

# New domain-wall fermion actions

Yigal Shamir\*

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

(Received 12 April 2000; published 11 August 2000)

In perturbation theory, the wave function of domain-wall quarks decreases exponentially with the fifth coordinate. We show that, regardless of the quark's own momentum, the fall-off rate of the one-loop wave function is equal to the *slowest* rate encountered at the tree level for any lattice four-momentum. We propose new domain-wall actions involving beyond-nearest-neighbor couplings in the four physical dimensions, for which the perturbative wave function decreases much faster. It is hoped that the new actions may preserve the good chiral properties of domain-wall fermions up to larger values of the lattice spacing.

PACS number(s): 11.15.Ha, 12.38.Gc

## I. INTRODUCTION

The coupling of the right-handed and the left-handed components of Wilson fermions through the QCD interaction leads in the continuum limit to the chiral anomaly [1], but for finite lattice spacing  $a$  it also leads to lattice-artifact violations of chiral symmetries. This results in an additive renormalization of the quark mass, as well as in a severe tuning problem for four-fermion operators which are needed for the computation of weak matrix elements. The mass renormalization is  $O(g_0^2)$  in lattice units. Since  $g_0^2 \sim [\log(a\Lambda_{\text{QCD}})]^{-1}$ , the mass renormalization diverges like  $[a\log(a\Lambda_{\text{QCD}})]^{-1}$  in the continuum limit  $a \rightarrow 0$ . In the challenging lattice calculation of non-leptonic kaon decays (e.g.,  $\epsilon'/\epsilon$ ) the tuning problem is formidable because of the large number of chirality-disallowed mixings.

In the domain-wall formulation of lattice QCD [2–6], the two chiral components arise as surface states on opposite boundaries of a five-dimensional lattice, and one expects their coupling to vanish when the size of the fifth dimension tends to infinity. (The fifth coordinate will be denoted  $s$ , and it takes values  $0 \leq s \leq N_s - 1$ .) A precise non-perturbative characterization of chiral symmetry violations can be given in terms of the transfer matrix for hopping in the  $s$ -direction [3,6]. Being the result of slow decay of correlations in the  $s$ -direction, chiral symmetry violations are associated with near-unity eigenvalues of that transfer matrix. For a given gauge-field configuration the approach to the chiral limit is exponential *iff* the spectrum of the transfer matrix has a gap, and the fall-off rate is determined by the size of the gap.

In full QCD there are several analytic results concerning the  $N_s \rightarrow \infty$  limit. The approach to the chiral limit is exponential in perturbation theory [5,7,8], and the same is true non-perturbatively if a constrained gauge action (believed to be in the same universality class as the standard plaquette action) is used [9].

For an unconstrained action one can also prove non-perturbatively that chiral symmetry is restored in the limit  $N_s \rightarrow \infty$ , provided the (finite!) number of sites in each of the four physical lattice dimensions is held fixed [6]. When near-

unity eigenvalues start playing a significant role, the chiral limit may be approached as slow as  $1/N_s$ . The proof that certain symmetries are restored in the limit  $N_s \rightarrow \infty$  is actually valid for *any* value of the coupling constant. But the identification of the restored symmetries as chiral ones depends on the fermion spectrum. It was recently shown [10] that within the strong-coupling expansion the massless spectrum of the domain-wall lattice Hamiltonian is either doubled or empty. Therefore the restored symmetries are not chiral at strong coupling (for further details see Appendix C 1).

Of major importance is the question of how close to the chiral limit one gets in Monte Carlo simulations. In trying to answer this question we rely on two sources. The first is the spectrum of the transfer matrix, or of the closely-related [3,11] Hermitian Wilson-Dirac operator. In the latter case, a key finding [12] is that the *spectral density* of near-zero modes (corresponding to near-unity eigenvalues of the transfer matrix) rises by two orders of magnitude as the (quenched) coupling changes from  $6/g^2 \equiv \beta = 6.3$  ( $a^{-1} \sim 4$  GeV) to  $\beta = 5.7$  ( $a^{-1} \sim 1$  GeV). We hope that more results on the eigenvalue spectrum will be available in the future.

More information is available through lattice computations of various correlation functions [13–20]. A detailed numerical study of the chiral limit of domain-wall fermions was first carried out in the Schwinger model [13]. In QCD the first domain-wall simulations were promising, and the results for weak matrix elements ( $B_K$ ,  $O_{LL}$ ) [14] and for the strange-quark mass [15] were in agreement with other methods. As of today more data is available. The pion-mass squared, which should extrapolate linearly to zero with the quark mass, is the most obvious measure of chiral symmetry. Using domain-wall fermions at quenched  $\beta = 6.0$  ( $a^{-1} \sim 2$  GeV) the extrapolated pion mass (for  $N_s \sim 20$ ) does not vanish exactly at zero quark mass but, rather, at a negative value of the order of few times  $10^{-3}$  in lattice units [16] (see also Ref. [17]). This value, however small, is in the same range as the light quark masses.

In a sense, chiral symmetry violations are worst for the pion mass, because the lattice-artifact term in the PCAC (partial conservation of axial current) relation is an ensemble average of a positive fermion correlator (see Appendix C 1;

\*Email address: shamir@post.tau.ac.il

we comment that that lattice-artifact term is a better measure of chiral symmetry violation compared to the extrapolated pion mass, since it does not suffer from theoretical uncertainties due to chiral perturbation theory and due to finite-volume effects).

No such positivity is encountered in the calculation of weak matrix elements [14,18], so chiral symmetry violations are expected to be smaller in this case. For example, in a recent simulation of four-fermion operators using the non-perturbative renormalization scheme, again at quenched  $\beta = 6.0$ , and using  $N_s = 16$ , it was found [18] that mixing into wrong-chirality operators was practically zero (in comparison with 10% for Wilson fermions in a typical example).

Going to a smaller value of the inverse lattice spacing, the situation at  $a^{-1} \sim 1$  GeV is unsatisfactory, as deviations from chiral symmetry are significant even for  $N_s$  as large as 50 or 100 (in both quenched and dynamical simulations) [19,20]. In the opposite direction, at  $a^{-1} \gtrsim 3$  GeV, no difficulties with the restoration of chiral symmetry have been reported [14,12,9].

We believe that the existing results, especially those for weak matrix elements at  $a^{-1} \sim 2$  GeV, do represent a breakthrough compared to the “pre domain-wall era.” On the other hand, the results for the pion mass at  $a^{-1} \sim 2$  GeV are not as good as one would hope for, and the present situation at  $a^{-1} \sim 1$  GeV makes scaling studies with domain-wall fermions very difficult.

Having summarized the situation in Monte Carlo simulations let us return to the underlying physics. The key question is what mechanism(s) determine the abundance of near-unity eigenvalues of the transfer matrix. It is known that a few exact-unity eigenvalues must occur during the transition from one topological sector to another on a finite lattice with periodic boundary conditions [3,4]. We believe, however, that the role of topology changing has been over-emphasized, for topological considerations alone do not explain the proliferation of near-unity eigenvalues nor the magnitude of the ensuing chiral symmetry violations.

A simple explanation may be that the observed chiral symmetry violations arise (mainly) from generic fluctuations of the gauge field [21]. The effect of fluctuations need not be small, because the coupling constant used in simulations is not small either. At the same time, as long as the coupling constant has not grown too much, perturbation theory should provide a reliable approximation of the leading quantum effects.

In this paper we calculate the fifth coordinate’s wave function of domain-wall quarks in the one-loop approximation (Sec. II). The results lead us to consider new classes of domain-wall actions (Sec. III). A preliminary account of this work was given in Ref. [22]. (An alternative-complementary approach, whose relative merits are discussed in Sec. IV, is to employ an improved gauge action [19].)

We now give an overview of the one-loop calculation. For large  $N_s$ , zero bare quark mass and with the right-handed quark field near the  $s=0$  boundary, the dressed fermion propagator near that boundary is

$$G_{s,s'}(p) = P_+ \chi(s) \frac{1}{i \not{p} (1 + \Sigma_K)} \chi(s') + \text{Reg.}, \quad s, s' \ll N_s, \quad (1.1)$$

where  $P_{\pm} = \frac{1}{2}(1 \pm \gamma_5)$ , “Reg.” stands for a continuous function of the four-momentum  $p$ , and  $\Sigma_K \approx \Sigma_K[g^2, g^2 \log(p^2)]$ . A unique feature of the domain-wall scheme is  $\chi(s)$ , the  $s$ -coordinate wave function for (right-handed) quark modes. At tree level one has

$$\chi_0(s) \propto q_0^s \equiv (1 - M)^s. \quad (1.2)$$

(The five-dimensional mass term  $M$  is often referred to as the domain-wall height, and should not be confused with the quark mass [5].) By choosing  $M=1$  the free wave function can be completely localized on the boundary

$$\chi_0(s) = \lim_{M \rightarrow 1} (1 - M)^s = \delta_{s,0}. \quad (1.3)$$

The result of the one-loop calculation of the wave function is

$$\chi_1(s) \sim s^{-2} q_1^s, \quad (1.4)$$

which contains also a power correction. Like an ordinary Wilson mass,  $M$  is renormalized additively. Making an optimal choice of  $M$  we find

$$q_1 = 0.5. \quad (1.5)$$

It should be noted that the difference between  $q_0$  and  $q_1$  is  $O(1)$ . In the full one-loop result [Eqs. (2.3) to (2.9) below]  $g^2$  occurs as a pre-factor of relatively little importance.

Let us now explain the physical origin of  $q_1$ . Consider the free domain-wall propagator  $G_{s,s'}^0(p)$  for a given four-momentum  $p$  in the vicinity of the  $s=0$  boundary at zero quark mass. The  $s$ -correlations described by this propagator are controlled by an exponent  $\alpha(p)$ . Each term in the propagator involves a factor  $\exp(-d\alpha(p))$  where  $d$  stands for either the separation  $|s - s'|$  or the sum of distances from the boundary  $s + s'$ . For the standard domain-wall action one has  $\max\{\exp(-\alpha)\} = 0.5$  for  $M=1$  where the maximum over the Brillouin zone is obtained at the “corner”  $p_{\pi} = (\pi, 0, 0, 0)$  and its three permutations. We will denote the set of global maxima by  $\mathcal{P}$ .

Now, at the tree level, a fermion eigenmode with momentum  $p$  propagates independently of all other eigenmodes. But for any non-zero gauge coupling the fermions propagate in non-trivial backgrounds, and these backgrounds allow any given momentum eigenmode to couple to all momentum eigenmodes. In particular, small-momentum quark modes couple to modes with  $p \in \mathcal{P}$ .

We arrive at the following physical picture. A four-dimensional fermion mode created on a given  $s$ -layer mixes on that layer with the modes of  $\mathcal{P}$  through the gauge field. As a mode with  $p \in \mathcal{P}$ , the fermion propagates with minimal suppression to some other layer  $s'$ , where the action of the gauge field turns it back into the original mode. Propagation in the  $s$ -direction is therefore dominated by the modes of  $\mathcal{P}$  leading to

$$q_1 = \max\{\exp(-\alpha(p))\}. \quad (1.6)$$

For  $M=1$  this reduces to Eq. (1.5). If we would momentarily regard the fifth direction as an imaginary-time direction, the above is recognized as the familiar result that propagation is always dominated by the lightest excitation in any given channel. The domain-wall case is particularly simple in that the gauge field is independent of the  $s$ -coordinate.

Under certain conditions (basically that the coupling constant is not too large, see Appendix C 2 for a more detailed discussion) it should be possible to describe the results of numerical simulations too in terms of an effective wave function  $\chi_{\text{eff}}(s) \sim s^{-1-\delta} q_{\text{eff}}^s$ . This means that every quark's wave function is assumed to be the product of a four-dimensional wave function and the universal fifth-coordinate wave function  $\chi_{\text{eff}}(s)$ . The exponential fall-off rate is accounted for by  $q_{\text{eff}}$ . At relatively weak coupling ( $a^{-1} \gtrsim 3$  GeV) there seems to be no problem with the restoration of chiral symmetry, suggesting that  $q_{\text{eff}} < 1$ . For  $a^{-1} \sim 2$  GeV, the rate at which chiral symmetry is restored depends sensitively on the observable. This, as well as other indications, suggest that  $q_{\text{eff}}$  is very close to one, and the restoration of chiral symmetry really follows a power-law behavior. (In Ref. [22] an estimate of  $q_{\text{eff}}$  was given which, however, is unjustified because the power-law correction was ignored.) Then, at  $a^{-1} \sim 1$  GeV the notion of a universal, localized,  $s$ -coordinate wave function breaks down.

Comparing the perturbative results with the numerical data shows that, not surprisingly, the optimal tree-level value  $q_0=0$  completely fails to describe that data. In comparison, the one-loop result  $q_1=0.5$  lies approximately “half-way” between the tree-level value and the close-to-one values of  $q_{\text{eff}}$  which seem to account for the results of simulations. Thus  $q_1$  gives at least some indication of the actual behavior of the system.

In this paper we adopt  $q_1$  as an analytic criterion for the quality of domain-wall actions. In Sec. III we consider new families of domain-wall actions involving beyond-nearest neighbor coupling. We compute the resulting  $q_1$ , and find that values much smaller than 0.5 can be achieved. Finally, in Sec. IV we discuss the relevance of our results to numerical simulations.

Some technicalities of the one-loop calculation are relegated to Appendix A. Higher order corrections are briefly discussed in Appendix B. An expanded discussion of some non-perturbative issues can be found in Appendix C.

## II. THE ONE-LOOP WAVE FUNCTION

In this section we calculate the one-loop wave function of domain-wall quarks, relegating some of the technicalities to Appendix A. The finite- $N_s$  tree-level propagator and the one-loop self energy for domain-wall fermions were calculated in Refs. [13,7] (see also Ref. [8]). Here we will be interested in the range  $1 \ll s \ll N_s$ , therefore we can use the simpler expressions for the tree-level propagator in the limit  $N_s \rightarrow \infty$  [5]. [The neglected terms are  $O(e^{-\alpha N_s})$ .] Assuming the right-handed quark is localized near the  $s=0$  boundary, the singular part of the tree-level propagator  $G_{s,t}^0(p)$  is

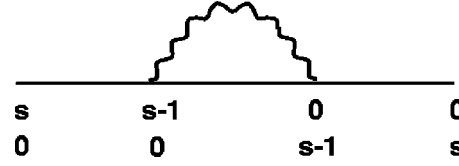


FIG. 1. The “setting-sun” diagram. The external lines are not amputated. The fifth coordinate of each point, as indicated on the first row below the diagram, corresponds to the case where  $1/\not{p}$  [Eq. (2.2)] occurs on the rightmost line, and the leftmost point is far off the boundary. This yields Eq. (2.4). The second row corresponds to Eq. (2.6).

$$P_+ \frac{M(2-M)(1-M)^{s+t}}{i\not{p}}. \quad (2.1)$$

For  $M \rightarrow 1$  this becomes

$$P_+ \frac{\delta_{s,0}\delta_{t,0}}{i\not{p}}, \quad (2.2)$$

which means that the massless right-handed fermion field is fully localized on the boundary layer.

The first quantum effect beyond the free theory is an additive correction,  $-\delta M$ , to the five-dimensional mass  $M$ . It arises in a mean-field approximation, or in perturbation theory from tadpole diagrams. It is well known that this effect must be treated non-perturbatively [23]. We thus discard the tadpole diagrams, absorbing them into the tree-level action via the replacement  $M \rightarrow M - \delta M$ .

The full one-loop wave function is

$$\chi_1(s) = |1 + \delta M - M|^s + \delta\chi_1(s), \quad (2.3)$$

where  $\delta\chi_1(s)$  comes from the “setting sun” diagram only. We will find  $\delta\chi_1(s)$  by matching the non-analytic piece of the dressed propagators with the RHS of Eq. (1.1). We are interested in the behavior of  $\delta\chi_1(s)$  when  $M$  is close to its optimal mean-field value (see also Sec. III). In the calculation below we thus set  $M=1+\delta M$  in tadpole-improved perturbation theory. Since the tadpole-improved free propagator is a function of  $M-\delta M$  (and not of  $M$  and  $\delta M$  separately) the resulting propagator is identical to the ordinary propagator with  $M=1$ . The setting-sun diagram will therefore be computed using the expression for the ordinary tree-level propagator for  $M=1$ .

The setting-sun diagram is depicted in Fig. 1. Notice that we have not amputated the external legs [24]. Except for  $s=t=0$  [see Eq. (2.2)], the tree-level propagator is not singular at  $p=0$ . To obtain a contribution to the right-hand side (RHS) of Eq. (1.1), at least one of the three fermion lines must coincide with expression (2.2). [The kinetic self-energy correction  $\Sigma_K$  in Eq. (1.1) arises when all three lines coincide with Eq. (2.2), see Ref. [7], and will not be discussed here any further.] Assume first that the rightmost propagator in Fig. 1 coincides with Eq. (2.2). Since the leftmost coordinate  $s$  is by assumption far from the boundary, we may take the limit  $p \rightarrow 0$  in the expressions for the self-energy part and for

the leftmost propagator. We thus arrive at a contribution to the dressed propagator, Eq. (1.1), given by  $P_+ \delta\chi_1(s) \times (i\hat{p})^{-1} \delta_{s',0}$  where

$$\delta\chi_1(s) = g^2 C_2 \sum_{s' \geq 0} G_{s,s'}^{0+} \Sigma_{s',0}^+ = g^2 C_2 \Sigma_{s-1,0}^+, \quad (2.4)$$

$C_2$  is the quadratic Casimir,

$$G_{s,t}^{0\pm} = \frac{1}{2} \text{tr} P_{\pm} G_{s,t}^0 (p=0), \quad \Sigma_{s,t}^{\pm} = \frac{1}{2} \text{tr} P_{\pm} \Sigma_{s,t} (p=0), \quad (2.5)$$

and  $\Sigma_{s,t}(p)$  is the 1PI self-energy obtained by amputating the external legs in Fig. 1. In the second equality of Eq. (2.4) we used the explicit expression for  $G_{s,t}^{0+}$  far from the boundary (see Appendix A).

When substituting Eq. (2.3) into Eq. (1.1) we find another term,  $\delta\chi_1(s')$ . This term is obtained when the leftmost propagator in Fig. 1 coincides with Eq. (2.2). Following similar steps we now find

$$\delta\chi_1(s') = g^2 C_2 \Sigma_{0,s'-1}^-. \quad (2.6)$$

Thanks to a ‘‘parity’’ symmetry (see Appendix A) one has  $\Sigma_{s,t}^+ = \Sigma_{t,s}^-$ . Hence Eqs. (2.4) and (2.6) agree.

It remains to compute the diagonal part of the self-energy. One can write

$$\Sigma_{s,0}^+ = \int_{-\pi}^{+\pi} \frac{d^4 k}{(2\pi)^4} h^+(k) \exp(-s\alpha(k)). \quad (2.7)$$

(Note that the external momentum is zero.) The  $s$ -dependence enters through the exponential. All other factors were lumped into  $h^+(k)$  (see Appendix A for more details). For large  $s$ , the above integral can be computed using a saddle-point approximation. As mentioned in the Introduction, the global maximum of  $\exp(-\alpha)$  corresponds to the lattice momentum  $p_{\pi}$  and its three permutations. In the computation we take  $h^+ = h^+(p_{\pi})$  outside the integral, and expand the exponent to second order around  $p_{\pi}$  where we define  $k = (\pi + k_{\parallel}, \vec{k}_{\perp})$ . Including a factor of four to account for the degeneracy of the global maximum we obtain

$$\begin{aligned} \Sigma_{s,0}^+ &= -\frac{3}{2} \left( \frac{1}{2} \right)^s \int \frac{dk_{\parallel} d^3 k_{\perp}}{(2\pi)^4} \exp \left( -\frac{s}{24} (7k_{\perp}^2 + k_{\parallel}^2) \right) \\ &= -\frac{54}{\pi^2 7^{3/2} s^2 2^s}. \end{aligned} \quad (2.8)$$

For the fundamental representation of SU(3) one has  $C_2 = 4/3$ . Substituting in Eq. (2.4) we finally find

$$\delta\chi_1(s) = -g^2 \frac{72}{\pi^2 7^{3/2} (s-1)^2 2^{s-1}} \approx -g^2 \frac{0.788}{s^2} \left( \frac{1}{2} \right)^s, \quad (2.9)$$

provided  $s \gg 1$ . Extrapolating Eq. (2.9) to smaller values of  $s$  suggests that as soon as (or shortly after) we move off the

boundary layer,  $\delta\chi_1(s)$  dominates over the tree-level term in Eq. (2.3). This is true even if  $g^2$  is small [or had the prefactor in Eq. (2.9) been numerically small]. The reason is that the relative magnitude of the two terms is proportional to  $[0.5/(1 + \delta M - M)]^s$ , and since  $1 + \delta M - M \ll 1$  this grows exponentially fast.

### III. NEW ACTIONS

We have found that the  $s$ -coordinate’s wave function of domain-wall quarks is dominated by quantum effects. The arguments of Sec. II show that the broadening of the wavefunction is controlled, in the one-loop approximation, by the maximum of  $\exp(-\alpha(p))$  over the Brillouin zone. This remains true for other domain-wall actions unless the  $s$ -couplings are drastically changed. So, if a different domain-wall action yields a smaller  $\max\{\exp(-\alpha)\}$ , namely a faster fall-off of the wave function at the one-loop level, it is plausible that that new action also performs better non-perturbatively (we return to this issue in Sec. IV).

The standard domain-wall action contains two parameters, the domain-wall height  $M$  and the Wilson parameter  $r$  (which is usually set equal to one). The  $M$ -dependence of  $\max\{\exp(-\alpha)\}$  was investigated in Ref. [25]. As mentioned in Sec. II, however, the additive renormalization of  $M$  must be treated non-perturbatively. The optimal value used in simulations ( $M \sim 1.8$ ) is nicely consistent with mean-field estimates. We will thus assume that the numerical optimization of  $M$  corresponds to setting  $M = 1 + \delta M$  in tadpole-improved perturbation theory. Again, this means that we should determine  $\max\{\exp(-\alpha)\}$  using the tree-level action with  $M = 1$ . As for the Wilson parameter, changing its value in the standard domain-wall action turns out to have little effect (see below).

We will depart from the standard domain-wall action by allowing for couplings not only between nearest neighbors. In view of the obvious increase in computer time needed for the inversion of the fermion matrix, we try to be as economic as possible in our beyond-nearest neighbor excursion. We allow only for coupling between sites  $x$  and  $x + n\hat{\mu}$  (but not e.g. for coupling between  $x$  and  $x + \hat{\mu} + \hat{\nu}$  for  $\mu \neq \nu$ ). In this paper we consider explicitly  $n=2$  and  $n=3$ , namely next-nearest and next-next-nearest couplings in the same direction. Also the modifications will be restricted to the four-dimensional part of the action, leaving the coupling in the fifth direction intact. (Note that we are interested in achieving a fast fall-off in the  $s$ -direction; any attempt to generate a smoother, continuum-like, behavior in the  $s$ -direction is thus the exact opposite of what we are aiming for.)

The domain-wall operators considered here will have the following general form for zero quark mass

$$D_{s,t}^{\text{d.w.}} = \delta_{s,t} D + (\delta_{s+1,t} - \delta_{s,t}) P_+ + (\delta_{s-1,t} - \delta_{s,t}) P_-, \quad (3.1)$$

with the understanding that  $0 \leq s, t \leq N_s - 1$  on a finite lattice. The inclusion of a quark mass can be done in the usual way [5]. The four-dimensional part of the action is



$$D(p) = i \sum_{\mu} \gamma_{\mu} f(p_{\mu}) - rW(p) + M. \quad (3.2)$$

This equation gives the tree-level operator in momentum space. The generalized Wilson term  $W(p)$  is a function of  $\cos(p_{\mu})$ . In the kinetic term,  $f(p_{\mu})$  is an odd function of its argument, which we take to be  $\sin(p_{\mu})$  times a polynomial in  $1 - \cos(p_{\mu})$ . Later we will give explicit expressions for  $W(p)$  and  $f(p_{\mu})$ . As explained earlier we set  $M=1$  in the tree-level action, but we will use the freedom in varying the Wilson parameter  $r$ . Our convention is that  $W(p)$  and  $r$  are both positive.

For any domain-wall action of the above form, the exponents  $\alpha(p)$  are determined by

$$2 \cosh(\alpha(p)) = \frac{1 + B^2(p) + \sum_{\mu} f^2(p_{\mu})}{B(p)}, \quad (3.3)$$

where  $B(p) = 1 - M + rW(p)$  and  $\alpha(p) \geq 0$  by convention. [ $B(p) = rW(p)$  for  $M=1$ ; for the standard domain-wall action Eq. (3.3) reduces to eq. (A8).] Lowering the global maximum of  $\exp(-\alpha(p))$  corresponds to raising the global minimum of Eq. (3.3).

Some insight about the features that control  $\min\{2\cosh(\alpha)\}$  can be obtained from very general considerations. One has  $W=0$  for  $p=0$ , and in all cases one aims for  $rW > 1$  at  $p_{\pi} = (\pi, 0, 0, 0)$ . As we gradually increase  $p_1$  from 0 to  $\pi$  (keeping  $p_2 = p_3 = p_4 = 0$ ) at some value  $p_c = (p_{1c}, 0, 0, 0)$  we will have  $rW = 1$ . Were it not for the  $f(p_1)$  term in Eq. (3.3), at  $p_c$  we would obtain  $\cosh(\alpha) = \exp(-\alpha) = 1$ , namely no exponential suppression at all. To avoid this dangerous situation, we would like to have  $f^2(p_1)$  as large as possible at  $p_1 = p_{1c}$ .

Another danger lurks at the (fifteen non-zero) corners of the Brillouin zone. There, by construction,  $f(p_{\mu}) = 0$ , and so  $\exp(-\alpha) = (rW)^{-1}$ . We will therefore also be interested in increasing  $rW$  at the corners of the Brillouin zone.

As a warm-up exercise let us consider the effect of varying  $r$  in the standard domain-wall action. In this case  $W = \sum_{\mu} (1 - \cos(p_{\mu}))$  and  $f(p_{\mu}) = \sin(p_{\mu})$ . For  $r \leq 1$ ,  $\min\{2\cosh(\alpha)\}$  occurs at  $p_{\pi}$ . We can increase  $rW$  at  $p_{\pi}$  by increasing  $r$ . But in that case the value  $rW = 1$  will occur at a smaller  $p_1$ , where  $\sin(p_1)$  is smaller. As can be seen from Fig. 2(a) there is a transition region around  $r \sim 1.2$ . For larger values of  $r$ ,  $\min\{2\cosh(\alpha)\}$  moves towards the point where  $rW = 1$ . As an example, for  $r = 2.0$  one has  $rW = 1$  at  $x = \cos(p_1) = 0.5$ , and  $\min\{2\cosh(\alpha)\}$  is at  $x \sim 0.42$ . The largest value of  $\min\{2\cosh(\alpha)\}$ , obtained for  $r \sim 1.3$ – $1.5$ , is around 2.8. This makes little improvement over the value 2.5 obtained at  $r = 1$ .

To avoid this dead-lock we take  $W$  to be a *non-linear* function of  $\cos(p_{\mu})$ . We define

$$W_n = \sum_{\mu} (1 - \cos(p_{\mu}))^n. \quad (3.4)$$

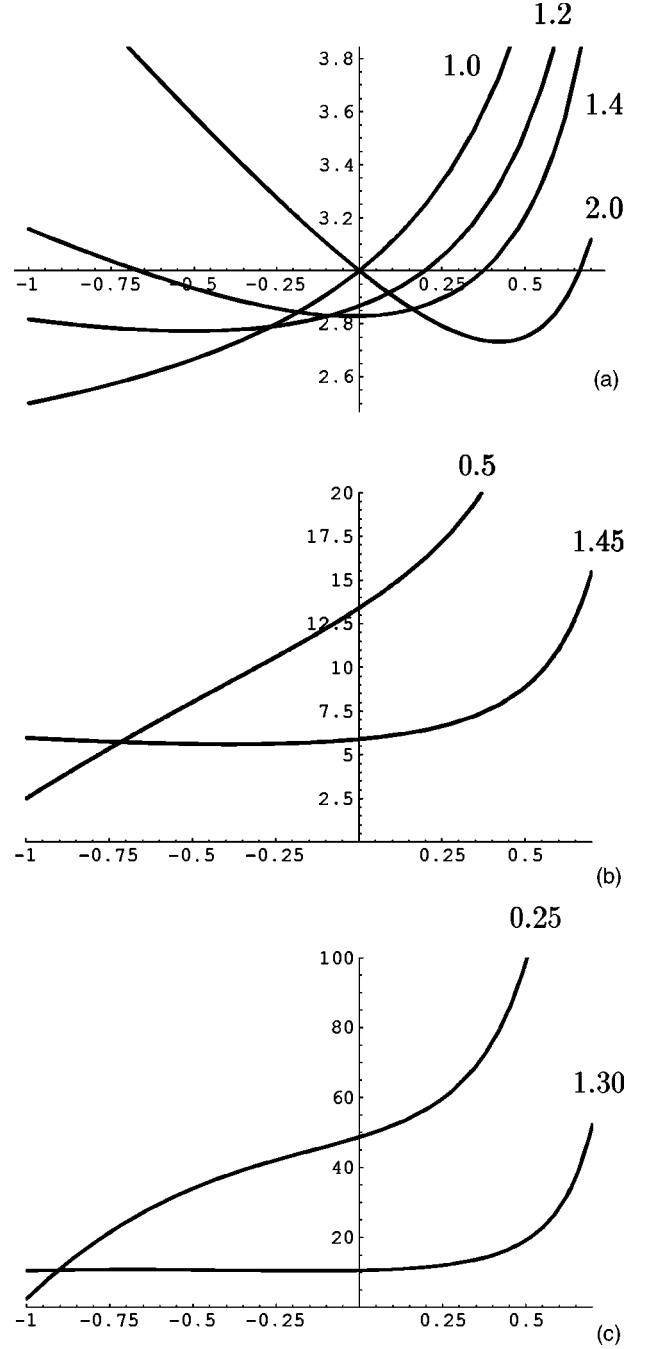


FIG. 2. (a)  $2\cosh(\alpha)$  as a function of  $x = \cos(p_1)$  and  $p_2 = p_3 = p_4 = 0$  for the standard domain-wall action. Results are shown for four values of the Wilson parameter  $r$ . In all the figures of this section, the minimum seen in the plot is also the global minimum over the entire Brillouin zone. [At larger values of  $x$ ,  $2\cosh(\alpha)$  is monotonically increasing, and eventually it diverges for  $x \rightarrow 1$ .] (b) Same as (a) for the action  $D_{23}$  with  $c_3 = 4/3$  for two values of  $r$ . The optimal value is  $r_{\text{opt}} = 1.45$ , whereas  $r_2 = 0.5$  is a “reference value” defined by the condition  $r_n W_n = 2$  at  $x = \cos(\pi) = -1$  (see Sec. IV). Notice the difference in vertical scales compared to (a). (c) Same as (a) for the action  $D_{35}$  with  $c_5 = 2$  for two values of  $r$ . The optimal value is  $r_{\text{opt}} = 1.30$ , whereas the “reference value” is  $r_3 = 0.25$ . Notice the flatness of  $\cosh(\alpha)$  for  $-1 \leq x \leq 0.25$ .

TABLE I.  $\max\{\exp(-\alpha)\}$  for the action  $D_{23}$  at  $M=1$  and for various values of  $c_3$  (see text for the definitions). The second column gives the first two terms in the expansion of  $f_3(p)$ . For each  $c_3$  we show the result for  $r=r_{\text{opt}}$  where  $\max\{\exp(-\alpha)\}$  is smallest.

$c_3$	$f_3(p)$	$r_{\text{opt}}$	$\min\{2\cosh(\alpha)\}$	$\max\{\exp(-\alpha)\}$
0	$p - \frac{1}{6}p^3$	1.46	2.83	0.414
1/3	$p$	1.14	3.40	0.326
2/3	$p + \frac{1}{6}p^3$	1.19	4.09	0.261
4/3	$p + \frac{1}{2}p^3$	1.45	5.62	0.184
7/3	$p + p^3$	1.98	8.06	0.126

$W_n$  requires couplings of sites  $x$  and  $x+n\hat{\mu}$ . Once such coupling have been introduced into the generalized Wilson term, we allow them also in the kinetic term. Further raising of the global minimum of  $\cosh(\alpha)$  will be made possible by choosing  $f(p_\mu)$  that increases faster than  $\sin(p_\mu)$ , and by adjusting the Wilson parameter  $r$ .

We now turn to the investigation of concrete actions. The minimization problem was solved numerically. Note that the RHS of Eq. (3.3) can be expressed as a function of  $x_\mu \equiv \cos(p_\mu)$  only. Using the invariance under permutations of the four components, it is enough to look for the global minimum over the range  $-1 \leq x_1 \leq x_2 \leq x_3 \leq x_4 \leq 1$ . [One can also study the minimization problem analytically. For any  $\mu$ ,  $\sin(p_\mu)=0$  always satisfies the extremality condition. For all momentum-components where  $\sin(p_\mu) \neq 0$  one finds a coupled algebraic equation in  $x_\mu$ . In all the cases we have studied, it turned out that the global minimum was either of the form  $(p_{\min}, 0, 0, 0)$  or else of the form  $(p_{\min}, p_{\min}, p_{\min}, p_{\min})$ .]

We first consider an action containing next-nearest neighbors:

$$D_{23} = i \sum_{\mu} \gamma_{\mu} f_3(p_{\mu}) - r W_2 + M, \quad (3.5)$$

$$f_3(p_{\mu}) = \sin(p_{\mu}) [1 + c_3(1 - \cos(p_{\mu}))]. \quad (3.6)$$

$W_2$  is defined in Eq. (3.4). In Table I we show the resulting values of  $\min\{2\cosh(\alpha)\}$  and  $\max\{\exp(-\alpha)\}$  for several values of  $c_3$ . For each  $c_3$  we looked for the best value of  $r$  which we denote  $r_{\text{opt}}$ . One sees that values of  $\max\{\exp(-\alpha)\}$  much smaller than 0.5 are feasible. A plot of  $2\cosh(\alpha)$  for  $c_3=4/3$  is shown in Fig. 2(b). Notice the flatness of  $\cosh(\alpha)$  for  $-1 \leq \cos(p_1) \leq 0$  at  $r=r_{\text{opt}}=1.45$ .

We now digress to discuss how the present work relates to the standard ‘‘improvement program’’ (see e.g. the review [26]). In the study of the hadron spectrum, only a single parameter (the bare quark mass) in the fermion action needs to be tuned. Once the correct continuum limit has been established, attention is focused on eliminating those lattice artifacts that vanish most slowly, that is, linearly with the

lattice spacing. However, in the calculation of weak matrix elements one has to first establish the correct continuum limit. This is very problematic with Wilson or staggered fermions because, due to the loss of full chiral and/or flavor symmetry, many subtraction coefficients must be tuned. Controlling those subtractions by having good chiral and flavor properties simultaneously is thus of higher priority than the removal of any other lattice error. Furthermore, in the massless-quark limit  $O(a)$  lattice artifacts are automatically excluded if chiral symmetry is maintained [14]. In that sense, approaching the chiral limit using domain-wall fermions encompasses the standard improvement program as well.

Coming back to the new domain-wall action, since the new Wilson term  $W_2$  starts off at order  $p^4$ , the first lattice deviation from a relativistic (tree-level) dispersion relation comes only from the kinetic term. This is shown in the second column of Table I. We observe that while increasing  $c_3$  from zero to 1/3 improves the dispersion relation, the opposite is true for  $c_3 > 1/3$ . Although the error is formally of order  $a^2$ , it might become significant if  $c_3$  is too large. To gain some idea on the magnitude of the error consider, say,  $p^2 \sim (400 \text{ MeV})^2$ , which is relevant for kaon physics, on a lattice with  $a^{-1} \sim 2 \text{ GeV}$ . This means  $a^2 p^2 \sim 1/25$ . For the last two rows of Table I, the effect is 2% and 4% respectively.

If next-next-nearest neighbors in the same direction are also allowed one can decrease  $\max\{\exp(-\alpha)\}$  further while maintaining a vanishing  $p^3$  term. Let

$$D_{35} = i \sum_{\mu} \gamma_{\mu} f_5(p_{\mu}) - r W_3 + M, \quad (3.7)$$

where again  $W_n$  is defined in Eq. (3.4) and where

$$f_5(p_{\mu}) = \sin(p_{\mu}) \left[ 1 + \frac{1}{3}(1 - \cos(p_{\mu})) + c_5(1 - \cos(p_{\mu}))^2 \right]. \quad (3.8)$$

Some values of  $\min\{2\cosh(\alpha)\}$  and  $\max\{\exp(-\alpha)\}$  are shown in Table II. A plot of  $2\cosh(\alpha)$  for  $c_5=2$  is shown in Fig. 2(c). Even for the last row in Table II ( $c_5=50$ ), the deviation from Lorentz covariance is at the level of  $(50/4)(a^2 p^2)^2 \sim 2\%$  for  $p^2 \sim (400 \text{ MeV})^2$ . For  $c_5=5$ , the deviation is below 2% up to  $(700 \text{ MeV})^2$ , and so on.

We conclude with a number of technical comments. Replacing the four-dimensional part of the domain-wall action by  $D_{23}$  ( $D_{35}$ ) approximately doubles (triples) the number of entries in the fermion matrix. Therefore one should expect a corresponding increase in the cost of a single inversion of the fermion matrix at fixed  $N_s$ .

In the continuum limit, both the standard domain-wall action and the new actions discussed above support a single quark (one Weyl field on each boundary) for  $|1-M| < 1$ . In the case of the standard action there is a four-quark zone (corresponding to the corner  $p_{\pi}$  and its permutations) for  $|3-M| < 1$ . When the Wilson term  $W_n$  is employed instead, the four-quark zone is at  $|1+r2^n-M| < 1$ . An additional benefit of the new actions is that the four-quark and the single-quark zones are separated by a large gap (as a function

TABLE II.  $\max\{\exp(-\alpha)\}$  for the action  $D_{35}$  for various values of  $c_5$  at  $M=1$  and  $r=r_{\text{opt}}$ . The second column gives the first two terms in the expansion of  $f_5(p)$ .

$c_5$	$f_5(p)$	$r_{\text{opt}}$	$\min\{2\cosh(\alpha)\}$	$\max\{\exp(-\alpha)\}$
1	$p + \frac{13}{60}p^5$	0.87	6.85	0.149
2	$p + \frac{14}{30}p^5$	1.30	10.50	0.096
3	$p + \frac{43}{60}p^5$	1.66	13.36	0.075
4	$p + \frac{29}{30}p^5$	1.96	15.74	0.064
5	$p + \frac{13}{60}p^5$	2.21	17.74	0.057
10	$p + 2\frac{14}{30}p^5$	3.18	25.44	0.039
30	$p + 7\frac{7}{15}p^5$	5.70	45.51	0.022
50	$p + 12\frac{7}{15}p^5$	7.48	59.84	0.017

of  $M$ ) for  $r \sim r_{\text{opt}}$ . We expect that a clear gap should be found in simulations too, even though its precise location will likely be different from the weak-coupling limit.

For the standard domain-wall action, the optimal value of  $M$  used in simulations agrees well with the mean-field estimate of  $1 + \delta M$ . One obtains  $\delta M$  by substituting a mean value  $u$  for each link variable in the Wilson term. We will assume that the new actions are gauged in the simplest way, namely using only products of link variables along straight lines. (E.g., the sites  $x$  and  $x + 2\hat{\mu}$  are connected via  $U_{x,\mu}U_{x+\hat{\mu},\mu}$  etc.) Using a mean link  $u \sim 0.8$  at  $\beta = 6.0$  (see e.g. Ref. [17]), the mean-field estimate is  $2r(3 - 4u + u^2) \sim 0.9r$  for  $D_{23}$  and  $r(10 - 15u + 6u^2 - u^3) \sim 1.3r$  for  $D_{35}$ .

Last, for any domain-wall operator with the form of Eqs. (3.1) and (3.2), the  $N_s \rightarrow \infty$  limit defines an overlap-Dirac operator [27,4] obeying the Ginsparg-Wilson relation (for a review see Ref. [28]) given by

$$D_{\text{GW}} = 1 - \gamma_5 \epsilon(\gamma_5 D), \quad (3.9)$$

where  $\epsilon(x) = \pm 1$  is the sign function acting on each of the eigenvalues of  $\gamma_5 D$ .

#### IV. DISCUSSION

In this paper we showed that the one-loop wave function of domain-wall fermions behaves like  $s^{-2}q_1^s$ , where  $q_1$  is determined by the free-fermion action. This is true for a wide class of domain-wall actions, including those defined in Eq. (3.1). For the standard action  $q_1 = 0.5$  provided  $M$  is chosen optimally, whereas the addition of beyond-nearest neighbor couplings allows for much smaller values of  $q_1$ .

Only in the weak-coupling limit does  $q_1$  fully control the

wave function. For finite  $g^2$  up to some critical value  $g_c^2$  we expect the wave function to be proportional to some  $q_{\text{pt}}^s$  (up to power corrections) with  $q_{\text{pt}} = q_{\text{pt}}(g^2)$ . The  $g^2$  dependence can be parametrized in various ways. In Appendix B we consider the role of higher-order diagrams, and the parametrization

$$q_{\text{pt}}^s = q_1^s \exp(s(g^2 \eta_1 + g^4 \eta_2 + \dots)), \quad (4.1)$$

is found to be natural. The Taylor expansion of  $\exp(sg^2 \eta_1)$  corresponds to a family of 1PI diagrams of all orders, where the first  $\eta_1$ -dependent terms are two-loop diagrams. (Analogous statements apply to  $\eta_2$  etc.) In terms of  $q_{\text{pt}}(g^2)$ , one can define  $g_c^2$  by the condition  $q_{\text{pt}}(g_c^2) = 1$ . The existing numerical results suggest that the (quenched) value of  $6/g_c^2$  is very close to 6.0 for the standard domain-wall action.

If both  $q_1$  and  $\eta_1$  were known for a given action, one could obtain a crude estimate of  $g_c^2$  via a linear extrapolation. We have computed only the  $q_1$  values, so we can only conjecture what trends are likely to affect  $g_c^2$ . First, in the lower rows in Table II,  $q_1$  is extremely small. Nevertheless, if  $\eta_1$  is large (and positive),  $g_c^2$  may end up being approximately the same as (or, for that matter, even smaller than) for the standard domain-wall action. Because of gauge invariance there are vertices that depend linearly on  $c_5$ . Since the lower rows in Table II come from actions with numerically large values of  $c_5$  this should, indeed, lead generically to a large  $\eta_1$  (and  $\eta_2$  and so on).

It is therefore safer to focus on the first few rows in Tables I and II, where one is less prone to the above risk. The following heuristic argument suggests that, in that range of parameters, the new domain-wall actions may indeed be superior to the standard one. When the Wilson parameter is equal to  $r_n \equiv 2^{1-n}$ , an action containing the Wilson term  $W_n$  [Eq. (3.4)] gives rise to  $\min\{2\cosh(\alpha)\} = 2.5$  (corresponding to  $q_1 = 0.5$ ) at  $p_\pi = (\pi, 0, 0)$  for  $M = 1$ . The behavior at  $r = r_n$  is therefore a common starting point over which we may try to improve by increasing  $r$ . As discussed in Sec. III, for the standard action  $\min\{2\cosh(\alpha)\}$  is relatively insensitive to  $r$ . Its largest value (which is 2.8) is obtained around  $r_{\text{opt}} \sim 1.3 - 1.5$ , namely  $r_{\text{opt}}$  is less than 50% above  $r_1$ . In comparison, for  $D_{23}$  at  $c_3 = 4/3$  (Fig. 2(b)) the largest value of  $\min\{2\cosh(\alpha)\}$  is achieved at  $r_{\text{opt}} = 1.45$  which is approximately three times  $r_2$ . For  $D_{35}$  at  $c_5 = 2$  (Fig. 2(c)) the best value is  $r_{\text{opt}} = 1.3$  which is more than five times  $r_3$ .

The ability to reach larger values of  $\min\{2\cosh(\alpha)\}$  is thus correlated with an enhanced sensitivity to the Wilson parameter, and with a bigger ratio  $r_{\text{opt}}/r_n$ . Now, while  $\eta_1, \eta_2, \dots$ , might in principle grow as  $r$  increases from  $r_n$  to  $r_{\text{opt}}$ , it is clear that as functions of the parameters of the theory their behavior will be very different from  $q_1$  (see Appendix B). Therefore it is plausible that there exist “windows” of parameters where  $q_{\text{pt}}(g^2)$  is controlled primarily by the decreasing  $q_1$ , implying that the exponential suppression holds up to a larger value of  $g^2$ .

In Appendix C 2 we discuss how different ways of approaching the chiral limit are related to different forms of the spectral-density function of the (normal-ordered) transfer

matrix. At weak coupling one expects to have a gap, namely almost all eigenvalues are smaller than some  $\lambda_0 < 1$ . The gap region  $\lambda_0 \leq \lambda \leq 1$  is not completely devoid of eigenvalues, but their total number is drastically smaller than just below  $\lambda_0$ . One also expects a big difference between the corresponding eigenfunctions. Those that lie outside the gap should be continuum-like modes that spread all over the lattice, while inside the gap the modes should be highly localized [29,9].

As the coupling constant increases a qualitative change takes place. Near-unity eigenvalues of the transfer matrix proliferate. For the standard domain-wall action this is a direct consequence [3,11] of the proliferation of approximate zero modes of the hermitian, four-dimensional, Wilson-Dirac operator [29,12,9]. The change (which seems to take place around quenched  $\beta=6.0$ ) shows the key features of the phenomenon known in condensed matter as localization [30]. Due to the randomness of generic gauge-field configurations, in any given part of the lattice there is a finite probability to find a localized (approximate) zero mode.

Viewing the (Hermitian) Wilson-Dirac operator as a Hamiltonian, under its action the fermions can hop only a single site. But with the new domain-wall actions the relevant hamiltonian is  $\gamma_5 D$  [see Eq. (3.2)]. Now the fermions may hop also two (or three) sites when the Hamiltonian acts just once on a given state. It should be more difficult to trap the new fermions inside a small potential well, as now they have more ways of escaping out of it. This consideration too suggests that the critical coupling, where the exponential suppression is lost, may be larger for the new actions.

The new domain-wall actions considered in this paper carry with them an obvious extra cost for a single inversion of the fermion matrix. One may hope to reduce chiral symmetry violations also by using improved gauge actions, because the latter tend to generate smoother configurations. If this goal is achieved, it may be a numerically much cheaper way to reduce chiral symmetry violations. Using the Iwasaki action [31] it was found that the residual large- $N_s$  pion-mass squared (extrapolated to zero quark mass) drops by about a factor of two for  $a^{-1} \sim 1$  GeV [19]. However, the new residual pion mass is still very big. Also, in thermodynamics, the Iwasaki action did not lead to any noticeable reduction in the residual pion mass (for a detailed discussion of various improvements see the first paper of Ref. [19]).

Our analysis suggests a possible explanation why the use of improved gauge actions has had only a limited success. If the (tree-level) domain-wall action is unchanged, the one-loop wave function (1.4) is still controlled by the same value of  $q_1 = 0.5$ . Only the numerical prefactor may change [cf. Eq. (2.9)]. Hence, in this approximation, the exponential fall-off rate is not getting any better for improved gauge actions. [The same reasoning applies to the use of “fat links.” One has to be careful, however, because this argument ignores higher-order corrections, cf. Eq. (4.1).]

In conclusion, in the one-loop approximation even a small reduction in  $q_1$  leads to a dramatic suppression of chiral symmetry violations for commonly used values of  $N_s$ . If the actual quark’s wave function is (even partly) correlated with

the one-loop one, the new actions could give rise to a significantly better chiral behavior, enough to justify their increased simulation cost.

## ACKNOWLEDGMENTS

The key results of this paper were worked out during a visit to Brookhaven National Laboratory. Extensive discussions with many members of the lattice groups in Brookhaven Lab. and Columbia Univ. were essential for the development of this work. In forming my view on the current state of domain-wall fermion simulations, as presented in this paper, I benefitted from discussions with the participants of the Dubna workshop on “Lattice Fermions and Structure of the Vacuum.” I thank Tom Blum, Maarten Golterman and Karl Jansen for their useful comments on this paper, and Karl Jansen also for discussing with me some as-yet unpublished results. This research is supported in part by the Israel Science Foundation.

## APPENDIX A: SOME TECHNICALITIES

In this appendix we collect a few useful formulas. The inverse tree-level propagator in momentum space is

$$(G^0)_{s,t}^{-1} = i \delta_{s,t} \sum_{\mu} \gamma_{\mu} \sin(p_{\mu}) + W_{s,t}^{+}(p) P_{+} + W_{s,t}^{-}(p) P_{-}, \quad (\text{A1})$$

where

$$W_{s,t}^{+}(p) = \delta_{s+1,t} - B(p) \delta_{s,t}, \quad (\text{A2})$$

$$B(p) = 1 - M + \sum_{\mu} (1 - \cos(p_{\mu})), \quad (\text{A3})$$

and  $W_{s,t}^{-}(p) = W_{t,s}^{+}(p)$ . We consider only the case of a zero quark mass. Also, for the calculation of the self-energy we set  $M=1$  in the tree-level action (see Sec. II). The tree-level propagator was computed in Refs. [5,13,7]. At the corners of the Brillouin zone [ $\sin(p_{\mu})=0$ , all  $\mu$ ] the two chiralities decouple in Eq. (A1). For  $p \rightarrow 0$  the limit is singular. But at the other fifteen corners the limit is regular, leading to

$$G_{s,t}^0 = (W^{+})_{s,t}^{-1}(p) P_{+} + (W^{-})_{s,t}^{-1}(p) P_{-}. \quad (\text{A4})$$

Specifically at  $p = p_{\pi}$  and with a semi-infinite coordinate  $s, t = 0, 1, 2, \dots$  one has

$$(W^{+})_{s,t}^{-1} = (W^{-})_{t,s}^{-1} = -\frac{1}{2} \theta(t \geq s) 2^{s-t}, \quad (\text{A5})$$

where  $\theta(t \geq s) = 1$  for  $t \geq s$  and  $\theta(t \geq s) = 0$  for  $t < s$ .

The explicit expression for the diagonal self-energy  $\Sigma_{s,t}^{+}$  of Eq. (2.5) is given by Eq. (36) of Ref. [7]. With a slight change of notation it reads (recall that the external momentum is zero)



$$\begin{aligned} \Sigma_{s,t}^+ &= \int_{-\pi}^{+\pi} \frac{d^4 k}{(2\pi)^4} \left( 4 \sum_{\nu} \sin^2(k_{\nu}/2) \right)^{-1} \sum_{\mu} \\ &\times \left( \cos^2(k_{\mu}/2) (W^+ G^-)_{s,t} - \sin^2(k_{\mu}/2) (W^- G^+)_{s,t} \right. \\ &\left. + \frac{1}{2} \sin^2(k_{\mu}) (G^+ + G^-)_{s,t} \right), \end{aligned} \quad (\text{A6})$$

where

$$(G^{\pm})^{-1}(p) = \sum_{\mu} \sin^2(p_{\mu}) + W^{\pm}(p) W^{\mp}(p). \quad (\text{A7})$$

Explicit expressions for  $G_{s,t}^{\pm}$  can be found in Refs. [5,13,7]. As explained in Sec. II we are interested in  $\Sigma_{s,0}^+$  for  $s \gg 1$ . In the saddle-point approximation we set the internal momentum on the fermion line to  $p_{\pi}$  (or its permutations) in all terms, except in the exponential  $\exp(-s'\alpha)$  that occurs inside  $G_{s',0}^{\pm}$ , which is expanded to second order around  $p_{\pi}$  using the definition

$$\frac{1 + B^2(p) + \sum_{\mu} \sin^2(p_{\mu})}{2 \cosh(\alpha(p))} = \frac{B(p)}{B(p)}. \quad (\text{A8})$$

This expansion gives rise to the integrand in Eq. (2.8). Since  $\sin(p_{\mu})=0$  for all  $\mu$  at  $p_{\pi}$ , the last term in Eq. (A6) is zero. At  $p_{\pi}$  one has  $W^{\pm} G^{\mp} = (W^{\mp})^{-1}$ . Since the matrices  $W^{\pm}$  and their inverses are triangular, the second term in Eq. (A6) gives zero too. Only the first term in Eq. (A6) contributes, and only for  $\mu=2,3,4$ , leading to eq. (2.8).

We also need the diagonal tree-level propagator  $G^{0+}$  [Eq. (2.5)] away from the boundaries at  $p=0$ . For  $M \rightarrow 1$  and  $p \rightarrow 0$ , the second-order operators  $(W^+ W^-)_{s,t}$  and  $(W^- W^+)_{s,t}$  tend to  $\delta_{s,t}$  [except  $(W^- W^+)_{s,t}$  at  $t=s=0$ ]. As a result  $G_{s,t}^{0+} = \delta_{s-1,t} = W_{s,t}^-$ . (For a semi-infinite  $s$ -coordinate  $W^-$  is a right-inverse of  $W^+$ .)

Finally we observe that the interacting domain-wall action has a generalized parity symmetry. Writing the action as

$$\sum_{s,s';\vec{x},\vec{y};x_4,y_4} \bar{\psi}_{s,\vec{x},x_4} D_{s,s';\vec{x},\vec{y};x_4,y_4}(U) \psi_{s',\vec{y},y_4} \quad (\text{A9})$$

one has

$$D_{s,s';\vec{x},\vec{y};x_4,y_4}(U) = \gamma_4 D_{s',s;\vec{x},-\vec{y};x_4,y_4}(U') \gamma_4, \quad (\text{A10})$$

where

$$\begin{aligned} U'_{x,x_4;4} &= U_{-\vec{x},x_4;4} \\ U'_{x,x_4;k} &= U_{-\vec{x}-\hat{k},x_4;k}, \quad k=1,2,3. \end{aligned} \quad (\text{A11})$$

This looks like an ordinary parity transformation, except that we have switched the fifth coordinates of  $\psi$  and  $\bar{\psi}$  ( $s$  and  $s'$ ). The above discrete symmetry implies

$$\Sigma_{s,t}(\vec{p}, p_4) = \gamma_4 \Sigma_{t,s}(-\vec{p}, p_4) \gamma_4 \quad (\text{A12})$$

and when  $\sin(p_{\mu})=0$  for all  $\mu$ , one has  $\text{tr } P_+ \Sigma_{s,t} = \text{tr } P_- \Sigma_{t,s}$ .

## APPENDIX B: BEYOND ONE LOOP

In this appendix we consider the role of higher-order corrections. Specifically the aim is to show how an  $O(g^2)$  correction to  $q_1$  is built. The new wave function is conveniently parametrized as

$$\delta\chi_{\text{pt}}(s) \propto s^{-2} q_{\text{pt}}^s,$$

$$q_{\text{pt}}^s = (q_1 \exp(g^2 \eta_1))^s = q_1^s \left( 1 + g^2 \eta_1 s + \frac{1}{2} (g^2 \eta_1)^2 s^2 + \dots \right). \quad (\text{B1})$$

The last expression suggests that, in the wave function  $\delta\chi_{\text{pt}}(s)$ , the  $O(g^2)$  correction to  $q_1$  arises from a resummation of perturbation theory. If true, at any finite order we should find contributions to the wave function whose structure is  $\delta\chi_1(s)$  times an increasing power of  $s$ . [We likewise expect additional  $O(g^4)$  etc. corrections to  $q_1$ , cf. Eq. (4.1); the arguments below are, however, too crude to tell how the power-law part of the wave function depends on  $g^2$ .]

We believe that a term  $g^2 \eta s \delta\chi_1(s)$  arises from the two-loop diagrams in Figs. 3(a) and 3(b) (and not from Figs. 3(c) or 3(d)). Following Sec. II, we will consider for definiteness the case where the rightmost fermion line in Fig. 3(a) corresponds to the singular part of the tree-level propagator, Eq. (2.2), and therefore the fifth coordinate of the rightmost vertex is zero. In Fig. 3(a) the momenta on all three internal fermion lines can be simultaneously equal (or close) to  $p_{\pi}$  [recall that the external momentum is (close to) zero]. Then, the one-loop exponent  $q_1 = \frac{1}{2}$  comes with a power  $|s-t'| + |t'-t| + t$  which, for given  $s$ , is minimal when the points are ordered:  $s \geq t' \geq t \geq 0$ . (Here we have ignored the difference between  $s$  and  $s-1$  which is negligible for  $s \gg 1$ .) When the points are not ordered we obtain an exponentially convergent series in the excessive length of the fermion's trajectory. For simplicity we will assume that the points are ordered, as we are only interested here in arguing that a contribution proportional to  $g^2 s \delta\chi_1(s)$  exists. (However, the full series has to be summed in order to obtain the numerical value of  $\eta$ .)

Next consider the integration over the loop momentum of the inner loop. As explained in Sec. II, this integration should give rise to a factor of  $|t'-t|^{-2}$ . Since the sum  $\sum_m m^{-2}$  is convergent, as a crude approximation one can say that the points  $t$  and  $t'$  are forced to be close together. Once this extra constraint has been taken into account, the remaining expression is independent of  $t$  (and  $t'$ ). Therefore the  $t$ -summation (approximated by an integral) gives  $\int_0^s dt = s$ . The Gaussian integration over the momentum of the outer loop gives rise roughly to the factor  $s^{-2}$  [associated with  $\delta\chi_1(s)$ ] as before. We conclude that a term proportional to  $g^2 s \delta\chi_1(s)$  may indeed arise from Fig. 3. [A technical complication is that, if the two loop-momenta are both equal to

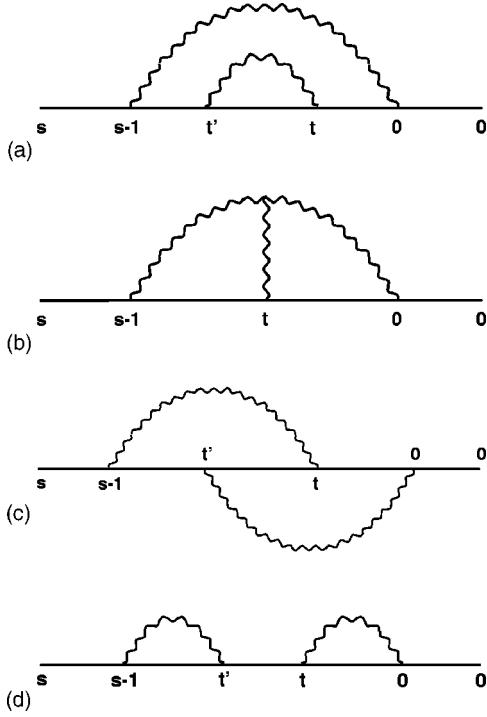


FIG. 3. (a) A two-loop diagram that (likely) contributes to Eq. (B1). In this figure and the following ones, the fifth coordinates correspond to the case where  $1/\not{p}$  occurs on the rightmost line, cf. Eqs. (2.2) and (2.4). (b) Another two-loop diagram that contributes to Eq. (B1). (c) This two-loop diagram does not contribute to Eq. (B1). (d) A reducible two-loop diagram that does not contribute to Eq. (B1).

$p_\pi$ , the momentum flowing through the inner gauge-boson line is zero. The singularity  $(k-p)^{-2}$  is integrable in four dimensions, but its existence makes the actual calculation quite complicated. We believe that the above considerations are robust enough to grasp the dominant behavior of Fig. 3(a).]

The above argument easily generalizes to the higher-loop diagrams of Fig. 4(a). At the  $n$ th-order one has  $n-1$  “inner” loops. The fifth coordinates are pair-wise close, but otherwise are constrained only by ordering. Hence one expects a contribution proportional to

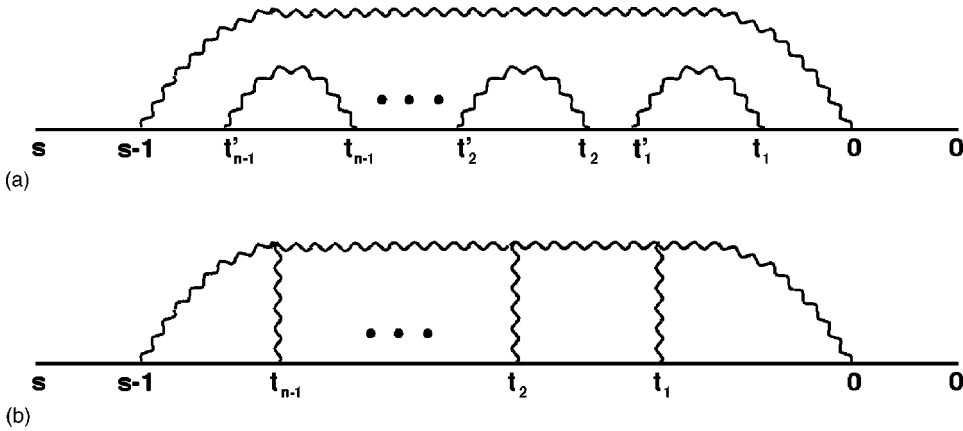


FIG. 4. (a) An  $n$ -loop diagram that (likely) contributes to Eq. (B1). (b) Another  $n$ -loop diagram that contributes to Eq. (B1).

$$\int_0^s dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_{n-2}} dt_{n-1} = \frac{s^{n-1}}{(n-1)!}, \quad (\text{B2})$$

in agreement with Eq. (B1). A similar reasoning implies that the two-loop and  $n$ -loop diagrams of Figs. 3(b) and 4(b) respectively also contribute to the Taylor series in Eq. (B1).

We now want to explain why the diagrams of Figs. 3(c) and 3(d) do not contribute to the RHS of Eq. (B1). Consider first the reducible diagram Fig. 3(d). The momentum on the middle fermion line is zero (being equal to the external momentum). Hence the points  $t$  and  $t'$  are very close (see Sec. II and Appendix A). Ignoring the difference between  $t$  and  $t'$  the two Gaussian integrations give rise to the product  $|s-t|^{-2}t^{-2}$ . The  $t$ -summation is then dominated by  $t$ -values which are either close to zero or to  $s$  (being the fifth coordinates of the rightmost and leftmost vertices). Therefore the result behaves like  $s^{-2}$  (and not like  $s^{-2}s = s^{-1}$  as in the case of Fig. 3(a)).

While the diagram in Fig. 3(c) is irreducible, the momenta on the internal fermion lines cannot be all equal to  $p_\pi$ . For example, if the momenta on the first and third internal lines are equal to  $p_\pi$ , then the momentum on middle line is zero. Hence, again, the points  $t$  and  $t'$  must be very close together, as well as close to either zero or  $s$ , and the final result is proportional to  $s^{-2}$  as in the case of Fig. 3(d).

## APPENDIX C: SOME NON-PERTURBATIVE OBSERVATIONS

### 1. Chiral-symmetry restoration

We outline here the proof of chiral-symmetry restoration given in Ref. [6] in the light of later works (in particular Refs. [9,10]). For definiteness we focus on the anomalous term in the lattice PCAC relation. (Its vanishing implies that the pion mass will be zero after taking the infinite-volume and massless-quark limits in that order.) The PCAC relation using domain-wall fermions reads [6]

$$\Delta_\mu \langle A_{5\mu}^a(x) J_5^b(y) \rangle = 2m_0 \langle J_5^a(x) J_5^b(y) \rangle + 2 \langle J_{5q}^a(x) J_5^b(y) \rangle + \text{contact term}. \quad (\text{C1})$$

Here  $\Delta_\mu$  is the backward lattice derivative and  $m_0$  the bare quark mass.  $A_{5\mu}^a(x)$  is the Noether current of a lattice trans-

formation that assigns opposite charges to fermions on the half-spaces  $0 \leq s < N_s/2$  and  $N_s/2 \leq s < N_s$ . This reduces to a chiral transformation on the quark states, as long as they are localized in their respective half-spaces. The pseudo-scalar density  $J_5^a(x)$  is composed of fermion variables situated on the two boundaries, and serves as the standard interpolating field for pions. The anomalous (lattice-artifact) term in this relation involves another pseudo-scalar density,  $J_{5q}^a(x)$ , which is localized on two  $s$ -layers exactly half-way in the fifth direction. [We are assuming an undoubled quark spectrum (one quark field for each five-dimensional fermion field, in most applications) which is true at weak coupling. As mentioned in the Introduction, the massless spectrum changes at strong coupling [10] in which case the above transformations are no longer chiral.]

Before we proceed with the discussion of the anomalous term we have to address a technical point. Let  $T = \sum_i |i\rangle \langle i|$  be the spectral decomposition of the (positive) first-quantized transfer matrix. We define a “normal-ordered” transfer matrix  $Q$  via its spectral decomposition  $Q = \sum_i |i\rangle \lambda_i \langle i|$  where  $\lambda_i = \min\{t_i, t_i^{-1}\}$ . Physically, the operation of replacing  $t_i > 1$  by its inverse amounts to filling the Dirac sea. The spectrum of  $Q$  lies in the interval  $0 \leq \lambda_i \leq 1$ . One can show [11] that the exact domain-wall propagator in a given background field is a sum of terms, each of which involves the matrix  $Q$  raised to a positive power which is a function of the fifth coordinates.

Let us now assume that, for a given background field, the spectrum of  $Q$  lies in the interval  $0 \leq \lambda_i \leq \lambda_0$  where  $\lambda_0 < 1$ . Coming back to the anomalous term in Eq. (C1), for a non-singlet current it involves the propagation of two fermions over an  $s$ -separation equal to  $N_s/2$  (or  $N_s/2 \pm 1$ ). The anomalous term is thus bounded by  $(\lambda_0^{N_s/2})^2 = \lambda_0^{N_s}$  times a constant. (Using the second-quantized transfer matrix formalism one can show that the proportionality constant is finite, being the norm of a product of bounded operators [6].) Moreover, if  $Q$  has no eigenvalues larger than  $\lambda_0$  for all gauge fields (a condition which is satisfied for a constrained gauge action [9]) one finds that the anomalous term falls exponentially after the functional averaging over the gauge field.

In Ref. [6] we showed that a very weak bound on the anomalous correlator exists even if there is no gap at all. The point is that exact-unity eigenvalues of the transfer matrix (hence of its normal-ordered version,  $Q$ , too) exist only on a submanifold of the lattice gauge-field space defined by the condition  $\det D_W = 0$  where  $D_W$  is the Hermitian four-dimensional Wilson-Dirac operator. As explained above, when a fermion propagates across an  $s$ -separation  $N_s/2$ , the propagator is bounded by (and, generically, falls like)  $\lambda_0^{N_s/2}$  where  $\lambda_0 \leq 1$  is the largest eigenvalue of  $Q$ . But  $\lambda_0^{N_s/2}$  is negligible unless  $\lambda_0 = 1 - O(1/N_s)$ . This condition, in turn, will be satisfied only for gauge-field configurations whose distance from the above submanifold does not exceed  $O(1/N_s)$ . The volume of the (compact) gauge-field subspace contributing to the anomalous correlator is therefore finite, and shrinks like  $1/N_s$ , implying a similar bound on the correlator itself. We comment that the restoration of chiral symmetry is consistent with the fact that the overlap operator

defined by the  $N_s \rightarrow \infty$  limit [see Eq. (3.9)] admits Lüscher’s chiral symmetry [32] (whose generators are functions of the gauge field).

## 2. Spectral density and effective wave function

In the infinite-volume limit one can define a spectral function  $\rho_Q(\lambda)$  associated with the normal-ordered transfer matrix  $Q$  introduced in the previous subsection, whose support is (contained in) the interval  $[0, 1]$ . Here we will not attempt to compute any spectral function. Instead, we adopt a “phenomenological” point of view. We will assume that a single, continuous, spectral density function  $\rho(\lambda)$  has been defined as a suitable configuration average of  $\rho_Q(\lambda)$ . Using the considerations of the previous subsection, the aim is to see how different forms of the spectral function lead to different ways of approaching the chiral limit. One can envisage three prototype scenarios which are listed below. [A recent treatment of domain-wall fermions based on spectral integrals can be found in Ref. [33]. A numerical study of a spectral quantity which is closely related to  $\rho(1)$  can be found in Ref. [12]. See also Ref. [9].]

(1) *Exponential suppression.* Assume that the support of  $\rho(\lambda)$  is the interval  $0 \leq \lambda \leq \lambda_0$  where  $\lambda_0 < 1$  and that, close to  $\lambda_0$ ,  $\rho(\lambda)$  vanishes like  $(\lambda_0 - \lambda)^\delta$  with  $\delta > 0$ . Consider the propagation of a single fermion from the boundary layer  $s' = 0$  to some other layer  $s$  (we assume  $1 \leq s \leq N_s/2$ ). This involves the integral

$$\int_0^{\lambda_0} d\lambda \rho(\lambda) \lambda^s \sim \lambda_0^s \int_0^{\lambda_0} d\lambda \rho(\lambda) \exp\left(s \frac{\lambda - \lambda_0}{\lambda_0}\right) \sim s^{-1-\delta} \lambda_0^s. \quad (C2)$$

In order to obtain the power-law correction we have used the assumed behavior of  $\rho(\lambda)$  close to  $\lambda_0$ , and wrote  $(\lambda/\lambda_0)^s = \exp(s \log[1 + (\lambda - \lambda_0)/\lambda_0])$ . (The one-loop result of Sec. II corresponds to  $\lambda_0 = 1/2$  and  $\delta = 1$ .) If  $n$  fermions propagate across a similar  $s$ -interval, we will obtain the factor  $s^{-1-\delta} \lambda_0^s$  for each of them. We may therefore consider  $\chi_{\text{eff}}(s) = s^{-1-\delta} \lambda_0^s$  as the effective  $s$ -coordinate wave function for all quark states. Since the anomalous divergence  $J_{5q}^a(x)$  is a fermion bilinear, chiral symmetry violations should fall like  $\chi_{\text{eff}}^2(N_s/2) \sim N_s^{-2(1+\delta)} \lambda_0^{N_s}$ . [The effective wave function describing the propagation of  $n$  fermions could be somewhat different from  $\chi_{\text{eff}}^n(s)$  due to interactions between the different particles. Since chiral symmetry violations are related to  $J_{5q}^a(x)$ , the relevant effective wave function is always the one extracted from the sector with one fermion and one antifermion.]

(2) *Power-law suppression.* Assume that  $\lambda_0 = 1$  but with a vanishing  $\rho(1)$ , namely  $\rho(\lambda) \sim (1 - \lambda)^\delta$  for  $\lambda \sim 1$  with  $\delta > 0$ . In that case the result of the spectral integral (C2) will be  $s^{-1-\delta}$ . We might still speak of an effective wave function  $\chi_{\text{eff}}(s) = s^{-1-\delta}$  and expect chiral symmetry violations to fall like  $N_s^{-2(1+\delta)}$ .

(3) *(Almost) no suppression.* Last assume that  $\lambda_0 = 1$  and that  $\rho(1)$  is non-zero. Remember now the submanifold discussed in the previous subsection of gauge fields supporting

an eigenvalue one of  $Q$ . If  $\rho(1)$  is finite, configurations close to that submanifold must have a non-negligible Boltzmann weight. As we have explained, in this case the only suppression of long-range  $s$ -correlations comes from phase space considerations [the need to pick a configuration located  $O(1/N_s)$  away from that submanifold]. As a result, chiral symmetry violations fall roughly like  $\rho(1)/N_s$ , and the concept of a localized, effective wave function breaks down. ( $n$ -fermion correlations fall like  $1/N_s$  too, and not like  $1/N_s^n$ .)

For clarity, we have presented above the three mathematically distinct scenarios. In reality, however, one is likely to encounter a more complicated behavior, characterized by the existence of a *crossover* region. Let us reexamine the exponential suppression scenario. As discussed in the previous subsection, in the ensemble of all gauge-field configurations (as opposed to the case where the plaquette is constrained to be everywhere small [9]) there is a non-zero probability of finding an eigenvalue arbitrarily close to one, for any  $g > 0$ . Consider the integrated spectral density  $\mathcal{I} = \int_{\lambda_0}^1 d\lambda \rho(\lambda)$ , and suppose that  $\mathcal{I}$  is comparable to  $\lambda_0^{N_0}$  for some  $N_0$  (up to power corrections). For  $N_s \lesssim N_0$ , chiral symmetry violations

will fall like  $N_s^{-2(1+\delta)} \lambda_0^{N_s}$ . The point is that for any  $\lambda_0$  significantly smaller than one, and  $N_0$  of the order of (few times) ten,  $\lambda_0^{N_0}$  will be so small that one will never have to use  $N_s \gtrsim N_0$  in a simulation. For practical purposes this scenario is therefore the same as the purely-exponential suppression scenario. If, nevertheless, very large values of  $N_s$  will be tried, then around  $N_s \sim N_0$  a crossover to some slower fall-off rate will be encountered. (For instance, a crossover to a slower exponential fall-off rate has been observed in the Schwinger model [13].)

Last consider the relation between the domain-wall actions discussed in Sec. III and the associated overlap operators [see Eq. (3.9)]. The spectral function that controls the approach to the chiral limit of domain-wall fermions also controls the localization range of the overlap operator. As is clear from the above discussion, problems start when there is no gap, namely when it is not possible to identify a range  $\lambda_0 \leq \lambda \leq 1$  (with  $\lambda_0 < 1$ ) where the eigenvalue density is (practically) zero. In that case both the overlap operator and the generators of the associated Lüscher symmetries become non-local. This is the counter-part of the loss of exponential suppression in the domain-wall case.

- 
- [1] L. H. Karsten and J. Smit, Nucl. Phys. **B183**, 103 (1981).
  - [2] D. B. Kaplan, Phys. Lett. B **288**, 342 (1992).
  - [3] R. Narayanan and H. Neuberger, Phys. Lett. B **302**, 62 (1993); Nucl. Phys. **B412**, 574 (1994).
  - [4] R. Narayanan and H. Neuberger, Nucl. Phys. **B443**, 305 (1995).
  - [5] Y. Shamir, Nucl. Phys. **B406**, 90 (1993).
  - [6] V. Furman and Y. Shamir, Nucl. Phys. **B439**, 54 (1995).
  - [7] S. Aoki and Y. Taniguchi, Phys. Rev. D **59**, 054510 (1999).
  - [8] Y. Kikukawa, H. Neuberger, and A. Yamada, Nucl. Phys. **B526**, 572 (1998).
  - [9] P. Hernández, K. Jansen, and M. Lüscher, Nucl. Phys. **B552**, 363 (1999).
  - [10] R. C. Brower and B. Svetitsky, Phys. Rev. D **61**, 114511 (2000).
  - [11] Y. Shamir, Phys. Rev. D **59**, 054506 (1999).
  - [12] R. G. Edwards, U. M. Heller, and R. Narayanan, Phys. Rev. D **60**, 034502 (1999).
  - [13] P. Vranas, Phys. Rev. D **57**, 1415 (1998).
  - [14] T. Blum and A. Soni, Phys. Rev. D **56**, 174 (1997); Phys. Rev. Lett. **79**, 3595 (1997).
  - [15] T. Blum, A. Soni, and M. Wingate, Phys. Rev. D **60**, 114507 (1999).
  - [16] See Table I of the second paper of Ref. [19].
  - [17] S. Aoki, T. Izubuchi, Y. Kuramashi, and Y. Taniguchi, Nucl. Phys. B (Proc. Suppl.) **83-84**, 624 (2000).
  - [18] RBC Collaboration, C. Dawson, Nucl. Phys. B (Proc. Suppl.) **83-84**, 854 (2000).
  - [19] P. Vranas, hep-lat/0001006; RBC Collaboration, L. Wu *et al.*, Nucl. Phys. B (Proc. Suppl.) **83-84**, 224 (2000); CP-PACS Collaboration, A. Ali Khan *et al.*, *ibid.* **83-84**, 591 (2000).
  - [20] J.-F. Lagae and D. K. Sinclair, Nucl. Phys. B (Proc. Suppl.) **83-84**, 405 (2000); RBC Collaboration, G. T. Fleming, *ibid.* **83-84**, 363 (2000).
  - [21] This possibility was first stressed to me by Tom Blum.
  - [22] Y. Shamir, hep-lat/9912027.
  - [23] G. P. Lepage and P. B. Mackenzie, Phys. Rev. D **48**, 2250 (1993).
  - [24] I thank Norman Christ for the suggestion to work out the wave function corrections directly from the resummed propagator.
  - [25] P. Vranas, I. Tziligakis, and J. Kogut, Phys. Rev. D **62**, 054507 (2000).
  - [26] S. Sharpe, hep-lat/9811006.
  - [27] H. Neuberger, Phys. Lett. B **417**, 141 (1998).
  - [28] F. Niedermayer, Nucl. Phys. B (Proc. Suppl.) **73**, 105 (1999).
  - [29] R. G. Edwards, U. M. Heller, and R. Narayanan, Nucl. Phys. **B535**, 403 (1998).
  - [30] This was pointed out to me by Martin Lüscher.
  - [31] Y. Iwasaki, Nucl. Phys. **B258**, 141 (1985).
  - [32] M. Lüscher, Phys. Lett. B **428**, 342 (1998).
  - [33] Y. Kikukawa, hep-lat/9912056.