PHYSICAL REVIEW D VOLUME 24, NUMBER 12 15 DECEMBER 1981

Perturbation theory for undoubled lattice fermions

Jeffrey M. Rabin
Stanford Linear Accelerator Center, Stanford University, Stanford, California 94305
(Received 24 March 1981)

I study a class of lattice versions of QED with fermions with the aim of clarifying the relationship between the fermion spectrum (doubled or not), continuous chiral symmetry, the range of the lattice interactions, and the validity of perturbation theory. Theories of this class with an undoubled spectrum, such as the formulation due to Drell, Weinstein, and Yankielowicz (DWY), have infinite-range interactions and infrared-divergent perturbation series. The infrared singularities can be removed by a resummation of the series. I then carry out a renormalization program for DWY QED after which the $a \to 0$ limit of the theory is finite and coincides with continuum QED. Finally, I consider the nonperturbative structure of DWY QED.

I. INTRODUCTION

Rigorous formulation of a continuum quantum field theory normally involves defining the theory as a singular limit of a cutoff or regularized theory. In perturbation theory many satisfactory regularization schemes exist, including Pauli-Villars, dimensional regularization, and others. However, for nonperturbative studies of gauge theories, interest has focused on the lattice regularization, which has the virtue of preserving exact local gauge invariance. Block-spin renormalization group,[1] Monte Carlo,[2] and rigorous mathematical methods[3] have provided a great deal of information concerning the phase structure and continuum limit of pure gauge theories on a lattice.

The extension of lattice techniques to realistic theories such as quantum chromodynamics (QCD) has been hindered by uncertainty regarding the proper treatment of lattice fermions. Straightforward transcription of the Dirac equation to the lattice by replacing derivatives by nearest-neighbor differences leads to the so-called spectrum-doubling problem: the continuum limit of the lattice Dirac equation describes 2^d fermions rather than just one, where d is the number of dimensions of space-time which are in the lattice formulation. Of the many proposed solutions for this problem, two will be discussed in this paper. Wilson[4] adds a term with no γ-matrix structure to the lattice Dirac equation. This term functions as a momentum-dependent "mass," giving the extra fermions masses on the order of the cutoff and removing them from the spectrum in the continuum limit. As an additional mass term, it also destroys the global chiral symmetry of the Dirac theory at $m=0$. The method of Drell, Weinstein, and Yankielowicz[5] (DWY) yields the correct fermion spectrum and preserves chiral symmetry by transcribing the continuum derivative as a nonlocal lattice difference operator. The definition is such that in momentum space the lattice derivative acts as multiplication by ip_μ.

Clearly, if the spectrum-doubling problem is connected with chiral symmetry, then it must be fully understood before lattice methods can give reliable information about the symmetry structure of QCD. Indeed, an important issue connected with chiral symmetry in any gauge theory is the axial anomaly. Any lattice gauge theory with continuous chiral symmetry must answer the following question. In consequence of the continuous symmetry, there will be a conserved axial-vector current on the lattice. The naive manipulations leading to a nonanomalous Ward identity for this current are valid in the presence of the lattice regularization. Does the continuum limit of this current exist? If so, does that not yield a continuum axial-vector current with no anomaly, and is that not impossible?

The straightforward transcription of the Dirac equation answers this question by doubling the spectrum: the anomaly is canceled between the different fermion species.[6] The Wilson formulation answers by explicitly breaking the lattice chiral symmetry. An extra term appears in the Ward identity and becomes the anomaly in the continuum limit.[6,7] In this paper I will show that the DWY theory encounters infrared divergences in perturbation theory which need careful treatment. Order by order the continuum limit of the conserved axial-vector current does not exist due to these infrared divergences.

It is becoming generally recognized that an undoubled spectrum, continuous chiral symmetry, and locality of interactions are incompatible though desirable properties of a lattice fermion scheme. Indeed, in the literature one can find the claim[6,8] that a lattice fermion theory with undoubled spectrum and continuous chiral symmetry is itself impossible, although the arguments in support of these claims involve additional assumptions. One purpose of the present work is to clarify the relations between these three properties of lattice fermion schemes.

Before using a particular regularization scheme for nonperturbative investigations, one would like to have confidence that it yields acceptable results in the familiar context of perturbation theory. Sharatchandra[9] has shown that Wilson's formulation of QED on a four-dimensional Euclidean lattice passes this test. He showed that in perturbation theory a multiplicative renormalization of fields and parameters suffices to remove all divergences in the $a \to 0$ limit of the S matrix, which then agrees with the S matrix of continuum QED. The main purpose of this paper is to give the corresponding analysis for the DWY version of QED. In this case multiplicative renormalization does not suffice: additional counterterms are required. This is to be expected, since once long-range interactions are admitted the DWY Lagrangian is by no means the most general one consistent with its symmetries. The analysis, like Sharatchandra's, should extend to QCD as well.

Perturbation theory with DWY lattice fermions has been studied by Karsten and Smit in the four-dimensional Euclidean lattice formulation.[6,10,11] They computed both the one-loop vacuum polarization and the *VVA* triangle diagrams. They concluded that the axial-vector current did not develop an anomaly in the continuum limit. Its matrix elements, along with the vacuum polarization, were nonlocal, not Lorentz covariant, and infrared singular in the continuum limit. Furthermore, the theory appeared nonrenormalizable in that infinitely many Green's functions were superficially divergent. Nakawaki[12] reached similar conclusions from a study of the DWY theory in Hamiltonian form. In this paper I show that the perturbation expansion of Karsten and Smit breaks down owing to the infrared singularities. I describe a resummation of the perturbation series which removes these divergences, and carry out a renormalization program to all orders of the modified expansion. The renormalized Green's functions at each order in this expansion go over, for $a \to 0$, to those of continuum QED to the same order.

The paper is organized as follows. Section II reviews the fermion-doubling problem and explores the reasons it occurs. The DWY solution to the problem is discussed, and the "topological" connection between spectrum doubling, chiral symmetry, and the range of interactions is explained. In Sec. III I summarize Sharatchandra's arguments for the renormalizability of Wilson's lattice QED, which form the basis for the arguments I subsequently apply to the DWY theory. In Sec. IV I show how continuum QED in a fixed gauge can be faithfully transcribed onto a lattice. The DWY derivative and long-range interactions appear automatically. Although this is not the DWY lattice gauge theory which has been discussed in the literature, it provides a simple counterexample to the claim that no lattice version of QED with undoubled spectrum and continuous chiral symmetry is possible. Section V begins the discussion of the DWY lattice gauge theory studied by Karsten and Smit. I derive the Feynman rules, check the classical continuum limit of the Lagrangian, and exhibit the conserved currents and Ward identities. The theory appears nonrenormalizable by power counting. However, the perturbation expansion is shown to be invalid due to infrared divergences which arise as a direct consequence of having an undoubled fermion spectrum. The summation of tadpole diagrams is shown to remove both the infrared divergences and the problems with power counting. Section VI begins the discussion of renormalization. The obstacle to direct application of Sharatchandra's methods is the inability to expand integrands in powers of external momenta. I divide the integrals into subregions, in each of which the Taylor expansion in external momenta is possible. I then give the prescription for order-by-order construction of counterterms, and show that in the presence of the counterterms the $a \to 0$ limit gives ordinary continuum QED. Section VII supplements this rather abstract discussion by applying the renormalization prescription to one- and two-loop examples. Although detailed calculations are not carried out, the form of the necessary counterterms is clarified. I consider to what extent the counterterms can be generated by rescaling fields and parameters. Finally, I show that in the renormalized perturbation expansion the conserved axial-vector current still has divergent matrix elements. These can be made finite by redefining the current, at the cost of introducing the usual anomaly. Section VIII summarizes the conclusions and points out remaining problems. In particular I consider whether the properties of the DWY lattice gauge theory established in perturbation theory will persist in the exact nonperturbative solution.

Notation: The Einstein summation convention is not used in this paper. Summations will be indicated explicitly.

II. LATTICE FERMIONS

This section reviews the spectrum-doubling problem of lattice fermions and motivates its solution via the "DWY derivative".

Consider the Klein-Gordon equation for a scalar field,

$$\left(\frac{\partial^2}{\partial t^2}-\nabla^2+m^2\right)\phi(\vec{x},t)=0\,. \qquad (2.1)$$

The problem of transcribing this equation onto a three-dimensional lattice with continuous time is solved by making ϕ a function on lattice sites indexed by $\vec{x}$ and replacing ∇^2 by an appropriate difference operator. Plausible choices are $\nabla_+{}^2$, $\nabla_-{}^2$, $\nabla_\pm{}^2$, and $\vec{\nabla}_-\cdot\vec{\nabla}_+$, where

$$\nabla_+^i f(\vec{x})=\frac{1}{a}[f(\vec{x}+\vec{a}^i)-f(\vec{x})]\,, \qquad (2.2a)$$

$$\nabla_-^i f(\vec{x})=\frac{1}{a}[f(\vec{x})-f(\vec{x}-\vec{a}^i)]\,, \qquad (2.2b)$$

$$\nabla_\pm^i f(\vec{x})=\frac{1}{2a}[f(\vec{x}+\vec{a}^i)-f(\vec{x}-\vec{a}^i)]\,. \qquad (2.2c)$$

(In this paper the variable $\vec{x}$ indexing lattice sites will always carry dimensions: $x_i=n_i a$ where a is the lattice spacing and n_i is an integer. $\vec{a}^i$ is a vector of length a in the i direction.) The spectrum of the lattice Klein-Gordon equation is found by seeking solutions of the form

$$\phi(\vec{x},t)\sim e^{-iEt}e^{i\vec{k}\cdot\vec{x}}\,,$$

leading to the dispersion relations

$$\nabla_+{}^2:\;\; E^2=\frac{4}{a^2}\sum_i e^{ik_i a}\sin^2\tfrac{1}{2}k_i a+m^2\,, \qquad (2.3a)$$

$$\nabla_-{}^2:\;\; E^2=\frac{4}{a^2}\sum_i e^{-ik_i a}\sin^2\tfrac{1}{2}k_i a+m^2\,, \qquad (2.3b)$$

$$\nabla_\pm{}^2:\;\; E^2=\frac{1}{a^2}\sum_i \sin^2 k_i a+m^2\,, \qquad (2.3c)$$

$$\vec{\nabla}_-\cdot\vec{\nabla}_+:\;\; E^2=\frac{4}{a^2}\sum_i \sin^2\tfrac{1}{2}k_i a+m^2\,. \qquad (2.3d)$$

(On an infinite lattice k_i is a continuous variable which can be chosen to run from $-\pi/a$ to $+\pi/a$. The notation $\Lambda=\pi/a$ will sometimes be used.)

All these expressions reduce to the usual continuum dispersion relation when $a\to 0$ with $\vec{k}$ fixed. However, $\vec{\nabla}_+$ and $\vec{\nabla}_-$ are not Hermitian: the energy in Eqs. (2.3a) and (2.3b) is not real. The remaining possibilities differ only in the period of the sine functions. Equation (2.3d) has the $2\pi/a$ periodicity of the lattice while Eq. (2.3c) has period π/a. This signals spectrum doubling. For an acceptable spectrum only the spatially constant ($\vec{k}=0$) solution should minimize the energy. For Eq. (2.3c) this solution is degenerate with seven others having $k_i=\pi/a$ for some values of i (ϕ alternates sign in some lattice directions). About each of these solutions there is a band of long-wavelength excitations, resulting in eight low-lying particle states in the continuum limit compared to one for Eq. (2.3d).

It is not coincidental that Eq. (2.3d) alone is satisfactory. The gradient of a function $f(\vec{x})$ on lattice sites is naturally defined as the function on links which is the sum (with sign changes for the orientation of the link) of the values of f at the sites bounding a given link. This is $\nabla_+^i f$. The divergence of a function $f_i(\vec{x})$ on links is a function on sites given by the sum (with sign changes for orientation) of the values of f_i on links impinging on a given site. This is $\sum_i \nabla_-^i f_i$. Hence the Laplacian is naturally given by $\vec{\nabla}_-\cdot\vec{\nabla}_+$. The different derivatives represent the lattice boundary and coboundary operators,[3] which are not equal.

From a more abstract point of view, what is happening is the following. Associated with a scalar, vector, or antisymmetric tensor field there is a differential 0-form, 1-form, or 2-form. A rotationally covariant differential operator acting on the field can be expressed in terms of the exterior differential operators d and δ acting on the form. A natural lattice formulation is available by associating n-forms with n-cochains (functions on sites, links, or plaquettes for $n=0,1,2$) and d and δ with the boundary and coboundary operators represented here by ∇_+^i and $\vec{\nabla}_-\cdot$. The problems with fermions arise because they fall into spinor rather than tensor representations of the rotation group and so have no associated n-forms.

Consider now the Dirac equation,

$$(i\gamma\cdot\partial-m)\psi=0\,, \qquad (2.4)$$

which is seen to have the same dispersion relation as the Klein-Gordon equation by applying $i\gamma\cdot\partial+m$ to both sides. Assume this equation is to be put on the lattice by substituting a difference operator for the spatial derivatives, ψ being defined at lattice sites. This assumption is by no means necessary, but it does guarantee that the lattice Dirac equation will have the usual chiral invariance when $m=0$. The fermion dispersion relation will be that of the Klein-Gordon equation whose Laplacian is the *square* of the Dirac difference operator. The acceptable dispersion relation (2.3d) cannot be obtained.

The Dirac equation requires a Hermitian difference operator whose square is an acceptable Laplacian. DWY[5] achieve this in terms of the Fourier transform of a lattice function $f(\vec{x})$,

$$\tilde{f}(\vec{p})=a^3\sum_{\vec{x}} e^{-i\vec{p}\cdot\vec{x}}f(\vec{x})\,,$$
$$f(\vec{x})=\frac{1}{(2\pi)^3}\int_{-\Lambda}^{\Lambda} d^3p\, e^{i\vec{p}\cdot\vec{x}}\tilde{f}(\vec{p})\,, \qquad (2.5)$$

by defining $\nabla_j f(\vec{x})$ as the inverse transform of $ip_j\tilde{f}(\vec{p})$. This leads to the exact relativistic spectrum $E^2=p^2+m^2$ on the lattice. In coordinate space the definition is

$$\nabla_j f(\vec{x})=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{na}[f(\vec{x}+n\vec{a}_j)-f(\vec{x}-n\vec{a}_j)]\,. \quad (2.6)$$

The nonlocality of this operator is essential for avoiding the spectrum doubling. Indeed, a general derivative operator may be written

$$\nabla_j f(\vec{x})=\sum_{\vec{y}} D_j(\vec{x}-\vec{y})f(\vec{y})\,, \quad (2.7a)$$

with Fourier transform

$$\nabla_j \tilde{f}(\vec{p})=i\tilde{D}_j(\vec{p})\tilde{f}(\vec{p})\,, \quad (2.7b)$$

where the factor i has been extracted for convenience. The fermion dispersion relation will be

$$E^2=\sum_j \tilde{D}_j^{\,2}(\vec{p})+m^2\,, \quad (2.8)$$

and spectrum doubling occurs if $\sum_j \tilde{D}_j^{\,2}(\vec{p})=0$ has solutions other than $\vec{p}=2\vec{n}\pi/a$. Usually $\tilde{D}_j(\vec{p})=\tilde{D}(p_j)$ is a function of p_j alone, but in any case one can fix $p_i=0$, $i\neq j$, and study the function $\tilde{D}(p_j)=\tilde{D}_j(\vec{p})$ alone. Hermiticity requires $\tilde{D}$ to be real, a satisfactory continuum limit requires $\tilde{D}(p_j)\to p_j$ as $a\to 0$ with p_j fixed, and on general grounds $\tilde{D}$ has period $2\pi/a$. It is evident from Fig. 1 that if $\tilde{D}$ is continuous, it has at least one zero for $0<p_j<2\pi/a$, with a band of low-lying states around this zero to become an extra fermion in the continuum limit. The DWY derivative (Fig. 2) escapes this conclusion due to its discontinuity at $p_j=\pi/a$. One recalls the Fourier coefficients of a discontinuous function fall off as $1/n$ or slower, so $D_j(\vec{x}-\vec{y})$ is necessarily nonlocal. This argument, which also appears in Ref. 6, is a simple and intuitive case of the more general topological theorem of Ref. 8.

It is amusing to note that, because a Fourier series converges to the mean at a point of discontinuity, the DWY function $\tilde{D}(p_j)=p_j$ for $p_j\in(-\pi/a,\pi/a)$, extended periodically, does have a zero at $p_j=\pi/a$. However, there is no band of low-lying states surrounding this point.

It is quite possible for $\tilde{D}(p_j)$ to have more than two zeros. The choice

$$\nabla_j f(\vec{x})=\frac{1}{4a}[f(\vec{x}+2\vec{a}_j)-f(\vec{x}-2\vec{a}_j)]\,,$$

for example, leads to "spectrum quadrupling."

It should be evident from this discussion that there are interesting geometric and topological issues connected with lattice fermions. Further research along these lines is in progress.

III. WILSON'S LATTICE QED

This section reviews Wilson's[4] lattice formulation of QED and Sharatchandra's[9] conclusions concerning its perturbative renormalizability. The method of Sharatchandra's proof is summarized in some detail since it provides a canonical set of arguments for establishing the perturbative equivalence of lattice and continuum theories. The analysis of the DWY lattice QED formulation in this paper will be based heavily on these arguments.

Throughout this paper, detailed discussions of lattice perturbation theory will be carried out in the four-dimensional Euclidean, rather than the Hamiltonian, formalism. This makes available the technical conveniences of the straightforward path-integral quantization and manifest symmetry between time and space coordinates characteristic of this formalism.

Wilson's lattice QED action is

$$I=a^4\sum_{x,\mu,\nu}\tfrac{1}{4}F_{\mu\nu}^{\,2}(x)+a^4\sum_{x,\mu}\frac{1}{2\lambda}[\nabla_\mu^- A_\mu(x)]^2-a^4\sum_{x,\mu}\bar{\psi}(x)\frac{1}{i}\gamma_\mu\frac{1}{2a}[\psi(x+a_\mu)e^{ieaA_\mu(x)}-\psi(x-a_\mu)e^{-ieaA_\mu(x-a_\mu)}]$$

$$-a^4\sum_{x,\mu}\bar{\psi}(x)\frac{1}{2a}[\psi(x+a_\mu)e^{ieaA_\mu(x)}+\psi(x-a_\mu)e^{-ieaA_\mu(x-a_\mu)}-2\psi(x)]-a^4\sum_x m\bar{\psi}(x)\psi(x)\,, \quad (3.1)$$

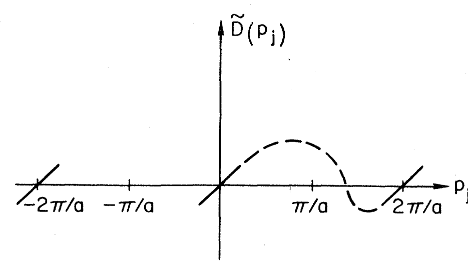

FIG. 1. General behavior of a continuous function $\tilde{D}(p_j)$ appearing in the fermion dispersion relation, illustrating the necessity of spectrum doubling.

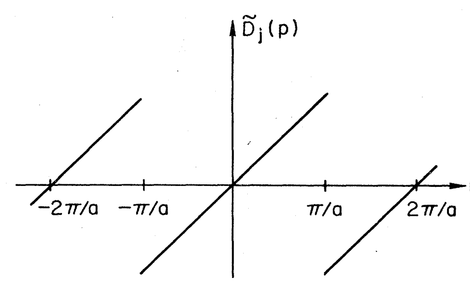

FIG. 2. The DWY derivative $\tilde{D}_j(p)$, which avoids spectrum doubling by virtue of discontinuities at $\pm\pi/a$.

where $F_{\mu\nu}(x)=\nabla_\mu^+ A_\nu(x)-\nabla_\nu^+ A_\mu(x)$. (The γ-matrix convention is $\{\gamma_\mu,\gamma_\nu\}=-2\delta_{\mu\nu}$.)

For $e=0$, the free fermion action is constructed using the derivative $\nabla_\mu^\pm$ and would therefore yield a doubled spectrum if not for the additional terms in the third line of Eq. (3.1). In momentum space these terms read

$$-\left(\frac{2}{a}\sin^2\tfrac{1}{2}p_\mu a\right)\bar{\tilde{\psi}}(p)\tilde{\psi}(p)\,,$$

and they vanish for $p\to 0$ or for $a\to 0$ with p fixed. However, they give masses of order Λ to the extra fermions with $p_\mu=\pi/a$, removing them from the spectrum in the continuum limit. They also explicitly break chiral symmetry, as is appropriate for a mass term.

The coupling to the gauge field is introduced in a manner consistent with invariance under the local gauge transformations

$$\begin{aligned}&\psi(x)\to e^{-ie\chi(x)}\psi(x)\,,\\&A_\mu(x)\to A_\mu(x)+\nabla_\mu^+\chi(x)\,.\end{aligned}\tag{3.2}$$

The second term in Eq. (3.1) serves to fix a "covariant" gauge. The form of the photon kinetic energy (not periodic in A_μ) identifies this as the noncompact formulation of QED; the compact formulation would replace $F_{\mu\nu}{}^2(x)$ by

$$-\frac{2}{e^2a^4}\left(e^{iea^2F_{\mu\nu}(x)}-1\right).$$

Finally, note that the lattice derivatives in Eqs. (3.1) and (3.2) are used "naturally" in the sense of Sec. II: ∇_μ^+ is used to create the plaquette variable $F_{\mu\nu}$ from the link variable A_μ while ∇_μ^- forms the scalar divergence of the vector A_μ.

Expanding the exponentials in Eq. (3.1) and introducing the Fourier-transformed fields permits one to read off the Feynman rules from the coefficients of the terms in the action. For the photon field it is convenient to define the Fourier transform by

$$A_\mu(x)=\int_{-\Lambda}^{\Lambda}\frac{d^4p}{(2\pi)^4}e^{ip\cdot(x+a_\mu/2)}\tilde{A}_\mu(p)\,,\tag{3.3}$$

so as to get real expressions for propagators and vertex functions. For example, the Fourier transform of $-i\nabla_\mu^- A_\mu(x)$ will be $(2/a)\sin(\frac{1}{2}p_\mu a)\,\tilde{A}_\mu(p)$ rather than $(1/ia)(1-e^{-ip_\mu a})\tilde{A}_\mu(p)$. Some of the resulting Feynman rules are given in Fig. 3.

Strictly speaking, the Feynman rules require an integration over the momentum of *each* internal line. In the continuum theory, many of these integrations are trivial because of the momentum-conserving δ functions. On the lattice, however, one has at each vertex a factor

$$a^4\sum_x \exp\left[i\left(\sum k\right)\cdot x\right]=(2\pi)^4\delta^4_{\rm per}\left(\sum k\right),$$

where

$$\delta_{\rm per}(q)\equiv\sum_{n=-\infty}^{\infty}\delta(q+2n\Lambda)\,.\tag{3.4}$$

It is shown in the Appendix that because the Feynman integrands are themselves periodic functions of momenta, the trivial integrations can still be done. Thus even on the lattice one can label Feynman graph lines with exactly conserved momenta and perform nontrivial integrations only over a set of loop momenta.

Sharatchandra showed that this set of Feynman rules defines a multiplicatively renormalizable lattice QED: fields and parameters can be rescaled so that when $a\to 0$ the Green's functions are finite and identical to those of ordinary QED. (In fact, Sharatchandra considered compact QED, which is technically more complicated.) This is demonstrated in four steps.

(1) The Feynman rules reduce to the continuum Feynman rules when $a\to 0$ with momenta fixed. Since the Feynman rules reflect the momentum-space coefficients in the action, this merely means that the action has the correct classical continuum limit. However, it does imply that if a normal diagram (one containing no multiphoton vertices) converges as $a\to 0$, it agrees with the continuum result for the diagram.

(2) The list of primitively divergent diagrams and their superficial degrees of divergence 0 (for $a\to 0$) coincides with the list for continuum QED. For normal diagrams this can be shown by bounding lattice quantities by continuum quantities. For example, for the photon propagator,

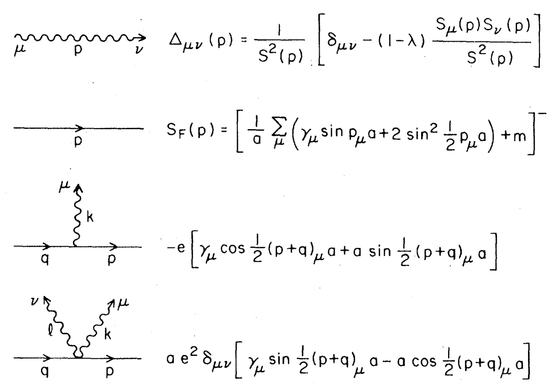

FIG. 3. Some of the Feynman rules for Wilson's lattice QED. There are n-photon vertices for all $n>0$.

$$S_\mu(p)\equiv\left(\frac{2}{a}\right)\sin\tfrac{1}{2}p_\mu a\,.$$

$$\frac{2}{\pi}p_\mu \leq \frac{2}{a}\sin\tfrac{1}{2}p_\mu a \leq p_\mu\,, \quad 0\leq p_\mu \leq \frac{\pi}{a}\,,$$

implies

$$\frac{1}{p^2}\leq \frac{1}{\sum_\mu (4/a^2)\sin^2(\frac{1}{2}p_\mu a)} \leq \frac{\pi^2}{4}\frac{1}{p^2}\,, \quad |p_\mu| \leq \frac{\pi}{a}\,.$$

Now imagine shrinking some internal fermion propagator to a point in a normal diagram. The loss of this propagator increases D by one unit, but a two-photon vertex is created which carries an explicit factor a according to Fig. 3. Hence D is unchanged. This argument generalizes to show that the presence of multiphoton vertices does not interfere with power counting.

(3) All Feynman integrands possess Taylor expansions in powers of their external momenta. Ignoring infrared problems, e.g., by assuming a photon mass, this means that the Bogolubov-Parasiuk-Hepp (BPH) procedure of subtracting the first $D+1$ terms in the Taylor expansions of divergent subgraphs, with combinatorics handled by a forest formula, can be implemented. It follows from point (1) that normal diagrams take on their continuum values when $a\to 0$ after the subtractions are done. If a divergent subgraph contains a multiphoton vertex then it has the form a^N times an integral of $O(1/a^{N/D})$, $N\geq 1$. After $D+1$ subtractions this becomes $a^N O(1/a^{N-1})$, so all such diagrams vanish when $a\to 0$.

(4) It remains to enumerate the counterterms which are required to implement the BPH subtractions. As in the continuum theory, the Ward identities are useful here. They are derived, as usual, by making a change of variables corresponding to an infinitesimal gauge transformation in the path integral for the vacuum functional in the presence of sources. The action proper is invariant under such a transformation but the gauge-fixing and source terms are not. The Ward identities state that the contribution of these terms does not affect the vacuum functional. It should be evident from Eqs. (3.1) and (3.2) that the Ward identities differ from their continuum versions only in the replacement of ∂_μ by ∇_μ^-. They read, in momentum space,

$$\sum_\mu S_\mu(k)\Gamma_\mu(p+k,p) = S_F^{-1}(p+k) - S_F^{-1}(p)\,, \quad (3.5)$$

$$\sum_\mu S_\mu(k)\Pi_{\mu\nu}(k) = 0\,, \quad (3.6)$$

$$\sum_\mu S_\mu(k_1) I_{\mu\nu\lambda\pi}(k_1,k_2,k_3,k_4) = 0\,, \quad (3.7)$$

where $S_\mu(k)\equiv(2/a)\sin\frac{1}{2}k_\mu a$ and $I_{\mu\nu\lambda\pi}$ is the photon-photon scattering amplitude. By substituting the Taylor expansions of the amplitudes into the Ward identities and using the lattice cubic symmetries one can show that $I_{\mu\nu\lambda\pi}$ is not divergent, the divergent terms are at worst logarithmic, and the momentum dependence and tensor structure of these terms is exactly as in continuum QED. Because the action differs by terms of order a from the continuum QED action, it follows that multiplicative renormalization of fields and parameters generates precisely the needed counterterms, plus additional terms of order $a\ln a$ which have no effect when $a\to 0$.

These arguments have been reviewed in detail so that the reader will understand exactly what ingredients go into a proof of perturbative equivalence of lattice and continuum field theories. In Sec. V I will discuss the problems that arise in applying the same arguments to the DWY version of lattice QED.

IV. FAITHFUL LATTICE TRANSCRIPTION OF QED

In Sec. II it was pointed out that with the DWY derivative one can construct a lattice free fermion theory with continuous chiral symmetry and a sensible spectrum. I now give an "existence proof," showing that in fact a lattice QED can be formulated which continues to enjoy these properties and makes sense in weak-coupling perturbation theory. This serves as a simple counterexample to statements in the literature that no such formulation is possible.[6,8]

The idea here is to make contact between continuum and lattice field theories via a momentum-space formulation which both share. This technique has been used by DWY[5] and by others[13] and in fact motivates the introduction of the DWY gradient.

The Euclidean action for ordinary continuum QED reads

$$I = \int d^4x \left\{ \sum_{\mu\nu} \tfrac{1}{4}F_{\mu\nu}{}^2 - \sum_\mu \bar\psi(x)\frac{1}{i}\gamma_\mu[\partial_\mu + ieA_\mu(x)]\psi(x) - m\bar\psi(x)\psi(x)\right\}. \quad (4.1)$$

The first step is to fix the Coulomb gauge and eliminate the dependent variable A_0 by means of its equation of motion:

$$I=\int d^4x\left[\tfrac{1}{2}(\partial_0\vec{A})^2+\tfrac{1}{2}(\vec{\nabla}\times\vec{A})^2-\sum_\mu\bar{\psi}(x)\frac{1}{i}\gamma_\mu\partial_\mu\psi(x)-m\bar{\psi}(x)\psi(x)-\sum_j e\bar{\psi}(x)\gamma_jA_j(x)\psi(x)\right.$$
$$\left.+e^2\int d^4x'\frac{\delta(t-t')}{8\pi|\mathbf{x}-\mathbf{x}'|}\psi^\dagger(x)\psi(x)\psi^\dagger(x')\psi(x')\right]. \quad (4.2)$$

It is to be emphasized that I is manifestly gauge invariant because it is written in terms of gauge-invariant fields: $\vec{A}$ is now the transverse photon field and ψ is the Coulomb-gauge (physical) electron field. All gauge degrees of freedom have been removed. These degrees of freedom are not true quantum variables and should not be included in the transcription to the lattice. The action (4.2) is now written in momentum space:

$$I=\int\frac{d^4k}{(2\pi)^4}\left[\tfrac{1}{2}k^2\tilde{\vec{A}}(k)\cdot\tilde{\vec{A}}(-k)-\tilde{\bar{\psi}}(k)\gamma\cdot k\tilde{\psi}(k)-m\tilde{\bar{\psi}}(k)\tilde{\psi}(k)-e\int\frac{d^4p\,d^4q}{(2\pi)^4}\tilde{\bar{\psi}}(k)\vec{\gamma}\cdot\tilde{\vec{A}}(p)\tilde{\psi}(q)\delta^4(p+q-k)\right.$$
$$\left.+\tfrac{1}{2}e^2\int\frac{d^4p\,d^4q\,d^4l}{(2\pi)^8}\frac{1}{|\vec{l}-\vec{k}|^2}\tilde{\psi}^\dagger(k)\tilde{\psi}(l)\tilde{\psi}^\dagger(p)\tilde{\psi}(q)\delta^4(k+p-l-q)\right]. \quad (4.3)$$

Next, impose a cutoff Λ on the magnitude of each component of momentum in Eq. (4.3). (This is why it was necessary to write I in terms of explicitly gauge-invariant variables. Had that not been done, gauge invariance would have been lost at this point.) The resulting action could equally well be interpreted as the momentum-space action of a lattice field theory, namely,

$$I_{\text{lattice}}=a^8\sum_{x,y}\tfrac{1}{2}d(x-y)\vec{A}(x)\cdot\vec{A}(y)-a^8\sum_{x,y,\mu}\bar{\psi}(x)\frac{1}{i}\gamma_\mu D_\mu(x-y)\psi(y)$$
$$-a^4\sum_x m\bar{\psi}(x)\psi(x)-a^{12}\sum_{x,y,z,j}ef(x,y,z)\bar{\psi}(x)\gamma_jA_j(y)\psi(z)+a^{16}\sum_{x,x',y,y'}\tfrac{1}{2}e^2g(x,x',y,y')\psi^\dagger(x)\psi(x')\psi^\dagger(y)\psi(y'), \quad (4.4)$$

where

$$d(x-y)=\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}k^2e^{ik\cdot(x-y)},$$

$$D_\mu(x-y)=\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}ik_\mu e^{ik\cdot(x-y)},$$

$$f(x,y,z)=\int_{-\Lambda}^{\Lambda}\frac{d^4k\,d^4p\,d^4q}{(2\pi)^8}e^{i(k\cdot x-p\cdot y-q\cdot z)}\delta^4(p+q-k),$$

$$g(x,x',y,y')=\int_{-\Lambda}^{\Lambda}\frac{d^4k\,d^4p\,d^4q\,d^4l}{(2\pi)^{12}}\frac{1}{|\vec{l}-\vec{k}|^2}e^{i(k\cdot x+p\cdot y-l\cdot x'-q\cdot y')}\delta^4(p+k-l-q).$$

The nonlocal coefficient functions here are all translation invariant and, except for $g(x,x',y,y')$ which contains the noncovariance associated with the Coulomb interaction, invariant under the lattice cubic symmetries. $D_\mu(x-y)$ is just the DWY derivative operator. Note also that in this formulation there is no possibility of assigning the photon field $A_j(x)$ to the links of the lattice: all fields are treated on an equal footing and may as well be situated on the sites.

The lattice theory (4.4) may be quantized by the path-integral technique provided one integrates only over transverse gauge fields with $\vec{k}\cdot\tilde{\vec{A}}(k)=0$. It is evident that in all respects—including perturbation theory—the theory is equivalent to Coulomb-gauge continuum QED regularized by a momentum cutoff. To each continuum operator there corresponds a lattice operator, obtained by a double Fourier transform, with the same regularized matrix elements. The fermion spectrum is sensible and there is chiral symmetry for $m=0$. Also, there are no umklapp processes: momentum conservation in Feynman diagrams is exact rather than periodic, and propagators and vertex functions are identical to those of continuum QED. The theory can be given a finite $a\to 0$ limit by including in the momentum-space action the counterterms needed to renormalize continuum QED. Because of the momentum-cutoff regularization, photon mass counterterms will be needed. For ϕ^4 theory in 1+1 dimensions all necessary counterterms are known exactly and this program has been carried out explicitly by Bronzan.[13]

Although this procedure provides a lattice theory which faithfully represents continuum QED, it is not a *lattice gauge theory*. A lattice gauge theory possesses a local gauge group on the lattice under which the action is invariant but the fields transform nontrivially. The above theory does not qualify because the gauge freedom in the fields was removed before transcription to the lattice. In the next section I discuss the lattice gauge theory constructed using the DWY derivative.

The lattice theory constructed above possesses neither a local gauge symmetry nor periodic momentum conservation. It is easy to understand qualitatively why these properties are connected. At a technical level, perturbative proofs of Ward identities require shifts of integration variables which are made possible by periodicity. More generally, consider a term in the lattice action

$$F(x_1,x_2,\ldots,x_n)\phi(x_1)\phi(x_2)\ldots\phi(x_n)\,,$$

where ϕ is a generic field. Assuming that F is translation invariant its Fourier transform $\tilde{F}(p_1,p_2,\ldots,p_n)$ can have support only when $\sum p_i = 0 \bmod 2\pi/a$. To obtain exact momentum conservation $\tilde{F}$ must be so chosen that its support lies in the subregion $\sum p_i = 0$: not all momenta can be allowed to become large simultaneously. This is the case for the coefficient functions in Eq. (4.4). However, a gauge symmetry which is local in coordinate space will affect the high-momentum components of fields. A gauge-invariant coupling term will couple high-momentum components of fields, so that in general the support of $\tilde{F}$ cannot be restricted to $\sum p_i = 0$.

V. DWY LATTICE GAUGE THEORY

A. Introduction

This section begins the discussion of the lattice gauge theory with action

$$I=a^4\sum_{x,\mu,\nu}\tfrac{1}{4}F_{\mu\nu}{}^2(x)+a^4\sum_{x,\mu}\frac{1}{2\lambda}[\nabla^-_\mu A_\mu(x)]^2 - a^8\sum_{x,y,\mu}\bar{\psi}(x)\gamma_\mu\frac{1}{i}D_\mu(x-y)\psi(y)\exp\Big[iea\sum_{z=x}^{y}A_\mu(z)\Big] - a^4\sum_x m\bar{\psi}(x)\psi(x)\,, \qquad (5.1)$$

where

$$F_{\mu\nu}(x)=\nabla^+_\mu A_\nu(x)-\nabla^+_\nu A_\mu(x)\,,$$

and

$$D_\mu(x)=\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}ik_\mu e^{ik\cdot x} = (-1)^{x_\mu/a}/a^4x_\mu \text{ if } x_\mu\neq 0 \text{ but } x_\nu=0 \text{ for all } \nu\neq\mu;\quad =0 \text{ otherwise}\,, \qquad (5.2)$$

and the notation $\sum_{z=x}^{y}A_\mu(z)$ means the following. Owing to the presence in Eq. (5.1) of the DWY derivative function $D_\mu(x-y)$, the summation need only be defined in case $x_\mu\neq y_\mu$ but $x_\nu=y_\nu$ for all $\nu\neq\mu$ (x and y are separated in the μ direction only). In that case it means the sum of the values of A_μ on the oriented links between x and y: $\sum_{z=x}^{y}A_\mu(z)$ means

$$\theta(y_\mu-x_\mu)\sum_{n=0}^{(y_\mu-x_\mu-a)/a}A_\mu(x+na_\mu) - \theta(x_\mu-y_\mu)\sum_{n=0}^{(x_\mu-y_\mu-a)/a}A_\mu(y+na_\mu)\,.$$

For $e=0$ the fermion action is that of the DWY formulation, with undoubled spectrum and continuous chiral symmetry for $m=0$. The action is invariant under the gauge transformations of Eq. (3.2). Since the photon action is exactly as in the Wilson formulation it should be clear that the Ward identities are still given by Eqs. (3.5)–(3.7). In particular, the nearest-neighbor derivative, not the DWY derivative, appears in Ward identities. (Nakawaki[12] has considered a lattice theory in which *all* derivatives are taken to be ik_μ in momentum space. This simply replaces $S_\mu(k)$ by k_μ everywhere without affecting the arguments to follow.) However, the *consequences* of the Ward identities are vastly different for the theories (3.1) and (5.1) due to the different fermion spectra. This will emerge shortly.

The theory (5.1) possesses a conserved electromagnetic current which can be identified by considering the coupling to an external field:

$$\sum_\mu \nabla^-_\mu j_\mu(z)=0\,,$$

$$j_\mu(z)=\frac{\delta I}{a^4\delta A^{\rm ext}_\mu(z)}\bigg|_{A^{\rm ext}_\mu=0} = -ea^5\sum_{\substack{x,y\\ x_\mu\le z_\mu<y_\mu\\ z_\nu=x_\nu,\nu\neq\mu}}\bar{\psi}(x)\gamma_\mu D_\mu(x-y)\psi(y)\exp\Big[iea\sum_{w=x}^{y}A_\mu(w)\Big]+\text{H.c.}\,. \qquad (5.3)$$

There is also an axial-vector current, conserved for $m=0$:

$$\begin{aligned} &\sum_\mu \nabla_\mu^- j_\mu{}^5(z) = 2im\bar\psi(z)\gamma_5\psi(z)\,, \\ &j_\mu{}^5(z) = a^5 \sum_{\substack{x,y \\ x_\mu \le z_\mu < y_\mu \\ z_\nu = x_\nu,\, \nu\ne\mu}} \bar\psi(x)\gamma_\mu\gamma_5 D_\mu(x-y)\psi(y)\exp\left[iea\sum_{w=x}^{y} A_\mu(w)\right] + \text{H.c.} \end{aligned} \tag{5.4}$$

Both these currents are gauge invariant.

By expanding the exponential in the action and introducing Fourier transforms, one derives what I shall call the naive Feynman rules. These are given in Fig. 4. Momentum conservation in this theory is once again modulo $2\pi/a$. The first point to observe is that the continuum Feynman rules are indeed recovered when $a\to 0$ with all momenta fixed. This verifies that the action has the correct classical continuum limit, a fact which is not immediately apparent from Eq. (5.1). The most striking feature of the naive Feynman rules, however, is the presence of infrared singularities in the vertex functions. The one-photon vertex, for example,

$$e\gamma_\mu \frac{\tilde D_\mu(p) - \tilde D_\mu(p+k)}{S_\mu(k)}\,,$$

behaves as $-2\pi e\gamma_\mu/ak_\mu$ as $k_\mu\to 0^+$ with $p_\mu\to\pi/a$ from below and $p_\mu+k_\mu\to\pi/a$ from above. This is a consequence of the discontinuity in $\tilde D_\mu(p)$ at $p_\mu=\pi/a$, and thus, indirectly, of the Ward identity (3.5) relating the vertex to the fermion propagator. These singularities have important consequences for the renormalization program in the manner of Sharatchandra. Due to the singularities and discontinuities in the vertices, naive Feynman integrands do not possess Taylor expansions in powers of external momenta. Furthermore, the singularities alter the results of naive power counting. A diagram with F external fermion lines and B external boson lines would normally have superficial degree of divergence $D=4-\frac{3}{2}F-B$. Here, however, for each external photon line there is a factor $1/S_\mu(k)$ which sits outside the integration and does not help to converge it. The integral is left with $D=4-\frac{3}{2}F$. The infinite class of diagrams with $F=0$ or 2 is superficially divergent. Due to the infrared singularities, then, the crucial steps (2) and (3) in the renormalization program of Sec. III do not go through for DWY fermions, and the theory indeed appears nonrenormalizable.

Karsten and Smit base their objections to the DWY lattice gauge theory on the above points, which they have explicitly verified in the example of the one-loop vacuum polarization.[11] They found that $\Pi_{\mu\nu}(k)$ had $D=2$ even *after* the cancellations due to gauge invariance. In the continuum limit there are infrared singular terms with unacceptable (nonlocal) tensor structure in both the divergent and finite terms, a typical structure being

$$\Pi_{\mu\nu}(k) \sim \frac{\delta_{\mu\nu}}{|k_\mu|}\sum_\lambda |k_\lambda| - \operatorname{sign} k_\mu \operatorname{sign} k_\nu \tag{5.5}$$
$$+ \text{other singular terms}\,.$$

(Note that the Ward identity $\sum_\mu k_\mu \Pi_{\mu\nu}=0$ is satisfied.) Furthermore, since the necessary Taylor expansions do not exist, there is no natural way to make the separation into divergent terms and finite remainders which defines the counterterms required. Since the Green's functions are not differentiable, the conventional normalization conditions do not make sense.

It is important to understand clearly the origin of the infrared singularities in the vertex functions. They come from the term in the action

$$-a^8 \sum_{x,y,\mu} \bar\psi(x)\gamma_\mu \frac{1}{i} D_\mu(x-y)\psi(y)\exp\left[iea\sum_{z=x}^{y} A_\mu(z)\right]. \tag{5.6}$$

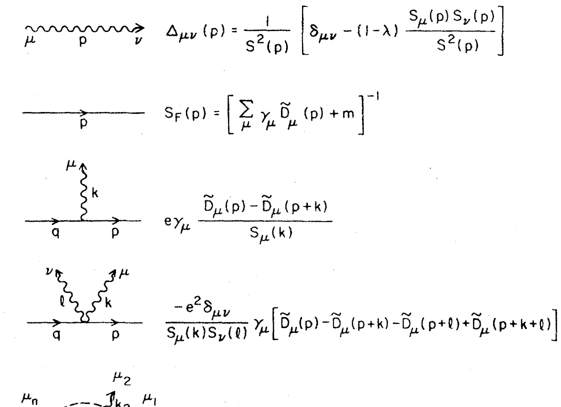

FIG. 4. Naive Feynman rules for DWY lattice QED. $S_\mu(p) \equiv \left(\frac{2}{a}\right)\sin\frac{1}{2}p_\mu a$.

The exponential factor, in momentum space, involves a geometric sum:

$$\exp\left[iea\sum_{z=x}^{y}\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}e^{ik\cdot z}e^{ik_\mu a/2}\tilde{A}_\mu(k)\right]$$
$$=\exp\left[iea\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}\frac{e^{ik\cdot x}-e^{ik\cdot y}}{1-e^{ik_\mu a}}e^{ik_\mu a/2}\tilde{A}_\mu(k)\right]$$
$$=\exp\left[-e\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}\frac{e^{ik\cdot x}-e^{ik\cdot y}}{S_\mu(k)}\tilde{A}_\mu(k)\right]. \quad (5.7)$$

The singular factors $1/S_\mu(k)$ enter the vertices via the expansion of this exponential in powers of e. However, consider the behavior of the integrand in the infrared region $k_\mu\to 0$; it is proportional to $i|x_\mu-y_\mu|$. Since x and y are summed over all lattice sites in (5.6), the distance between them is unbounded. This means that the expansion of the exponential to any finite order n in e cannot be a uniformly valid approximation over the entire range of values of $|x_\mu-y_\mu|$. If the expansion is attempted anyway, its nth term will behave as $|x_\mu-y_\mu|^{n-1}$. Since the function $D_\mu(x-y)$ in (5.6) falls off only as $|x_\mu-y_\mu|^{-1}$, the individual terms in the perturbation expansion will be divergent in the infrared. The conclusion is that the infrared singularities in the naive Feynman rules are symptomatic of an invalid perturbation expansion which does not accurately represent the infrared behavior of the theory. I emphasize that the fault lies with the perturbative expansion rather than with any inconsistency in the theory. If the expansion in powers of e is avoided then the exponential enters the sum (5.6) as a rapidly oscillating phase when $|x_\mu-y_\mu|$ is large. Such a phase factor actually improves convergence of the sum. Finally, note that perturbation theory can fail even when the fermion spectrum is doubled. If $D_\mu(x-y)$ has a power-law falloff faster than $|x_\mu-y_\mu|^{-1}$ then as pointed out in Sec. II the spectrum is doubled, but singularities will still appear at sufficiently high order in perturbation theory. The equivalent momentum-space statement is that even if $\tilde{D}_\mu(p)$ is continuous, a discontinuity in its nth derivative induces a singularity in the $(n+1)$-photon vertex function. This follows from the recursion relation for the vertices in Fig. 4. A nonsingular perturbation expansion is obtained only if $D_\mu(x-y)$ falls faster than any power of $|x_\mu-y_\mu|$. Such a $D_\mu(x-y)$ strongly suppresses the contributions from the region of large $|x_\mu-y_\mu|$ where the expansion of the exponential is invalid.

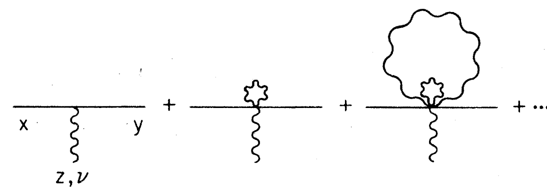

FIG. 5. A class of diagrams whose summation removes the infrared singularity from the vertex function and permits a nonsingular perturbation expansion.

The failure of naive perturbation theory discussed above becomes evident from the structure of $\Pi_{\mu\nu}$ in Eq. (5.5). Consider a diagram in which the one-loop $\Pi_{\mu\nu}(k)$ appears as a subgraph. The integration over k encounters a $1/k_\mu$ singularity. Such a singularity is not integrable, in contrast to the usual infrared singularities which often are, e.g., $\int d^4k/k^2$. Since the singularity arises from a vertex function rather than a propagator, it also is not regularized by a photon mass, and simply leads to a divergent amplitude indicating the breakdown of perturbation theory.

B. Removal of the infrared problem

Now that the origin of the infrared problems which plague naive perturbation theory is clear, how can they be circumvented? The most obvious approach is simply to impose a cutoff on $|x_\mu-y_\mu|$ in the nonlocal interaction Lagrangian:

$$\sum_{x,y,\mu}\bar\psi(x)\gamma_\mu\frac{1}{i}D_\mu(x-y)\psi(y)\exp\left[iea\sum_{z=x}^{y}A_\mu(z)\right]$$
$$\to\sum_{x,y,\mu}\bar\psi(x)\gamma_\mu\frac{1}{i}D_\mu(x-y)\psi(y)+\sum_{\substack{x,y,\mu\\|x_\mu-y_\mu|<Na}}\bar\psi(x)\gamma_\mu\frac{1}{i}D_\mu(x-y)\psi(y)\left[\exp\left(iea\sum_{z=x}^{y}A_\mu(z)\right)-1\right].$$

The cutoff permits a nonsingular expansion in powers of e but destroys manifest gauge invariance. Therefore the cutoff must be imposed in the fixed gauge in which quantization is performed. This should be a physical gauge, since otherwise the loss of the Ward identities will jeopardize unitarity.

I now show that in fact an *ad hoc* cutoff is unnecessary since the theory generates its own cutoff. Consider for example the bare one-photon vertex function, and add to it all diagrams in which additional photons are emitted and absorbed at the same vertex (Fig. 5). The sum gives the vertex function computed to lowest order in the interaction Lagrangian rather than lowest order in e. The diagrams are most easily summed in coordinate space, where they yield

$$\sum_{x',y',\mu}\sum_{z'=x'}^{y'} ieaS_F(x-x')\gamma_\mu D_\mu(x'-y')S_F(y'-y)\Delta_{\mu\nu}(z-z')$$
$$\times\left[1-\tfrac{1}{2}e^2a^2\sum_{w_1=x'}^{y'}\sum_{w_2=x'}^{y'}\Delta_{\mu\mu}(w_1-w_2)+\cdots+\frac{(-1)^n}{(2n)!}(2n-1)(2n-3)\cdots 1(ea)^{2n}\right.$$
$$\left.\times\sum_{w_1,\ldots,w_{2n}=x'}^{y'}\Delta_{\mu\mu}(w_1-w_2)\cdots\Delta_{\mu\mu}(w_{2n-1}-w_{2n})+\cdots\right], \tag{5.8}$$

where the combinatorial factor $(2n-1)(2n-3)\cdots 1$ is the number of ways of pairing the points $w_1,\ldots,w_{2n}$ in the photon propagators. The sum in brackets exponentiates, giving

$$\exp\left[-\tfrac{1}{2}e^2a^2\sum_{w_1,w_2=x'}^{y'}\Delta_{\mu\mu}(w_1-w_2)\right]=\exp\left[-\tfrac{1}{2}e^2a^2\sum_{w_1,w_2=x'}^{y'}\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}e^{ik\cdot(w_1-w_2)}\Delta_{\mu\mu}(k)\right]$$
$$=\exp\left[-\tfrac{1}{2}e^2\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}\left|\frac{e^{ik\cdot x'}-e^{ik\cdot y'}}{S_\mu(k)}\right|^2\Delta_{\mu\mu}(k)\right]. \tag{5.9}$$

A similar calculation applies to the multiphoton vertex functions. The inclusion of these photon tadpole contributions to the vertex functions thus generates effective Feynman vertices which differ from the naive ones of Fig. 4 only in the replacement

$$D_\mu(x-y)\to\mathfrak{D}_\mu(x-y)\equiv D_\mu(x-y)\exp\left[-\tfrac{1}{2}e^2\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}\left|\frac{e^{ik\cdot x}-e^{ik\cdot y}}{S_\mu(k)}\right|^2\Delta_{\mu\mu}(k)\right]$$
$$=D_\mu(x-y)\exp\left[-\tfrac{1}{2}e^2\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}a^2\frac{\sin^2\frac{1}{2}k_\mu(x-y)_\mu}{\sin^2\frac{1}{2}k_\mu a}\Delta_{\mu\mu}(k)\right]. \tag{5.10}$$

At issue is the large-distance behavior of $\mathfrak{D}_\mu(x)$. Since[14]

$$\frac{\sin^2(\frac{1}{2}nx)}{n\sin^2(\frac{1}{2}x)}\xrightarrow[n\to\infty]{}2\pi\delta_{\text{per}}(x),$$

one has

$$\mathfrak{D}_\mu(x)\xrightarrow[x_\mu\to\infty]{}D_\mu(x)\exp\left[-\pi e^2\int_{-\Lambda}^{\Lambda}\frac{d^4k}{(2\pi)^4}\Delta_{\mu\mu}(k)\delta(k_\mu)x_\mu\right], \tag{5.11}$$

and $\mathfrak{D}_\mu(x)$ falls off exponentially fast. It follows that the Fourier transform $\tilde{\mathfrak{D}}_\mu(p)$ and all its derivatives are continuous, and that there are no infrared singularities in any of the modified vertices. Although $\tilde{D}_\mu(p)$ as a function of p_μ has unit slope at $p_\mu=0$, there is no reason for $\tilde{\mathfrak{D}}_\mu(p)$ to share this property. This means that ultimately a finite renormalization will be required to express the theory in terms of a charge defined by the static limit of the electron-photon vertex rather than the parameter e. This is discussed more fully below. Figure 6 shows the expected behavior of $\tilde{\mathfrak{D}}_\mu(p)$.

In general, the summation of a selected class of diagrams is not a gauge-invariant procedure. This is reflected in the explicit appearance of the photon propagator in Eq. (5.10). $\mathfrak{D}_\mu(x-y)$ is thus a gauge-dependent function. It will be shown, however, that the S matrix has a gauge-invariant continuum limit order by order in the modified perturbation expansion.

Summing the diagrams of Fig. 7 effects the replacement $D_\mu(x-y)\to\mathfrak{D}_\mu(x-y)$ in the fermion propagator, resulting in a doubled spectrum according to the analysis in Sec. II. Since we wish to develop a perturbation expansion about the free field theory with undoubled fermion spectrum, this replacement must be un-

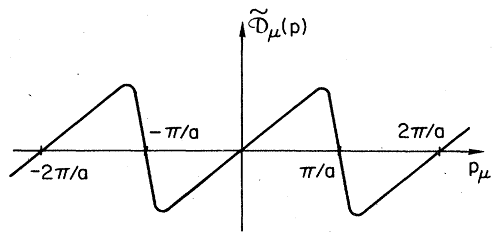

FIG. 6. Qualitative behavior of the function $\tilde{\mathfrak{D}}_\mu(p)$ appearing in the effective Feynman rules.

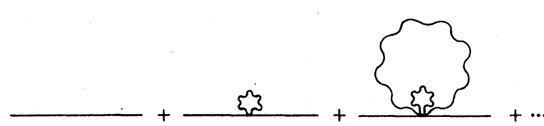

FIG. 7. A class of diagrams whose summation would double the fermion spectrum.

done by the addition of a counterterm

$$\sum_{x,y,\mu} \bar{\psi}(x)\gamma_\mu \frac{1}{i}[D_\mu(x-y) - \mathfrak{D}_\mu(x-y)]\psi(y)\,, \tag{5.12}$$

again in the fixed, physical, quantization gauge. Of course this amounts to an assumption that the interacting theory (5.1) has the same qualitative spectrum as the noninteracting $e=0$ theory. The validity of this assumption is discussed further in Sec. VIII.

The resummation of perturbation theory discussed above is most clearly understood in the Hamiltonian formulation of the theory in the physical Coulomb gauge.[15] The Hamiltonian (now on a three-dimensional lattice) is

$$\begin{aligned} H = a^3 \sum_{\vec{x}} [\tfrac{1}{2}E_T{}^2(\vec{x}) + \tfrac{1}{2}B^2(\vec{x}) + m\bar{\psi}(\vec{x})\psi(\vec{x})] + a^6 \sum_{\vec{x},\vec{y}} \tfrac{1}{2}e^2\varphi(\vec{x}-\vec{y})\psi^\dagger(\vec{x})\psi(\vec{x})\psi^\dagger(\vec{y})\psi(\vec{y}) \\ + a^6 \sum_{\vec{x},\vec{y},j} \bar{\psi}(\vec{x})\gamma_j \frac{1}{i} D_j(\vec{x}-\vec{y})\psi(\vec{y}) \exp\left[iea\sum_{\vec{z}=\vec{x}}^{\vec{y}} A_j^T(\vec{z})\right], \quad \vec{\nabla}_- \cdot \vec{\nabla}_+\varphi(\vec{x}) = -\delta_{\vec{x},0}, \quad \tilde{\varphi}(\vec{k}) = 1/\vec{S}^2(\vec{k})\,, \quad \vec{\nabla}_- \cdot \vec{A}^T = 0\,. \end{aligned} \tag{5.13}$$

The summation of photon tadpole diagrams simply corresponds to normal ordering the exponential in Eq. (5.13). Including the counterterm analogous to (5.12) the Hamiltonian becomes

$$\begin{aligned} H = a^3 \sum_{\vec{x}} [\tfrac{1}{2}E_T{}^2(\vec{x}) + \tfrac{1}{2}B^2(\vec{x}) + m\bar{\psi}(\vec{x})\psi(\vec{x})] + a^6 \sum_{\vec{x},\vec{y}} \tfrac{1}{2}e^2\varphi(\vec{x}-\vec{y})\psi^\dagger(\vec{x})\psi(\vec{x})\psi^\dagger(\vec{y})\psi(\vec{y}) \\ + a^6 \sum_{\vec{x},\vec{y},j} \left[\bar{\psi}(\vec{x})\gamma_j \frac{1}{i} D_j(\vec{x}-\vec{y})\psi(\vec{y}) + \bar{\psi}(\vec{x})\gamma_j \frac{1}{i}\mathfrak{D}_j(\vec{x}-\vec{y})\psi(\vec{y}) : \exp\left(iea\sum_{\vec{z}=\vec{x}}^{\vec{y}} A_j^T(\vec{z})\right) - 1 :\right], \\ \mathfrak{D}_j(\vec{x}-\vec{y}) = D_j(\vec{x}-\vec{y}) \exp\left\{-\tfrac{1}{2}e^2 \int_{-\Lambda}^{\Lambda} \frac{d^3k}{(2\pi)^3 2|\vec{S}(\vec{k})|} \left|\frac{e^{i\vec{k}\cdot\vec{x}} - e^{i\vec{k}\cdot\vec{y}}}{S_j(\vec{k})}\right|^2 \left[1 - \frac{S_j{}^2(\vec{k})}{\vec{S}^2(\vec{k})}\right]\right\}. \end{aligned} \tag{5.14}$$

H is gauge invariant because the fields appearing in it are, but Ward identities which state that $S_\mu(k)$ terms in the photon propagator do not contribute to physical quantities do not hold. This may be understood as follows. In a more general gauge, related to the Coulomb gauge by a time-independent gauge transformation, a structure $\bar{\psi}(\vec{x})D_j(\vec{x}-\vec{y})\psi(\vec{y})$ in Eq. (5.14) appears as $\bar{\psi}(\vec{x})D_j(\vec{x}-\vec{y})\psi(\vec{y})\exp[iea\sum_{\vec{z}=\vec{x}}^{\vec{y}} A_j^L(\vec{z})]$. Thus $\vec{A}_L$ is coupled to the conserved current

$$j_i(\vec{z}) = -ea^4 \sum_{\substack{\vec{x},\vec{y} \\ x_i \le z_i < y_i \\ x_j = z_j,\, j\ne i}} \left[\bar{\psi}(\vec{x})\gamma_i D_i(\vec{x}-\vec{y})\psi(\vec{y}) + \bar{\psi}(\vec{x})\gamma_i \mathfrak{D}_i(\vec{x}-\vec{y})\psi(\vec{y}) : \exp\left(iea\sum_{\vec{w}=\vec{x}}^{\vec{y}} A_i^T(\vec{w})\right) - 1 :\right] + \text{H.c.}$$

(in the Coulomb gauge) as required by gauge invariance, while $\vec{A}_T$ couples to the nonconserved

$$j_i'(\vec{z}) = -ea^4 \sum_{\substack{\vec{x},\vec{y} \\ x_i \le z_i < y_i \\ x_j = z_j,\, j\ne i}} \bar{\psi}(\vec{x})\gamma_i \mathfrak{D}_i(\vec{x}-\vec{y})\psi(\vec{y}) : \exp\left[iea\sum_{\vec{w}=\vec{x}}^{\vec{y}} A_i^T(\vec{w})\right] : + \text{H.c.}$$

[$j_0(\vec{z}) = -e\bar{\psi}(\vec{z})\psi(\vec{z})$ in either case.] In continuum QED $\vec{A}_L$ and $\vec{A}_T$ enter the action only through the local field $\vec{A}$, so both couple to the same current.

The effective vertices possess all the properties required for a proof of renormalizability as in Sec. III. The functions involved are C^∞ and possess the required Taylor expansions. Furthermore, naive power counting now works properly. A diagram with F external fermions and B external photons is $1/S_\alpha(k_1)S_\beta(k_2)\cdots S_\mu(k_B)$ times an integral with superficial $D = 4 - \frac{3}{2}F$. But the absence of infrared singularities requires that the Taylor expansions of the vertex functions in the integrand begin with the term of order $k_{1\alpha}k_{2\beta}\cdots k_{B\mu}$, reducing D to $4 - \frac{3}{2}F - B$. Similarly the numerator of an n-photon vertex must go as $k_{1\mu}k_{2\mu}\cdots k_{n\mu}$ when the k's are small, and this must be accompanied by a factor a^{n-1} on dimensional grounds. Hence multiphoton vertices are accompanied by factors of a as required in the arguments of Sec. III. However, one obstacle remains to the application of Sharatchandra's arguments to the DWY lattice gauge theory: the presence in the fermion propagator of the discontinuous function $\tilde{D}_\mu(p)$. This problem is addressed next.

VI. PROOF OF RENORMALIZABILITY

So far it has been established that in the modified perturbation expansion for the DWY lattice gauge theory, the vertices are infinitely differentiable functions of momenta and naive power counting correctly gives the degree of divergence of Feynman integrals. In general, diagrams will actually have their full superficial degrees of divergence since the Ward identities which normally reduce D do not hold order by order in this expansion. However, Feynman integrands still do not possess Taylor expansions because the fermion propagators contain the discontinuous function $\tilde{D}_\mu(p)$. This difficulty exists in any lattice field theory in which there is (a) periodic momentum conservation, and (b) fermions with undoubled spectrum. In this section I explain how to carry out a subtractive renormalization program for such theories. The next section considers the form of the counterterms required to implement the subtractions.

Consider an arbitrary Feynman diagram. The corresponding amplitude takes the form

$$A(k)=\left[\prod_{\text{lines}}\int_{-\Lambda}^{\Lambda}\frac{d^4l}{(2\pi)^4}\right]I(l,k)\times\left[\prod_{\text{vertices}}(2\pi)^4\delta^4_{\text{per}}\left(\sum\text{momenta}\right)\right], \qquad (6.1)$$

where k denotes the external momenta and I is written using the Feynman rules. At this point $I(l,k)$ possesses an expansion in powers of k because for $|l_\mu|<\Lambda$, $\tilde{D}_\mu(l)=l_\mu$ which is perfectly continuous. $A(k)$ does not have an expansion, though, because the periodic δ functions contain additional dependence on k.

Choose now a subset $\{q\}$ of the momenta $\{l\}$ to act as independent loop momenta. According to the Appendix the trivial integrations over $\{l\}-\{q\}$ may be done provided $I(l,k)$ is a periodic function; provided, in other words, the fermion propagators are written in terms of the discontinuous $\tilde{D}_\mu(l)$ instead of simply l_μ. The integrations then result in a discontinuous integrand $I(q,k)$. However, since $\tilde{D}_\mu(l)$ is piecewise continuous, the domain of integration can be divided into subregions with $I(q,k)$ continuous in each.

An efficient way to do this is to return to Eq. (6.1) and to substitute for the periodic δ functions

$$\delta^4_{\text{per}}(p)=\prod_\mu\sum_{n_\mu=-\infty}^{\infty}\delta(p_\mu+2n_\mu\Lambda). \qquad (6.2)$$

Since only finitely many lines enter each vertex of the graph, and all lines are restricted by $|l_\mu|<\Lambda$, only finitely many terms in the sum can actually contribute. Doing trivial integrations then yields

$$A(k)=\sum_j\left[\prod_{\{q\}}\int_{-\Lambda}^{\Lambda}\frac{d^4q}{(2\pi)^4}\right]I_j(q,k)\times\left\{\prod_{\text{lines},\mu}\theta[\Lambda-|l_j^\mu(q,k)|]\right\}, \qquad (6.3)$$

i.e., a sum of integrals indexed by j. The integrands $I_j(q,k)$ are generally all different, as are the functions $l_j^\mu(q,k)$ which give the μth component of the momentum in line l in terms of q and k. In writing $I_j(q,k)$, $\tilde{D}_\mu(l)$ is to be replaced by l_μ as is permitted by the θ functions. Each integrand $I_j(q,k)$ thus has a Taylor expansion in the variables k. Let $j=0$ label the integral with no umklapps—$n_\mu=0$ in Eq. (6.2) for every periodic δ function in Eq. (6.1). The terms $j\neq0$ are diagrams in which momentum components in multiples of 2Λ enter vertices "from nowhere" in all possible ways.

Consider one particular integral labeled by j. The integral will be made finite in the limit $a\to0$ by replacing $I_j(q,k)$ by a renormalized integrand $R_j(q,k)$ via the following prescription. As in ordinary BPH renormalization, lay down forests of nonoverlapping boxes on the diagram, each box surrounding a renormalization part—a two-, three-, or four-point function. Make the usual subtractions of the first $D+1$ terms of the Taylor expansions of the boxed subgraphs, with the following exception. If a box contains an umklapp process (if the external momenta of the boxed subgraph do not sum to zero, but to a multiple of 2Λ, which can happen only for three- and four-point functions) then no subtractions need be made for that box. The reason for this exception is the following. According to the usual criterion a Feynman integral converges if all subintegrations have $D<0$, a subintegration being an integral over a subset of the q's with all other momenta held *fixed* as $a\to0$. The integration over the internal momenta of a boxed umklapp process does not count as a subintegration because the external momenta cannot be held fixed when $a\to0$. Renormalized Green's functions are not required to be finite when their external momenta approach infinity.

After the subtractions are made, the jth integral is guaranteed to be finite when $a\to0$, even ignoring the θ-function constraints in Eq. (6.3). The θ functions impose additional restrictions on the region of integration, so including them does not make a formerly finite integral diverge. As in Sec. III, if the diagram under consideration includes a multiphoton vertex then the explicit factors of a in such a vertex cause the renormalized diagram to vanish as $a\to0$. For a normal diagram, the integrand $I_0(q,k)$ of the no-umklapp term in Eq. (6.3) becomes the continuum Feynman integrand for the same diagram when $a\to0$ (provided the continuum parameter e is identified as the coefficient of γ_μ in the zero-momentum limit of the lattice one-photon vertex). The θ functions make a negligible contribution in the limit $a\to 0$, so the renormalized $j=0$ integral at $a=0$ equals the corresponding renormalized continuum integral. Finally, consider the $j\neq 0$ contributions to a normal diagram. The integral of $R_j(q,k)$ is finite. Now consider the effect of the θ functions. There is a vertex of the graph at which some components of the three entering momenta sum to $2n\Lambda$, $n\neq 0$. Since no momentum exceeds Λ (θ functions), at least two momenta are large on the scale Λ (and incidentally $n=\pm 1$). These large momenta may be traced through the graph; eventually a large momentum must flow through a line carrying one of the integration momenta q. But if one has an integral from $-\Lambda$ to $+\Lambda$, finite when $\Lambda\to\infty$, and adds a θ function requiring the integration variable to be of order Λ, the result vanishes for $\Lambda\to\infty$. Hence all $j\neq 0$ terms vanish for $a\to 0$.

It has now been shown that in the modified perturbation expansion for the DWY lattice gauge theory the subtracted Feynman integrals yield the usual results of continuum QED order by order when $a\to 0$. It follows trivially that the $a\to 0$ limit of the S matrix is in fact gauge-invariant despite the gauge dependence of the lattice expansion due to the summation of photon tadpoles. It is clear that the subtractions described above can be implemented by counterterms in the action, but the structure of these counterterms is not as simple as in the case of Wilson's QED. This is discussed next.

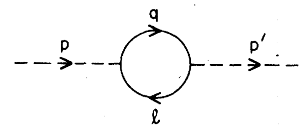

FIG. 8. The scalar self-energy in the lattice $\bar\psi\psi\phi$ theory.

VII. STRUCTURE OF COUNTERTERMS

A. Examples

This section presents some examples of the renormalization program just discussed for lattice theories with undoubled fermion spectra, with the purpose of exhibiting the types of counterterms to be expected. Since the Ward identities are not maintained order by order in the modified perturbation expansion for lattice QED, there is no formal difference between the renormalization program for lattice theories with and without local gauge invariance. Therefore, to save indices, the examples here are taken from a theory of DWY fermions interacting with scalar mesons via a $g\bar\psi(x)\psi(x)\phi(x)$ coupling.

1. Scalar self-energy

The one-loop scalar self-energy (Fig. 8) is given by

$$\begin{aligned}\Pi &= g^2\mathrm{Tr}\int_{-\Lambda}^{\Lambda} d^4q\, d^4l\, S_F(q)S_F(l)\delta^4_{\mathrm{per}}(p+l-q)\delta^4_{\mathrm{per}}(q-l-p')\\ &= g^2\mathrm{Tr}\sum_{n,m}\int_{-\Lambda}^{\Lambda} d^4q\, d^4l\, S_F(q)S_F(l)\delta^4(p+l-q+2n\Lambda)\delta^4(q-l-p'+2m\Lambda)\\ &= g^2\mathrm{Tr}\sum_{n,m}\int_{-\Lambda}^{\Lambda} d^4l\, S_F(p+l+2n\Lambda)S_F(l)\prod_\mu \theta(\Lambda-|p_\mu+l_\mu+2n_\mu\Lambda|)\delta^4[p-p'+2(m+n)\Lambda]\,,\end{aligned} \tag{7.1}$$

where m and n are four-vectors with integer components. The fact that all momentum components are bounded in magnitude by Λ imposes the restrictions $m=-n$ and $n_\mu=0,\pm 1$. Extracting the overall momentum conserving δ function gives

$$\Pi(p)=g^2\mathrm{Tr}\sum_{n_\mu=0,\pm 1}\int_{-\Lambda}^{\Lambda} d^4l\, S_F(p+l+2n\Lambda)S_F(l)\prod_\mu \theta(\Lambda-|p_\mu+l_\mu+2n_\mu\Lambda|), \tag{7.2}$$

where $S_F(q)$ now means $(\gamma\cdot q+m)^{-1}$, $\tilde D_\mu(q)$ no longer appearing.

Consider first the no-umklapp ($n=0$) contribution:

$$g^2\prod_\mu\left[\int_{\max(-\Lambda,-\Lambda-p_\mu)}^{\min(\Lambda,\Lambda-p_\mu)} dl_\mu\right]\mathrm{Tr}\,S_F(p+l)S_F(l)=g^2\prod_\mu\left[\theta(p_\mu)\int_{-\Lambda}^{\Lambda-p_\mu} dl_\mu+\theta(-p_\mu)\int_{-\Lambda-p_\mu}^{\Lambda} dl_\mu\right]\mathrm{Tr}S_F(p+l)S_F(l)\,. \tag{7.3}$$

It is clear that apart from the θ functions the integrals have expansions in powers of p_μ of which the terms up to $\mathcal{O}(p^2)$ may be divergent, while subsequent terms must give the continuum results when $a\to 0$ with p fixed. The discontinuous behavior of the integrand has been isolated in the θ functions which appear because one must know the sign of p_μ to tell whether $p_\mu+l_\mu>\Lambda$ or $p_\mu+l_\mu<-\Lambda$ is possible for $|l_\mu|<\Lambda$. The required counterterms will have the form

$$\theta(\pm p_0)\theta(\pm p_1)\theta(\pm p_2)\theta(\pm p_3)\Big(A+\sum_\mu B_\mu p_\mu+\sum_{\mu,\nu}C_{\mu\nu}p_\mu p_\nu\Big)\tilde\phi(p)\tilde\phi(-p)\,, \tag{7.4}$$

with $A, B_\mu, C_{\mu\nu}$ divergent constants. Indeed, one can say more: since $\Pi(p)$ has definite symmetry under $p\to -p$, $\theta(p_\mu)$ must appear in the even and odd combinations $\theta(p_\mu)+\theta(-p_\mu)=1$ and $\theta(p_\mu)-\theta(-p_\mu)=\text{sign }p_\mu$, giving counterterms

$$\Big(A+\sum_\mu B_\mu|p_\mu|+\sum_{\mu\nu}C_{\mu\nu}p_\mu p_\nu+\sum_{\mu\nu}D_{\mu\nu}|p_\mu|\,|p_\nu|\Big)\tilde\phi(p)\tilde\phi(-p)\,, \tag{7.5}$$

which may be further restricted by the lattice cubic symmetries. These counterterms will be nonlocal when expressed in position space, but this is to be expected since the bare action was nonlocal as well. It would be wrong to conclude from this nonlocality that infinitely many counterterms are required (counting separately the nearest-neighbor, next-nearest-neighbor, etc., terms) since in momentum space there are clearly finitely many divergent constants.

Next, examine a typical contribution to $\Pi(p)$ containing an umklapp ($n_0=1$, $\vec n=0$):

$$g^2\theta(-p_0)\int_{-\Lambda}^{-\Lambda-p_0}dl_0\left[\prod_i\int_{\max(-\Lambda,-\Lambda-p_i)}^{\min(\Lambda,\Lambda-p_i)}dl_i\right]\mathrm{Tr}S_F(p_0+l_0+2\Lambda,\vec p+\vec l)S_F(l)\,. \tag{7.6}$$

Evidently counterterms of the form (7.4) will suffice to make this finite for $a\to 0$. After the removal of the terms up to $O(p^2)$ in the expansion of the above integral, the remaining terms vanish because the umklapp restricts the l_0 integration to a small region near $-\Lambda$, as expected from the arguments of Sec. VI. All umklapp contributions vanish similarly and when $a\to 0$ the continuum result is recovered from the no-umklapp term.

2. *Vertex function*

The one-loop vertex correction (Fig. 9) reads

$$\Gamma=g^3\bar v(q)\int_{-\Lambda}^{\Lambda}d^4k\,d^4k'\,d^4l\,S_F(k)S_F(k')\Delta(l)\delta^4_{\text{per}}(p+k'-k)\delta^4_{\text{per}}(k-l-q)\delta^4_{\text{per}}(l-k'-q')v(-q')\,, \tag{7.7}$$

where $\Delta(l)=1/S^2(l)$. This becomes

$$\begin{aligned}\Gamma&=g^3\bar v(q)\sum_{n,n',n''}\int_{-\Lambda}^{\Lambda}d^4k\,d^4k'\,d^4l\,S_F(k)S_F(k')\Delta(l)\\&\qquad\times\delta^4(p+k'-k+2n\Lambda)\delta^4(k-l-q+2n'\Lambda)\delta^4(l-k'-q'+2n''\Lambda)v(-q')\\&=g^3\bar v(q)\sum_{n,n',n''}\int_{-\Lambda}^{\Lambda}d^4l\,S_F(l+q-2n'\Lambda)S_F(l-q'+2n''\Lambda)\Delta(l)\\&\qquad\times\prod_\mu\theta(\Lambda-|l_\mu+q_\mu-2n'_\mu\Lambda|)\theta(\Lambda-|l_\mu-q'_\mu+2n''_\mu\Lambda|)\delta^4[p-q-q'+2(n+n'+n'')\Lambda]v(-q')\,.\end{aligned} \tag{7.8}$$

According to Sec. VI subtractions are only required in the case of overall momentum conservation $n+n'+n''=0$. Consider the no-umklapp term $n=n'=n''=0$:

$$g^3\bar v(q)\left[\prod_\mu\int_{\max(-\Lambda,-\Lambda+q'_\mu,-\Lambda-q_\mu)}^{\min(\Lambda,\Lambda+q'_\mu,\Lambda-q_\mu)}dl_\mu\right]S_F(l+q)S_F(l-q')\Delta(l)\delta^4(p-q-q')v(-q')\,. \tag{7.9}$$

The conditions on the range of integration can be expressed using θ functions, but this is not necessary: since the integral is only logarithmically divergent, the limits of integration can be taken as $-\Lambda$ to Λ with vanishing error as $\Lambda\to\infty$. The integrand requires only a subtraction of its value at $q=q'=0$, which can evidently be effected by a counterterm of the same form as in the cutoff continuum theory.

For a typical umklapp term, $n_0=0$, $n'_0=-1$, $n''_0=+1$, $\vec n=\vec n'=\vec n''=0$,

$$\begin{aligned}&g^3\bar v(q)\theta(q'_0)\theta(-q_0)\int_{-\Lambda}^{\min(-\Lambda-q_0,-\Lambda+q'_0)}dl_0\left[\prod_i\int_{\max(-\Lambda,-\Lambda+q'_i,-\Lambda-q_i)}^{\min(\Lambda,\Lambda+q'_i,\Lambda-q_i)}dl_i\right]\\&\qquad\times S_F(l_0+q_0+2\Lambda,\vec l+\vec q)S_F(l_0-q'_0+2\Lambda,\vec l-\vec q)\Delta(l)\delta^4(p-q-q')v(-q')\,,\end{aligned} \tag{7.10}$$

the situation is even better. Since the integrand has $D=0$, the limited range of the l_0 integral causes it to vanish as $a \to 0$ and no counterterm is needed.

3. Two-loop scalar self-energy

This is included as an example of the vanishing of umklapp contributions beyond one-loop order. The only diagram which is not simply an insertion of the one-loop fermion propagator gives (Fig. 10)

$$\Pi^{(4)} = g^4 (2\pi)^{-4} \operatorname{Tr} \int_{-\Lambda}^{\Lambda} d^4k\, d^4k' d^4l\, d^4l' d^4q\, S_F(k) S_F(k') S_F(l') S_F(l) \Delta(q) \times \delta^4_{\rm per}(p+k'-k)\delta^4_{\rm per}(k-l-q)\delta^4_{\rm per}(l-l'-p')\delta^4_{\rm per}(l'+q-k') . \tag{7.11}$$

$$\Pi^{(4)}(p) = g^4 \operatorname{Tr} \sum_{m,m',n} \int_{-\Lambda}^{\Lambda} d^4k\, d^4l\, S_F(k) S_F(k-p-2m\Lambda) S_F(l-p+2n\Lambda) S_F(l) \Delta(k-l+2m'\Lambda) \times \prod_\mu \theta(\Lambda - |k_\mu - p_\mu - 2m_\mu \Lambda|)\theta(\Lambda - |l_\mu - p_\mu + 2n_\mu\Lambda|)\theta(\Lambda - |k_\mu - l_\mu + 2m'_\mu \Lambda|) . \tag{7.12}$$

In addition to the overall $D=2$ integration there are various subintegrals having $D=0$. The overlapping divergences in the no-umklapp term are handled exactly as in the continuum theory: the overall subtractions plus the inclusion of the vertex counterterms discussed above yield a finite result. Consider now the umklapp contribution $m=n=0$, $m'_0=-1$, $\vec{m}'=0$:

$$g^4 \left[\prod_\mu \int_{\max(-\Lambda,\, -\Lambda+p_\mu)}^{\min(\Lambda,\, \Lambda+p_\mu)} dk_\mu dl_\mu \right] \operatorname{Tr} S_F(k) S_F(k-p) S_F(l-p) S_F(l) \Delta(k_0 - l_0 - 2\Lambda, \vec{k}-\vec{l})\, \theta(k_0 - l_0 - \Lambda) \prod_i \theta(\Lambda - |k_i - l_i|) . \tag{7.13}$$

Here the explicitly indicated range of integration is not particularly small. However, there is the θ-function restriction $k_0 - l_0 > \Lambda$. The subintegral over k at fixed l is therefore restricted to a small region near $k_0 = \Lambda$, which causes it to vanish as $\Lambda \to \infty$ since it had $D=0$, and similarly for the l subintegral at fixed k. Finally, a subintegral over $k+l$ at fixed $k-l$ vanishes as $\Lambda \to \infty$ since a fixed $k-l$ will fail to satisfy $k_0 - l_0 > \Lambda$. Then, after counterterms of the form (7.4) have removed the terms up to $O(p^2)$ in the integrand's Taylor expansion the result must vanish since $k_0 - l_0 > \Lambda$ requires the integration variables to be large.

B. Summary

From these examples it appears that in lattice theories with undoubled fermions one must expect momentum-space counterterms which are polynomials in the momenta, plus sign p_μ functions times such polynomials. The dependence on sign p_μ reflects the fact that although the lattice Green's functions do not have Taylor expansions about $p_\mu = 0$, they do possess "one-sided" Taylor expansions valid when $p_\mu > 0$ or $p_\mu < 0$. The counterterms thus serve to impose appropriate normalization conditions on the left and right limits and derivatives of the Green's functions at $p_\mu = 0$. Only finitely many types of counterterms arise although they are nonlocal in position space. Some of the counterterms which are simple polynomials *and* only logarithmically divergent can be generated by rescaling fields and parameters, as in Wilson's QED, but others must be added by hand.

For DWY lattice QED, Eq. (5.1), the prescription is as follows. First rescale fields and parameters in Eq. (5.1), writing it as a renormalized action plus counterterms. Next sum the photon tadpole diagrams to produce an infrared finite set of Feynman rules. Third, execute the renormalization program of this and the preceding sections. This both determines the multiplicative renormalization constants and requires additional counterterms. In particular, photon mass and photon-photon scattering counterterms will be needed due

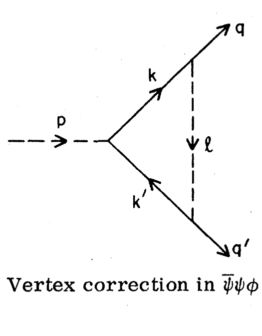

FIG. 9. Vertex correction in $\bar{\psi}\psi\phi$ theory.

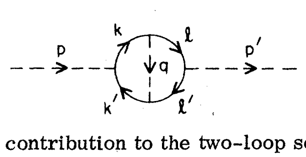

FIG. 10. A contribution to the two-loop scalar self-energy in $\bar{\psi}\psi\phi$ theory.

to the absence of Ward identities. Finally, to make contact with continuum QED a finite charge renormalization is needed to express the theory in terms of a charge defined by the static limit of the effective one-photon vertex: $e_{\text{physical}} = \tilde{\mathfrak{D}}'_\mu(0)e$.

C. The axial-vector current

The fates of the axial-vector current and the axial-vector Ward identity (5.4) in the present treatment are easy to see. The axial-vector current couples to vertices exactly like the photon but with an extra factor γ_5. In naive perturbation theory its matrix elements, like most Green's functions, are infrared divergent. The divergences can be removed, along with the Ward identities, by dressing the vertices at which the current couples with photon tadpoles. In the absence of the Ward identities the *VVA* triangle diagram will be linearly divergent. To obtain a finite continuum limit obeying Bose symmetry and the vector Ward identity it will be necessary to modify the axial-vector current by the addition of counterterms which introduce the anomaly. There seems to be no way to arrange perturbation theory so that one maintains order by order both nonanomalous Ward identities and infrared finiteness.

The important point to abstract from perturbation theory is that there is no reason to expect a nonlocal operator such as the conserved axial-vector current (5.4) to have a finite continuum limit. In view of the anomaly, there is every reason not to.

VIII. CONCLUDING REMARKS

A. Summary

This has been a long and somewhat technical paper. Three major points deserve emphasis in addition to the overall claim that the DWY derivative permits the formulation of a completely satisfactory gauge- and chiral-invariant lattice QED.

(1) All theories with DWY fermions require a special renormalization prescription because the fermion propagator does not admit a Taylor expansion. Equivalently, the usual normalization conditions cannot be applied because the bare Green's functions need not be continuous or differentiable at $p_\mu = 0$. One must normalize the left and right limits $p_\mu \to 0^\pm$ such that renormalized Green's functions are continuous and differentiable. This is just the restoration of rotational symmetry in the continuum limit. It requires nonpolynomial counterterms, nonlocal in position space but finite in number.

(2) Gauge theories with DWY fermions do not admit an immediate perturbative expansion. Such an expansion creates spurious infrared problems. At least some effects must be treated nonperturbatively, e.g., by normal ordering. This creates an infrared cutoff which permits a perturbative treatment of the remaining effects. Note that what is perturbative in the continuum may be nonperturbative on the lattice. In this paper some processes involving multiphoton vertices were summed to all orders, but the continuum limit was just the usual perturbative expansion.

(3) The relation between gauge invariance and Ward identities in a nonlocal theory is rather subtle. Results normally ascribed to gauge invariance are actually consequences of gauge invariance plus locality. A Ward identity is a statement that some component of A_μ—such as the longitudinal part $\vec{A}_L$—does not contribute to physical processes. In a local theory $\vec{A}_L$ cannot appear in the Hamiltonian alone, but only in the combination $\vec{A}_L + \vec{A}_T = \vec{A}$ which is a local field. A constraint on the coupling of $\vec{A}_L$ is then a constraint on the coupling of $\vec{A}$ and hence $\vec{A}_T$. Thus Ward identities put constraints on physical processes involving real transverse photons, but only via the locality assumption.

In Hamiltonian Coulomb gauge DWY QED a counterterm was required to keep the fermion spectrum undoubled. In the form

$$\sum_{\vec{x},\vec{y},j} \bar{\psi}(\vec{x})\gamma_j \frac{1}{i}[D_j(\vec{x}-\vec{y}) - \mathfrak{D}_j(\vec{x}-\vec{y})]\,\psi(\vec{y}) \times \exp\left[iea \sum_{\vec{z}=\vec{x}}^{\vec{y}} A_L^j(\vec{z})\right]$$

it is fully invariant under time-independent gauge transformations. In a local theory it would be ruled out because $\vec{A}_L$ is a nonlocal function of $\vec{A}$, but on the lattice it is acceptable. (Of course it is only *useful* in the Coulomb gauge where $\vec{A}_L = 0$.) With this counterterm there is still a Ward identity insuring that $\vec{A}_L$ does not contribute to physical processes, but it no longer directly constrains $\vec{A}_T$. The Ward identity has been "lost" in the sense that it no longer functions to reduce the degree of divergence of a Coulomb-gauge Feynman diagram. Ward identity constraints on physical processes are recovered only in the continuum limit where the theory becomes local if properly renormalized. This is how the nonlocal lattice theory escapes the problem of the axial anomaly. The triangle diagram is divergent, and the Ward identity it obeys in the continuum limit is determined by the renormalization prescription.

B. Beyond perturbation theory

The results of this paper are rather formal in that they show what can be done with DWY lattice QED in perturbation theory and what counterterms are needed to do it. Continuum QED at this time is defined by its renormalized perturbation series, but a lattice theory presumably has a meaning even beyond the region of validity of perturbation theory. As remarked earlier, perturbation theory cannot predict a qualitative spectrum, but must instead be constructed around a zeroth-order approximation which already has the correct qualitative spectrum. It is important to ask whether the perturbation theory constructed in this paper accurately reflects the exact solution to the theory (5.1). In principle this should be determined by an exact renormalization-group treatment and analysis of the fixed points. The renormalization-group transformation should generate an action containing the counterterms required in perturbation theory. What can be said in the absence of such information?

There seem to be two possible scenarios based on the Ward identity

$$\sum_\mu S_\mu(k)\Gamma_\mu(p+k,p)=S_F^{-1}(p+k)-S_F^{-1}(p), \quad (8.1)$$

which is an exact property of the theory. If the exact fermion propagator describes an undoubled spectrum then S_F^{-1} has a discontinuity at some point p_0. Letting $p\to p_0$ and $k\to 0$ in Eq. (8.1) shows that Γ_μ must have a singularity there. This in itself is not a disaster since p_0 is normally of order $1/a$. A disaster occurs only if this singularity propagates down into the low-momentum (continuum) limit of some Green's function. This happens in naive perturbation theory where loops of high-momentum particles contribute to the low-momentum behavior of, for example, $\Pi_{\mu\nu}(p)$. If it happens in general then the theory has problems. If it does not happen, so that singularities are confined to high momenta, then the continuum limit may be as described perturbatively in this paper. The high-momentum singularities would be generated from the sum to all orders of the order-by-order nonsingular effective theory of Sec. V B. The conserved lattice axial current has no continuum limit due probably to singular contributions to its matrix elements.

If no infrared singularities arise at any momentum, then S_F^{-1} must be continuous and the fermion spectrum doubles. This happens nonperturbatively since the spectrum is undoubled at $e=0$. This scenario is suggested by the summation of the photon tadpole contributions to S_F (Fig. 7). Summing perturbation theory to all orders would not introduce any singularities but would merely restore gauge invariance, which was lost order by order. The axial current could have a nonanomalous continuum limit, the anomaly being cancelled between the doubled fermion species. It is even possible that both these scenarios could occur, each characterizing a different phase of the lattice theory. The DWY lattice gauge theory (5.1) could thus have an extremely rich and interesting structure beyond perturbation theory. In my opinion it is extremely important, though difficult, to learn which of these cases occurs. The possibility that the fermion spectrum multiplicity is determined dynamically does not seem to have been previously suggested, and would add a new dimension to our understanding of the realization of chiral symmetry in lattice theories.

ACKNOWLEDGMENTS

I thank Sidney Drell, Luuk Karsten, Helen Quinn, and Marvin Weinstein for useful conversations throughout the course of this work. This work was supported by the Department of Energy, Contract No. E-AC03-76SF00515.

APPENDIX

I prove that, given a set of Feynman rules periodic in all momenta, periodic δ functions can be used to do trivial momentum integrations just as ordinary δ functions are used in continuum theories.

It suffices to show that if

$$I\equiv\int_{-\Lambda}^{\Lambda}dk_1\cdots dk_n F(k_1,\ldots,k_n)\,\delta_{\text{per}}[k_1-G(k_2,\ldots,k_n)], \quad (\text{A1})$$

where F is periodic in k_1 with period 2Λ, then

$$I=\int_{-\Lambda}^{\Lambda}dk_2\cdots dk_n F[G(k_2,\ldots,k_n),k_2,\ldots,k_n]. \quad (\text{A2})$$

To do this, write (A1) as

$$I=\sum_{m=-\infty}^{\infty}\int_{-\Lambda}^{\Lambda}dk_1\cdots dk_n F(k_1,\ldots,k_n)\,\delta[k_1-G(k_2,\ldots,k_n)+2m\Lambda]. \quad (\text{A3})$$

In the mth term change variables from k_1 to $k_1'=k_1+2m\Lambda$, giving

$$I=\sum_{m=-\infty}^{\infty}\int_{(2m-1)\Lambda}^{(2m+1)\Lambda}dk_1'\int_{-\Lambda}^{\Lambda}dk_2\cdots dk_n F(k_1'-2m\Lambda,k_2,\ldots,k_n)\delta[k_1'-G(k_2,\ldots,k_n)]$$
$$=\int_{-\infty}^{\infty}dk_1'\int_{-\Lambda}^{\Lambda}dk_2\cdots dk_n F(k_1',k_2,\ldots,k_n)\delta[k_1'-G(k_2,\ldots,k_n)]=\int_{-\Lambda}^{\Lambda}dk_2\cdots dk_n F[G(k_2,\ldots,k_n),k_2,\ldots,k_n], \tag{A4}$$

by periodicity.

If the function F is initially defined only for $-\Lambda<k_1<\Lambda$ then the above holds if F is extended periodically.

---

[1]D. Horn and S. Yankielowicz, Nucl. Phys. B161, 533 (1979); B. Svetitsky, S. D. Drell, H. R. Quinn, and M. Weinstein, Phys. Rev. D 22, 490 (1980).

[2]M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. D 20, 1915 (1979); Phys. Rev. Lett. 42, 1390 (1979); L. McLerran and B. Svetitsky, Phys. Lett. 98B, 195 (1981).

[3]A. Guth, Phys. Rev. D 21, 2291 (1980), and references therein.

[4]K. G. Wilson, Phys. Rev. D 10, 2445 (1974); in *New Phenomena in Subnuclear Physics*, edited by A. Zichichi (Plenum, New York, 1977).

[5]S. D. Drell, M. Weinstein, and S. Yankielowicz, Phys. Rev. D 14, 487 (1976); 14, 1627 (1976). Actually the "DWY derivative" seems to have been introduced by G. Wentzel, Helv. Phys. Acta. 13, 269 (1940).

[6]L. H. Karsten and J. Smit, Nucl. Phys. B183, 103 (1981).

[7]W. Kerler, Phys. Rev. D 23, 2384 (1981).

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

[9]H. S. Sharatchandra, Phys. Rev. D 18, 2042 (1978).

[10]L. H. Karsten and J. Smit, Nucl. Phys. B144, 536 (1978).

[11]L. H. Karsten and J. Smit, Phys. Lett. 85B, 100 (1979).

[12]Y. Nakawaki, Prog. Theor. Phys. 59, 248 (1978); 61, 1197 (1979).

[13]J. B. Bronzan, Phys. Rev. D 21, 2270 (1980).

[14]W. Rudin, *Principles of Mathematical Analysis* (McGraw-Hill, New York, 1964).

[15]I thank Marvin Weinstein for stressing this point until I finally understood it.