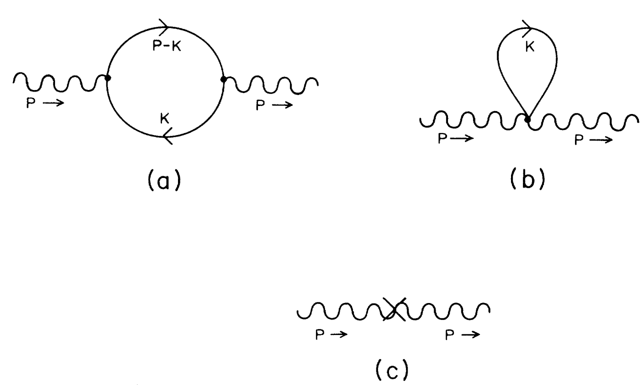

FIG. 4. (a) Vacuum polarization at one loop: contribution of V_1. (b) Vacuum polarization at one loop: contribution of V_2. (c) Vacuum polarization at one loop: measure factor contribution.

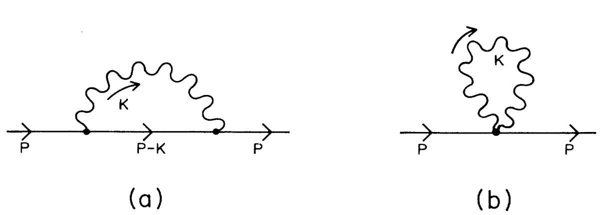

FIG. 3. (a) Electron self-energy at one loop: contribution of V_1. (b) Electron self-energy at one loop: contribution of V_2.

these measure factor terms are the *only* corrections needed to make the functional formalism valid for arbitrarily high-loop effects, and that they preserve the finiteness of higher loops by virtue of the factors of $\mathcal{E}_0$ which reside on each external leg.

We close with a discussion of the axial vector anomaly. A curious feature of our theory is that there is not one: the Lagrangian (3.8) is chirally invariant in the limit of zero mass, and the measure factor does not disturb this. A consequence is that Noether's theorem gives a fully regulated axial-vector current which is conserved in the massless limit. *We emphasize that there is no fermion doubling in our theory.* Since the Nielsen-Ninomiya theorem[27] can be evaded by sacrificing locality on the lattice, it should not be surprising that we can make a chirally invariant theory by giving up continuum locality. After all, an anomaly can always be expressed as the presence in the effective action of a finite, noninvariant and nonlocal term. In a formalism where such terms are allowed as classical interactions, their appearance in the effective action could always be renormalized away. We do not possess quite this much freedom because our higher classical interactions are constrained to formally vanish in the limit of infinite Λ—and this is why we can see genuine gauge anomalies through the nonexistence of an invariant measure. However, it is perfectly possible for one of our classical interactions to suffer quantum ordering corrections which engender nonvanishing terms of the desired sort. Consider, for example, the diagram which results from contracting the two Fermi fields in $\mathcal{L}_2$ of Eq. (3.6).

A more serious issue is reproducing the correct rate for $\pi^0 \rightarrow \gamma\gamma$. Although often cited as a physical proof for the existence of a chiral anomaly in QED, this is not quite true owing to an ambiguity in the correct axial-vector current to use for PCAC (partial conservation of axial-vector current). The conserved axial-vector current which follows from Noether's theorem is heavily contaminated with higher photon couplings which are essential to maintaining its conservation. These purely electromagnetic terms do not seem to have any business in coupling to an eigenstate of the strong interaction such as the pion. If we drop them, then a more plausible identification emerges for "the axial-vector current:"

$$J^5_\mu \equiv \bar{\psi}^\Lambda \gamma_5 \gamma_\mu \psi^\Lambda \ . \tag{3.44}$$

One of course loses conservation. By taking vacuum expectation values (VEV's), it is easy to check the following weak operator identity in the massless limit:

$$\partial^\mu J^5_\mu = -\frac{e^2}{32\pi^2}\epsilon^{\alpha\beta\gamma\delta}F^\Lambda_{\alpha\beta}F^\Lambda_{\gamma\delta} + O\left[\frac{\ln(\Lambda^2)}{\Lambda^2}\right] . \tag{3.45}$$

By basing PCAC on (3.44), we get the right rate for $\pi^0 \rightarrow \gamma\gamma$ in a theory which is nonetheless chiral in the massless limit. A final point to note is that it would not have been possible to gauge the global chiral invariance of our model. The analog of (3.38) would reveal an order e^3 term coming from the A-A-A anomaly, which could not be absorbed in the variation of a measure factor.

IV. NON-ABELIAN GAUGE THEORIES AND GRAVITY

Our strategy for nonlocalizing non-Abelian gauge theories and gravity is in broad outline the same as for QED. One first constructs an invariant nonlocal classical action and then searches for a measure factor which will make the functional formalism invariant under the nonlocal symmetry. The classical action is built up iteratively as for QED, but one now has to face theories which are nontrivial without matter coupling.

Our conventions for the Yang-Mills Lagrangian and field strength are

$$\mathcal{L} = -\tfrac{1}{4}F_{a\mu\nu}F_a^{\mu\nu} \ , \tag{4.1a}$$

$$F_{a\mu\nu} \equiv \partial_\mu A_{a\nu} - \partial_\nu A_{a\mu} - g f_{abc} A_{b\mu} A_{c\nu} \ , \tag{4.1b}$$

where g is the coupling constant and f_{abc} are the structure constants. It is invariant under the familiar transformation

$$\delta A_{a\mu} = -\partial_\mu \theta_a + g f_{abc} A_{b\mu}\theta_c \ . \tag{4.2}$$

For gravity we shall take the Hilbert action, partially integrated to remove second derivatives:

$$\mathcal{L} = \frac{1}{\kappa^2}(\Gamma^{\beta\gamma}_{\alpha}\Gamma^{\alpha}_{\beta\gamma} - \Gamma^{\alpha\beta}_{\beta}\Gamma^{\gamma}_{\gamma\alpha})\sqrt{-g} \ , \tag{4.3a}$$

where $\kappa^2 = 16\pi G$ and the affine connection is

$$\Gamma^{\alpha}_{\beta\gamma} \equiv \tfrac{1}{2} g^{\alpha\rho}(g_{\rho\beta,\gamma} + g_{\gamma\rho,\beta} - g_{\beta\gamma,\rho}) \ . \tag{4.3b}$$

A comma on the index list of a tensor denotes ordinary differentiation with respect to the subsequent indices. Because our analysis is of necessity perturbative, to avoid the nonperturbative problems of Ref. 3, the field we shall actually use is the graviton:

$$h_{\mu\nu} \equiv \frac{1}{\kappa}(g_{\mu\nu} - \eta_{\mu\nu}) \ . \tag{4.4}$$

As always in perturbation theory, around flat space the indices of $h_{\mu\nu}$ are raised and lowered using the Minkowski metric. The action obtained by integrating (4.3a) is invariant under general coordinate transformations:

$$\delta h_{\mu\nu} = -\epsilon_{\mu,\nu} - \epsilon_{\nu,\mu} - \kappa(\epsilon^{\rho}{}_{,\mu}h_{\rho\nu} + \epsilon^{\rho}{}_{,\nu}h_{\rho\mu} + h_{\mu\nu,\rho}\epsilon^{\rho}) \ , \tag{4.5}$$

where one must assume that the parameters ϵ^μ generate a Poincaré transformation at infinity. Actually, it would be possible to develop our method around any background, but we shall assume flat space for convenience.

The initial nonlocalization is provided by decorating each field in the interaction terms with a factor of $\mathcal{E}_0$:

$$A_{a\mu} \rightarrow A^\Lambda_{a\mu} \ , \tag{4.6a}$$

$$h_{\mu\nu} \rightarrow h^\Lambda_{\mu\nu} \ . \tag{4.6b}$$

This destroys the local gauge invariance at order g (or κ for gravity), but since the current (stress energy) is still conserved at this order the theory is invariant up to g^2 (κ^2) under a simple nonlocal extension:

$$\delta A_{a\mu} = -\partial_\mu \theta_a + g f_{abc} \mathcal{E}_0 A^{\Lambda}_{b\mu} \theta^{\Lambda}_c \ . \tag{4.7a}$$

$$\delta h_{\mu\nu} = -\epsilon_{\mu,\nu} - \epsilon_{\nu,\mu} - \kappa \mathcal{E}_0 [(\epsilon^\rho{}_{,\mu})^\Lambda h^\Lambda_{\rho\nu} + (\epsilon^\rho{}_{,\nu})^\Lambda h^\Lambda_{\rho\mu} + h^\Lambda_{\mu\nu,\rho} (\epsilon^\rho)^\Lambda] \ . \tag{4.7b}$$

The physical problem comes at order g^2 (κ^2) with a breakdown of decoupling. This is restored by adding a new four-point interaction which is subject to the same restrictions as for QED: manifest Poincaré invariance, vanishing with infinite Λ, exponential convergence in Euclidean space, Hermiticity, and analyticity. Just as for QED, there are many such interactions, if there is one, and there is certainly one at this order. We expect that there will be a natural choice at each order, as there was for QED, in which the higher interactions simply extend the range of Schwinger parameter integration for diagrams generated by the interactions of the original nonlocalization (4.6). At any rate we will prove at the close of this section that there is at least one solution to all orders.

Now recall the argument from Sec. II A that decoupling implies some form of gauge invariance. Since we made no assumption about the S matrix, this result applies as well to any successful nonlocalization of Yang-Mills or gravity. In fact, the transformations can be built up perturbatively as we enforce decoupling. At Nth order the modification would be a term of the form $g^N \theta A^N$ ($\kappa^N \epsilon h^N$). Though we have not proved it, we suspect that the resulting algebra will be unchanged on shell as was the care for QED. We also expect that the final transformation will be interpretable in terms of nonlocal and field-dependent representation operators, also as in QED, and that there will be a similar mixing between gauge indices and spacetime ones.

Matter can be added if desired and the method is the same: First, nonlocalize by attaching appropriate factors of $\mathcal{E}_M$ to the fields of the interactions, then enforce decoupling by adding acceptable higher interactions, and finally infer the modified gauge transformation rule. However, one must have first obtained a successful nonlocalization for the pure gluon (graviton) theory. To see why, consider the process of enforcing decoupling on an amplitude with two external matter lines and N gluon (graviton) ones. This amplitude will receive crucial contributions from the attachment of pure gluon (graviton) diagrams with $n < N$ external lines to diagrams with two external matter lines and $(N-n+2)$ external gluon (graviton) ones. Indeed, such terms are required for decoupling even in the local theory. We cannot ignore them and we cannot account for them until after the pure gluon (graviton) theory has been decoupled, at least to the penultimate order. A similar effect occurs for gauge invariance since the matter transformation laws will depend nonlinearly upon the gluon (graviton) fields, whose transformation laws are unknown before a successful nonlocalization. We believe it will be possible to find transformations which are still linear in the matter fields.

Assuming a successful classical nonlocalization has been achieved, quantization is carried out in functional formalism, as for QED. One does this by finding a measure factor $\mu[gA]$ ($\mu[\kappa h]$) to make $[dA]$ ($[dh]$) invariant under the classical symmetry. If there are matter fields, we must add measure factors for them as well. Potential gauge anomalies will reveal themselves through the nonexistence of such factors, an approach pioneered by Fujikawa.[28] Assuming the matter transformation laws are still linear in the matter fields, the entire measure factor will be a functional of gA (κh), if it exists at all. Should the matter fields transform nonlinearly in themselves, then the measure factor will depend upon them as well.

An effect not apparent in QED is the more complicated paraphernalia of gauge fixing. Since the gluon (graviton) transformation becomes both nonlocal and nonlinear, so too will the Faddeev-Popov ghost action. Although it may seem an additional complication, this is really a very *desirable* feature in that it ensures that ghost loops receive the same exponential suppression as physical particle loops. Note also that we need waste no effort fretting over potential Gribov ambiguities, because these are nonperturbative and our nonlocalizations in any case succumb to more serious problems when pushed beyond perturbation theory.[3]

Though it seems perfectly possible, we have not yet attempted to obtain a classically successful nonlocalization for either Yang-Mills or gravity. Our proof that this can be done is therefore indirect. It consists of proper time regulating the one-loop amplitudes in a covariant gauge and then using the method of Ref. 29 to find a nonlocal classical action which gives these amplitudes and agrees on shell and at tree order with the local theory. (Recall from Sec. II C that there is no inconsistency between the usual tree amplitudes and finite-loop ones because nonlocal field theories do not obey Feynman's tree theorem.) We will sketch this procedure shortly, but let us first explore its consequences. Although the S matrix obtained from this action does not decouple at loop order, it obviously does so at the tree level, since it reproduces the usual tree amplitudes. Hence we may conclude that this action is the gauge-fixed version of an invariant one. If a measure factor can be found which makes the functional formalism invariant under the associated nonlocal gauge symmetry, then this invariance can be preserved in the quantum theory. The measure factor will induce corrections to the original regulated loop amplitudes which make them decouple. It will be obvious from the construction that these corrections prejudice neither finiteness nor manifest Poincaré invariance, and of course they *restore* perturbative unitarity on the physical space. It will also be obvious that these properties extend to all higher loops as well.

We will review the method of Ref. 29 in the context of a scalar field ϕ whose VEV is assumed to be zero. Let $S[\phi]$ be the possibly nonlocal Euclidean action and define the differential operator $F[\bar{\phi}](x)$ as

$$\frac{\delta^2 S[\bar{\phi}]}{\delta\phi(x)\delta\phi(y)} \equiv F[\bar{\phi}](x)\delta^4(x-y) \ . \tag{4.8}$$

A useful and elegant expression for the one-loop effective action is

$$\Gamma^{(1)}[\bar{\phi}] \equiv -\frac{1}{2}\hbar\int_0^\infty \frac{dt}{t}\mathrm{Tr}[\exp(-tF[\bar{\phi}])] \ . \tag{4.9}$$

Note that the ultraviolet divergences of local field theory are associated with the lower limit, while cut structure of perturbative unitarity derives from the upper limit. Note also that the one-loop contributions to all on-shell amplitudes can be obtained from $\Gamma^{(1)}[\phi_\infty]$, where ϕ_∞ is the general *classical* scattering solution defined in (2.18) and (2.19).[17,18] This suggests that we define a set of ultraviolet finite-loop amplitudes through the following on-shell effective action:

$$\Gamma_N^{(1)}[\phi_\infty] \equiv -\tfrac{1}{2}\hbar\int_{1/\Lambda^2}^\infty \frac{dt}{t}\mathrm{Tr}[\exp(-tF_L[\phi_\infty])] \ , \tag{4.10}$$

where F_L is defined according to (4.8) from a local action $S_L[\phi]$, and ϕ_∞ is the general scattering field of this action. The resulting amplitudes cut to give the trees of S_L.

To find the nonlocal classical action S_N which gives (4.10) *and* the trees of S_L, we define the operator $F_N[\phi_\infty]$ by the relation

$$\int_0^\infty \frac{dt}{t}\mathrm{Tr}[\exp(-tF_N[\phi_\infty])] \equiv \int_{1/\Lambda^2}^\infty \frac{dt}{t}\mathrm{Tr}[\exp(-tF_L[\phi_\infty])] \ . \tag{4.11}$$

Functional variation allows us to absorb the proper time integral:

$$\mathrm{Tr}\left[\frac{\delta F_{N\infty}}{F_{N\infty}}\right] = \mathrm{Tr}\left[\frac{\delta F_{L\infty}}{F_{L\infty}}\exp\left[-\frac{F_{L\infty}}{\Lambda^2}\right]\right] \ . \tag{4.12}$$

Adding the variation of an arbitrary trace-free operator $H[\phi_\infty]$ allows us to dispense with the trace:

$$\frac{\delta F_{N\infty}}{F_{N\infty}} = \frac{\delta F_{L\infty}}{F_{L\infty}}\exp\left[-\frac{F_{L\infty}}{\Lambda^2}\right] + \delta H_\infty \ . \tag{4.13}$$

Functional integration then gives a curious result in terms of our old friend the exponential integral:

$$F_{N\infty} = \exp\left[H_\infty - E_1\left[\frac{F_{L\infty}}{\Lambda^2}\right]\right] \ . \tag{4.14}$$

Note that F_N is an entire function of the derivative operator and also that the nonlocal kinetic operator $F_N[0]$ is just the local one $F_L[0]$ multiplied by a manifestly nonzero and entire function of $F_L[0]$.

One now constructs S_N from F_N. The first step is to choose the trace-free operator $H[\phi_\infty]$ to enforce on-shell integrability. What this means is that the functional

$$\frac{\delta}{\delta\xi(\mathbf{k}_1)}\int d^4x \frac{\delta\phi_\infty(x)}{\delta\xi(\mathbf{k}_2)}F_N[\phi_\infty]\frac{\delta\phi_\infty(x)}{\delta\xi(\mathbf{k}_3)} \tag{4.15}$$

is symmetric under any interchange of $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$. The next step is to extend $F_N[\phi_\infty]$ off shell so as to preserve this integrability. That is, we define $F_N[\bar{\phi}]$ for arbitrary $\bar{\phi}$ so that it agrees with $F_N[\phi_\infty]$ and causes the variation

$$\frac{\delta F_N[\bar{\phi}](y)}{\delta\bar{\phi}(x)}\delta^4(y-z) \tag{4.16}$$

to be symmetric under any interchange of x, y, and z. On-shell integrability implies that this can be done; in fact, there are many ways to do it. The final step is to functionally integrate F_N to obtain S_N. The resulting action reproduces (4.10) and agrees with the tree of S_L as a consequence of its manifest perturbative unitarity.

In fact, S_N is necessarily a nonlocal field redefinition of S_L:

$$S_N[\phi] = S_L[\Phi[\phi]] \ . \tag{4.17}$$

This guarantees that the two S matrices agree at tree order. One way of viewing the differences that appear in loops is by changing variables in the functional formalism for the local theory:

$$\int [d\phi]\exp(iS_L[\phi]) = \int [d\phi]\det\left[\frac{\delta\Phi}{\delta\phi}\right]\exp(iS_N[\phi]) \ . \tag{4.18}$$

The functional formalism for the nonlocal theory does not contain the change-of-variables Jacobian. Although there would need to be a measure factor for gauge theories, it would still not be interpretable as resulting from a change-of-variables Jacobian.

We call this method "proper time nonlocalization." It is a special case of the previous method which we might term "iterative nonlocalization." The classical actions produced by the proper time method are not simple. For example, even in QED they contain arbitrarily high orders of the Fermi fields. Section III proves that certain iterative solutions avoid this for QED, and we can probably avoid inducing higher matter couplings for Yang-Mills and gravity as well. Our reasons for describing the proper time method are, first, to prove the existence of solutions for gravity and Yang-Mills and, second, to emphasize that gauge and matter degrees of freedom can be treated symmetrically, if this is desired. Such a symmetric treatment is necessary in order to preserve a global supersymmetry.

V. CONCLUSIONS

We have described a method for distorting gauge theories and gravity into nonlocal field theories which are finite, Poincaré invariant, and perturbatively unitary—barring anomalies. The procedure has two stages, classical and quantum. In the first we make the theory finite by nonlocalizing its interactions. The concomitant disruption of decoupling is then repaired at each order by adding an appropriate new interaction. Since the resulting tree amplitudes decouple, it follows that the modified action possesses some form of gauge invariance. For QED this symmetry can be viewed on shell as a nonlocal and nonlinear representation of U(1). We believe that a similar interpretation will be possible for Yang-Mills and gravity.

A significant point is that enforcing decoupling at Nth order does not uniquely determine the new Nth-order interaction. It is ambiguous up to terms which decouple among themselves. However, since such terms induce their own violations of decoupling at higher orders, and since there is no guarantee that we can cancel these violations, it seems possible that making a poor choice might lead to an eventual breakdown of the process. That there is at least one consistent set of choices for QED was shown by exhibiting the model in Sec. III. The analogous result was obtained for Yang-Mills and gravity by an indirect method at the end of Sec. IV.

The quantum stage of our method consists of finding a measure factor to make the functional formalism invariant. This cannot always be done, although we were able to prove that a solution exists for QED. If the theory possesses a genuine anomaly, it will show up here.

That is what we actually did, but in retrospect an interesting alternative suggests itself: One might bypass decoupling and try to discover nonlocal gauge symmetries directly. For QED this would take the form of attempting to solve Eq. (2.3). The obvious ansatz is

$$V[eA](x,z)=\int d^4y\ A(y)\mathcal{T}[eA](x,y,z)\ . \qquad (5.1)$$

One would then search for representation operators which obeyed the resulting equation. It would be fascinating to extract the general condition for integrability. Any complete treatment of the representation theory would probably need to be approached from this perspective.

Nonlocalization can be viewed as a regularization of the original, local theory. Although it might seem to be a cumbersome one, what with all the higher interactions and measure factor corrections, we stress that is not necessarily so. In our explicit QED model, the higher interactions resulted in diagrams which merely extended the range of parameter integration on diagrams coming from the cubic vertex. Indeed, our result could almost be read off from the analogous result of dimensional regularization by replacing the Γ function with an incomplete Γ function. We expect that similarly simple solutions exist for gravity and Yang-Mills.

Our method reproduces the logarithmic divergences of dimensional regularization according to the correspondence (3.23), but we also see higher divergences which are lost to the automatic subtraction of dimensional regularization. Such a technique is essential when one wishes to avoid implicit renormalizations. In gravity this is particularly important, since the automatic subtraction may surreptitiously generate higher-derivative counterterms that one would never tolerate explicitly. One can also study canonical ordering problems using our method, though indirectly in the guise of the functional measure.

Another nice feature of our method is that it does not entail changing either the dimension of spacetime or the particle content, yet it preserves global Poincaré invariance. This strongly suggests that the method could be made to preserve global supersymmetry as well. The most promising candidate for this would seem to be the proper time method outlined at the end of Sec. IV.

As we have stated it, our method regulates noncoincident Green's functions, but not the vacuum energy. This is easily remedied by the field redefinition

$$\phi\rightarrow\mathcal{E}_\mu^{-1}\phi\ . \qquad (5.2)$$

This strips the convergence factors from the vertices and places them on the propagators. Polchinski has already regulated ϕ^4 theory in this manner.[4]

Nonlocal regularization derives much of its motivation from invariant string field theory. The magnitude of our debt can be seen from the persistent failure of previous attempts,[1,11] which were undertaken without the benefit of knowning how string field theory resolves the key problems of enforcing tree-order decoupling and of quantization. But all is not well with our paradigm. We are referring, of course, to the great unsolved problem of quantizing the new invariant closed-string action.[14] The many similarities between string field theory and our nonlocalizations suggest that it is now time for insight to flow the other way.

An obvious advantage our models possess is that they involve only a finite number of fields. One can therefore distinguish the effects of nonlocality from those due to the infinite number of component fields. Our models obey the Cutkosky rules; the only singularities come from exchange propagators going on shell. In string field theory singularities can also arise from the sum over component fields. These two effects are hopelessly entangled in the measure factor of closed-string field theory. For example, the fact that there even *is* such a factor derives entirely from nonlocality. But it is the infinite number of fields which explains why choosing the measure factor to enforce gauge invariance at one loop restores neither perturbative unitarity nor even finiteness. A possibly useful insight from our work is that invariance alone does not uniquely determine the measure factor. Might it not be possible to exploit this freedom to make the higher loops come out right without additional corrections?

Another suggestive insight comes from the correspondence between our finite model of QED and Efimov's[26] incompletely regulated one. Because the two S matrices agree at tree level, there must be a field redefinition which relates the actions on shell, but this transformation cannot be fully invertible because our action possesses a larger off-shell symmetry than Efimov's. One way of describing the problem of quantizing Efimov's theory is that the wrong off-shell symmetry is being enforced. Off-shell physics can matter for nonlocal theories because they do not generally obey Feynman's tree theorem. The symptom of having obtained the off-shell action incorrectly is discrepancies in loop amplitudes even when the trees are correct. Enforcing the wrong off-shell symmetry means that terms proportional to the field equations are missing from the measure factor. If we include these and then absorb them into field redefinitions, the result is a measure factor containing higher powers of $\hbar$. This is precisely what is seen in unitarizing invariant closed-string field theory.[30]

A related issue is gauge invariance. Our symmetries exhibit a peculiar mixing of spacetime and group indices which is very reminiscent of the symmetries one encounters in string field theory.[13] This is often taken as a reflection of the inherent string unification between gauge interactions and gravitational ones. The presence of a similar phenomenon in our model of QED puts the lie to such a simple view, although the appearance of general coordinate invariance among the string symmetries undoubtedly *is* significant. Another point of similarity between our models and invariant string field theory is the frustrating absence of a gauge-covariant directional derivative. This makes constructing gauge-invariant observables very difficult, and it is conceivable that we can learn something of relevance to string theory by solving the much simpler problem in nonlocal QED.

Nonlocal regularization is unique in that it produces a regulated theory which is perturbatively viable provided that attention is restricted to the on-shell S matrix. No other method does this, and the temptation is well nigh irresistible to consider elevating our nonlocalizations to the status of candidate fundamental theories. A case could even be made that our models are better than their local ancestors. For example, the nonlocal extension of QED described in Sec. III is not only finite, it is also free of the Landau ghost and possesses chiral invariance in the massless limit without fermion doubling.

We take a cautious view of this sort of claim owing to the same nonperturbative instabilities which appear, for precisely the same reason, in string theory.[3] Even on the perturbative level one encounters off-shell noncausality in any field which interpolates the usual S matrix. However, if one chooses to ignore the analogous problems of string theory, then we insist that our nonlocalizations must be taken seriously as well. There is nothing wrong with them that is not also wrong with strings.

The only case that *we* feel merits serious consideration is when the nonlocal action is a field redefinition of a local one. Microcausality is necessarily forfeit in this way,[31] but nonlocal instabilities can be avoided if the classical agreement between the two theories is nonperturbative. (We would perhaps remind the reader that "perturbation theory" refers to an asymptotic expansion in some coupling constant and can be applied as easily to building up classical solutions as to computing scattering amplitudes.) The proper time method of Sec. IV generates models which agree classically with local theories on the perturbative level. If this agreement were somehow to persist in the nonperturbative regime, these models might conceivably be acceptable as fundamental theories.

ACKNOWLEDGMENTS

We wish to thank P. Ramond, A. Strominger, and C. B. Thorn for stimulating conversations. This work was partially supported by the Natural Sciences and Engineering Research Council of Canada and by DOE Contract No. DE-FG05-86ER-40272.

---

[1]G. Wataghin, Z. Phys. **86**, 92 (1934); H. McManus, Proc. R. Soc. London **A195**, 323 (1948); R. P. Feynman, Phys. Rev. **74**, 939 (1948); **74**, 1430 (1948); A. Pais and G. E. Uhlenbeck, *ibid.* **79**, 145 (1950); P. Kristensen and C. Møller, K. Dan Vidensk. Selsk. Mat. Fys. Medd. **27**, 668 (1952); M. Chrétien and R. Peierls, Proc. R. Soc. London **A223**, 468 (1954).

[2]D. J. Gross and A. Jevicki, Nucl. Phys. **B283**, 1 (1987); E. Cremmer, A. Schwimmer, and C. B. Thorn, Phys. Lett. B **179**, 57 (1986).

[3]D. A. Eliezer and R. P. Woodard, Nucl. Phys. **B325**, 389 (1989).

[4]J. Polchinski, Nucl. Phys. **B231**, 269 (1984).

[5]S. Mandelstam, Nucl. Phys. **B213**, 149 (1983).

[6]G. 't Hooft and M. Veltman, Nucl. Phys. **B44**, 189 (1972).

[7]J. S. Dowker and R. Critchley, Phys. Rev. D **13**, 224 (1976); **13**, 3224 (1976); **16**, 3390 (1977).

[8]W. Pauli and F. Villars, Rev. Mod. Phys. **21**, 434 (1949).

[9]Of course, no method which leaves the theory with an infinite number of degrees of freedom can provide a strong operator regularization. The nonlocal regularization of (1.1) suffices to make all the usual Green's functions finite, but one can still find operators with divergent VEV's.

[10]J. Schwinger, Phys. Rev. **82**, 664 (1951).

[11]C. Hayashi, Prog. Theor. Phys. **10**, 533 (1953); **11**, 226 (1954); R. Marnelius, Phys. Rev. D **8**, 2472 (1973); **10**, 3411 (1974).

[12]X. Jaén, J. Llosa, and A. Molina, Phys. Rev. D **34**, 2302 (1986).

[13]E. Witten, Nucl. Phys. **B268**, 253 (1986); **B276**, 291 (1986).

[14]M. Saadi and B. Zwiebach, Ann. Phys. (N.Y.) **192**, 213 (1989); T. Kugo, H. Kunitomo, and K. Suehiro, Phys. Lett. B **226**, 48 (1989); Nucl. Phys. **B337**, 434 (1990).

[15]J. W. Moffat, Phys. Rev. D **41**, 1177 (1990).

[16]S. Mandelstam, Phys. Rep. **13**, 259 (1974).

[17]B. S. DeWitt, in *Relativity Groups and Topology II*, proceedings of the Les Houches Summer School, Les Houches, France, 1983, edited by B. S. DeWitt and R. Stora (Les Houches Summer School Proceedings, Vol. 40) (North-Holland, Amsterdam, 1984).

[18]M. Srednicki and R. P. Woodard, Nucl. Phys. **B293**, 612 (1987).

[19]N. C. Tsamis and R. P. Woodard, Class. Quantum Grav. **2**, 841 (1985).

[20]One could quibble over imaginary terms that vanish on shell. This usually only postpones the breakdown of unitarity to a higher tree amplitude or to a loop. In any case it is *safer* to insist upon real, analytic interactions, and we have been able to find at least one solution which obeys this restriction.

[21]W. Pauli, Nuovo Cimento **10**, 648 (1953).

[22]To see this, recall that the functional integral of an operator gives the VEV of its T^*-ordered product. Now consider a series of N local operators evaluated on a surface of simultaneity. Because they are spacelike separated, their temporal sequence is not Lorentz invariant; hence the time-ordered product varies as we change frames. In a microcausal theory this reordering is irrelevant because any two spacelike separated operators commute; hence we can freely commute the N operators to achieve a standard order $\Phi_1\Phi_2\cdots\Phi_N$. In a fundamentally nonlocal theory, microcausality is forfeited (Refs. 3 and 21). Spacelike separated operators do not generally commute, and the price of imposing our standard or-

dering upon the time-ordered product is a frame-dependent commutator term. If the result in one frame gives the standard ordering

$$T^*(\Phi_1 \cdots \Phi_N) = \Phi_1 \cdots \Phi_N ,$$

then a Lorentz transformation can be found for which the temporal sequence is different and

$$T^*(\Phi'_1 \cdots \Phi'_N) = \Phi'_1 \cdots \Phi'_N + (\text{nonzero, frame dependent commutators}) .$$

Now take the VEV of both equations. The two left-hand sides are Green's functions; if we quantize so that they are Lorentz invariant, it follows that the VEV of the N operators in standard order cannot be invariant. Hence the operator formalism lacks manifest Lorentz invariance. Similarly, if we make the operator formalism invariant, it follows that the functional formalism cannot be.

[23]R. P. Feynman, Acta Phys. Pol. **24**, 697 (1963); in *Magic Without Magic,* edited by J. Klauder (Freeman, New York, 1972), p. 355.

[24]One also sometimes hears the argument that dispersion relations imply the equality of two S matrices which agree at tree order. This is not so. Dispersion relations only determine the next-order amplitudes up to entire functions. Since all one-loop ultraviolet divergences reside on precisely such terms, it is perfectly possible for the amplitudes of one theory to be finite, while those of the other diverge. The proper time nonlocalization discussed at the end of Sec. IV results in precisely this situation.

[25]There is also a field redefinition which carries it to $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ plus (1.4).

[26]G. V. Efimov, Fiz. Elem. Chastits At. Yadra [Sov. J. Part. Nucl. **5**, 89 (1974)].

[27]H. B. Nielsen and M. Ninomiya, Nucl. Phys. **B185**, 20 (1981).

[28]K. Fujikawa, Phys. Rev. D **21**, 2848 (1980).

[29]R. P. Woodard, Phys. Lett. B **213**, 144 (1988).

[30]H. Hata, Nucl. Phys. **B329**, 698 (1990); **B339**, 663 (1990).

[31]It has become orthodox among string theorists to dismiss any concern over the loss of microcausality as an example of atavistic obscurantism. In fact, the effect, which is a necessary consequence of imposing any sort of initial value formalism upon a nonlocal theory, poses a considerable problem of principle. It is perhaps easier to perceive this in our models than in string theory where the discussion tends to become mired in the mystique of extended objects. One might reconcile the usual time travel paradoxes by denying our ability to explore them with sufficiently localized observables. It is more difficult to understand what it means to formulate a nonlocal theory in a finite volume or to accept the fact that joining two such volumes gives a slightly different result than one obtains from formulating the theory directly on the larger volume.