SUMMER SCHOOL ON PARTICLE PHYSICS

16 June - 4 July 2003

NON-PERTURBATIVE METHODS IN QUANTUM FIELD THEORY

Lectures I, II & III

C. REBBI
Dept. of Physics
Boston University
Boston, MA 02215
U.S.A.

Non-perturbative Methods in Quantum Field Theory.

Claudio Rebbi

Boston University - Department of Physics
590 Commonwealth Avenue, Boston MA 02215
USA
e-mail: rebbi@bu.edu

1 Introduction

These are very schematic notes, meant to provide the students with a record of the equations and main notions covered in my lectures. They should not be considered formal lecture notes.

The topic of non-perturbative methods in quantum field theory is very vast and I could not possibly cover it in its entirety, even in cursory manner, in four lectures, leaving aside that a lot of it would go beyond my own domain of expertise.

In these lectures I will focus on

- Lattice gauge theory, mainly lattice QCD.

- (time permitting) Semiclassical techniques for tunneling in high-energy collisions.

Insofar as lattice gauge theory is concerned, in my first lecture I plan to review the basics of the formalism, while in my second lecture I will talk about lattice fermions, with emphasis on the overlap discretization of lattice fermions, which preserves chiral symmetry. These

notes cover the topics that will be treated in the first two lectures. In my third lecture I will provide an illustration of current lattice techniques and results by describing a calculation of light quark masses and of the quark condensate with overlap fermions, performed in collaboration with L. Giusti and Ch. Hoelbling. References [1, 2] provide background material for this lecture. (In order to give proper credit, I would like to mention that this investigation is now continuing, with extended scope, in collaboration with F. Berruto, N. Garron, Ch. Hoelbling, L. Lellouch and N. Shoresh.) If time permits, in the final lecture I will illustrate the use of semiclassical techniques for high-energy processes involving tunneling. Reference [3], which gives a rather detailed account of an investigation done in collaboration with F. Bezrukov, D. Levkov, V. Rubakov and P. Tinyakov, may be used as background material for this lecture.

2 Lattice Gauge Theory.

General references: books by M. Creutz [4, 5], I. Montvay and G. Münster [6], H. J. Rothe [7]; lecture notes by R. Gupta [8]; the proceedings of the yearly International Symposia on Lattice Gauge Theory, published during the last several years in the Proceeding Supplements series of Nuclear Physics.

2.1 Lattice regularization.

Gauge invariant regularization, not relying on a perturbative expansion (Wilson, 1974 [9]).

Gauge invariant: avoids the occurrence of undesirable extra operators and the need to adjust the corresponding coupling constants.

Discretize the field degrees of freedom in configuration space, after continuation ($t \to it$) to Euclidean space-time.

Introduce a hypercubical lattice, with uniform lattice spacing a

$$x_\mu = n_\mu a \qquad n_\mu \text{ integer} \tag{1}$$

Notation: $\hat{\mu} \equiv$ unit vector in direction μ.

(Note: for some applications people use different lattice spacings a_s, a_t, generally with $a_t \ll a_s$ in the space and time directions. Lattices with different geometry or non-uniform lattices have been studied as well, but in general have not proven useful.)

Consider specifically **lattice QCD:**

Matter fields ($\equiv$ quark fields) are defined over the sites of the lattice:

$$\psi_{c,s,f}(x) \qquad \text{c=color index, s=spin index, f=flavor index} \tag{2}$$

Gauge fields provide color transport factors between neighboring lattice sites. They are **finite** elements of the $SU(3)$ color group:

$$U_\mu(x) \equiv U_{\mu,cc'}(x) \tag{3}$$

defined over the oriented link of the lattice, from x to $x + \hat{\mu}a$.

Nearest neighbor coupling

$$\bar{\psi}(x)U_\mu(x)\psi(x + \hat{\mu}a) \tag{4}$$

(with implicit color, spin and flavor indices).

Correspondence with continuum theory:

$$U_\mu(x) = e^{igA_\mu(x)a} \tag{5}$$

$$
\begin{aligned}
\bar{\psi}(x)U_\mu(x)\psi(x + \hat{\mu}a) &= \bar{\psi}(x)(1 + igA_\mu(x) + \ldots)(\psi(x) + a\partial_\mu\psi(x) + \ldots) \\
&= \bar{\psi}(x)\psi(x) + a\bar{\psi}(x)(\partial_\mu + igA_\mu)\psi(x)
\end{aligned} \tag{6}
$$

where we recover the gauge covariant derivative $D_\mu = \partial_\mu + igA_\mu$.

Gauge transformations are defined over the sites of the lattice

$$
\begin{aligned}
G(x) \quad &: \\
\psi(x) \quad &\Rightarrow \quad G(x)\psi(x) \tag{7} \\
U_\mu(x) \quad &\Rightarrow \quad G(x)U_\mu(x)G^\dagger(x + \hat{\mu}a) \tag{8}
\end{aligned}
$$

$[\psi(x)U_\mu(x)\psi(x + \hat{\mu}a)$ is obviously invariant; derive the transformation properties of $A_\mu(x)$ in the limit $a \to 0$.]

The gauge field action: the *plaquette* variables

$$U_{\mu\nu}(x) = U_\mu(x)U_\nu(x + \hat{\mu}a)U_\mu^\dagger(x + \hat{\nu}a)U_\nu^\dagger(x) \tag{9}$$

$a \to 0$:

$$
\begin{aligned}
e^{igA_\mu(x)a}e^{igA_\nu(x+\hat{\mu}a)a}e^{-igA_\mu(x+\hat{\nu}a)}e^{-igA_\nu(x)} &= (1 + igA_\mu a + \ldots) \times \\
(1 + igA_\nu a + ig\partial_\mu A_\nu a^2 + \ldots)(1 - igA_\mu a - ig\partial_\nu A_\mu a^2 + \ldots)(1 - igA_\nu a + \ldots) &= \\
= 1 + iga^2(\partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu] + \ldots) &= e^{iga^2 F_{\mu\nu}}
\end{aligned} \tag{10}
$$

measure the "curvature" of the gauge field. (Check that the neglected terms of $O(a^2)$ cancel.)

$$\text{tr}\, U_{\mu\nu} = 3 - \frac{g^2 a^4}{2}\text{tr}\, F_{\mu\nu}F_{\mu\nu} + \ldots \tag{11}$$

Continuum gauge action

$$\int dx \frac{1}{2} \sum_{\mu\nu} \text{tr } F_{\mu\nu} F_{\mu\nu} \approx a^4 \sum_{x,\mu<\nu} \text{tr } F_{\mu\nu} F_{\mu\nu} \tag{12}$$

Lattice gauge action

$$S_G = \frac{6}{g^2} \sum_{x,\mu<\nu} [1 - \frac{1}{3} \text{Re tr } U_{\mu\nu}(x)] \tag{13}$$

Notation:

$$\frac{6}{g^2} = \beta \tag{14}$$

inspired by analogy with statistical mechanics, **but** β **is not** $\beta_T = 1/kT$.

Fermionic action, (**naive, not correct, see later lecture**):

$$\begin{aligned} S_F &= \frac{1}{2a} \sum_{x,\mu} \bar{\psi}(x) \gamma_\mu [U_\mu(x)\psi(x+\hat{\mu}a) - U_\mu^\dagger(x-\hat{\mu}a)\psi(x-\hat{\mu}a)] + \sum_x \bar{\psi}(x) M \psi(x) \\ &\equiv \bar{\psi}(D+M)\psi \end{aligned} \tag{15}$$

(Note, r.h.s. in eq. above should be multiplied by a^4 to keep continuum normalization for ψ. In practice, l.g. theorists often work with $a = 1$, i.e. express dimensionful quantities in units of a.)

2.2   Observables.

$$\begin{aligned} Z &= \int \prod_{x,\mu} dU_\mu(x) \prod_x (d\psi(x)\, d\bar{\psi}(x))\, e^{-S_G(U,g)-S_F(U,\psi,\bar{\psi})} \\ &= \int \prod_{x,\mu} dU_\mu(x)\, e^{-S_G(U,g)} \text{Det}(D+M) \end{aligned} \tag{16}$$

$dU=$ invariant measure over group manifold. Example, for $SU(2)$ with

$$U = \cos\alpha + i\hat{n} \cdot \vec{\sigma} \sin\alpha \tag{17}$$

$$dU = \sin^2\alpha \sin\theta\, d\alpha\, d\theta\, d\phi \tag{18}$$

$$\begin{aligned} \langle \mathcal{O} \rangle &= Z^{-1} \int \prod_{x,\mu} dU_\mu(x) \prod_x (d\psi(x)\, d\bar{\psi}(x))\, \mathcal{O}(U,\psi,\bar{\psi})\, e^{-S_G(U,g)-S_F(U,\psi,\bar{\psi})} \\ &= Z^{-1} \int \prod_{x,\mu} dU_\mu(x)\, \langle\langle \mathcal{O} \rangle\rangle_U\, e^{-S_G(U,g)} \text{Det}(D+M) \end{aligned} \tag{19}$$

where $\langle\langle\mathcal{O}\rangle\rangle_U$ stands for the average over fermion fields with fixed gauge field background.

Example, with $\mathcal{O} = \bar{\psi}(x)\gamma_5\psi(x)\bar{\psi}(y)\gamma_5\psi(y)$

$$G_\pi(x,y) \equiv \langle\mathcal{O}\rangle = Z^{-1}\int\prod_{x,\mu}dU_\mu(x)\,\mathrm{tr}\,\left(-\gamma_5 P(x,y)\gamma_5 P(y,x)\right) \tag{20}$$

where

$$P(x,y) \equiv P(x,y,U) = (D(U)+M)^{-1}(x,y) \tag{21}$$

is the quark propagator in the background gauge field configuration given by U.

Expanding the meson propagator into a sum over physical states one can recover the lowest mass from its rate of decay

$$\sum_{\vec{x}} G_\pi(x_0\vec{x},0) = \sum_n \langle 0|\bar{\psi}(0)\gamma_5\psi(0)|n\rangle\frac{e^{-m_n x_0}}{2m_n}\langle n|\bar{\psi}(0)\gamma_5\psi(0)|0\rangle \tag{22}$$

An important "pure gauge" observable: rectangular Wilson loop $W(r,t)=$ trace over group indices of the product of gauge variables $U_\mu(x)$ over a rectangular path of sides of length r and t. For $t\to\infty$

$$\langle W(r,t)\rangle \propto e^{-V(r)t} \tag{23}$$

with $V(r) = $ potential of two static quark sources at separation r. If $V(r)\sim\sigma r$ (σ = string tension) then, for $r,t\to\infty$, $\langle W(r,t)\rangle \propto e^{-\sigma A}$, with $A = rt$ = area enclosed by the loop (area law).

2.3 Continuum limit.

The continuum limit relies on the existence of a "critical point" g_{cr} such that for $g\to g_{\mathrm{cr}}$ correlation lengths in lattice units go to infinity: $f(g) = l/a \to \infty$. Then we can fix $a = a(g)$ demanding

$$af(g) = l_{\mathrm{phys}} = \text{fixed} \tag{24}$$

for some definite correlation length (or equivalently for some mass $m = l^{-1}$ or some other physical constant). If the fixed point has universal scaling properties, all other physical quantities will tend to finite limits for $g\to g_{\mathrm{cr}}, a\to 0$.

For QCD, asymptotic freedom tells us that the critical point we want is $g_{\mathrm{cr}} = 0$ ($\beta_{\mathrm{cr}} = \infty$). Moreover it also tells us that

$$a = \frac{1}{\Lambda_{\mathrm{QCD}}}e^{-1/2\beta_0 g^2}(\beta_0 g^2)^{-\beta_1/2\beta_0^2}[1 + O(g^2)] \tag{25}$$

with β_0, β_1 the first two coefficient in the expansion of the β function.

Note: often one needs to calculate matrix elements of operators which are not physical quantities per se, but will produce physical quantities, such as a decay rate etc., when combined with other expressions. These matrix elements may then need multiplicative and possibly also additive renormalizations.

2.4 Computational techniques.

The r.h.s. in Eqns. 16, 19 can be calculated by a perturbative expansion in powers of g. This is similar to continuum perturbation theory, made much more complicated by the loss of Euclidean space-time rotational and translational symmetry. It is used only to relate renormalization constants evaluated on the lattice and in some continuum scheme.

The great advantage of the lattice discretization is that it allows one to do non perturbative calculations, like strong coupling expansions, mean field calculations and, especially, numerical simulations. It also makes it possible to define a gauge theory with discrete gauge group, although this is not as important as it was years ago, when computers were much less powerful.

As an example of a strong coupling expansion, let us evaluate $\langle W(r,t) \rangle$ in the pure gauge theory $(S_F = 0)$ to leading order in β. The integral over the gauge group elements $U_\mu(x)$ elements will vanish unless we bring down from the $\exp(-S_G)$ one gauge group element $U_\mu^\dagger(x)$ for each $U_\mu(x)$ in W. This in turn requires that we bring down from $\exp(-S_G)$ one plaquette variable for every plaquette inside the rectangle spanned by W. As a result, the expectation value will contain a factor $(\text{const} \times \beta)^{rt/a^2}$, where rt/a^2 is the number of plaquettes inside the rectangle, and the expectation value will obey an area law with a string tension $\sigma \propto -\log(\beta)$ or $\sigma \propto 2\log(g)$. Of course, this proof of confinement at strong coupling is of little value for the continuum limit, which must be recovered at $g \to 0$. Numerical calculations, however, have proven capable of bridging the gap to the continuing limit, providing valuable information in the intermediate coupling domain $g \approx 1$, which turns out to correspond to a cut-off a^{-1} roughly of order of 2GeV.

Numerical calculations typically proceed through a stochastic algorithm (deterministic, molecular dynamics simulations can also be done), whose outcome is the generation of gauge field configurations, $C_i \equiv \{U_\mu(x)\}_i$, distributed according to the measure

$$P(C) = Z^{-1} e^{-S_G + \log \text{Det}(D+M)} \tag{26}$$

The phase space integral over all values of $U_\mu(x)$ in Eq. 19 is then approximated with an average over the configurations C_i. A discussion of the algorithms used to generate configurations goes beyond the scope of these lectures. The inclusion of the fermionic determinant in the measure is particularly challenging from a computational point of view. This has prompted the approximation, called "quenched" of "valence" approximation, of neglecting

6

the determinant altogether, using just $\exp(-S_G)$ as measure for the gauge field. This is justified with the fact that the determinant, form a diagrammatic point of view, contributes terms with internal quark loops, i.e. the effects of "sea quarks" and that to a first approximation the binding of quarks in nucleons is due to the dynamics of the gauge field. The term valence approximation conveys well the notion that one only keeps in the expression for the observables the propagators for the constituent quarks, neglecting the sea quarks. The term quenched was borrowed from condensed matter terminology, and somehow remained in prevalent use. Calculations that do include the effects of the fermionic determinant are referred to as dynamical QCD or unquenched simulations. Computationally, the cost of calculating the quark propagators is a couple of orders of magnitude larger than the cost of generating the corresponding quenched configurations (although the use of a chirally invariant formulation of lattice fermions entails much higher computational costs), and the cost of generating unquenched configurations is another couple of orders of magnitude higher than the cost of calculating the propagators themselves. Current supercomputer resources allows one to perform simulations on lattices as large as $32^3 \times 64$ (one often uses lattices with larger extent in one direction for a better calculation of the rate of decay of correlation functions) or larger, but of course a variety of algorithmic and computational considerations (e.g. trade-offs between size and statistics) go into the specific choice of a lattice size.

3   Lattice Fermions.

3.0.1   Fermion doublers.

The fermionic action of Eq. 15 has unwanted low frequency excitations. Consider the free theory ($U_\mu(x) = I$). In momentum space the operator D takes the form

$$D = i \sum_\mu \frac{1}{a} \gamma_\mu \sin(a p_\mu) \tag{27}$$

where the momentum components vary over a range of width $2\pi/a$, which, for convenience, we will take to be:

$$-\frac{\pi}{2a} \le p_\mu \le \frac{3\pi}{2a} \tag{28}$$

We see that for $p_\mu \approx 0$ D reduces to the proper continuum form

$$D = i \sum_\mu \gamma_\mu p_\mu \tag{29}$$

However, the shift of any momentum component by π/a will produce a continuum limit contribution to D of similar form. Equivalently, if we define a new field $\tilde\psi(x) = (\prod_\mu s_\mu^{x_\mu/a}) \psi(x)$,

where the factors s_μ can be 1 or -1, the lattice Dirac operator D in the new basis will take the form

$$D = \imath \sum_\mu \frac{1}{a} \gamma_\mu s_\mu \sin(ap_\mu) \tag{30}$$

with a $p_\mu \approx 0$ limit

$$D = \imath \sum_\mu \gamma_\mu s_\mu p_\mu \tag{31}$$

Since the low frequency components of ψ and $\tilde{\psi}$ clearly do not overlap, we conclude that our fermionic action describes one independent continuum fermionic excitation for each choice of s_μ. Thus, in addition to the original fermion, the action will describe 15 extra fermionic modes (doublers). Also, since the continuum limit action for the doublers has γ_μ replaced by $s_\mu \gamma_\mu$, we see that the chirality properties of the doubles will be identical or opposite to those of the original fermion according to whether $\sum s_\mu$ is even or odd. In particular, if we use a projection operator $(1 - \gamma_5)/2$ in the original action to define, in our hopes, a chiral theory, the action will actually describe 8 left-handed fermions and 8 fermions of opposite chirality.

3.0.2 Wilson fermions.

One can remove the unwanted doublers by giving them masses of order $1/a$. This can be done [9] by adding to the action of Eq. 15 a term

$$
\begin{aligned}
S'_F &= \frac{1}{2a} \sum_{x,\mu} \bar{\psi}(x)[2\psi(x) - U_\mu(x)\psi(x + \hat{\mu}a) - U_\mu^\dagger(x - \hat{\mu}a)\psi(x - \hat{\mu}a)] \\
&\equiv \bar{\psi}D'\psi
\end{aligned} \tag{32}
$$

In the free theory, in momentum space D' takes the form

$$D' = \frac{1}{a} \sum_\mu [1 - \cos(ap_\mu)] \tag{33}$$

and we see that the doublers takes an extra mass $\sum_\mu (1 - s_\mu)/a$.

We will denote by D_W (Wilson lattice Dirac operator) the combination

$$D_W = D + D' \tag{34}$$

While the Wilson discretization of lattice fermions has no doublers, it violates chiral symmetry in an intrinsic manner. In the continuum, one can start with a massive theory and go to the chiral limit by taking $M \to 0$. With Wilson fermions, sending $M \to 0$ in $D_W + M = D + D' + M$ does not produce a chirally symmetric theory, since D' breaks

8

chiral symmetry. The difficulty in obtaining a chiral formulation of lattice fermions is a manifestation of a theorem due to Nielsen and Ninomiya [10], which states that a strictly local (i.e. only containing couplings between fermionic fields at a separation that does not exceed some fixed number of lattice sites) lattice discretization of chiral fermions is impossible.

There are ways to alleviate the consequences of the above theorem. For example, in the Kogut-Susskind or staggered discretization of lattice fermions [11] one reduces the number of doublers to 4 (the original fermion plus 3) and one can maintain a non-diagonal chiral symmetry. The real breakthrough has come, however, with the closely related domain wall [12, 13] or overlap [14, 15, 16] discretizations of lattice fermions.

3.0.3 Overlap fermions.

The continuum Dirac operator $D_c = D = \imath \sum_\mu \gamma_\mu D_\mu$ is antihermitian and thus its eigenvalues lie on the imaginary axis. Because of the property $D_c^\dagger = \gamma_5 D_c \gamma_5$ its eigenvalues different from 0 occur in complex conjugate pairs, while the eigenvectors with zero eigenvalue are or can be chosen to be eigenstates of chirality. When the background gauge field has non-vanishing topological number Q, the operator D_c has an index $n = n_+ - n_- = Q$, where n_+, n_- are the numbers of zero eigenstates with positive, negative chirality. The Wilson Dirac operator $D_W = D + D'$ is neither Hermitian nor antihermitian, since D is antihermitian and D' is Hermitian, and its eigenvalues are in general complex. D_W also satisfy the property

$$D_W^\dagger = \gamma_5 D_W \gamma_5 \tag{35}$$

and so its complex eigenvalues occur in complex conjugate pairs. D_W can occasionally have real eigenvalues, but the corresponding eigenvector are not bound to have definite chirality. It is easy to see that, if $\psi(x)$ is an eigenvector of D_W with eigenvalue λ, $\psi'(x) = (-1)^{(x+y+z+t)/a)}\psi(x)$ is an eigenvector with eigenvalue $8/a - \lambda$. Thus the spectrum of eigenvalues of D_W is symmetric with respect to reflections about the real axis and about a vertical axis of intercept $4/a$. The eigenvalues are contained in the strip $0 \leq \mathrm{Re}\lambda \leq 8/a$ and fill, rather irregularly, a region with roughly the shape of an oval, with four pronounced voids, again of roughly oval shape, between the values 0 and $2/a$, $2/a$ and $4/a$, $4/a$ and $6/a$, $6/a$ and $8/a$. Near the points $0, 2/a, 4/a, 6/a, 8/a$ the spectrum reaches toward the real axis. These are the regions where one recovers the 16 doublers of the theory.

It will be useful to consider the Wilson Dirac operator with a negative mass

$$D_W(\rho) = D_W - \frac{\rho}{a} \tag{36}$$

and the corresponding "Hermitian Wilson operator"

$$H(\rho) = \gamma_5 D_W(\rho) = \gamma_5 D_W - \gamma_5 \frac{\rho}{a} \tag{37}$$

(It is easy to check that H is Hermitian: $H^\dagger = D_W^\dagger \gamma_5 = \gamma_5 \gamma_5 D_W^\dagger \gamma_5 = \gamma_5 D_W = H$.)

The spectrum of $D_W(\rho)$ is, of course, just the spectrum of D_W shifted to the right by ρ/a. (The spectrum of H lies of course on the real axis, but its dependence on ρ is non-trivial. We will return to it later.) A value of ρ equal to one or in the vicinity of one will produce a shift that places the origin of the complex λ plane in the middle of the first void in the spectrum. The overlap Dirac operator, intuitively speaking, is obtained by projecting first the spectrum of of $D_W(\rho)$ from this point onto a circle, and then by shifting this circle to the right, to make it tangent to the imaginary axis. The eigenvalues of the overlap Dirac operator will thus lie on a circle of radius ρ/a tangent at the origin to the imaginary axis. In the limit $a \to 0$ the circle will tend toward the imaginary axis. Of course the spectrum of an operator cannot just be projected eigenvalue by eigenvalue. Rather we perform the projection by calculating first the unitary component V in the polar representation of $D_W(\rho)$:

$$D_W(\rho) = V[D_W(\rho)^\dagger D_W(\rho)]^{1/2} \tag{38}$$

that is

$$V = D_W(\rho)[D_W(\rho)^\dagger D_W(\rho)]^{-1/2} \tag{39}$$

and with this we form the overlap Dirac operator (or Neuberger Dirac operator) [14, 15]:

$$D = \frac{\rho}{a}(1 + V) \tag{40}$$

(This is for massless fermions, we will add a mass term later.)

V can also be expressed in terms of the Hermitian Dirac operator:

$$V = D_W(\rho)[D_W(\rho)^\dagger D_W(\rho)]^{-1/2} = \gamma_5 H(\rho)[(\gamma_5 H(\rho))^\dagger \gamma_5 H(\rho)]^{-1/2} = \gamma_5 H(\rho)[H(\rho)^2]^{-1/2} \tag{41}$$

In this last expression we encounter the interesting operator

$$\epsilon[H(\rho)] = H(\rho)[H(\rho)^2]^{-1/2} \tag{42}$$

It is the sign function of the operator $H(\rho)$, namely the projection operator over the subspace spanned by its positive eigenvectors minus the projection operator over the subspace of negative eigenvectors. Thus, in terms of H, the overlap Dirac operator is given by

$$D = \frac{\rho}{a}\{1 + \gamma_5 \epsilon[H(\rho)]\} \tag{43}$$

The above derivation does not respect the historical path of development, which followed a different route. Kaplan showed that one could avoid the problem of doubling by formulating a theory of fermions extended to a fifth dimension where a suitable domain wall is introduced to trap four-dimensional solutions of definite chirality [12]. Shamir provided an alternative

10

formulation where the two chiral components of Dirac fermions where instead localized at two boundaries in the fifth dimension. Narayanan and Neuberger [14, 15] reinterpreted the fifth dimension in terms of a tower of states, and, in the case of a vector theory, integrating out these extra degrees of freedom produces the expressions of 40, 43.

D satisfies a remarkable identity, the Ginsparg-Wilson relation [17]:

$$\gamma_5 D + D\gamma_5 = \frac{a}{\rho} D\gamma_5 D \tag{44}$$

Proof: substituting from 40, and after trivial cancellation of the ρ/a factors, 44 becomes

$$\gamma_5(1+V) + (1+V)\gamma_5 = (1+V)\gamma_5(1+V) \tag{45}$$

Expanding this becomes

$$2\gamma_5 + \gamma_5 V + V\gamma_5 = \gamma_5 + \gamma_5 V + V\gamma_5 + V\gamma_5 V \tag{46}$$

But from the identity 35 and the fact that V is unitary, it follows

$$V\gamma_5 V = V\gamma_5 V\gamma_5\gamma_5 = VV^\dagger\gamma_5 = \gamma_5 \tag{47}$$

and we see that the r.h.s in Eq. 45 is indeed equal to the l.h.s.

In terms of the propagator D^{-1} the Ginsparg-Wilson relation reads

$$D^{-1}\gamma_5 + \gamma_5 D^{-1} = \frac{a}{\rho} \tag{48}$$

and this turns out to be the minimal generalization to the lattice of the chiral identity satisfied by the continuum propagator:

$$D_c^{-1}\gamma_5 + \gamma_5 D_c^{-1} = 0 \tag{49}$$

The identity 44 implies the symmetry of the action under the transformation [18]

$$\delta\psi = \hat{\gamma}_5\psi, \qquad \delta\bar{\psi} = \bar{\psi}\gamma_5 \tag{50}$$

where

$$\hat{\gamma}_5 = \gamma_5(1 - \frac{a}{\rho}D) \tag{51}$$

which can be interpreted as a lattice chiral symmetry. (Note: from 44 it follows that $\hat{\gamma}_5^2 = 1$.) The corresponding chirally projected fields are

$$\psi_\pm = \frac{1}{2}\left(1 \pm \hat{\gamma}_5\right)\psi = \frac{1}{2}\left(1 \pm \gamma_5 \mp \frac{a}{\rho}\gamma_5 D\right)\psi, \qquad \bar{\psi}_\pm = \frac{1}{2}\bar{\psi}(1 \mp \gamma_5) \tag{52}$$

11

These projections should be taken into account when forming properly transforming bilinears. In particular, the scalar bilinear is

$$\bar{\psi}_+\psi_- + \bar{\psi}_-\psi_+ = \bar{\psi}\Big(1 - \frac{a}{2\rho}D\Big)\psi \tag{53}$$

We can use this to add a mass term to the overlap operator, which thus becomes

$$D_M = D + M\Big(1 - \frac{a}{\rho}D\Big) = \Big(1 - \frac{aM}{2\rho}\Big)D + M \tag{54}$$

The overlap operator also matches the chiral properties of the continuum operator in regard to zero modes. In order to describe these briefly, it is useful to start by considering the spectrum of V. This consists of eigenvectors $|\lambda\rangle$ with strictly complex (i.e. not real) eigenvalues λ, and of possible eigenvectors with eigenvalue -1 and/or +1. The latter two can be taken to be eigenstates of chirality. The eigenvectors of V with eigenvalue -1 are of course zero modes of D (i.e. eigenvectors with zero eigenvalue) and we will denote them by $|0\rangle_\pm$, $\pm$ referring to chirality. We will denote the eigenvectors of V with eigenvalue +1 by $|\infty\rangle_\pm$, because the corresponding eigenvalue of D, equal to $2\rho/a$, recedes to infinity in the continuum limit. Let us consider the trace of γ_5. Here we talk about the full trace of γ_5, which is a matrix defined over the lattice, of large but finite dimension (assuming that we are working with a lattice of large but finite volume). Obviously $\mathrm{Tr}\,\gamma_5 = 0$. Let us evaluate this trace in the basis formed by the eigenvectors of V. Since $\gamma_5|\lambda\rangle = |\lambda^*\rangle$ (this follows from $\gamma_5 V\gamma_5 = V^\dagger$ and the fact that V is unitary) $\mathrm{Tr}\,\gamma_5$ receives 0 contribution from the complex eigenvalues of V. Thus

$$\mathrm{Tr}\,\gamma_5 = n_+^{(0)} - n_-^{(0)} + n_+^{(\infty)} - n_-^{(\infty)} = 0 \tag{55}$$

where $n_\pm^{(0,\infty)}$ stands for the number of corresponding eigenvectors. Let us now consider $\mathrm{Tr}\,\gamma_5(1 + V)$. We have

$$\mathrm{Tr}\,\gamma_5(1 + V) = 2(n_+^{(\infty)} - n_-^{(\infty)}) = \mathrm{Tr}\,\gamma_5 V = \mathrm{Tr}\,\epsilon[H(\rho)] = \Delta H(\rho) \tag{56}$$

where $\Delta H(\rho)$ is the difference between the dimensionality of the subspace spanned by the positive eigenvectors of $H(\rho)$ and the one spanned by the negative eigenvectors. Combining 55 and 56 we find that

$$n_+^{(0)} - n_-^{(0)} = -\frac{1}{2}\Delta H(\rho) \tag{57}$$

On the other hand one can prove that, if the background gauge field configuration is sufficiently smooth, i.e. if $\mathrm{tr}\,[1 - \frac{1}{3}\mathrm{Retr}\,U_{\mu\nu}(x)]$ does not exceed a certain bound, one can define a topological number of the gauge field configuration $Q(U)$ as in the continuum [19] and that $\Delta H(\rho) = Q(U)$ [20]. Thus we see that for the overlap operator there is a direct relation between number of zero modes and topology, $n_+^{(0)} - n_-^{(0)} = Q$, very much as in the continuum.

12

For the above it should be apparent that the overlap discretization, or equivalently the domain wall discretization, provides a fully satisfactory formulation of lattice fermions, which preserves maximally the properties of the continuum theory. It circumvents the bound of the Nielsen-Ninomiya theorem by having a lattice Dirac operator which is not strictly local, or ultralocal, using a more technical term. The inverse square root appearing in 40 has indeed matrix elements between any two lattice sites, however remote. Still the operator is local in the sense that these matrix elements drop-off exponentially with separation and in such a way as to go to a local operator in the continuum limit. Unfortunately calculating the inverse square root of the matrix $D_W(\rho)^\dagger D_W(\rho)$ or, equivalently, the sign function of $H(\rho)$ is very demanding from a numerical point of view. Thus the practical use of the overlap formulation requires major algorithmic and computational efforts. Yet these new formulations of lattice fermions have been used in actual calculations, with very gratifying results.

References

[1] L. Giusti, Ch. Hoelbling, C. Rebbi, Phys. Rev. D64, 114508, 2001, hep-lat/0108007.

[2] L. Giusti, Ch. Hoelbling, C. Rebbi, Nucl. Phys. Proc. Suppl. 106, 739, 2002, hep-lat/0110184.

[3] F. Bezrukov, D. Levkov, C. Rebbi, V. Rubakov, P. Tinyakov, Phys. Rev. D, to be published, 2003, hep-ph/0304180.

[4] M. Creutz, *Quarks, Gluons and Lattices*, Cambridge University Press, 1983.

[5] M. Creutz, *Quantum Fields on the Computer*, World Scientific, 1992.

[6] I. Montvay and G. Münster, *Quantum Fields on the Computer*, Cambridge University Press, 1994.

[7] *Quantum Gauge Theories: An Introduction*, World Scientific, 1997.

[8] R. Gupta, *Introduction to Lattice QCD*, hep-lat/9807028, 1998.

[9] K. G. Wilson, Phys. Rev. D10, 2445, 1974.

[10] H. B. Nielsen, M. Ninomiya, Nucl. Phys. B185, 20, 1981, Erratum Nucl. Phys. B195, 541, 1982, Nucl. Phys. B193, 173, 1981.

[11] J. Kogut, L. Susskind, Phys. Rev.D11, 395, 1975.

[12] D. Kaplan, Phys. Lett. B288, 342, 1992.

[13] Y. Shamir, Nucl. Phys. B406, 90, 1993.

[14] R. Narayanan, H. Neuberger, Phys. Lett. B302, 62, 1993, Nucl. Phys. B443, 305, 1995.

[15] H Neuberger, Phys. Lett. B417, 141, 1998, Phys. Lett. B427, 353, 1998.

[16] S. Randjbar-Daemi, J. Strathdee, Nucl. Phys. B443, 386, 1995, hep-lat/9501027.

[17] P. H. Ginsparg, K. G. Wilson, Phys. Rev. D25, 2649, 1982.

[18] M. Lüscher, Phys. Lett. B428, 342, 1998.

[19] M. Lüscher, Comm. Math. Phys. 85, 39, 1982.

[20] D. H. Adams, Chin. J. Phys. 38 , 633 ,2000, hep-lat/0001014.