PHYSICAL REVIEW D **73,** 114511 (2006)

Observations on staggered fermions at nonzero lattice spacing

Claude Bernard,[1] Maarten Golterman,[2,*] and Yigal Shamir[3]

[1]*Department of Physics, Washington University, St. Louis, Missouri 63130, USA*

[2]*Grup de Física Teòrica and IFAE Universitat Autònoma de Barcelona, 08193 Barcelona, Spain*

[3]*School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Ramat Aviv, 69978 Israel*

(Received 30 April 2006; published 22 June 2006)

We show that the use of the fourth-root trick in lattice QCD with staggered fermions corresponds to a nonlocal theory at nonzero lattice spacing, but argue that the nonlocal behavior is likely to go away in the continuum limit. We give examples of this nonlocal behavior in the free theory, and for the case of a fixed topologically nontrivial background gauge field. In both special cases, the nonlocal behavior indeed disappears in the continuum limit. Our results invalidate a recent claim that at nonzero lattice spacing an additive mass renormalization is needed because of taste-symmetry breaking.

DOI: 10.1103/PhysRevD.73.114511 PACS numbers: 12.38.Gc, 11.15.Ha

I. INTRODUCTION

Staggered fermions [1] have long been in use as a method for formulating the quark sector of lattice QCD. The main advantages are that they are relatively inexpensive when it comes to including sea-quark effects in lattice computations, and that they have an exact chiral symmetry in the limit of vanishing bare quark mass. The combination of these two advantages makes it possible to reach rather low quark masses, which are essential for any serious phenomenological applications of lattice QCD.

These benefits come at a price, however. A theory with one flavor of staggered fermion on the lattice yields a theory with four quarks in the continuum limit. This is a consequence of fermion species doubling, which is unavoidable in any situation in which an exact chiral symmetry is preserved on the lattice. In modern language, these four quarks per flavor of lattice staggered fermion are referred to as "tastes." Only in the continuum limit does the theory recover a full $SU(4)$ taste symmetry, whereas at any nonzero value of the lattice spacing this group is broken to a smaller discrete subgroup [2].

In principle, the four tastes can be given different masses [2], but this is not what is done in practice.[1] Instead, each staggered flavor (up, down, or strange) is given a single mass, leading to four tastes of degenerate quarks per flavor. In order to obtain a theory with only one quark per flavor appearing in sea-quark loops, one reduces the number of tastes by taking the fourth root[2] of the degenerate-mass staggered determinant for each flavor [3].

This formulation of the sea-quark sector of QCD does not necessarily correspond to a local field theory at nonzero lattice spacing a. The potential lack of locality has been the cause for much concern recently [4] about the application of staggered fermions to high-precision hadron phenomenology. At issue is: (1) whether the theory is local at $a \neq 0$, and (2) whether the theory, if nonlocal at $a \neq 0$, becomes local in the continuum limit. An alternative way to phrase the second question is to ask whether the theory is in the correct universality class.

In this paper, we will argue (Sec. II) that the theory with the fourth root of the staggered determinant is indeed nonlocal at nonzero a, but that this does not imply that the answer to the second question is negative. We connect the issue of locality to the role of taste and chiral symmetries. In Sec. III, we give some simple examples that show how the correct local continuum theory may indeed be obtained. In addition, we demonstrate that recent claims about the properties of staggered fermions at nonzero a, in particular, about the renormalization of the bare mass [5], are incorrect. A concluding section summarizes our arguments and results, while the Appendix collects some useful properties of various Dirac operators in the taste basis.

II. GENERAL CONSIDERATIONS

We begin by giving our general argument. Suppose that the theory with the fourth root *did* correspond to a local field theory on the lattice at nonzero a. By definition, this would require that the two theories differ only by a local functional. In other words,

$$\mathrm{Det}^{1/4}(D_{\mathrm{stag}}) = \mathrm{Det}(D)\exp(-\tfrac{1}{4}\delta S_{\mathrm{eff}}), \tag{2.1}$$

where D_{stag} is the staggered Dirac operator, D is a local lattice Dirac operator that describes one quark field in the continuum limit, and δS_{eff} is a local effective action for the

*Permanent address: Department of Physics and Astronomy, San Francisco State University, San Francisco, CA 94132, USA.

[1]One reason is that breaking the taste degeneracy requires additional hopping terms in the lattice action, which, for a generic choice, make the fermion determinant complex. Also, the existence of a partially conserved continuous chiral symmetry depends on the choice of mass term.

[2]In the isospin limit, the up-down sector is represented by a square root of a staggered determinant with the common light quark mass.

gauge field.[3] Saying that $\delta S_{\rm eff}$ is local means that it produces only effects at the scale of the cutoff. This would imply that, apart from a renormalization of the gauge coupling constant, the presence of $\delta S_{\rm eff}$ would not affect the behavior at any physical length scale that is to be held constant as the lattice spacing is taken to zero.

It is rather easy to see that this set of assumptions leads to a conflict with what we know to be true about the original staggered theory, i.e. the one without the fourth root of the determinant. Taking the fourth power of Eq. (2.1), we have

$$\mathrm{Det}\,(D_{\rm stag}) = \mathrm{Det}^4(D)\exp(-\delta S_{\rm eff}). \qquad (2.2)$$

Under our assumption $\delta S_{\rm eff}$ is local, and it therefore cannot change the long-distance behavior of any correlation function. In particular, it cannot have any effect on the Goldstone-boson (GB) masses predicted by the staggered theory defined by $D_{\rm stag}$, and those predicted by the theory defined by

$$D_{4t} = D \otimes \mathbf{1}, \qquad (2.3)$$

where the second factor is a unit 4×4 matrix, to be interpreted as the identity matrix in taste space. The operator D describes a lattice theory with one taste; in a finite volume, the size of the matrix D is in fact 4 times smaller than the size of $D_{\rm stag}$. Clearly, we have that $\mathrm{Det}(D_{4t}) = \mathrm{Det}^4(D)$, and the lattice theory defined by D_{4t} has a continuous $SU(4)$ taste symmetry.

We can now compare what we know about the GB spectrum of the two theories. In the theory defined by D_{4t}, there will be 15 GBs, transforming in the adjoint representation of $SU(4)$, with possibly a common nonvanishing mass if the operator D violates chiral symmetry and/or is not massless. Under our assumption described above, $\delta S_{\rm eff}$ does not change this fact: all long-distance physics would be contained in D_{4t}.

The GB spectrum of the staggered theory is completely different, irrespective of the value of the staggered bare quark mass. Of course, in the continuum limit, one recovers 15 degenerate (pseudo)GBs, but at nonzero lattice spacing, they split up into at least four [7,8], and up to seven [9], nondegenerate irreducible representations, consistent with the lattice symmetry group of the staggered theory. Indeed, at strong coupling [10,11], there is only one exact GB (at zero quark mass), because of the exact $U(1)_\epsilon$ axial symmetry [10].

It thus becomes clear that our assumption on $\delta S_{\rm eff}$ cannot be correct. The effective action $\delta S_{\rm eff}$ has to know about the long-distance effects of taste-symmetry breaking, and cannot be a local functional of the lattice gauge field. Of course, given a local operator D, one can always define $\delta S_{\rm eff}$ through Eq. (2.2) or Eq. (2.1) (as long as we consider gauge fields on which D has no exact zero modes, cf. Sec. III B), but what we find is that $\delta S_{\rm eff}$ cannot be local. This shows that the theory defined by taking the fourth root of the staggered determinant must be nonlocal at $a \neq 0$.

It also follows that the staggered theory without the fourth root cannot be written as an $SU(4)$-symmetric local theory at $a \neq 0$. In Ref. [5], it was assumed that Eq. (2.2) held with $\delta S_{\rm eff}$ local.[4] However, we have shown that such a decomposition is not possible.

What might be confusing is that the left-hand side of Eq. (2.2) is the determinant of a local operator, $D_{\rm stag}$. Clearly, the determinant, or equivalently the effective action $S_{\rm eff} = -\mathrm{Tr}\log(D_{\rm stag})$, is a nonlocal object. What we observe is simply the fact that the nonlocality of $S_{\rm eff}$ cannot be reproduced entirely by the effective action for the operator D_{4t}, because of a conflict between the symmetries of $D_{\rm stag}$ and D_{4t} at nonzero lattice spacing. It *is* true that $D_{\rm stag}$ itself can be written as the sum of taste-invariant and taste-breaking local operators:

$$D_{\rm stag} = D\otimes\mathbf{1} + \sum_A D_A \otimes \Xi_A, \qquad (2.4)$$

with the Ξ_A a set of 15 $SU(4)$-algebra valued (Hermitian) generators in taste space,[5] with D and D_A all local.[6] Considering the determinant, however, one has that

$$\begin{aligned} S_{\rm eff} &= -\log\mathrm{Det}(D_{\rm stag}) \\ &= -4\log\mathrm{Det}(D) - \log\mathrm{Det}\Big(1 + \sum_A D^{-1}D_A\otimes\Xi_A\Big). \end{aligned} \qquad (2.5)$$

This split of the effective action corresponds to choosing a specific D in Eq. (2.2). Because of the presence of D^{-1}, the second term produces a nonlocal $\delta S_{\rm eff}$, even though the taste-breaking part of the Dirac operator in Eq. (2.4) is local. What we have argued above, on the basis of the GB spectrum of the staggered theory without fourth root, is that no split of the form of Eq. (2.2) exists for which $\delta S_{\rm eff}$ is local. While it is generally accepted that the taste-breaking effects of the operator $\sum_A D_A\otimes\Xi_A$ vanish in the continuum limit, it is precisely the nonlocality of $\delta S_{\rm eff}$ that causes the 15 GBs of the staggered theory to be nondegenerate at $a \neq 0$.

While our argument demonstrates that no local lattice theory exists with a fermion determinant equal to the fourth root of the staggered determinant, it leaves open the question of whether the nonlocal behavior persists in the continuum limit. Nevertheless, Eq. (2.5) lends support to the

[3]Adams [6] has recently emphasized that Eq. (2.1) with $\delta S_{\rm eff}$ local is indeed the proper definition of locality of the rooted theory at $a \neq 0$; requiring $\delta S_{\rm eff} = 0$ would be too strong.

[4]$\exp(-\delta S_{\rm eff})$ was written as $\mathrm{Det}(T)$ in Ref. [5].

[5]We may choose this set to be $\{\xi_\mu, i\xi_\mu\xi_\nu, i\xi_\mu\xi_5, \xi_5\}$ with ξ_μ a set of 4×4 matrices satisfying $\{\xi_\mu, \xi_\nu\} = 2\delta_{\mu\nu}$.

[6]Lattice symmetries, such as $U(1)_\epsilon$ symmetry, further restrict which Ξ_A can appear, as well as the form the D_A can take.

conjecture that the nonlocalities vanish in this limit. Although $\delta S_{\rm eff}$ is nonlocal, the operator $\sum_A D_A \otimes \Xi_A$ is of order a. Thus, the effects of $\delta S_{\rm eff}$ should vanish when the limit $a \to 0$ is taken while keeping physical momenta fixed.[7]

In closing this section we comment on Refs. [12,13]. Rather than looking at the fourth root of the staggered determinant, these references directly study the locality properties of a certain fourth root[8] of the staggered Dirac operator itself. Since a huge number of operators can have the same determinant, demanding that a fourth root of the Dirac operator itself be local is much stronger than the sufficient condition that the fourth root of the staggered determinant correspond to some local operator in the sense of (2.1), in the continuum limit. Further, taking the fourth root of the Dirac operator does not satisfy the intuitive requirement of reducing the number of degrees of freedom by a factor of 4. Even starting from the operator $D_{4t} = D \otimes \mathbf{1}$, the procedure of Refs. [12,13] would fail to find the local operator D, which has four times fewer degrees of freedom than D_{4t}, and instead would find the nonlocal $\sqrt[4]{D} \otimes \mathbf{1}$. (See Ref. [14], Sec. IX D 7, for a discussion of this point.) It is therefore not surprising that Ref. [12] finds a nonlocal result even in the continuum limit of the free theory, while there are two explicit local constructions in that case: Ref. [15] and Ref. [6].

III. EXAMPLES

To make the discussion more concrete, we now give a possible prescription for the construction of the operator D_{4t} in Eq. (2.4). We begin with a massive staggered Dirac operator $D_{\rm stag}(m) = D_{\rm stag}(0) + m$ with bare quark mass m in the one-component formalism.[9] There exists a gauge-covariant unitary transformation $Q^{(0)}$ which puts the theory into the taste representation of Refs. [16,17].[10] We may however carry out this transformation as a Gaussian renormalization-group (RG) blocking, leading to a staggered Dirac operator in the taste representation $D_{\rm taste}(m)$ given by [15]

$$D_{\rm taste}^{-1}(m) = \frac{1}{\alpha} + Q^{(0)} D_{\rm stag}^{-1}(m) Q^{(0)\dagger}, \qquad (3.1)$$

where α is a parameter which appears in the Gaussian blocking kernel. We then have that

$$\mathrm{Det}\,(D_{\rm stag}(m)) = \mathrm{Det}(G^{-1})\mathrm{Det}(D_{\rm taste}(m)), \qquad (3.2)$$

with

$$G^{-1} = \frac{1}{\alpha} D_{\rm stag}(m) + Q^{(0)\dagger} Q^{(0)} = \frac{1}{\alpha} D_{\rm stag}(m) + 1, \qquad (3.3)$$

where in the last step we have used the fact that the kernel $Q^{(0)}$ is unitary for this "RG blocking." For $\alpha \to \infty$, one recovers a transformation of the type considered in Ref. [17], but we will take α to be finite here. Because G^{-1} is a Dirac operator with a mass of order α in lattice units, all the long-distance physics should be contained in $D_{\rm taste}(m)$.

Again following Ref. [15], one may use $D_{\rm taste}$ as the input for n true RG blocking steps (in which actual thinning out of fermionic degrees of freedom occurs) with an RG blocking kernel $Q^{(n)}$. The nth blocking step takes us from a lattice with spacing a_{n-1} to a lattice with spacing $a_n = 2a_{n-1}$; a_0 is defined to be the spacing of the lattice associated with $D_{{\rm taste},0} \equiv D_{\rm taste}$ and is twice the spacing of the original lattice on which $D_{\rm stag}$ is defined. Blocked operators $D_{{\rm taste},n}$ and G_n^{-1} result from this process, with, recursively,

$$\begin{aligned} D_{{\rm taste},n}^{-1}(m) &= \frac{1}{\alpha} + Q^{(n)} D_{{\rm taste},n-1}^{-1}(m) Q^{(n)\dagger}, \\ G_n^{-1} &= \frac{1}{\alpha} D_{{\rm taste},n-1}(m) + Q^{(n)\dagger} Q^{(n)}, \\ Q^{(n)} Q^{(n)\dagger} &= c\mathbf{1}, \end{aligned} \qquad (3.4)$$

where c is a positive constant, and here "$\mathbf{1}$" stands for the Kronecker delta on the coarse lattice. One expects that the long-distance physics is entirely carried by $D_{{\rm taste},n}^{-1}$, which is manifestly the sum of a smeared quark propagator and a contact term, while $\mathrm{Tr}\log(G_n^{-1})$ is a local functional of the gauge field. The determinants are related by

$$\mathrm{Det}\,(D_{\rm stag}(m)) = \mathrm{Det}(D_{{\rm taste},n}(m)) \prod_{k=0}^{n} \mathrm{Det}(G_k^{-1}), \qquad (3.5)$$

with $G_0^{-1} \equiv G^{-1}$ from Eq. (3.3). While Eq. (3.5) resembles Eq. (2.2), it is fundamentally different. In Eq. (3.5), both $\mathrm{Det}(D_{\rm stag}(m))$ and $\mathrm{Det}(D_{{\rm taste},n}(m))$ describe the same long-distance physics, and the factor $\prod_{k=0}^{n} \mathrm{Det}(G_k^{-1})$ is expected to be a local functional of the gauge field. For any finite n, both $D_{\rm stag}(m)$ and $D_{{\rm taste},n}(m)$ break taste symmetry, consistent with our general arguments above.

The massless one-component action is invariant under $U(1)_\epsilon$ transformations [10],

$$\delta\chi(x) = i\epsilon(x)\chi(x), \qquad \delta\bar{\chi}(x) = i\epsilon(x)\bar{\chi}(x), \qquad (3.6)$$

because $\epsilon(x) \equiv (-1)^{x_1+x_2+x_3+x_4}$ anticommutes with $D_{\rm stag}(0)$. From

$$Q^{(0)}\epsilon = (\gamma_5 \otimes \xi_5) Q^{(0)}, \qquad (3.7)$$

it follows [15] that $D_{\rm taste} = D_{{\rm taste},0}$ satisfies a Ginsparg-

[7]We expect that the continuum limit will have to be taken before the theory is continued to Minkowski space.

[8]These references actually study the square root rather than the fourth root, but for simplicity we do not bother making this distinction here since it is not relevant to the points under discussion.

[9]We will make the dependence on the quark mass explicit for the rest of this paper.

[10]The transformation $Q^{(0)}$ is not unique; see Ref. [18] for details.

Wilson (GW) relation [19]

$$\{\gamma_5 \otimes \xi_5, D^{-1}_{\rm taste}(0)\} = \frac{2}{\alpha}(\gamma_5 \otimes \xi_5), \tag{3.8}$$

if the original operator $D_{\rm stag}$ is massless.[11] Using Eqs. (3.1) and (3.7) one can show that $(\gamma_5 \otimes \xi_5)D_{\rm taste}(0)$ is Hermitian. Equation (3.8) then implies that the eigenvalues of $D_{\rm taste}(0)$ lie on a circle in the complex plane crossing the real axis at 0 and α, with center at $\alpha/2$.

If we start with a massive staggered Dirac operator $D_{\rm stag}(m)$ in the one-component formalism, we obtain a corresponding massive operator $D_{\rm taste}(m)$ in the taste representation. Using the fact that $D_{\rm taste}(m) = D_{\rm taste}(0) + m$ for $\alpha = \infty$, it is straightforward to show for finite α that

$$D_{\rm taste}(m) = \frac{D_{\rm taste}(0) + m(1 - \frac{1}{\alpha}D_{\rm taste}(0))}{1 + \frac{m}{\alpha}(1 - \frac{1}{\alpha}D_{\rm taste}(0))}. \tag{3.9}$$

This operator is local, because the second term in the denominator is small compared to the 1 (as long as $m \ll 1$ in lattice units). The eigenvalues still lie on a circle, now with center $(\alpha/2 + m)/(1 + m/\alpha)$ and radius $(\alpha/2)/(1 + m/\alpha)$. In particular, the two possible real eigenvalues are $m/(1 + \frac{m}{\alpha})$ and α.

In general, $D_{{\rm taste},n}(0)$ satisfies a GW relation for any n, since the RG kernels $Q^{(n)}$ for $n = 1, \ldots$ are trivial with respect to Dirac and taste indices. Explicitly, we have that [15]

$$\{\gamma_5 \otimes \xi_5, D^{-1}_{{\rm taste},n}(0)\} = \frac{2}{\alpha_n}(\gamma_5 \otimes \xi_5), \qquad \alpha_n = \frac{1-c}{1-c^{n+1}}\alpha. \tag{3.10}$$

$D_{{\rm taste},n}$ is not invariant under the full taste $SU(4)$ for any finite n. We may construct an $SU(4)$ taste-invariant operator by simply taking the trace in taste space:

$$D_{{\rm inv},n}(m) = \tfrac{1}{4}{\rm tr}(D_{{\rm taste},n}(m)) \otimes \mathbf{1}, \tag{3.11}$$

where tr denotes a trace over taste only. This operator is not necessarily massless if we set $m = 0$, but whatever quark mass the theory defined by $D_{{\rm inv},n}(m)$ has, it is proportional to the unit matrix in taste space. It is also clear that $D_{{\rm inv},n}(0)$ does not satisfy a GW relation.

However, it is straightforward to construct an operator that does obey a GW relation. In order to do this, we note that $D_{{\rm inv},n}$ has no fermion species doublers for finite α. (We will show this explicitly in Sec. III A.) Furthermore, the fact that ϵ anticommutes with $D_{\rm stag}(0)$, combined with anti-Hermiticity of $D_{\rm stag}(0)$, implies that

$$(D^{-1}_{\rm stag}(m))^\dagger = \epsilon D^{-1}_{\rm stag}(m)\epsilon. \tag{3.12}$$

Using Eqs. (3.1), (3.4), (3.7), and (3.11), it is then easy to see that $\gamma_5 D_{{\rm inv},n}(m)$ is Hermitian. We may thus construct a taste-invariant overlap operator, just as when one starts with a Wilson-Dirac operator [20]:

$$D_{{\rm ov},n} \equiv \frac{\alpha_n}{2}\left(1 - \gamma_5 {\rm sign}\left(\gamma_5\left(1 - \frac{2}{\alpha_n}D_{{\rm inv},n}(0)\right)\right)\right), \tag{3.13}$$

with α_n given in Eq. (3.10). Since this operator is taste invariant, it satisfies a GW relation for any taste matrix Ξ:

$$\{\gamma_5 \otimes \Xi, D^{-1}_{{\rm ov},n}\} = \frac{2}{\alpha_n}(\gamma_5 \otimes \Xi). \tag{3.14}$$

It follows that $D_{{\rm ov},n}$ is a massless operator.[12]

The operator $D_{{\rm ov},n}$ can be written as $D \otimes \mathbf{1}$ as in Eq. (2.3), and the resulting D is a possible choice for use in Eqs. (2.1) and (2.2). Obviously, we can only have that $D_{{\rm ov},n} \to D_{{\rm taste},n}(m)$ for $n \to \infty$ if we take the original one-component staggered operator to be massless, so that Eq. (3.10) coincides with Eq. (3.14) for $\Xi = \xi_5$. This overlap operator is "natural," because it has been constructed such that the difference between $D_{{\rm ov},n}$ and $D_{{\rm taste},n}(0)$ is expected to be of order $a_0^2/a_n^2 = 1/2^{2n}$ [18]. The distinction is that, by construction, $D_{{\rm ov},n}$ has exact $SU(4)$ taste symmetry [in fact a full chiral $SU(4)_L \times SU(4)_R$], while $D_{{\rm taste},n}(0)$ does not. The expectation that the difference decreases like $1/2^{2n}$ arises from the similar expectation that taste symmetry is restored in the unrooted staggered theory as we take $n \to \infty$, i.e. as the lattice spacing of the original (unblocked) theory is sent to zero.

The sequence of overlap operators can be made massive by choosing

$$D_{{\rm ov},n}(m) = D_{{\rm ov},n}(0) + D_{{\rm inv},n}(m) - D_{{\rm inv},n}(0), \qquad m \neq 0, \tag{3.15}$$

with m the original bare staggered mass, and $D_{{\rm ov},n}(0) \equiv D_{{\rm ov},n}$ of Eq. (3.13). Our choice is different from the massive overlap operator commonly used in the literature. The reason is that, this way, we maintain the above naturalness property for $m \neq 0$ as well. For details, see Appendix A 1. Unless the $n \to \infty$ limit is taken, the two theories defined by $D_{{\rm taste},n}(m)$ and $D_{{\rm ov},n}(m)$ will not have the same renormalized mass; but since the mass in both theories renormalizes multiplicatively, both theories are massless for $m = 0$. This follows from the fact that both $D_{{\rm taste},n}(0)$ and $D_{{\rm ov},n}(0)$ have a Ginsparg-Wilson-Lüscher (GWL) chiral symmetry [21]. Any of the operators $D_{{\rm ov},n}(m)$ is a possible choice for D_{4t} in Eq. (2.3).

[11]Note that this reduces to an ordinary chiral symmetry for $\alpha \to \infty$.

[12]This is true even if the original operator $D_{\rm stag}(m)$ is not massless, i.e. if $D_{{\rm inv},n}(0)$ is replaced by $D_{{\rm inv},n}(m)$ on the right-hand side of Eq. (3.13).

A. The free case

The free case provides an explicit example of Eq. (2.2), with $\delta S_{\rm eff}$ nonlocal. We choose $n = 0$ and use $D_{\rm inv,0}(m)$ for $D_{4t} = D \otimes \mathbf{1}$ on the right-hand side of this equation. In the free case, a $Q^{(0)}$ exists such that [16,17]

$$Q^{(0)} D_{\rm stag}(m) Q^{(0)\dagger} = \sum_\mu (i(\gamma_\mu \otimes \mathbf{1}) \sin p_\mu + (\gamma_5 \otimes \xi_\mu \xi_5) \times (1 - \cos p_\mu)) + (\mathbf{1} \otimes \mathbf{1}) m \quad (3.16)$$

in momentum space. Using this $Q^{(0)}$ in Eqs. (3.1) and (3.11), we obtain

$$D_{\rm taste,0}(m) = \frac{\sum_\mu (i(\gamma_\mu \otimes \mathbf{1}) \bar{p}_\mu + \frac{1}{2}(\gamma_5 \otimes \xi_\mu \xi_5) \hat{p}_\mu^2) + (\mathbf{1} \otimes \mathbf{1})(m + \frac{1}{\alpha}(\hat{p}^2 + m^2))}{1 + \frac{2m}{\alpha} + \frac{1}{\alpha^2}(\hat{p}^2 + m^2)},$$
$$D_{\rm inv,0}(m) = \frac{\sum_\mu i(\gamma_\mu \otimes \mathbf{1}) \bar{p}_\mu + (\mathbf{1} \otimes \mathbf{1})(m + \frac{1}{\alpha}(\hat{p}^2 + m^2))}{1 + \frac{2m}{\alpha} + \frac{1}{\alpha^2}(\hat{p}^2 + m^2)}, \quad (3.17)$$

where

$$\bar{p}_\mu \equiv \sin p_\mu, \qquad \hat{p}_\mu \equiv 2\sin(p_\mu/2), \qquad \hat{p}^2 \equiv \sum_\mu \hat{p}_\mu^2. \quad (3.18)$$

We see that $D_{\rm inv,0}(m)$ is a Wilson-like Dirac operator, and thus has no fermion doubling as long as α is finite. The massless overlap operator of Eq. (3.13) in the free case is

$$D_{\rm ov,0}(0) = \frac{\alpha}{2}\left(1 - \frac{1 - \frac{2}{\alpha} D_{\rm inv,0}(0)}{\sqrt{1 - \frac{\alpha^2 \sum_\mu \hat{p}_\mu^4}{(\alpha^2 + \hat{p}^2)^2}}}\right) = D_{\rm inv,0}(0) + O(p^4). \quad (3.19)$$

The argument of the square root is strictly positive as long as $\alpha < 2$.

We may now calculate $\delta S_{\rm eff}$ for the free case from Eqs. (2.2) and (2.5), choosing $D \otimes \mathbf{1} = D_{\rm inv,0}(m)$ and using Eq. (3.17). We find

$$e^{-\delta S_{\rm eff}} = \prod_p \left(1 + \frac{\frac{1}{4}\sum_\mu \hat{p}_\mu^4}{\bar{p}^2 + (m + \frac{1}{\alpha}(\hat{p}^2 + m^2))^2}\right)^8. \quad (3.20)$$

Defining $\delta \mathcal{L}_{\rm eff}$ by $\delta S_{\rm eff} = -{\rm Tr}(\delta \mathcal{L}_{\rm eff})$, we have $\delta \mathcal{L}_{\rm eff}(p) \sim (\sum_\mu p_\mu^4)/(p^2 + m^2)$ at small p (and $am \ll 1$). This implies that the Fourier transform $\delta \mathcal{L}_{\rm eff}(x - y)$ decays like inverse powers of the separation $x - y$ (or its components) times a factor $e^{-m|x-y|}$. Because m is a physical scale, $\delta S_{\rm eff}$ is nonlocal, Choosing $D = D_{\rm ov,0}(m)$ instead in Eq. (2.2) gives a similar result. While this only demonstrates the nonlocality of $\delta S_{\rm eff}$ in the free case, it is clear that in the interacting case the nonlocality would be gauge-field dependent [see the discussion around Eq. (2.5)].

If we choose to consider the case of $D_{{\rm taste},n}(m)$ for $n > 0$, the exact expressions become more cumbersome. However, using the free-theory results of [15], it is possible to show that

$$D_{{\rm taste},n}(m) = \sum_\mu \left(i(\gamma_\mu \otimes \mathbf{1}) p_\mu + \frac{1}{2^{n+1}}(\gamma_5 \otimes \xi_\mu \xi_5) p_\mu^2 + (\mathbf{1} \otimes \mathbf{1}) m + O\left(\frac{m^2}{2^n}, \frac{p^3}{2^{2n}}\right)\right), \quad (3.21)$$

for small p, leading to

$$\delta S_{\rm eff} = -8 \sum_p \frac{1}{2^{2(n+1)}} \frac{\sum_\mu p_\mu^4}{p^2 + m^2} + \cdots \quad (3.22)$$

for small p. This shows explicitly how $\delta S_{\rm eff} \to 0$ for $n \to \infty$, but also how $\delta S_{\rm eff}$ is nonlocal for any fixed n.

The free case is rather special in that there are no pions, so the argument of Sec. II does not apply. This allows for the possibility that there may be other choices for the operator D in Eq. (2.2) for which $\delta S_{\rm eff}$ is local. Indeed, Adams [6] has constructed such an operator, which has range $\sqrt{a/m}$ and $\delta S_{\rm eff} = 0$. However, the general features of the GB spectrum show that a similar construction is not possible in the interacting case.

B. Background with nonzero topological charge

Another example is provided by the staggered Dirac operator in the background of a smooth gauge field with fixed topological charge $Q = 1$. Here, we take the operator $D_{4t} = D \otimes \mathbf{1}$ of Eq. (2.3) to be an overlap operator, and use it (setting $m = 0$) to define the topological charge of the gauge field under consideration. As already mentioned, a possible choice would be one of the $D_{{\rm ov},n}$ of Eq. (3.13), but in principle any overlap operator will do.

Our choice of gauge field implies that the operator D will have one exact zero mode for quark mass $m = 0$, and thus one eigenvalue proportional to m when $m \neq 0$. It follows that ${\rm Det}^4(D)$ on the right-hand side of Eq. (2.2) will be proportional to m^4 and vanish as $m \to 0$. The operator $D_{\rm stag}$ on the left-hand side of Eq. (2.2) will not have any exact zero modes for a generic gauge-field configuration (in any topological sector) at nonzero lattice spacing. Instead, it will have four nondegenerate corresponding eigenvalues

$$\lambda_i = m + c_i a^\gamma, \qquad i = 1, \ldots, 4, \tag{3.23}$$

with γ a positive exponent.[13] In general none of the c_i will be exactly zero: If we consider for instance an instanton with radius ρ, the c_i will be proportional to $\rho^{-\gamma-1}$. It follows that

$$e^{-\delta S_{\rm eff}} \propto \prod_i \left(1 + \frac{c_i a^\gamma}{m}\right). \tag{3.24}$$

This is just the zero-mode contribution to

$$\delta S_{\rm eff} = -{\rm Tr}\log\left(1 + \sum_A D^{-1} D_A \otimes \Xi_A\right), \tag{3.25}$$

in Eq. (2.5), where now D has been chosen to be an overlap operator.[14] Thus, the $1/m$ signals the dependence of $\delta S_{\rm eff}$ on the nonlocal D^{-1}. Note also that $\delta S_{\rm eff}$ diverges in the chiral limit for any nonzero lattice spacing, exhibiting the well-known fact that the chiral and continuum limits do not commute [22].

C. Consequences for Ref [5]

Our results invalidate the basic assumption made in Ref. [5], which was that $\delta S_{\rm eff}$ defined by Eq. (2.2) cannot affect long-distance physics even at nonzero a. Instead, we find that $\delta S_{\rm eff}$ has to contain long-distance physics at $a \neq 0$ because of the mismatched symmetries of $D_{\rm stag}$ and D_{4t} of Eqs. (2.2) and (2.3). Contrary to what was suggested in Ref. [5], it is not possible to reconcile the theories described by $D_{\rm stag}$ and D_{4t} by an additive shift in the quark mass. Unlike the theory defined by $D_{\rm stag}$, the theory defined by D_{4t} has to contain 15 (pseudo) Goldstone bosons, which remain degenerate even if they pick up a mass due to the presence of an explicit ($SU(4)$-symmetric) quark mass. Based on a comparison of zero modes of $D_{\rm stag}$ and D_{4t}, Ref. [5] furthermore argues that the quark masses of the two theories have to be related by an $O(a^2)$ additive quark mass renormalization. That argument fails, however, precisely because the relation between the two theories is nonlocal. Our construction of the overlap operators $D_{{\rm ov},n}(m)$ demonstrates that in fact $SU(4)$-symmetric lattice Dirac operators exist which become exactly massless when the staggered quark mass m is set equal to zero. We emphasize however that any overlap operator can be used to invalidate the claim of Ref. [5], as discussed in Sec. III B. A correct description of the approach of the continuum limit as far as the physics of GBs is concerned is provided by staggered chiral perturbation theory [7,23].

[13]For generic eigenvalues, one expects that $\gamma = 1$. It could be that $\gamma = 2$ for zero modes. The precise value of γ does not affect our argument.

[14]In this case the sum over A includes a term with $\Xi_A = \mathbf{1}$.

IV. CONCLUSION

Our main result is a proof in Sec. II that the theory defined by the fourth root of the staggered fermion determinant does not correspond to a local theory at nonzero lattice spacing a. This follows from the fact that $SU(4)$ taste symmetry is broken at nonzero a in the unrooted staggered theory. If a local theory corresponding to the fourth-root theory existed, one could take four copies of it and construct a local theory with exact $SU(4)$ taste symmetry, cf. the theory defined by D_{4t} in Eq. (2.3). The $SU(4)$ symmetry implies that the 15 pseudo-Goldstone bosons of this theory must be degenerate. On the other hand, it is well known that the 15 pseudo-Goldstone bosons in the staggered theory at nonzero a are nondegenerate because of taste violations. There is thus a mismatch in the long-distance physics of the staggered and $SU(4)$ theories when $a \neq 0$. The contradiction implies that the rooted theory cannot be local at nonzero a: $\delta S_{\rm eff}$, defined through Eq. (2.2), must be nonlocal.

The key issue is then whether the nonlocality persists in the continuum limit. While this remains an open question, the argument given around Eq. (2.5) suggests that the theory is in the desired universality class as long at the continuum limit is taken before the chiral limit. In other words, it appears that locality will be restored for $a \to 0$ at any $m \neq 0$ (cf. Sec. III B). For recent theoretical results supporting this conjecture, we refer to Refs. [18,24].

While the main argument summarized above stands alone, we have discussed two examples that make our reasoning more concrete. The examples are provided by the staggered theory in the free case (Sec. III A) and in the background of a smooth gauge field with nonzero topological charge (Sec. III B). Starting from the staggered Dirac operator, we constructed a sequence of overlap operators $D_{{\rm ov},n}$ in Sec. III, which can be used to give a fermionic definition of topological charge suited to our arguments. In both examples, we find that $\delta S_{\rm eff}$ is explicitly nonlocal, but that the nonlocal behavior disappears in the continuum limit.

ACKNOWLEDGMENTS

We thank Anna Hasenfratz and Tom DeGrand for discussions, and are grateful to the Institute for Nuclear Theory at the University of Washington for its hospitality. M. G. also thanks the Physics Department of the University of Rome "La Sapienza" for hospitality. M. G. was supported in part by the Generalitat de Catalunya under the program PIV1-2005; both C. B. and M. G. were supported in part by the U.S. Department of Energy. Y. S. was supported by the Israel Science Foundation under Grant No. 173/05.

Note added.—Recently, Hasenfratz and Hoffmann [25] have posted a paper that discusses staggered fermions in the context of the Schwinger model. They present numerical evidence that the staggered determinant (on both unrooted and rooted ensembles) can be made approximately equal to an overlap determinant by adjusting the overlap mass appropriately, up to a local effective action. When the quark mass is large compared to the taste violations, it is not inconsistent with the arguments given here that the physics of the overlap and staggered fermions could be approximately the same. However, at low quark mass the properties of the GBs guarantee that the physics of the two theories must be drastically different; indeed, numerically the matching of determinants deteriorates. In QCD, current simulations [8] are in this "low mass" region ($m \sim a^2\Lambda_{\mathrm{QCD}}^3$), where staggered chiral perturbation theory [7,23,24] is the appropriate tool.

APPENDIX A: SELECTED PROPERTIES OF TASTE-BASIS DIRAC OPERATORS

In this appendix we collect a number of useful results pertaining to the three families of taste-basis Dirac operators considered in the text: $D_{\mathrm{taste},n}(m)$, $D_{\mathrm{inv},n}(m)$, and $D_{\mathrm{ov},n}(m)$.

1. Construction of $D_{\mathrm{ov},n}(m)$

Consider a massless overlap operator D_{ov} that satisfies the GW relation

$$\{\gamma_5, D_{\mathrm{ov}}\} = \frac{2}{\alpha} D_{\mathrm{ov}}\gamma_5 D_{\mathrm{ov}}. \tag{A1}$$

Here $\alpha = O(1/a)$, where a is the lattice spacing. The choice of a massive overlap operator most common in the literature is

$$D_{\mathrm{ov}}(m) = (1 - m/\alpha)D_{\mathrm{ov}} + m, \tag{A2}$$

where $D_{\mathrm{ov}}(0) = D_{\mathrm{ov}}$ is a solution of Eq. (A1). In fact, as we will explore, there is a large freedom in extending the definition of an overlap operator to the massive case.

Let us spell out the requirements that a massive overlap operator should meet. First, the definition (A2) satisfies

$$D_{\mathrm{ov}}(m) = D_{\mathrm{ov}}(0) + Zm + O(m^2a, mpa). \tag{A3}$$

This is an obvious requirement for any sensible $D_{\mathrm{ov}}(m)$. The $O(m^2a, mpa)$ irrelevant terms cannot reintroduce any fermion doublers because $ma \ll 1$. Since m is a bare mass, we have allowed for an $O(1)$ multiplicative renormalization factor Z. In the case of Eq. (A2) one has $Z = 1$, but, anticipating less explicit definitions, there is nothing wrong in principle with having $Z \neq 1$. Either way, the value of m must be adjusted to reproduce the desired renormalized mass.

The second requirement has to do with the algebraic transformation properties under the GWL chiral symmetry [19,21] (for reviews see Refs. [26,27]). The GW relation (A1) implies that the operator

$$\hat{\gamma}_5 = \gamma_5(1 - (2/\alpha)D_{\mathrm{ov}}) \tag{A4}$$

satisfies $\hat{\gamma}_5^2 = 1$. In all relevant cases it will further be true that $\hat{\gamma}_5$ (or its generalization) is Hermitian. A possible choice of the GWL chiral transformation is then given by $\delta\psi = \hat{\gamma}_5\psi$, $\delta\bar{\psi} = \bar{\psi}\gamma_5$. The GW relation can be rewritten as $\gamma_5 D_{\mathrm{ov}} + D_{\mathrm{ov}}\hat{\gamma}_5 = 0$, which implies that the fermion action $S_{\mathrm{ov}} = \bar{\psi}D_{\mathrm{ov}}\psi$ is invariant under the GWL transformation (see also Sec. A 3 below).

In the massive case the fermion action cannot be invariant under the GWL transformation. Instead, in analogy with an ordinary mass term, and assuming that parity is a symmetry, one requires that the mass term be a scalar density that transforms into a pseudoscalar density under the GWL transformation. In fact, this requirement can be rather trivially satisfied. Consider a general bilinear fermion action $S_F = \bar{\psi}D\psi$, assuming only that S_F is hypercubic and parity invariant. Assume also a given GW operator D_{ov} (with in general $D_{\mathrm{ov}} \neq D$). We introduce the standard chiral projectors $P_{R,L} = \frac{1}{2}(1 \pm \gamma_5)$ as well as "hatted" chiral projectors $\hat{P}_{R,L} = \frac{1}{2}(1 \pm \hat{\gamma}_5)$, and define $\psi_{R,L} = \hat{P}_{R,L}\psi$, $\bar{\psi}_{R,L} = \bar{\psi}P_{L,R}$. Note that hatted projectors are used for ψ while ordinary projectors are used for $\bar{\psi}$. One can now split the action into two parts,

$$S_F = \bar{\psi}(D_\chi + D_{\mathrm{mass}})\psi, \tag{A5}$$

where

$$D_\chi = P_R D\hat{P}_L + P_L D\hat{P}_R, \tag{A6}$$

$$D_{\mathrm{mass}} = P_R D\hat{P}_R + P_L D\hat{P}_L. \tag{A7}$$

Under the chiral GWL transformation, D_χ is invariant, whereas D_{mass} transforms as required for a mass term.

While the decomposition (A5) is possible for *any* D, clearly this does not imply that any D would qualify as a massive overlap operator. In accordance with Eqs. (A3) and (A5), we require that a massive overlap operator satisfy

$$D_{\mathrm{ov},\chi}(m) = D_{\mathrm{ov}}(0) + O(mpa), \tag{A8}$$

$$D_{\mathrm{ov,mass}}(m) = Zm + O(m^2a, mpa), \tag{A9}$$

where $D_{\mathrm{ov},\chi}(m)$ and $D_{\mathrm{ov,mass}}(m)$ are defined by substituting $D_{\mathrm{ov}}(m)$ into Eqs. (A6) and (A7) respectively.[15] Note that corrections of $O(m^2a)$ are absent in Eq. (A8) because the difference between $\hat{\gamma}_5$ and γ_5 is $O(pa)$, and ordinary chiral symmetry (as opposed to the GWL type) would forbid mass terms in $D_{\mathrm{ov},\chi}(m)$. Like Eq. (A3), this asserts that $D_{\mathrm{ov}}(m)$ satisfies a GW relation in the limit $m \to 0$; that the difference $D_{\mathrm{ov}}(m) - D_{\mathrm{ov}}(0)$ is $O(m)$; and that to leading

[15] The hatted projectors are always defined with respect to $D_{\mathrm{ov}}(0) = D_{\mathrm{ov}}$. In the case of Eq. (A2), one has $D_{\mathrm{ov},\chi}(m) = D_{\mathrm{ov}}$ and $D_{\mathrm{ov,mass}}(m) = m(P_R\hat{P}_R + P_L\hat{P}_L) = m + O(mpa)$.

order, this difference is actually linear in m. What Eqs. (A8) and (A9) add is that $D_{\mathrm{ov,mass}}(m)$ transforms as expected under the GWL symmetry; the above discussion clarifies that this additional requirement can always be met for any operator that already satisfies Eq. (A3). These properties ensure that the mass parameter will be renormalized multiplicatively.[16]

Here, we will add one new requirement. Under a certain scaling assumption to be discussed in Sec. A 2, we demand that the sequences $D_{\mathrm{inv},n}(m)$ and $D_{\mathrm{ov},n}(m)$ both have the same $n \to \infty$ limit as the original RG-blocked operators $D_{\mathrm{taste},n}(m)$, for any m. With $D_{\mathrm{ov},n}(0) = D_{\mathrm{ov},n}$ of Eq. (3.13), a massive overlap operator that satisfies all the above requirements is (the following is identical to Eq. (3.15) in the main text)

$$D_{\mathrm{ov},n}(m) = D_{\mathrm{ov},n}(0) + D_{\mathrm{inv},n}(m) - D_{\mathrm{inv},n}(0), \qquad m \neq 0. \tag{A10}$$

Of course, we now define the GWL transformation and the hatted projectors using $\hat{\gamma}_{5,n} = \gamma_5(1 - (2/\alpha_n)D_{\mathrm{ov},n}(0))$. Equations (A8) and (A9) follow because, similarly to Eq. (A3), one has $D_{\mathrm{inv},n}(m) = D_{\mathrm{inv},n}(0) + Zm + O(m^2a, mpa)$. Note that the proportionality constant Z is necessary in this case, because $D_{\mathrm{inv},n}(m)$ was defined such that m is the value of the mass in the original one-component staggered operator. The $n \to \infty$ convergence properties will be established in the following subsection.

Last, we briefly comment on the construction of the low-energy effective theories: the Symanzik action and the chiral Lagrangian. In the case of Eq. (A2), the GW chiral Lagrangian has the same internal symmetries as the continuum chiral Lagrangian. The situation is slightly more involved in the more general case of Eqs. (A8) and (A9). There, terms proportional to powers of ma appear in the chirally invariant part of the Dirac operator, $D_{\mathrm{ov},\chi}(m)$. This feature will carry over to the chirally invariant part of the Symanzik action. In constructing the corresponding chiral theory, one therefore has to include a chirally invariant spurion proportional to ma. The spurion would, in effect, make the low-energy constants (LECs) of the chiral theory functions of am. Such mass dependence in the LECs could present a practical difficulty in extracting chiral physics from a simulation that used $D_{\mathrm{ov}}(m)$ as the fundamental Dirac operator. However, there is no theoretical problem in considering $D_{\mathrm{ov}}(m)$, and all the standard implications of chiral symmetry are preserved. In particular, the masses of Goldstone pions vanish in the chiral limit, for any value of the lattice spacing.

[16]The (finite) ratio of continuum and lattice Z factors (both evaluated at the same scale) will generically be a function of ma.

2. Scaling and convergence for $n \to \infty$

A basic hypothesis of the RG treatment of staggered fermions (with or without the fourth root) is that the taste-breaking terms of the RG-blocked operator $D_{\mathrm{taste},n}(m)$ tend to zero in the limit of infinitely many RG blocking steps [15,18]. The taste-breaking part Δ_n is given explicitly by writing

$$D_{\mathrm{taste},n}(m) = D_{\mathrm{inv},n}(m) + \Delta_n(m), \tag{A11}$$

where $D_{\mathrm{inv},n}(m)$ is given by Eq. (3.11). We will hold fixed the coarse-lattice spacing $a_c \equiv a_n$ obtained after n blocking steps, implying that $a_0 = 2^{-n}a_n$ goes to zero when n is taken to infinity. In the free theory [15], one can prove that $\|a_c\Delta_n\| = O(2^{-n})$. In the interacting case no proofs can be given; we will assume that Δ_n scales in the same way, up to logarithmic corrections in a_0/a_c (that we suppress below). We refer to Ref. [18] for a discussion of the status of this assumption, as well as a more precise statement about the gauge fields for which it is expected to apply.

Under this scaling hypothesis it is trivial that $D_{\mathrm{taste},n}(m)$ and $D_{\mathrm{inv},n}(m)$ have a common $n \to \infty$ limit, for any m. Furthermore, by Eq. (A10), the same will be true for $D_{\mathrm{ov},n}(m)$, provided $D_{\mathrm{ov},n}(0)$ has the same $n \to \infty$ limit as $D_{\mathrm{inv},n}(0)$. We will now prove this. In the rest of this subsection we set $m = 0$ and drop the mass argument. We begin by substituting Eq. (A11) into

$$\{\gamma_5 \otimes \xi_5, D_{\mathrm{taste},n}\} = \frac{2}{\alpha_n} D_{\mathrm{taste},n}(\gamma_5 \otimes \xi_5)D_{\mathrm{taste},n}, \tag{A12}$$

which is equivalent to Eq. (3.10). We then multiply both sides of the resulting equation by $\mathbf{1} \otimes \xi_5$, take the trace over taste indices only, and form the tensor product with an arbitrary taste matrix Ξ, obtaining

$$\{D_{\mathrm{inv},n}, (\gamma_5 \otimes \Xi)\} - \frac{2}{\alpha_n} D_{\mathrm{inv},n}(\gamma_5 \otimes \Xi)D_{\mathrm{inv},n} = \frac{1}{2\alpha_n}\,\mathrm{tr}((\mathbf{1} \otimes \xi_5)\Delta_n(\gamma_5 \otimes \xi_5)\Delta_n) \otimes \Xi. \tag{A13}$$

We used that Δ_n is traceless on the taste index [compare Eq. (3.11)]. By the scaling hypothesis, the right-hand side of Eq. (A13) is $O(2^{-2n})$, which tells us by how much $D_{\mathrm{inv},n}$ fails to satisfy the GW relation (3.14). Now introducing

$$\tilde{\gamma}_{5,n} = \gamma_5(1 - (2/\alpha_n)D_{\mathrm{inv},n}), \tag{A14}$$

it follows from Eq. (A13) that $\tilde{\gamma}_{5,n}^2 = 1 + O(2^{-2n})$. Hence, $\hat{\gamma}_{5,n} \equiv \mathrm{sign}(\tilde{\gamma}_{5,n}) = \tilde{\gamma}_{5,n} + O(2^{-2n})$. Finally, inserting this into Eq. (3.13) we find $D_{\mathrm{ov},n} = (\alpha_n/2)(1 - \gamma_5\hat{\gamma}_5) = D_{\mathrm{inv},n} + O(2^{-2n})$.

3. Index of $D_{\mathrm{taste},n}$

Here we address the following issue. The one-component staggered theory has an exact chiral symmetry for $m = 0$, the $U(1)_\epsilon$ symmetry. The corresponding chiral transformations of the continuum four-taste theory are generated by $\gamma_5 \otimes \xi_5$, and they form a nonanomalous subgroup of $SU(4)_L \times SU(4)_R$.[17] In contrast, after any number of RG blocking steps, we obtain the operator $D_{\mathrm{taste},n}(m)$ which, for $m = 0$, only satisfies the GW relation (3.10). While the RG-blocked action is invariant under the corresponding GWL transformation, this is not enough to establish that it is a symmetry. One must further check that the measure term, arising from this change of variables, vanishes. Here we show that this is indeed the case. Again we will set $m = 0$ and drop the mass argument.

The variation of the measure is given by [21]

$$-\operatorname{tr}((\gamma_5 \otimes \xi_5) D_{\mathrm{taste},n}/\alpha_n) = \operatorname{index}(D_{\mathrm{taste},n}). \tag{A15}$$

We note that, loosely speaking, one expects the index of $D_{\mathrm{taste},n}$ to vanish in the continuum limit, because taste symmetry is recovered in this limit, and $\operatorname{tr}(\xi_5) = 0$. We will establish the stronger result that the index of $D_{\mathrm{taste},n}$ is actually zero on the lattice. The precise statement is that the index is zero except possibly on a subspace $\mathcal{U}_{00} \subset \mathcal{U}_0 \subset \mathcal{U}$, where $\mathcal{U}$ is the (finite-volume) gauge-field space, $\mathcal{U}_0$ is the (proper) subspace where $D_{\mathrm{taste},n}$ has at least one exact zero mode, and $\mathcal{U}_{00}$ is a proper subspace of $\mathcal{U}_0$ defined below. Further, $\mathcal{U}_0$ is a measure zero subset of $\mathcal{U}$, and $\mathcal{U}_{00}$ is a measure zero subset of $\mathcal{U}_0$.

One can always choose a basis for the exact zero modes of $D_{\mathrm{taste},n}$ such that each zero mode ψ_0 has a definite chirality,

$$(\widehat{\gamma_5 \otimes \xi_5})\psi_0 = (\gamma_5 \otimes \xi_5)\psi_0 = \pm\psi_0, \tag{A16}$$

where, analogous to Eq. (A4),

$$(\widehat{\gamma_5 \otimes \xi_5}) = (\gamma_5 \otimes \xi_5)(1 - (2/\alpha_n) D_{\mathrm{taste},n}). \tag{A17}$$

By Eq. (A17), ordinary and hatted projectors coincide when acting on a zero mode. Also, on a zero mode, the Dirac operator $D_{\mathrm{taste},n}$ commutes with the chiral generator, as usual. Therefore it is enough to show that the index of $D_{\mathrm{taste},n}$ is zero with respect to $\gamma_5 \otimes \xi_5$ chirality. This can be done by relating the zero modes of $D_{\mathrm{taste},n}$ to those of D_{stag}. Iterating Eq. (3.4) we have

$$D^{-1}_{\mathrm{taste},n} = 1/\alpha_n + Q_n D^{-1}_{\mathrm{stag}} Q_n^\dagger, \tag{A18}$$

where $Q_n = Q^{(n)} Q^{(n-1)} \cdots Q^{(1)} Q^{(0)}$. If we gradually vary the gauge field so as to approach a configuration where $D_{\mathrm{taste},n}$ has an exact zero mode, the norm of $D^{-1}_{\mathrm{taste},n}$ on the left-hand side diverges. This is possible only if the norm of D^{-1}_{stag} diverges too. Thus, not surprisingly, any exact zero mode of $D_{\mathrm{taste},n}$ must be obtained via RG blocking from an exact zero mode of D_{stag}.

Because of $U(1)_\epsilon$ symmetry, the spectrum of D_{stag} consists of imaginary pairs $\pm i\lambda$, and the corresponding eigenmodes are related by multiplication with $\epsilon(x)$. Since the eigenvalues are continuous functions of the gauge fields, and since there are no zero modes in the free case, any zero modes that appear must also be paired. We choose a chiral basis for the two zero modes, which is always possible. Then, as the gauge field changes, the off-diagonal matrix element of D_{stag} between the modes is not forbidden by $U(1)_\epsilon$ symmetry, and is thus generically nonzero. This suggests—in accordance with standard lore—that exact zero modes exist only on a zero measure subspace $\mathcal{U}_0$.

Using Eq. (3.7) it follows from the above discussion that, given a pair of zero modes of D_{stag}, then $D_{\mathrm{taste},0}$ must have a corresponding pair of zero modes, with one zero mode of each $\gamma_5 \otimes \xi_5$ chirality. The index of both D_{stag} and $D_{\mathrm{taste},0}$ is, thus, *always* zero. The index of $D_{\mathrm{taste},n}$ could only be nonzero if the blocking transformation $Q^{(n)} Q^{(n-1)} \cdots Q^{(1)}$ exactly annihilated one of the definite-chirality zero modes of $D_{\mathrm{taste},0}$ but not the other. Generically this will not happen, and the subspace $\mathcal{U}_{00}$ where this does happen therefore has measure zero with respect to $\mathcal{U}_0$. (We leave it open whether or not $\mathcal{U}_{00}$ is an empty set.) Assuming that no (interesting) QCD observable has a δ-function support on $\mathcal{U}_{00}$, the GWL transformation is then a symmetry of the RG-blocked theory.

[17] This is true when the staggered mass term is introduced as $D_{\mathrm{stag}}(m) = D_{\mathrm{stag}}(0) + m$ [2].

[1] J. B. Kogut and L. Susskind, Phys. Rev. D **11**, 395 (1975).

[2] M. Golterman and J. Smit, Nucl. Phys. **B245**, 61 (1984).

[3] This trick to eliminate the unwanted degrees of freedom was first proposed in E. Marinari, G. Parisi, and C. Rebbi, Nucl. Phys. **B190**, 734 (1981).

[4] K. Jansen, Nucl. Phys. B, Proc. Suppl. **129**, 3 (2004); T. A. DeGrand, Int. J. Mod. Phys. A **19**, 1337 (2004); A. D. Kennedy, Nucl. Phys. B, Proc. Suppl. **140**, 190 (2005); S. Dürr, Proc. Sci., LAT2005 (2005) 021 [hep-lat/0509026].

[5] A. Hasenfratz, hep-lat/0511021.

[6] D. H. Adams, Phys. Rev. D **72**, 114512 (2005).

[7] W. J. Lee and S. R. Sharpe, Phys. Rev. D **60**, 114503 (1999).

[8] For numerical evidence, see, for example, C. Aubin *et al.*, Phys. Rev. D **70**, 094505 (2004).

[9] M. Golterman, Nucl. Phys. **B273**, 663 (1986).

[10] N. Kawamoto and J. Smit, Nucl. Phys. **B192**, 100 (1981).

[11] H. Kluberg-Stern, A. Morel, and B. Petersson, Nucl. Phys. **B215**, 527 (1983); T. Jolicoeur, H. Kluberg-Stern, M. Lev, A. Morel, and B. Petersson, Nucl. Phys. **B235**, 455 (1984).

[12] B. Bunk, M. Della Morte, K. Jansen, and F. Knechtli, Nucl. Phys. **B697**, 343 (2004).

[13] A. Hart and E. Müller, Phys. Rev. D **70**, 057502 (2004).
[14] C. Aubin *et al.* (MILC Collaboration), Phys. Rev. D **70**, 114501 (2004).
[15] Y. Shamir, Phys. Rev. D **71**, 034509 (2005).
[16] F. Gliozzi, Nucl. Phys. **B204**, 419 (1982); A. Duncan, R. Roskies, and H. Vaidya, Phys. Lett. **114B**, 439 (1982).
[17] H. Kluberg-Stern, A. Morel, O. Napoly, and B. Petersson, Nucl. Phys. **B220**, 447 (1983).
[18] Y. Shamir (to be published).
[19] P. H. Ginsparg and K. G. Wilson, Phys. Rev. D **25**, 2649 (1982).
[20] H. Neuberger, Phys. Lett. B **417**, 141 (1998).
[21] M. Lüscher, Phys. Lett. B **428**, 342 (1998).
[22] J. Smit and J. C. Vink, Nucl. Phys. **B286**, 485 (1987); S. Dürr and C. Hoelbling, Phys. Rev. D **69**, 034503 (2004); **71**, 054501 (2005); hep-lat/0604005; C. Bernard, Phys. Rev. D **71**, 094020 (2005).
[23] C. Aubin and C. Bernard, Phys. Rev. D **68**, 034014 (2003); **68**, 074011 (2003).
[24] C. Bernard, Phys. Rev. D **73**, 114503 (2006).
[25] A. Hasenfratz and R. Hoffmann, hep-lat/0604010.
[26] F. Niedermayer, Nucl. Phys. B, Proc. Suppl. **73**, 105 (1999).
[27] M. Golterman, Nucl. Phys. B, Proc. Suppl. **94**, 189 (2001).