This document is downloaded from DR-NTU (https://dr.ntu.edu.sg)
Nanyang Technological University, Singapore.

Fourth root prescription for dynamical staggered fermions

Adams, David H.

2005

Adams, D. H. (2005). Fourth root prescription for dynamical staggered fermions. Physical Review D, 72(11), 114512.

https://hdl.handle.net/10356/101102

https://doi.org/10.1103/PhysRevD.72.114512

© 2005 The American Physical Society. This paper was published in Physical Review D and is made available as an electronic reprint (preprint) with permission of The American Physical Society. The paper can be found at the following official DOI: http://dx.doi.org/10.1103/PhysRevD.72.114512. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.

Downloaded on 09 Dec 2022 13:41:18 SGT

PHYSICAL REVIEW D **72,** 114512 (2005)

Fourth root prescription for dynamical staggered fermions

David H. Adams*

Instituut-Lorentz for Theoretical Physics, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, The Netherlands, National Center for Theoretical Sciences, National Taiwan University, Taipei 106, Taiwan, Republic of China, and Physics Department, Florida International University, 11200 SW 8th Street, Miami, Florida 33199, USA†

(Received 10 September 2005; published 22 December 2005)

With the aim of resolving theoretical issues associated with the fourth root prescription for dynamical staggered fermions in lattice QCD simulations, we consider the problem of finding a viable lattice Dirac operator D such that $(\det D_{\text{staggered}})^{1/4} = \det D$. Working in the flavor field representation we show that in the free field case there is a simple and natural candidate D satisfying this relation, and we show that it has acceptable locality behavior: exponentially local with a localization range vanishing $\sim\sqrt{a/m}$ for lattice spacing $a \to 0$. Prospects for the interacting case are also discussed, although we do not solve this case here.

DOI: 10.1103/PhysRevD.72.114512 PACS numbers: 11.15.Ha

I. INTRODUCTION

The development in recent years of an improved staggered fermion formulation [1] has made unquenched numerical lattice QCD simulations possible at realistically small quark masses. The resulting impressive agreement between the calculated parameters of QCD phenomenology and their experimental values [2,3] (along with predictions for quantities not yet measured experimentally [4]) indicates that the long-time dream of being able to do high-precision lattice QCD calculations is now becoming reality. However, the advantageous properties of staggered fermions for numerical implementation are currently offset by unresolved problematic issues at the conceptual/theoretical level. In particular, there is concern [5–7] about the use of the fourth root of the staggered fermion determinant to represent the fermion determinant of a single dynamical (sea) quark. A number of works have appeared recently addressing this concern via theoretical considerations [8–15], derivations of predictions that can be used to test its viability [16], and various numerical investigations [17–19]. The present paper is intended as another theoretical contribution in this direction.

A staggered fermion is a lattice formulation of four continuum fermion flavors, nowadays called ‘‘tastes’’ (to distinguish them from the actual quark flavors). The fermion determinant for a single quark flavor in this framework is represented by a rooted determinant $\det(D_{\text{staggered}})^{1/4}$. While this formally goes over to the determinant for a single quark flavor in the continuum limit, the concern regarding this prescription is that it does not fit in an obvious way into the framework of local lattice field theory at nonzero lattice spacing. The lattice model might therefore not be in the right universality class to reproduce QCD. This raises the question of whether the dynamical staggered fermion formulation is a first principles approach to QCD or simply a phenomenological model which describes QCD very well in a certain regime.

One way to establish that the universality class is the right one would be to show that there is a viable (and, in particular, local) single-flavor lattice Dirac operator D such that [5]

$$(\det D_{\text{staggered}})^{1/4} = \det D. \tag{1.1}$$

This would imply equivalence between the dynamical staggered fermion formulation and the manifestly local formulation with sea quarks described by D. (To avoid unitarity issues, D, with suitably adjusted bare mass, should then also be used as the Dirac operator for the valence quarks.) We will refer to (1.1) as the staggered determinant relation (SDR) in the following.

The most direct attempt at a solution to the SDR is simply to take $D = (D_{\text{staggered}})^{1/4}$. This is essentially the approach taken by Jansen and collaborators in Ref. [9]. More precisely, they considered the operator

$$M = ((D_{\text{staggered}})^{\dagger} D_{\text{staggered}})^{1/2}|_{\text{even sites}} \tag{1.2}$$

for which $\det M = \det(D_{\text{staggered}})^{1/2}$ since $(D_{\text{staggered}})^{\dagger} D_{\text{staggered}}$ couples lattice sites by even to even and odd to odd. Thus M is a candidate operator for D in the case where $1/4 \to 1/2$ in (1.1), i.e. the case of two degenerate quark flavors. However, this operator was found to have unacceptable locality behavior: it is exponentially local (for bare mass $m > 0$), but the localization range is $\sim m^{-1}$ and thus fails to vanish in the limit of vanishing lattice spacing, $a \to 0$ [9].

This negative result of Ref. [9] is unsurprising, since the operator (1.2) does not take account of the staggered fermion taste structure. The staggered fermion action can be viewed as consisting of naive fermion actions for four fermion species (the tastes), together with terms that couple these, with the latter formally vanishing for $a \to 0$. This suggests that, in attempting to find a local, single-flavor lattice Dirac operator satisfying the SDR, one should

*Email address: adams@lorentz.leidenuniv.nl
†Current address.

 © 2005 The American Physical Society

consider operators of the form

$$D = \nabla\!\!\!/ + W + m \tag{1.3}$$

where $\nabla\!\!\!/$ is the (massless) naive lattice Dirac operator, and W is a term which formally vanishes for $a \to 0$ and whose role is to take account of the taste mixing in $D_{\text{staggered}}$. In this paper, working in the flavor field representation of staggered fermions [20], we show that, in the *free field case*, there is a natural candidate D of the form (1.3) which satisfies the SDR. The W operator, although formally vanishing for $a \to 0$, turns out to involve a square root, so its locality status is not immediately clear. We show, however, that the operator does have acceptable locality behavior: exponentially local, with localization range vanishing $\sim\sqrt{a/m}$ for $a \to 0$.

Our operator can be gauged (i.e. coupled to the link variables) in a variety of ways. However, for reasons which we will discuss later, it is most unlikely that a gauging of this operator exists such that the SDR continues to hold in the interacting case. Our operator should therefore be regarded as a prototype, i.e. a first step on a path to constructing more sophisticated operators which have a chance of satisfying the SDR in the full interacting case.

Of course, there is no *a priori* guarantee that a viable lattice Dirac operator satisfying the SDR actually exists in the interacting case, so other approaches should also be considered. One possibility is the following: If there is a single-flavor D such that the effect of including the determinant ratio $\det(D_{\text{staggered}})^{1/4}/\det D$ in the lattice QCD functional integral is simply to renormalize the bare coupling constant (just as dynamical heavy quarks do [21]), then representing the sea quark determinant by $\det(D_{\text{staggered}})^{1/4}$ is equivalent to representing it by $\det D$ together with a shift in the bare coupling. Since the latter description is manifestly local, this is another way in which the locality issue could be positively resolved. The prospects for this, and the properties that such a D would be expected to have, are also discussed in some detail in this paper. The problem of finding such a D is seen to be essentially equivalent to the problem of finding a solution to a generalized version of the SDR.

The paper is organized as follows. After a general discussion of the problem of finding viable solutions to the SDR, we arrive at our free field candidate D in Sec. II. In Sec. III we prove that this operator has the good locality behavior mentioned earlier. Our argument is entirely analytic and the techniques are of a generally applicable nature; we also apply them to give a new derivation of the negative locality result for the operator considered previously by Jansen and collaborators [9]. (Their argument in the free field case had a numerical as well as analytic component.) We conclude in Sec. IV with a discussion of the issues and prospects for the interacting case.

II. CONSIDERATIONS FOR FINDING A CANDIDATE D

For concreteness we specialize to 4 spacetime dimensions in this section (everything generalizes straightforwardly to arbitrary even dimensions). The usual staggered fermion action, obtained via spin diagonalization of the naive action [22], is $S_{\text{staggered}} = a^4\sum_x \bar{\psi}(x)D_{\text{st}}\psi(x)$ where the staggered Dirac operator is given by

$$D_{\text{st}} = \eta^\mu \frac{1}{a}\nabla_\mu + m \tag{2.1}$$

with $\eta^\mu(x) = (-1)^{(x_1+\cdots+x_{\mu-1})/a}$ and $\nabla_\mu = \frac{1}{2}(\nabla_\mu^+ + \nabla_\mu^-)$ the symmetrized gauge-covariant difference operator. The usual flavor (i.e. taste) identification comes about by considering the free field propagator: it has 4 poles, and the momentum space Brillouin zone is divided into 4 subregions, each containing a single pole, with the momenta in each of these subregions being interpreted as the momenta of different fermion tastes. An alternative, and conceptually more appealing way to identify the tastes is provided by the flavor (taste) field representation of the staggered fermion action derived in [20]. In this representation the taste fields are manifest from the beginning in the fermion action. The taste fields live on the blocked lattice (spacing $= 2a$), whereas the lattice paths and link variables which specify the gauging of the action are those of the original lattice. The action in general gauge background does not have a simple expression in this setting though, making it more difficult to work with in practice. However, in the free field case the action does have a simple expression. Denoting the staggered Dirac operator in the taste field representation by D_{stt}, it can be written in the free field case as [20][1]

$$D_{\text{stt}}^{\text{free}} = (\gamma^\mu \otimes \mathbf{1})\frac{1}{2a}\nabla_\mu + i(\gamma_5 \otimes \Gamma^\nu)\frac{1}{2(2a)}\Delta_\nu + m \tag{2.2}$$

where now the (free field) difference operators are on the blocked lattice; $\Delta_\nu = \nabla_\nu^- - \nabla_\nu^+$ [so that $\Delta = \sum_\nu \Delta_\nu$ is $(2a)^2$ times the blocked lattice Laplace operator], and $\{\Gamma^\nu\}$ is a Hermitian representation of the Dirac γ algebra on taste space $\mathbf{C}^4$.[2]

The importance of taking account of taste structure when attempting to find a solution D to the SDR can now be seen in the free field case as follows. The γ-matrix representation $\{\Gamma^\nu\}$ in (2.2) can be chosen such that the diagonal

[1] The free field version of the staggered Dirac operator in the taste representation also arises from a first principles approach to constructing the Dirac operator on the lattice [23].

[2] The part $i\gamma_5 \otimes \Gamma^\nu$ in (2.2) is usually written as $\gamma_5 \otimes \tau_5\tau^\nu$ where $\{\tau^\nu\}$ is a Hermitian representation of the Dirac γ algebra on taste space. Note that $\Gamma^\nu = -i\tau_5\tau^\nu$ defines another (equivalent) Hermitian representation of this algebra.

matrix elements in each of the Γ^ν's all vanish. Then the taste-mixing terms in the free field Lagrangian correspond to the terms with Γ^ν's in the free field Dirac operator (2.2). Therefore, if the taste-mixing terms are "turned off" the free field Dirac operator reduces to $\not\nabla \otimes \mathbf{1} + m$, where $\not\nabla = \gamma^\mu \frac{1}{2a}\nabla_\mu$ is the (massless) free field naive Dirac operator and $\mathbf{1}$ is the identity matrix on taste space. Consequently,

$$(\det D_{\rm st}^{\rm free})^{1/2} \to \det\begin{pmatrix} \not\nabla + m & 0 \\ 0 & \not\nabla + m \end{pmatrix}, \qquad (\det D_{\rm st}^{\rm free})^{1/4} \to \det(\not\nabla + m) \tag{2.3}$$

or, alternatively,

$$(\det D_{\rm st}^{\rm free})^{1/2} \to \det\begin{pmatrix} \sqrt{(\not\nabla + m)^\dagger(\not\nabla + m)} & 0 \\ 0 & \sqrt{(\not\nabla + m)^\dagger(\not\nabla + m)} \end{pmatrix}, \qquad (\det D_{\rm st}^{\rm free})^{1/4} \to \det(\sqrt{(\not\nabla + m)^\dagger(\not\nabla + m)}). \tag{2.4}$$

In the former case the fractional powers of $\det D_{\rm st}^{\rm free}$ become determinants of ultralocal lattice Dirac operators, while in the latter case they become determinants of operators which cannot be expected to have good locality properties. If we now imagine turning back on the taste-mixing terms, there is reason to hope that there will be corresponding deformations of $\begin{pmatrix} \not\nabla + m & 0 \\ 0 & \not\nabla \pm m \end{pmatrix}$ or $\not\nabla + m$ into some two-taste lattice Dirac operator $\tilde{D}$ or single-taste D, respectively, which continues to have good locality behavior, such that $(\det D_{\rm st}^{\rm free})^{1/2} = \det\tilde{D}$ and $(\det D_{\rm st}^{\rm free})^{1/4} = \det D$. On the other hand, if a solution D to the SDR, or a solution $\tilde{D}$ to the version of the SDR with fractional power $1/2$ of the staggered fermion determinant, has been constructed "blindly" without taking account of the taste structure of the staggered fermion formulation, it can happen that when taste-mixing terms are turned off in the free field case the scenario (2.4) arises; then it is to be expected that the D or $\tilde{D}$ have bad locality behavior. In fact this is essentially the situation for the solution $\tilde{D} = M$ considered in [9], and the negative locality result found there is therefore unsurprising. However, the possibility (2.3) gives hope of doing better than this, at least in the free field case.[3]

In light of (2.3), when attempting to find a viable D in the free field case it is natural to consider Dirac operators of the form

$$D = \gamma^\mu \frac{1}{2a}\nabla_\mu + \frac{1}{2a}W + m \tag{2.5}$$

on the blocked lattice, where the purpose of $\frac{1}{2a}W$ is to take account of the taste-mixing terms in the staggered Dirac operator. In particular, W should formally vanish $\sim a^2$ for $a \to 0$, and should lift the species doubling of the naive Dirac operator. In other words, $\frac{1}{2a}W$ is to be a Wilson-type term.

A feature of the free field staggered Dirac operator (2.2) in the taste field representation, which is very useful in this context, is that $(D_{\rm stt}^{\rm free})^\dagger D_{\rm stt}^{\rm free}$ is trivial in spinor $\otimes$ flavor space:

$$(D_{\rm stt}^{\rm free})^\dagger D_{\rm stt}^{\rm free} = \frac{1}{(2a)^2}\left(-\nabla^2 + \sum_\nu\left(\frac{1}{2}\Delta_\nu\right)^2 + (2am)^2\right) \times (\mathbf{1}\otimes\mathbf{1}). \tag{2.6}$$

On the other hand, for a free field operator of the form (2.5) we have

$$D^\dagger D = \frac{1}{(2a)^2}(-\nabla^2 + (W + 2am)^2)\mathbf{1} \tag{2.7}$$

trivial in spinor space. Comparing (2.6) and (2.7), and noting that $\det D = \det(D^\dagger D)^{1/2}$ (assuming $\frac{1}{2a}W + m \geq 0$) and $\det(D_{\rm st}) = \det(D_{\rm st}^\dagger D_{\rm st})^{1/2}$ (assuming $m \geq 0$), we immediately see that a sufficient criteria for the desired determinant relation $\det D = (\det D_{\rm st}^{\rm free})^{1/4}$ to be satisfied is

$$(W + 2am)^2 = \sum_\nu\left(\frac{1}{2}\Delta_\nu\right)^2 + (2am)^2. \tag{2.8}$$

This has the solution

$$W = \sqrt{(2am)^2 + \sum_\nu\left(\frac{1}{2}\Delta_\nu\right)^2} - 2am, \tag{2.9}$$

which clearly has the required properties for $\frac{1}{2a}W$ to be a Wilson-type term (i.e. W lifts species doubling and formally vanishes $\sim a^2$ for $a \to 0$). Thus we have arrived at a free field solution D to the SDR (1.1). Substituting (2.9) into (2.5) we get the expression

$$D = \gamma^\mu \frac{1}{2a}\nabla_\mu + \frac{1}{2a}\sqrt{(2am)^2 + \sum_\nu\left(\frac{1}{2}\Delta_\nu\right)^2}. \tag{2.10}$$

Note that turning off the taste-mixing terms in the free field staggered fermion action, which, as pointed out previously, corresponds to putting $\Gamma^\nu \to 0$ in (2.2), has the same effect as putting $\Delta_\nu \to 0$. By (2.10) this gives $D \to \not\nabla + m$ (for $m \geq 0$); thus we have a realization of the scenario (2.3). However, because of the square root in (2.10), it is not immediately clear that good locality behavior of D, anticipated in our earlier discussion, is realized. In fact this square root operator has some similarity with the free field square root operator considered by Jansen and collabora-

[3] The interacting case is more difficult, since the taste field representation of the staggered Dirac operator is not given simply by some gauging of the ∇_μ's and Δ_ν's in (2.2) but has a more complicated structure [20].

tors in [9], which turned out to have unacceptable locality behavior. Nevertheless, we show in the next section that our operator does have good locality behavior. The reason why its behavior is different from the operator in [9] is a bit subtle, and to elucidate this we also provide in the next section a new derivation of the negative locality result of [9] which reveals the origin of the different behaviors.

We remark that other, ultralocal solutions to the SDR exist in the free field case. Using $-\nabla_\nu^2 + (\frac{1}{2}\Delta_\nu)^2 = \Delta_\nu$ (2.6) reduces to $(D_{\rm stt}^{\rm free})^\dagger D_{\rm stt}^{\rm free} = ([1/(2a)^2]\Delta + m^2)(\mathbf{1}\otimes\mathbf{1})$ and it follows that the free field SDR is satisfied, e.g., by $D = ([1/(2a)^2]\Delta + m^2)^2$ acting on scalar Grassmann fields on the lattice. Other examples of ultralocal solutions are easily constructed. However, these are unattractive options since they do not have the form of a lattice Dirac operator.

III. FREE FIELD LOCALITY RESULT

In this section we work in arbitrary spacetime dimension d and show that the free field operator

$$\sqrt{(am)^2 + \sum_\nu \Delta_\nu^2} \tag{3.1}$$

on lattice with spacing a is exponentially local with localization range vanishing $\sim\sqrt{a/m}$ for $a \to 0$; then D in (2.10) obviously has this same locality behavior on the blocked lattice. The argument proceeds in several steps. First, we specialize to $d = 1$ dimension and write

$$\begin{aligned}\sqrt{(am)^2 + \Delta^2}(x,y) &= \frac{1}{a}\int_{-\pi}^{\pi}\frac{dp}{2\pi}\sqrt{(am)^2 + \Delta(p)^2}e^{ip(x-y)/a}\\ &= \frac{1}{2\pi i a}\oint_{|z|=1}\frac{dz}{z}\\ &\quad\times\sqrt{(am)^2 + (2-(z+z^{-1}))^2}z^{|x-y|/a}\\ &= \frac{1}{2\pi i a}\oint_{|z|=1}\frac{dz}{z}\\ &\quad\times\sqrt{z^{-2}((am)^2z^2 + (z-1)^4)}z^{|x-y|/a}\end{aligned} \tag{3.2}$$

(the integral is counterclockwise around the unit circle in the complex plane and we have set $z = e^{\pm ip}$ with "+" if $x - y > 0$ and "−" if $x - y < 0$; for $x = y$ either choice can be used). The square root $z \mapsto \sqrt{z}$ is holomorphic after making a cut in $\mathbf{C}$; we choose the cut to be the half line of negative real numbers $\mathbf{R}_-$. Then, by the residue theorem, the circle around which the integral in (3.2) is performed can be shrunk to a closed loop around the region containing the z's for which $f(z) \in \mathbf{R}_-$, where

$$f(z) = z^{-2}((am)^2z^2 + (z-1)^4) \tag{3.3}$$

is the function inside the square root in (3.2), since outside this region (and away from $z = 0$) $z \mapsto \sqrt{f(z)}$ is holomorphic. The excluded z's are found as follows:

$$f(z) = -\lambda,\quad \lambda\in\mathbf{R}_+ \Leftrightarrow (z-1)^4 + ((am)^2+\lambda)z^2 = 0 \Leftrightarrow ((z-1)^2 + i\sqrt{(am)^2+\lambda}z)((z-1)^2 - i\sqrt{(am)^2+\lambda}z) = 0. \tag{3.4}$$

For given $\lambda \in \mathbf{R}_+$ there are 4 solutions; we are only interested in the ones with $|z| \le 1$ and these are $z = z_\pm(s = \sqrt{(am)^2 + \lambda})$ where

$$z_\pm = 1 \pm \frac{is}{2} - \sqrt{s}\sqrt{\pm i - s/4}, \qquad s \in]0,\infty[. \tag{3.5}$$

Thus the z's for which $f(z) \in \mathbf{R}_-$ and $|z| \le 1$ form curves inside the unit circle in $\mathbf{C}$, parametrized by (3.5) with $s \in [am,\infty[$. It is useful to reparametrize these curves as follows. We introduce

$$t = 1 - \sqrt{\frac{s/2}{s/4 + \sqrt{1+(s/4)^2}}}; \tag{3.6}$$

note that this is a strictly decreasing function of s with $t = 1$ for $s = 0$ and $t \to 0$ for $s \to \infty$. After a little calculation (3.5) can be reexpressed in terms of t as

$$z_\pm(t) = t \mp i(1-t)\sqrt{\frac{t}{2-t}}, \qquad t \in]0,1[. \tag{3.7}$$

The z's for which $f(z) \in \mathbf{R}_-$ and $|z| \le 1$ are now parametrized by the curves $z_+(t)$ and $z_-(t)$ for $t \in]0, t_{am}]$ where t_{am} is given by setting $s = am$ in (3.6); we write this out explicitly for future reference:

$$t_{am} = 1 - \sqrt{\frac{am/2}{am/4 + \sqrt{1+(am/4)^2}}}. \tag{3.8}$$

These curves, which we denote by C_+ and C_-, lie in the lower and upper half planes of $\mathbf{C}$, respectively. They have a common limit point at $z_+(0) = z_-(0) = 0$. See Fig. 1.

According to the residue theorem, the integral (3.2) remains unchanged when the unit circle is shrunk to a closed curve C around $C_+ \cup \{0\} \cup C_-$. In the limit this reduces to an integral over $C_+ \cup \{0\} \cup C_-$ itself, with a factor of 2 to take account of the fact that C goes along $C_+ \cup C_-$ twice, with opposite orientations. (This is assuming that the argument in $z^{|x-y|/a}$ is sufficiently large to avoid a divergence of the limit integral due to singularity

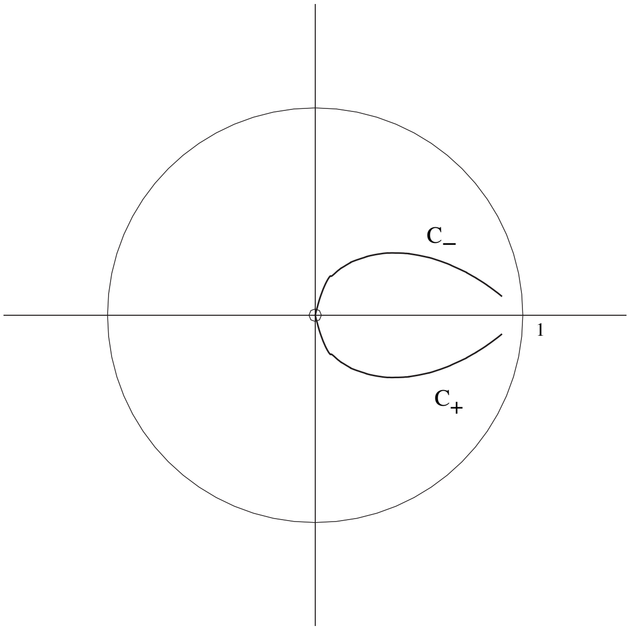

FIG. 1. The "exclusion curves" $C_\pm$. The locations of the end points near 1 depend on am and converge to 1 for $a \to 0$.

at $z = 0$; an explicit criterion for this will be given further below.)[4] Then, using obvious symmetries, the integral can be seen to be 2× the *real part* of the integral over C_-. (The integrals over C_+ and C_- are complex conjugate, so the imaginary parts cancel out as they should.) The square root $\sqrt{f(z)}$ in the integral then reduces to $\pm i\sqrt{\lambda(t)}$ with the explicit expression for $\lambda(t)$ determined below and the sign determined to be $-$. Thus (3.2) reduces to

$$\sqrt{(am)^2 + \Delta^2}(x, y) = \frac{-2}{\pi a} \int_0^{t_{am}} dt \left| \frac{dz_-}{dt} \right| \times \sqrt{\lambda(t)}\, \mathrm{Re}(z_-(t)^{|x-y|/a-1}). \tag{3.9}$$

Recalling that the solution to $f(z) = -\lambda$ can be written as (3.5) with $s = \sqrt{(am)^2 + \lambda}$, and noting that the relation between s and t in (3.6) can be inverted to give $s = 2(1-t)^2/\sqrt{t(2-t)}$, we find

$$\sqrt{\lambda(t)} = \sqrt{\frac{4(1-t)^4}{t(2-t)} - (am)^2}. \tag{3.10}$$

The sign in $\pm i\sqrt{\lambda(t)}$ can be determined by considering $\sqrt{f(z)} \approx \sqrt{z^{-2}}$ for z near zero. Writing $z = \epsilon + i\delta$ we have $z^{-2} \approx (-\delta^2 - i2\delta\epsilon)/(\delta^4 + (2\delta\epsilon)^2)$; hence, recalling that we have chosen the cut $\mathbf{R}_-$ to define the square root, $z^{-2} \xrightarrow{\epsilon\to 0_+} -i/\delta$. From this it is straightforward to see that the sign in $\pm i\sqrt{\lambda(t)}$ in the integral over C_- is $-$, and this is the origin of the minus sign in (3.9). Explicit expressions for the remaining ingredients in the integrand in (3.9) are readily found from (3.7):

$$|z_\pm(t)| = \sqrt{\frac{t}{2-t}}, \tag{3.11}$$

$$\left| \frac{dz_\pm}{dt} \right| = \frac{1}{2-t} \sqrt{\frac{1 + t(2-t)}{t(2-t)}}. \tag{3.12}$$

Note that the divergences $\sim 1/\sqrt{t}$ for $t \to 0$ in $\sqrt{\lambda(t)}$ and $|dz_-/dt|$ are compensated in (3.9) by powers of $\sqrt{t}$ in $z_-(t)^{|x-y|/a-1}$ provided $|x-y| > a$, which we henceforth assume to be the case. (This is the criterion alluded to above.) We can now use (3.9) to draw conclusions about the exponential decay of $\sqrt{(am)^2 + \Delta^2}(x, y)$. Explicit evaluation of the integral in (3.9) will not be needed for this, so we do not attempt to perform it here.

For fixed $|x-y| > 0$ and given $t \in [0, t_{am}]$ the integrand in (3.9) is dominated in the $a \to 0$ limit by the exponential factor $z_-(t)^{|x-y|/a}$. From (3.11) we see that $|z_-(t)|$ increases with t for $t \in [0, t_{am}]$ (recall $t_{am} \le 1$); therefore there can be no cancellation between the exponential factors for different t in the integral (3.9) and it follows that $\sqrt{(am)^2 + \Delta^2}(x, y)$ decays exponentially $\sim z_-(t_{am})^{|x-y|/a}$ for small a. From (3.7) we see that

$$t_{am} = 1 - \sqrt{am/2} + O(am). \tag{3.13}$$

Consequently, using (3.11), the magnitude of the exponential decay of $\sqrt{(am)^2 + \Delta^2}(x, y)$ for small a (i.e. $am \ll 1$) is found to be

$$\begin{aligned} |z_-(t_{am})|^{|x-y|/a} &= (1 - \sqrt{am/2} + O(am))^{|x-y|/2a} \\ &= [(1 - \sqrt{m/2}\sqrt{a} + O(am))^{1/\sqrt{a}}]^{|x-y|/2\sqrt{a}} \\ &\overset{a\to 0}{\approx} (e^{-\sqrt{m/2}})^{|x-y|/2\sqrt{a}} \\ &= e^{-(1/2)(\sqrt{m/2a})|x-y|}. \end{aligned} \tag{3.14}$$

Thus the localization range for the exponential decay of $\sqrt{(am)^2 + \Delta^2}(x, y)$ is seen to be $2\sqrt{2a/m}$.

We now supplement the preceding with a bound on $|\sqrt{(am)^2 + \Delta^2}(x, y)|$ which allows one to check that the integral in (3.9) does not give rise to other factors which mask the exponential decay when $|x-y|$ is of the same order of magnitude as $\sqrt{a/m}$. From (3.10) and (3.11) we see that for $t \in [0, 1]$

$$\left| \frac{dz_-}{dt} \right| \le \frac{\sqrt{2}}{\sqrt{t}}, \qquad \sqrt{\lambda(t)} \le \frac{2}{\sqrt{t}}, \qquad |z_-(t)| \le \sqrt{t}, \tag{3.15}$$

[4] If the argument in $z^{|x-y|/a}$ is not sufficiently large, e.g. if $x = y$, then the curve around which the integral is performed cannot be completely shrunk to $C_+ \cup \{0\} \cup C_-$—a small detour around $z = 0$ must be included. This case is more subtle, and we do not consider it here since it is not needed to derive the advertised locality result.

and it follows from (3.9) that

$$|\sqrt{(am)^2+\Delta^2}(x,y)| \le \frac{4\sqrt{2}}{\pi a}\int_0^{t_{am}} dt(\sqrt{t})^{|x-y|/a-3} = \frac{8\sqrt{2}}{\pi\sqrt{t_{am}}(|x-y|-a)}(t_{am})^{|x-y|/2a}. \tag{3.16}$$

For $am \ll 1$ the exponential factor here reduces as in (3.14) to give the same decay found earlier. The factor $1/\sqrt{t_{am}}$ has no effect since by (3.13) it is ≈ 1. On the other hand, the factor $1/(|x-y|-a)$ blows up for $|x-y|\approx a$; however it does not mask the exponential decay once $|x-y|\ge 2a$ (and enhances the locality when $|x-y|$ is large).[5] When a is sufficiently small, the localization range ($\sim\sqrt{a/m}$) of the exponential decay is much larger than a and therefore does not get masked by this factor.

We now proceed to the case of arbitrary spacetime dimension d and consider

$$\sqrt{(am)^2+\sum_\nu \Delta_\nu^2}(x,y) = \frac{1}{(2\pi a)^d}\int_{[-\pi,\pi]^d} d^dp \times\sqrt{(am)^2+\sum_\nu \Delta_\nu(p_\nu)^2} \times e^{i\sum_\mu p_\mu(x_\mu-y_\mu)/a}. \tag{3.17}$$

Writing $x=(x_1,\mathbf{x})$, $p=(p_1,\mathbf{p})$ and setting

$$M(\mathbf{p}) = \sqrt{(am)^2+\sum_{\nu=2}^d \Delta_\nu(p_\nu)^2} \tag{3.18}$$

we have

$$\sqrt{(am)^2+\sum_\nu \Delta_\nu^2}(x,y) = \frac{1}{(2\pi a)^{d-1}} \times\int_{[-\pi,\pi]^{d-1}} d^{d-1}\mathbf{p}e^{i\mathbf{p}\cdot\mathbf{x}/a}\int_{-\pi}^{\pi}\frac{dp_1}{2\pi a} \times\sqrt{M(\mathbf{p})^2+\Delta_1(p_1)^2}e^{ip_1(x_1-y_1)/a}. \tag{3.19}$$

The integral over p_1 here is the same as the previous $d=1$ integral (3.2) except that m is replaced here by $M(\mathbf{p})$. It can therefore be rewritten as (3.9) with this replacement. By our previous argument this integral decays exponentially $\sim z_-(t_{M(\mathbf{p})})^{|x_1-y_1|/a}$. The decay is slowest when $t_{M(\mathbf{p})}$ is largest, i.e. when $M(\mathbf{p})$ is smallest, and this happens when $\mathbf{p}=(0,\ldots,0)$ in which case $M=M(0)=am$. The same reasoning which led to (3.14) then implies that for $am\ll 1$ the operator kernel $\sqrt{(am)^2+\sum_\nu\Delta_\nu^2}(x,y)$ decays $\sim e^{-(1/2)\sqrt{m/2a}|x_1-y_1|}$ along the $\mu=1$ axis. Obvious modifications in the preceding show that the same decay holds along any other coordinate axis. Thus we see that the localization range is *no smaller than* $2\sqrt{2}\sqrt{a/m}$. It could however be *larger* along directions which are not parallel to a coordinate axis. To derive an upper bound on the localization range we use bounds similar to those leading to (3.16) to get[6]

$$|\sqrt{(am)^2+\sum_\nu\Delta_\nu^2}(x,y)| \le \frac{8\sqrt{2}}{a^{d-1}\pi\sqrt{t_{am}}(|x_\mu-y_\mu|-a)} \times(t_{am})^{|x_\mu-y_\mu|/2a} \tag{3.20}$$

holding for each $\mu=1,2,\ldots,d$. It follows that

$$|\sqrt{(am)^2+\sum_\nu\Delta_\nu^2}(x,y)|^d \le \prod_{\mu=1}^d \frac{8\sqrt{2}}{a^{d-1}\pi\sqrt{t_{am}}(|x_\mu-y_\mu|-a)} \times(t_{am})^{|x_\mu-y_\mu|/2a}$$

which in turn gives

$$|\sqrt{(am)^2+\sum_\nu\Delta_\nu^2}(x,y)| \le \frac{8\sqrt{2}}{a^{d-1}\pi\sqrt{t_{am}}(\prod_\mu(|x_\mu-y_\mu|-a))^{1/d}} \times(t_{am})^{\|x-y\|_1/2da}, \tag{3.21}$$

where $\|x-y\|_1=\sum_\mu|x_\mu-y_\mu|$ is the "taxi-driver" norm. A calculation analogous to (3.14) gives

$$(t_{am})^{\|x-y\|_1/2da} \underset{a\to 0}{\approx} e^{-(1/2d)\sqrt{m/2a}\|x-y\|_1}. \tag{3.22}$$

Since $\|x-y\|\le\|x-y\|_1$ it follows from this and (3.21) that the localization range is *no bigger* than $2d\sqrt{2a/m}$, i.e. it lies between this value and the previously derived lower limit $2\sqrt{2a/m}$. This completes the demonstration of exponential locality, with localization range vanishing $\sim\sqrt{a/m}$, claimed at the beginning of this section.

It is interesting to compare this result with the free field locality result derived in [9] for the operator

$$\sqrt{(am)^2+\sum_\nu(\nabla_\nu)^\dagger\nabla_\nu}. \tag{3.23}$$

This operator was shown there to be exponentially local but with the localization range remaining finite in the $a\to 0$ limit. The argument involved a mixture of analytic and numerical calculations[7]; however, the result can be established by purely analytic means, using the techniques introduced in the preceding, as we now demonstrate.

[5]Recall that the derivation of (3.9), and hence also (3.16), assumes $|x-y|>a$.

[6]The factor $1/a^{d-1}$ originates from the first integral in (3.19): $[1/(2\pi a)^{d-1}]\int_{[-\pi,\pi]^{d-1}} d^{d-1}\mathbf{p}|e^{i\mathbf{p}\cdot\mathbf{x}/a}|=1/a^{d-1}$.

[7]Specifically, the locality behavior of the *continuum version* of this operator was analytically determined and numerical calculations were then performed to check that the lattice operator kernel reduced to the continuum expression in the $a\to 0$ limit—see Part 3 of Appendix B in [9].

This will also show the origin of the difference in locality behavior between our operator and this one. (Note $\sum_\nu \nabla_\nu^\dagger \nabla_\nu = -\nabla^2$; we use the latter expression in the following.)

In the $d = 1$ case,

$$\sqrt{(am)^2 - \nabla^2}(x, y) = \frac{1}{2\pi a}\int_{-\pi}^{\pi} dp\sqrt{(am)^2 - \nabla^2(p)}e^{ip(x-y)/a} = \frac{1}{2\pi i a}\oint_{|z|=1}\frac{dz}{z}\sqrt{(am)^2 - (z - z^{-1})^2}z^{|x-y|/a}. \tag{3.24}$$

Setting $g(z) = (am)^2 - (z - z^{-1})^2 = z^{-2}((am)^2z^2 - (z^2 - 1)^2)$ we proceed as before by determining the z's satisfying $g(z) \in \mathbf{R}_-$ and $|z| \le 1$:

$$g(z) = -\lambda, \quad \lambda \in \mathbf{R}_+ \Leftrightarrow -(z^2 - 1)^2 + ((am)^2 + \lambda)z^2 = 0 \Leftrightarrow (z^2 - 1 + \sqrt{(am)^2 + \lambda}z)(z^2 - 1 - \sqrt{(am)^2 + \lambda}z) = 0. \tag{3.25}$$

The solutions with $|z| \le 1$ are $z_\pm(s = \sqrt{(am)^2 + \lambda})$ where

$$z_+(s) = -s/2 + \sqrt{1 + (s/2)^2} = -z_-(s), \qquad s \in [0, \infty[. \tag{3.26}$$

Note that $z_\pm(s) \to 0$ for $s \to \infty$. Hence the solutions form curves C_+ and C_- inside the unit circle in $\mathbf{C}$, parametrized, respectively, by $z_+(s)$ and $z_-(s)$, $s \in [am, \infty[$. The curves in this case are simply intervals on the real axis: $C_+ =]0, z_+(am)]$ and $C_- = [-z_+(am), 0[$ (see Fig. 2). The circle around which the integration in (3.24) is carried out can now be shrunk to a closed curve around $C_+ \cup \{0\} \cup C_-$, leading in the limit to an integral over these curves. By arguments similar to those in the previous case one then finds that $\sqrt{(am)^2 - \nabla^2}(x, y)$ decays exponentially $\sim |z_+(am)|^{|x-y|/a}$. From (3.26) we see that for $am \ll 1$ the magnitude of the decay factor becomes

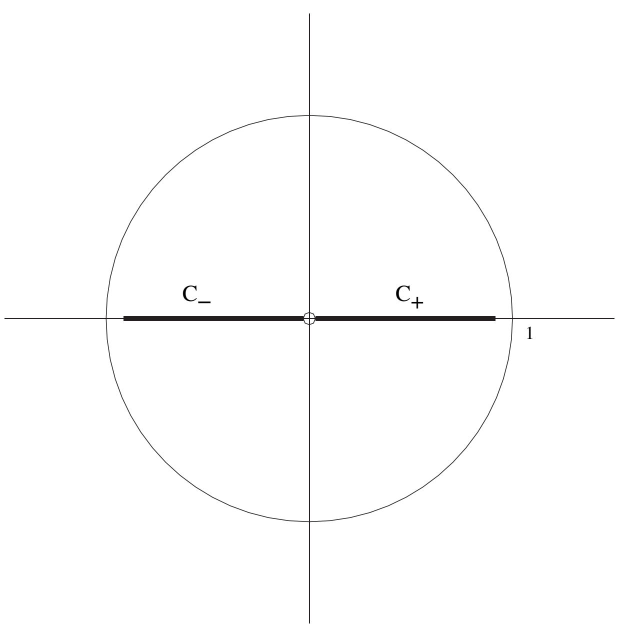

FIG. 2. The exclusion curves $C_\pm$ for the operator (3.24). Compare with Fig. 1 for the operator (3.2).

$$|z_+(am)|^{|x-y|/a} = ((1 - am/2 + O((am)^2))^{1/a})^{|x-y|} \overset{a\to 0}{\approx} e^{-(m/2)|x-y|}. \tag{3.27}$$

Thus the localization range in this case, $2/m$, is independent of a and remains finite in the $a \to 0$ limit. The general dimension d case can now be dealt with by an argument analogous to our earlier one; this leads to the result that $\sqrt{(am)^2 - \nabla^2}(x, y)$ is exponentially local with lower and upper bounds on the localization range being $2/m$ and $2d/m$, respectively, showing that the range is also finite $\sim m^{-1}$ in the $a \to 0$ limit in the general dimension d case. Thus we reproduce the general finding of [9] for the locality behavior of this free field operator.[8]

The origin of the different locality behavior of our operator and the free field operator considered in [9] is now apparent: The exponential decay in $d = 1$ dimension, which, as we have seen, is the same as the decay along a coordinate axis in general d dimensions, is given in both cases by $|z_{\max}|^{|x-y|/a}$, where $z_{\max}$ is the point on the "exclusion curves" in Fig. 1 (our case) or Fig. 2 (the case of Ref. [9]) which is closest to the unit circle. In our case,

$$|z_{\max}| \approx (t_{am})^{1/2} \approx (1 - \sqrt{am/2} + O(am))^{1/2}, \tag{3.28}$$

resulting in decay $\sim e^{-(1/2)\sqrt{m/2a}|x-y|}$, whereas in the case of Ref. [9],

$$|z_{\max}| \approx 1 - am/2 + O((am)^2), \tag{3.29}$$

resulting in decay $\sim e^{-(m/2)|x-y|}$. The essential difference is that the leading a-dependent term inside the bracket in (3.28) is $\sim\sqrt{am}$ whereas the corresponding leading term in (3.29) is $\sim am$. This leads to the localization range being $\sim\sqrt{a/m}$ in the former case and $\sim m^{-1}$ (independent of a) in the latter. The former tends to zero for $a \to 0$ while the latter stays constant in this limit.

[8]In the expression for the free field operator kernel $G(x, y)$ in Eq. (3.11) of [9] the integration range for dp_μ (after a change of variables $p_\mu \to p_\mu/a$) is $[-\pi/2, \pi/2]$. But since $-\nabla_\mu^2(p_\mu) = \sin^2(p_\mu)$ this gives precisely 1/2 of what the integration over $[-\pi, \pi]$ would give. Hence $G(x, y) = (1/2^d)\sqrt{(am)^2 - \nabla^2}(x, y)$ so the locality result derived above applies.

We remark that the technique of writing the kernel $\mathcal{O}(x, y)$ of a free field lattice operator as an integral around a closed curve in the complex plane, and then attempting to shrink the curve as a way of deriving locality properties, was used previously in a different context in Ref. [24].

IV. DISCUSSION

Working in the flavor (taste) representation for staggered fermions, we have shown that in the free field case there is a simple and natural Wilson-type lattice Dirac operator D on the blocked lattice, given by (2.10), which satisfies $\det D = (\det D_{\rm staggered})^{1/4}$ and is exponentially local with localization range vanishing $\sim\sqrt{a/m}$ for $a \rightarrow 0$. The techniques developed to derive the free field locality result are of a generally applicable nature, and we also used them to give a new, purely analytic derivation of the negative locality result in [9]. They can also be used to study free field locality properties of other lattice operators of current interest, in particular, the overlap Dirac operator [25], which is treated in a forthcoming paper [26].

Our free field operator can be gauged, i.e. coupled to the link variables of the original lattice, in a variety of ways. The simplest way is to define link variables V_μ on the blocked lattice in terms of the link variables U_μ on the original lattice by

$$V_\mu(2x) = U_\mu(2x)U_\mu(2x + a\hat{\mu}) \qquad (4.1)$$

($\hat{\mu}$ = unit vector in the positive μ direction); then the difference operators ∇_μ, Δ_ν on the blocked lattice in (2.10) can be coupled to V_μ in the usual way—this specifies the "minimal gauging" of our D. However, the resulting operator cannot be expected to satisfy the SDR in the interacting case. Our argument in Sec. II does not carry over to this case; it is specific to the free field case.

Regarding the possibility of gauging our operator such that the SDR does continue to hold in the interacting case, we note the following. The taste-mixing part of the staggered Dirac operator leaves unbroken a U(1) subgroup of the continuum U(4) axial flavor symmetry.[9] In the taste representation, this symmetry can be expressed in the free field case as

$$\{\gamma_5 \otimes \Gamma_5, D_{\rm stt}\} = 0, \qquad (m = 0) \qquad (4.2)$$

with notations as in (2.2). This chiral symmetry protects staggered fermions against additive mass renormalization [27–29]. The sea quark effective action in dynamical staggered fermion simulations is $\log\det(D_{\rm st})^{1/4} = \frac{1}{4}\,{\rm Tr}\log D_{\rm st}$, the same as for a usual staggered fermion modulo an overall factor 1/4. Thus the protection against additive mass renormalization is also present in this case. This situation would be difficult, if not impossible, to reconcile with the existence of a single-flavor D satisfying the SDR unless the D is also protected against additive mass renormalization: if the bare mass is small then for staggered fermions the physical mass will also be small, whereas for a lattice fermion described by a D which does not have a chiral symmetry the physical mass will be large due to the additive mass renormalization induced by radiative corrections. Since our D is of Wilson-Dirac form, any gauged version of it will be afflicted with additive mass renormalization, unless there is a very special choice of gauging which endows this D with a new, hitherto undiscovered type of chiral symmetry. The latter seems very unlikely though, so most probably a gauged version of our D satisfying the SDR simply does not exist.

The symmetry corresponding to (4.2) for $(D_{\rm stt})^\dagger D_{\rm stt}$ in the free field case is

$$[\gamma_5 \otimes \Gamma_5, (D_{\rm stt})^\dagger D_{\rm stt}] = 0, \qquad (m = 0). \qquad (4.3)$$

But, as noted in Sec. II, $(D_{\rm stt})^\dagger D_{\rm stt} \sim \mathbf{1} \otimes \mathbf{1}$ in the free field case, so (4.3) is trivially satisfied. This explains why it was possible to find a single-flavor D without chiral symmetry but nevertheless satisfying the SDR in the free field case. In the interacting case $(D_{\rm stt})^\dagger D_{\rm stt}$ is no longer $\sim\mathbf{1} \otimes \mathbf{1}$ and the gauged version of the symmetry (4.3) is a nontrivial property.

Although a gauged version of our D satisfying the SDR is unlikely to exist, the free field locality result for it is still relevant as a general indication of the possibility of having local single-flavor Dirac operators satisfying the SDR, and as a first step toward constructing more sophisticated operators which have a chance to be local (also for $m = 0$) and satisfy the SDR in the full interacting case. At present it is the only analytic positive locality result derived for a solution of the SDR in any gauge background (the background in our case being the trivial one).

The above discussion indicates that, for a viable D to satisfy the SDR in the full interacting case, it should have an exact chiral-type symmetry (presumably corresponding in some way to the aforementioned chiral symmetry of staggered fermions). The only such symmetry currently known for single-flavor lattice Dirac operators is the lattice-deformed chiral symmetry [30] possessed by operators satisfying the Ginsparg-Wilson (GW) relation [31–33] and its generalizations [34]. This suggests to look for free field solutions to the SDR which also satisfy the GW relation, in the hope that among these there may be a D which can be gauged such that the SDR continues to hold in the interacting case. In fact this was already investigated by numerical means on finite lattices in Ref. [13]. The numerical results there appear to be encouraging. However, the problem can also be addressed analytically

[9]We are assuming that the mass term of the staggered fermion is of the standard form $m(\mathbf{1} \otimes \mathbf{1})$. [If the mass matrix is not proportional to the identity operator then the interpretation of the U(1) symmetry is different.] Note that the U(1) symmetry is *not* the diagonal U(1) subgroup in U(4). The latter, associated with the axial anomaly, is explicitly broken by the taste-mixing part of $D_{\rm staggered}$.

and when this is done difficulties become apparent. Setting the lattice spacing of the blocked lattice to unity for convenience, the GW relation in its broad form is

$$\gamma_5 D + D\gamma_5 = 2D\gamma_5 RD \tag{4.4}$$

where R is an arbitrary local scalar Hermitian operator. As noted at the end of Sec. II, a sufficient condition for the free field D to satisfy the SDR is

$$D^\dagger D = \Delta + m^2. \tag{4.5}$$

In the free field case, assuming γ_5-Hermiticity $D^\dagger = \gamma_5 D \gamma_5$ and exploiting the fact that solutions to (4.4) are of the form $D = (2R)^{-1}(1 + \gamma_5 \epsilon)$ where $\epsilon^2 = 1$ and $\epsilon^\dagger = \epsilon$, the most general solution to (4.4) is seen to be of the form

$$D_{\mathrm{GW}} = \frac{1}{2R}\left(1 + \frac{\gamma^\mu \tilde{\nabla}_\mu + W}{\sqrt{-\tilde{\nabla}^2 + W^2}}\right), \tag{4.6}$$

where $\tilde{\nabla}_\mu$ and W are sums of scalar operators multiplied by an even number of γ matrices. Straightforward algebra now shows that requiring this operator to satisfy (4.5) fixes the W such that

$$D_{\mathrm{GW}} = \gamma^\mu \tilde{\nabla}_\mu \frac{R}{|R|}\sqrt{\frac{(\Delta + m^2)(1 - \frac{1}{4}(2R)^2(\Delta + m^2))}{-\tilde{\nabla}^2}} + R(\Delta + m^2). \tag{4.7}$$

This is the general solution to (4.4) and (4.5) in the free field case. The numerical solutions investigated in Ref. [13] are particular cases of this operator (or more precisely, approximations to it on finite lattices) with the $\tilde{\nabla}_\mu$'s being scalar operators. The main interest here is in the case $m = 0$, since it is in the chiral limit that D should have the chiral symmetry implied by the GW relation (4.4).

In the simplest case where $R = 1/2$ and $\tilde{\nabla}_\mu = \nabla_\mu$, i.e. the usual symmetrized difference operator, D_{GW} reduces in the $m = 0$ case to

$$\gamma^\mu \nabla_\mu \sqrt{1 - \frac{1}{4}\sum_{\mu \neq \nu}\left(\frac{\Delta_\mu \Delta_\nu}{-\nabla^2}\right)} + \frac{1}{2}\Delta. \tag{4.8}$$

While this operator correctly reproduces the continuum free field Dirac operator in the $a \to 0$ limit, it is most unlikely to have acceptable locality behavior. The presence of the $(-\tilde{\nabla}^2)^{-1}$ inside the square root in the general operator (4.7) makes it difficult to envisage that there exist $\tilde{\nabla}_\mu$'s and a local R for which this operator has acceptable locality behavior either, in spite of the numerical indications from Ref. [13]. Thus it would seem that the condition (4.5), which is sufficient, but not necessary, for the SDR to be satisfied, is actually too restrictive to lead to a local operator D satisfying both the SDR and GW relations.

The preceding considerations indicate that finding a viable exact solution to the SDR in the interacting case is a difficult problem. However, to resolve the fourth root issue is not actually necessary to have an exact solution; it suffices to find a viable lattice Dirac operator which satisfies the SDR *approximately* in the sense that the effective action difference

$$d(U) = \tfrac{1}{4}\mathrm{logdet}D_{\mathrm{st}} - \mathrm{logdet}D \tag{4.9}$$

is effectively just a lattice Yang-Mills action for the gauge field. In this case, representing the quark determinant by $\det(D_{\mathrm{st}})^{1/4}$ is physically equivalent to representing it by $\det D$ together with a renormalization of the bare coupling constant (i.e. a shift in β). In other words, $d(U)$ has the same effect as the fermion determinant for dynamical heavy quarks [21]. In connection with this it is useful to note that the perturbative expansion of a general single-flavor lattice fermion determinant has the form [35]

$$\mathrm{logdet}D = \left(-\frac{1}{8\pi^2}\log(am)^2 + c_D\right)S_{\mathrm{YM}}(A) + \sum_{n=2}^{\infty}(I_n(A;m) + v_n(A;am)), \tag{4.10}$$

where $S_{\mathrm{YM}}(A)$ is the continuum Yang-Mills action, $I_n(A;m)$ is a nonlocal continuum functional of order n in A, and the $v_n(A;am)$'s (also nonlocal and of order n in A) are terms which vanish for $am \to 0$. The dependence on the specific choice of lattice Dirac operator D enters only through the numerical coefficient c_D and the functions $v_n(A;am)$. [A gauge field-independent term which diverges for $am \to 0$ has been ignored in (4.10).] In fact the perturbative expansion of $\frac{1}{4}$ logdetD_{st} has the same form (4.10) as a single-flavor Dirac operator [35]. Letting c_{st} denote the coefficient c_D in this case, it follows that the perturbative expansion of the effective action difference (4.9) has the form

$$d(A) = (c_{\mathrm{st}} - c_D)S_{\mathrm{YM}}(A) + \sum_{n=2}^{\infty} w_n(A;am) \tag{4.11}$$

with each $w_n(A;am)$ vanishing for $am \to 0$. Thus it would seem that, in the perturbative setting at least, for $am \ll 1$ the effective action difference is indeed just a Yang-Mills action for the gauge field, for any sensible choice of single-flavor D. The situation is not this simple though—although they vanish for $a \to 0$, the functions v_n in (4.10) and w_n in (4.11) still do affect the quantum continuum limit. At the perturbative level this is manifested in that Feynman diagrams with vertices from these terms can be nonvanishing, in fact divergent. To see this, recall that the terms in the perturbative expansion of logdet$D$ are given by 1-(fermion)-loop gluonic n-point functions. The internal propagators and vertices in these receive radiative corrections. In particular, unless D is protected by a chiral symmetry, the radiative corrections to the internal propagators give rise to a large additive mass renormalization. Since there is no corresponding effect from D_{st} to cancel this, it will manifest itself in the divergence (for $am \to 0$) of various Feynman diagrams involving vertices from the "irrelevant" terms $w_n(A; am)$ in the effective action difference (4.11). Thus the importance of D having a chiral symmetry becomes clear in this context as well.

In perturbative (lattice) QCD the renormalizations of interaction vertices are independent of the renormalization of the fermion propagator and bare mass. Thus (4.11) and the observations above would suggest that, for $am \ll 1$, the effective action difference $d(A)$ is indeed essentially a Yang-Mills action when the single-flavor lattice Dirac operator D has an exact chiral symmetry; e.g., when D is the overlap Dirac operator. Support for this hypothesis comes from a numerical study carried out in two dimensions in Ref. [17]. In two spacetime dimensions the perturbative expansion of $\log\det D$ is completely universal, modulo terms which vanish for $am \to 0$ [35] (this is a reflection of the fact that QCD in two dimensions is super renormalizable). Thus the first term in the right-hand side of the effective action difference (4.11) is absent in this case, and the hypothesis then states that representing the quark determinant by $\det(D_{\mathrm{st}})^{1/2}$ is equivalent to using $\det D_{\mathrm{ov}}$ (with D_{ov} being the overlap Dirac operator) *without* any renormalization of the bare coupling. If this hypothesis holds, then $\det(D_{\mathrm{st}})^{1/2}$ should coincide with $\det D_{\mathrm{ov}}$ for equilibrium gauge configurations of an ensemble generated by taking the probability weight to be $e^{-\beta S_{\mathrm{YM}}(U)} \det(D_{\mathrm{st}})^{1/2}$. And this is precisely what was found to good accuracy in a numerical study in Ref. [17].

It must be remembered though that the perturbative picture is not the full picture. Low-lying eigenvalues of the Dirac operator are associated with long-range, low energy dynamics in QCD which is not captured by the perturbative framework. Indeed, numerical studies of the Wilson fermion determinant in Ref. [36] show that the log of the determinant cannot be modeled by a linear combination of local loop functionals [i.e. the functional of the form $\mathrm{Tr}U(\sigma)$ where $U(\sigma)$ is the product of the link variables around a closed lattice path σ]; in particular, it cannot be modeled by a local lattice YM action. However, the product of the Dirac eigenvalues of magnitude $\geq \Lambda_{\mathrm{QCD}}$ *does* admit such a description [36]. Thus, the aforementioned hypothesis, coming from the perturbative considerations above, should be regarded as applying to the *truncations* of $\det(D_{\mathrm{st}})^{1/4}$, $\det D$, and the effective action difference $d(U)$, given by excluding the eigenvalues of magnitude $<\Lambda_{\mathrm{QCD}}$. (A way to implement and study this truncation in the perturbative setting is mentioned in [37].)

Specifically, defining $\det D_{\mathrm{high}}$ and $\det D_{\mathrm{low}}$ to be the products of the eigenvalues of D of magnitudes $\geq \Lambda_{\mathrm{QCD}}$, and $<\Lambda_{\mathrm{QCD}}$, respectively, and splitting up the effective action difference into $d = d_{\mathrm{high}} + d_{\mathrm{low}}$ in the obvious way, the hypothesis can be stated as follows: "When D is the overlap Dirac operator and $am \ll 1$ then $d_{\mathrm{high}}(U)$ is essentially a local lattice YM action in 4 dimensions, and essentially vanishing in 2 dimensions." The question of whether $d(U)$ itself is effectively a local YM action (which would give a positive resolution of the fourth root issue if the answer is affirmative) is then reduced to a question of whether or not $d_{\mathrm{low}}(U)$ is effectively zero. Thus it would be highly desirable to numerically study $\det(D_{\mathrm{st,low}})^{1/4}$, $\det D_{\mathrm{ov,low}}$, and thereby $d_{\mathrm{low}}(U)$, in equilibrium gauge backgrounds in 4 dimensions. We remark that, in 2 dimensions, combining the hypothesis that d_{high} vanishes (in 2 dimensions) with the numerical agreement [17] between the full rooted staggered and overlap determinants implies that d_{low} does indeed vanish in this case.

In 4 dimensions numerical studies have found that, after applying a UV-filtering procedure, there is good agreement between the low-lying eigenvalues of D_{ov} and D_{st} (modulo a fourfold degeneracy in the latter)[17].[10] Comparisons of the spectrum of D_{st} with predictions of random matrix theory also back up this picture [18,19]. While this does not by itself prove that d_{low} vanishes, it is certainly compatible and suggestive of it.

While the numerical work in this direction may lead to a resolution of the fourth root issue at a practical level, one should ask whether it is possible to also get a resolution at the theoretical level in this approach. If it is possible it will probably happen as follows: (i) Use renormalization group arguments to justify a perturbative treatment of d_{high} and verify the hypothesis that it is effectively a local lattice YM action. (ii) By applying random matrix theory and theoretical implications of UV filtering to the low-lying spectra of D_{ov} and D_{st} show that d_{low} is effectively zero when taking the quantum continuum limit.

Another interesting and promising approach to the fourth root issue has been given recently by Shamir [14]. A renormalization group argument is used to express the free field staggered fermion action in the flavor (taste) representation on a lattice spacing a_0 as an action on a coarse lattice of spacing $a = 2^n a_0$. This results in a decomposition of the staggered fermion determinant in the form

$$\det(D_{\mathrm{st}}) = \det(D_n)\det(G_n^{-1}). \qquad (4.12)$$

The operator D_n encodes the low energy/long-range dynamics of staggered fermions; it decays exponentially with localization range $\sim a$, satisfies a GW relation when $m = 0$, and becomes proportional to the identity matrix in flavor

[10] In fact it is only when this UV filtering is applied that the aforementioned agreement between the rooted staggered and overlap determinants in 2 dimensions holds [17]. Without the filtering the agreement breaks down, just as the agreement between the low-lying eigenvalues does. So it is tempting to ascribe the breakdown in the agreement between the determinants to the breakdown in the agreement between the low-lying eigenvalues. This gives a further hint that vanishing of d_{low} is intimately connected with having agreement between the low-lying eigenvalues of D_{ov} and D_{st}.

space in the large n limit: $\lim_{n\to\infty} D_n = D_{\rm rg} \otimes \mathbf{1}_{\rm flavor}$. Hence, in this limit

$$\det(D_{\rm st})^{1/4} = \det(D_{\rm rg}) \det(G_\infty^{-1/4}). \qquad (4.13)$$

The magnitude of the spectrum of each G_n^{-1} has a lower bound $\sim a$; in units of the fine lattice spacing a_0 this blows up for large n, so the expectation is that in a gauged version of this setting the effect of $\det(G_\infty^{-1/4})$ in (4.13) is exactly the same as that of the determinant of a heavy dynamical fermion: to simply renormalize the bare coupling parameter.

It should be pointed out though that the $n \to \infty$ limit leading to (4.13) cannot actually be taken in practise—it corresponds to $a \to 0$—but a must remain nonzero since it is the spacing of the lattice on which the staggered fermion lives and the lattice QCD simulations are performed. Therefore, in this approach one needs to remain at finite n, i.e. the setting of (4.12). For large finite n the operator D_n is close to being diagonal in flavor space, but is not exactly diagonal. This is different from the situation in the present paper where we obtain a single-flavor candidate Dirac operator already at nonzero lattice spacing. To fully resolve the fourth root issue in Shamir's approach it is necessary to find a single-flavor lattice Dirac operator D' such that adding $\mathrm{logdet} D' - \frac{1}{4}\mathrm{logdet} D_n$ to the lattice gauge field action does not affect the quantum continuum limit. Shamir has a proposal for this operator D' [38]. Moreover, his approach has a definite possibility of being extended to the interacting case, although this remains a difficult challenge for future work.

An appealing feature of Shamir's approach is that the GW chiral symmetries of D_n and $D_{\rm rg}$ (at $m = 0$) originate in a clear and direct way from the chiral symmetry (4.2) of the staggered Dirac operator. This also raises intriguing questions. The chiral symmetry of a GW Dirac operator is generally anomalous—it gets broken by the fermion integration measure [30]. On the other hand, the chiral symmetry of $D_{\rm st}$ gets broken spontaneously in the $m \to 0$ limit (at least at strong coupling), and there is an associated Goldstone meson [39,40]. In connection with this we mention a potentially troubling aspect of the fourth root prescription which has been pointed out already by Creutz [41]: When the determinant for a single quark is represented by $\det(D_{\rm st})^{1/4}$, what becomes of the Goldstone meson associated with the spontaneous breaking of the chiral symmetry of $D_{\rm st}$? Single-flavor Dirac operators are not supposed to have spontaneously broken chiral symmetries. This and other intriguing issues for the fourth root prescription remain as an urgent topic for future work.

Finally, we mention that a completely different approach to this issue, involving relating the fourth root prescription to local theories via a parameter deformation in a family of lattice theories in 6 spacetime dimensions, has been described by Neuberger in Ref. [12].

ACKNOWLEDGMENTS

I thank Mike Peardon for an interesting discussion and for telling me about Ref. [13] prior to publication. Thanks also go to the National Center for Theoretical Sciences and the nuclear/particle physics group at National Taiwan University for kind hospitality and support during visits where part of this work was done, and, in particular, to Ting-Wai Chiu for making the arrangements. For interesting discussions on topics related to this work I would like to thank Mike Creutz, Urs Heller, Taku Izubuchi, Tony Kennedy, Peter Orland and, in particular, Rajamani Narayanan, whom I also thank for drawing my attention to Refs. [36,37]. Last but not least I thank Yigal Shamir for useful correspondence. At Leiden the author was supported by a Marie Curie fellowship from the European Commission, Contract No. HPMF-CT-2002-01716. Some of the work on this manuscript was done during the KITP program "Modern Challenges for Lattice Field Theory" and I thank the KITP for support under NSF Grant No. PHY99-0794. At FIU the author is supported by NSF Grant No. PHY-0400402.

[1] C. W. Bernard *et al.*, Nucl. Phys. Proc. Suppl. **60A**, 297 (1998); G. P. Lepage, Nucl. Phys. Proc. Suppl. **60A**, 267 (1998); C. W. Bernard *et al.* (MILC Collaboration), Phys. Rev. D **58**, 014503 (1998); G. P. Lepage, Phys. Rev. D **59**, 074502 (1999); K. Orginos and D. Toussaint (MILC Collaboration), Nucl. Phys. B, Proc. Suppl. **73**, 909 (1999); K. Orginos, D. Toussaint, and R. L. Sugar (MILC Collaboration), Phys. Rev. D **60**, 054503 (1999); C. W. Bernard *et al.* (MILC Collaboration), Phys. Rev. D **61**, 111502 (2000); K. Orginos and D. Toussaint (MILC Collaboration), Phys. Rev. D **59**, 014501 (1999).
[2] C. T. H. Davies *et al.* (HPQCD, UKQCD, MILC, and Fermilab Collaborations), Phys. Rev. Lett. **92**, 022001 (2004).
[3] C. Aubin *et al.* (MILC Collaboration), Phys. Rev. D **70**, 114501 (2004).
[4] I. F. Allison *et al.* (HPQCD, UKQCD, and Fermilab Collaborations), Phys. Rev. Lett. **94**, 172001 (2005).
[5] K. Jansen, Nucl. Phys. B, Proc. Suppl. **129**, 3 (2004).
[6] T. DeGrand, Int. J. Mod. Phys. A **19**, 1337 (2004).
[7] H. Neuberger, hep-ph/0402148.
[8] D. H. Adams, Phys. Rev. Lett. **92**, 162002 (2004); Nucl. Phys. B, Proc. Suppl. **140**, 148 (2005).
[9] B. Bunk, M. Della Morte, K. Jansen, and F. Knechtli,

Nucl. Phys. **B697**, 343 (2004).
[10] B. Bunk, M. Della Morte, K. Jansen, and F. Knechtli, hep-lat/0408048.
[11] A. Hart and E. Muller, Phys. Rev. D **70**, 057502 (2004).
[12] H. Neuberger, Phys. Rev. D **70**, 097504 (2004).
[13] F. Maresca and M. J. Peardon, hep-lat/0411029.
[14] Y. Shamir, Phys. Rev. D **71**, 034509 (2005).
[15] J. Giedt, hep-lat/0507002.
[16] S. R. Sharpe and R. S. van der Water, Phys. Rev. D **71**, 114505 (2005).
[17] S. Dürr, C. Hoelbling, and U. Wenger, Phys. Rev. D **70**, 094502 (2004); hep-lat/0409108; S. Dürr and C. Hoelbling, Phys. Rev. D **69**, 034503 (2004); hep-lat/0408039; Phys. Rev. D **71**, 054501 (2005).
[18] E. Follana, A. Hart, and C. T. H. Davies (HPQCD Collaboration), Phys. Rev. Lett. **93**, 241601 (2004); E. Follana, A. Hart, C. T. H. Davies, and Q. Mason, Phys. Rev. D **72**, 054501 (2005).
[19] K. Y. Wong and R. M. Woloshyn, Phys. Rev. D **71**, 094508 (2005).
[20] F. Gliozzi, Nucl. Phys. **B204**, 419 (1982); A. Duncan, R. Roskies, and H. Vaidya, Phys. Lett. **114B**, 439 (1982); H. Kluberg-Stern, A. Morel, O. Napoly, and B. Petersson, Nucl. Phys. **B220**, 447 (1983).
[21] A. Hasenfratz and T. A. DeGrand, Phys. Rev. D **49**, 466 (1994).
[22] N. Kawamoto and J. Smit, Nucl. Phys. **B192**, 100 (1981).
[23] P. Rossi and U. Wolff, Phys. Rev. D **30**, 2233 (1984).
[24] K. Fujikawa and M. Ishibashi, Nucl. Phys. **B605**, 365 (2001).
[25] H. Neuberger, Phys. Lett. B **417**, 141 (1998).
[26] D. H. Adams and P. Bicudo (unpublished).
[27] H. S. Sharatchandra, H. J. Thun, and P. Weisz, Nucl. Phys. **B192**, 205 (1981).
[28] P. Mitra and P. Weisz, Phys. Lett. **126B**, 355 (1983).
[29] M. F. L. Golterman and J. Smit, Nucl. Phys. **B245**, 61 (1984).
[30] M. Lüscher, Phys. Lett. B **428**, 342 (1998).
[31] P. H. Ginsparg and K. G. Wilson, Phys. Rev. D **25**, 2649 (1982).
[32] P. Hasenfratz, V. Laliena, and F. Niedermayer, Phys. Lett. B **427**, 125 (1998).
[33] H. Neuberger, Phys. Lett. B **427**, 353 (1998).
[34] K. Fujikawa, Nucl. Phys. **B589**, 487 (2000).
[35] D. H. Adams (unpublished).
[36] A. Duncan, E. Eichten, R. Roskies, and H. Thacker, Phys. Rev. D **60**, 054505 (1999).
[37] A. Duncan, E. Eichten, and H. Thacker, Phys. Rev. D **59**, 014505 (1999).
[38] Y. Shamir (private communication).
[39] P. Rossi and U. Wolff, Nucl. Phys. **B248**, 105 (1984).
[40] D. H. Adams and S. Chandrasekharan, Nucl. Phys. **B662**, 220 (2003).
[41] M. Creutz, presentation during the KITP program ‘‘Modern Challenges for Lattice Field Theory’’; (private communication).