Lattice fermions, exact chiral symmetry & unitarity

Martin Lüscher August 2000

1. Introduction

In euclidean space chiral symmetry can be preserved on the lattice by choosing a lattice Dirac operator D that satisfies the Ginsparg-Wilson relation [1–5]

$$\gamma_5 D + D\gamma_5 = aD\gamma_5 D. \tag{1.1}$$

The known solutions of this equation are relatively complicated expressions that involve arbitrary powers of the lattice derivatives. Unitarity is thus expected to be violated at energy scales on the order of the momentum cutoff, as is generally the case in theories with higher-derivative actions.

In the present note we consider free fermions with D given by [3]

$$aD = 1 - A(A^{\dagger}A)^{-1/2}, \qquad A = 1 - aD_{\rm w}, \tag{1.2}$$

$$D_{\rm w} = \tfrac{1}{2}\left\{\gamma_\mu(\partial_\mu^* + \partial_\mu) - a\partial_\mu^*\partial_\mu\right\} \tag{1.3}$$

(see appendix A for unexplained notations). We then show that the fermion propagator admits a Källén–Lehmann spectral representation

$$\langle\psi(x)\overline{\psi}(y)\rangle \underset{x_0>y_0}{=} \int_0^\infty {\rm d}E \int_{-\pi/a}^{\pi/a} \frac{{\rm d}^3\mathbf{p}}{(2\pi)^3}\,\varrho(E,\mathbf{p})\,{\rm e}^{-E(x_0-y_0)+i\mathbf{p}(\mathbf{x}-\mathbf{y})} \tag{1.4}$$

such that

$${\rm d}E\,{\rm d}^3\mathbf{p}\;\zeta^{\dagger}\gamma_0\varrho(E,\mathbf{p})\zeta \tag{1.5}$$

is a *non-negative measure* for all complex Dirac spinors ζ. Contrary to expectations, this theory is thus unitary and it could be reformulated in terms of operators that act in a Hilbert space of physical states.

2. Preliminaries

For notational simplicity the lattice spacing a is set to unity in the following. The fermion propagator is then given by

$$\langle\psi(x)\overline{\psi}(y)\rangle = \tfrac{1}{2}\,\delta_{xy} - \tfrac{1}{2}\int_{-\pi}^{\pi}\frac{\mathrm{d}^4p}{(2\pi)^4}\,\mathrm{e}^{ip(x-y)}\,\frac{i\gamma_\mu\mathring{p}_\mu}{F(p)}, \tag{2.1}$$

$$F(p) = \left\{\mathring{p}^2 + \left(1 - \tfrac{1}{2}\hat{p}^2\right)^2\right\}^{1/2} - \left(1 - \tfrac{1}{2}\hat{p}^2\right), \tag{2.2}$$

where the square root is to be taken with positive sign. For small momenta we have

$$F(p) = \tfrac{1}{2}\,p^2 + \mathrm{O}(p^4), \tag{2.3}$$

and it is not difficult to show that $F(p)$ does not vanish anywhere else in the Brillouin zone.

To derive the spectral representation (1.4), we shall fix the spatial components of the momentum and shall evaluate the integral over p_0 by deforming the integration contour to the complex domain

$$R = \left\{p_0 \,\middle|\, -\pi \leq \mathrm{Re}\,p_0 \leq \pi,\ \mathrm{Im}\,p_0 \geq 0\right\}. \tag{2.4}$$

The contributions from the pole and cut singularities of the integrand then yield the spectral density. Formal manipulations at $p = 0$ can be avoided by excluding the ball $|\mathbf{p}| \leq \epsilon$ from the integral and taking the limit $\epsilon \to 0$ at the end of the calculation. This gives the correct result since the integral (2.1) is absolutely convergent.

We now first need to study the analytic continuation of the integrand to the complex region R at fixed spatial momentum in the range $|p_k| \leq \pi$, $|\mathbf{p}| > 0$.

3. Analyticity properties of $F(p)$

It is easy to check that

$$\mathring{p}^2 + \left(1 - \tfrac{1}{2}\hat{p}^2\right)^2 = 1 + \tfrac{1}{4} \sum_{\mu \neq \nu} \hat{p}_\mu^2 \hat{p}_\nu^2. \tag{3.1}$$

The argument of the square root in eq. (2.2) is thus a linear function of $\hat{p}_0^2$ and we now first need to understand the dependence of this variable on p_0 in the complex domain R. If we set

$$p_0 = r + is, \qquad \hat{p}_0^2 = u + iv, \tag{3.2}$$

we have

$$u = 2\left(1 - \cos r \cosh s\right), \qquad v = 2 \sin r \sinh s. \tag{3.3}$$

For fixed $s > 0$ and $-\pi \leq r \leq \pi$ these equations imply that $\hat{p}_0^2$ moves once around the ellipse given by

$$\left(\frac{u-2}{2\cosh s}\right)^2 + \left(\frac{v}{2\sinh s}\right)^2 = 1 \tag{3.4}$$

(see fig. 1). The region R is thus mapped to the whole plane and the mapping is one-to-one in the range $s > 0$ if the lines $r = \pm\pi$ are identified with each other.

It follows from these remarks that $F(p)$ extends to an analytic function in R with a cut along the imaginary axis from $E_{\mathbf{p}}$ to infinity, where $E_{\mathbf{p}}$ is given by

$$\cosh E_{\mathbf{p}} = 1 + \frac{1}{\hat{\mathbf{p}}^2}\left\{1 + \tfrac{1}{4}\sum_{k \neq l} \hat{p}_k^2 \hat{p}_l^2\right\}, \qquad E_{\mathbf{p}} > 0. \tag{3.5}$$

The function evaluates to

$$F(p) = \pm i \left\{\hat{\mathbf{p}}^2 \left(\cosh E - \cosh E_{\mathbf{p}}\right)\right\}^{1/2} - \cosh E + \tfrac{1}{2}\hat{\mathbf{p}}^2 \tag{3.6}$$

along the cut $p_0 = iE \pm \epsilon$, $E \geq E_{\mathbf{p}}$. Another property of the cut, which will turn out to be important later, is that

$$\sinh E > |\mathring{\mathbf{p}}| \quad \text{for all} \quad E > E_{\mathbf{p}}. \tag{3.7}$$

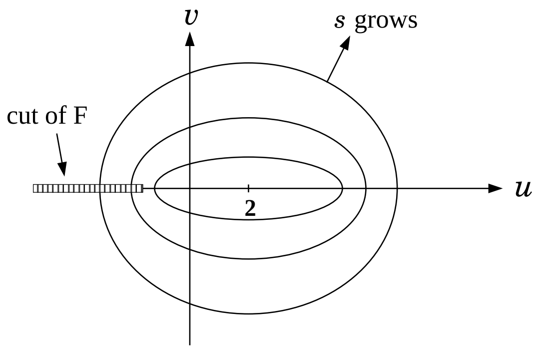

Fig. 1. The transformation $p_0 \to \hat{p}_0^2$ maps the lines at fixed $\mathrm{Im}\, p_0 = s$ to concentric ellipses with monotonically growing half-axes. In the limit $s \to 0$ the ellipses collapse to the interval $0 \leq \hat{p}_0^2 \leq 4$. The points where the ellipses cross the negative real axis correspond to the imaginary axis $r = 0, s > 0$.

One can easily prove this by noting that the argument of the square root in eq. (2.2) is real and negative along the cut (excluding the branch point $E = E_{\mathbf{p}}$). Since $\hat{p}^2$ is real for these values of p, the inequality

$$\mathring{p}^2 \leq \mathring{p}^2 + \left(1 - \tfrac{1}{2}\hat{p}^2\right)^2 < 0, \tag{3.8}$$

is obtained which implies (3.7).

4. Zeros of $F(p)$ in the region R

We first note that $F(p)$ does not vanish along the cut $p_0 = iE$, $E > E_{\mathbf{p}}$ [cf. eq. 3.6)]. Elsewhere the function is analytic and single-valued. It vanishes if and only if

$$\left\{\mathring{p}^2 + \left(1 - \tfrac{1}{2}\hat{p}^2\right)^2\right\}^{1/2} = \left(1 - \tfrac{1}{2}\hat{p}^2\right). \tag{4.1}$$

This equation implies $\mathring{p}^2 = 0$ which is equivalent to

$$\mathring{p}_0 = \pm i|\mathring{\mathbf{p}}|. \tag{4.2}$$

In particular, the real part of p_0 has to be equal to 0 or $\pm\pi$. The latter possibility is excluded since $\hat{p}_0^2 = 2\,(1 + \cosh s)$ in this case. The two sides of eq. (4.1) then have opposite sign and cannot match.

Our discussion so far shows that the zeros of $F(p)$ in the region R have to be on the imaginary axis below the cut. The square root is non-negative there and the same thus has to be true for the right-hand side of eq. (4.1). In other words, the function vanishes at $p_0 = is$ if and only if

$$\sinh s = |\mathring{\mathbf{p}}|, \qquad \cosh s \geq \tfrac{1}{2}\hat{\mathbf{p}}^2. \tag{4.3}$$

Note that the equality implies $s \leq E_{\mathbf{p}}$ as we have previously remarked (cf. discussion at the end of sect. 3).

We now define the set

$$\mathcal{B}_< = \left\{ \mathbf{p} \;\middle|\; -\pi \leq p_k \leq \pi, 1 + \mathring{\mathbf{p}}^2 \geq \tfrac{1}{4}\left(\hat{\mathbf{p}}^2\right)^2 \right\}, \tag{4.4}$$

which will be referred to as the *inner region of the Brillouin zone*. It can be shown that there is a non-zero distance between $\mathcal{B}_<$ and the boundaries $p_k = \pm\pi$ of the integration range. In particular, $\mathring{\mathbf{p}}$ does not vanish in $\mathcal{B}_<$ except at $\mathbf{p} = 0$.

From eq. (4.3) it is now obvious that $F(p)$ has a zero at

$$p_0 = i\omega_{\mathbf{p}}, \qquad \sinh \omega_{\mathbf{p}} = |\mathring{\mathbf{p}}|, \tag{4.5}$$

if $\mathbf{p}$ is in the inner region of the Brillouin zone and no zero otherwise. The residue

$$\left(\frac{\partial F(p)}{\partial p_0}\right)_{p_0 = i\omega_{\mathbf{p}}} = i\,\frac{\sinh \omega_{\mathbf{p}} \cosh \omega_{\mathbf{p}}}{\cosh \omega_{\mathbf{p}} - \frac{1}{2}\hat{\mathbf{p}}^2}, \tag{4.6}$$

does not vanish (if we exclude $\mathbf{p} = 0$) and diverges at the boundary of $\mathcal{B}_<$.

5. Computation of the spectral density $\varrho(E, \mathbf{p})$

For $x_0 > y_0$ the integrand in eq. (2.1) decreases exponentially at large s since

$$F(p) = -\tfrac{1}{2}\,\mathrm{e}^{s-ir} + \mathrm{O}(\mathrm{e}^{s/2}). \tag{5.1}$$

The integration over the range $-\pi \leq p_0 \leq \pi$ can thus be extended to a contour integral along the boundary of the region R (see fig. 2). Since the integrand is periodic in r, the integrals along the imaginary axes at $r = \pm\pi$ cancel.

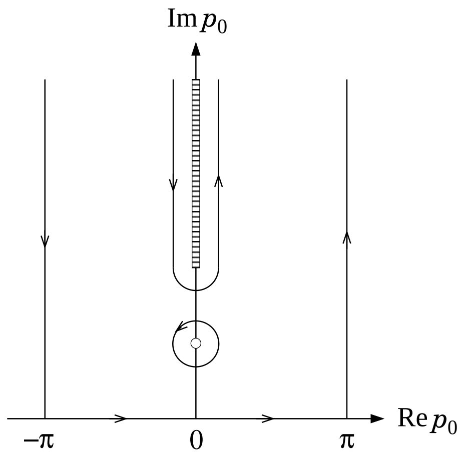

Fig. 2. The integration contour in the complex p_0 plane runs initially along the boundary of the region R. It may then be deformed until it runs around the left- and right rims of the imaginary axis, avoiding the pole (small open circle) and the cut singularity (shaded vertical bar) of the integrand.

We can now deform the integration contour towards the imaginary axis at $r = 0$. As a result the propagator assumes the form (1.4) with

$$\varrho(E, \mathbf{p}) = (\gamma_0 \sinh E - i\gamma_k \mathring{p}_k)$$

$$\times \Bigg\{ \delta\big(E - \omega_{\mathbf{p}}\big)\theta\big(\cosh E - \tfrac{1}{2}\hat{\mathbf{p}}^2\big)\frac{\cosh E - \frac{1}{2}\hat{\mathbf{p}}^2}{\sinh 2E}$$

$$+\frac{1}{2\pi}\theta\big(E - E_{\mathbf{p}}\big)\frac{\big\{\hat{\mathbf{p}}^2\big(\cosh E - \cosh E_{\mathbf{p}}\big)\big\}^{1/2}}{\hat{\mathbf{p}}^2\big(\cosh E - \cosh E_{\mathbf{p}}\big) + \big(\cosh E - \frac{1}{2}\hat{\mathbf{p}}^2\big)^2}\Bigg\}, \tag{5.2}$$

where $E_{\mathbf{p}}$ and $\omega_{\mathbf{p}}$ are given by eqs. (3.5) and (4.5). Note that the expression in the curly bracket is a completely well-behaved and non-negative density. In particular, the θ function in the pole term excludes the momenta $\mathbf{p}$ that are not in the inner region of the Brillouin zone.

We finally need to show that the measure (1.5) is non-negative. To this end we first remark that $\sinh E \geq |\mathring{\mathbf{p}}|$ at all points $(E, \mathbf{p})$ in the support of the spectral density. For the continuous part of the spectrum this follows from the discussion at the end of sect. 3, and in the case of the pole term, the inequality is an immediate

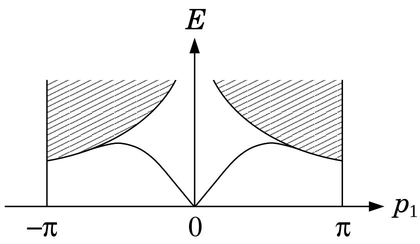

Fig. 3. Support of the spectral density (5.2) at $p_2 = p_3 = 0$. The physical pole (curves starting from the origin) disappears in the continuum (shaded areas) precisely at the boundary of the inner region of the Brillouin zone.

consequence of the definition (4.5) of $\omega_{\mathbf{p}}$. Now since

$$\zeta^\dagger \gamma_0 \left(\gamma_0 \sinh E - i\gamma_k \mathring{p}_k\right) \zeta \geq 0 \tag{5.3}$$

for all complex Dirac spinors ζ if $\sinh E \geq |\mathring{\mathbf{p}}|$, and since this is guaranteed in the support of the spectral density, it follows that the measure (1.5) is non-negative.

6. Remarks

It seems likely that the unitarity of the theory considered in this note derives from the existence of a transfer matrix in the 4+1 dimensional (domain-wall) formulation of the theory. Evidently the transfer matrix has to go in the physical time direction in this case, not in the fifth dimension. The inclusion of the gauge field should not interfere with this, i.e. one may have both exact chiral symmetry and unitarity in lattice QCD (needs to be checked).

In the case of chiral lattice gauge theories the situation is more complicated and it is not clear at present whether unitarity is preserved in the current set-up. One may, however, be able to construct these theories directly in a hamiltonian framework, starting with a hamiltonian formulation of domain-wall fermions. The continuous spectrum that we have found above can presumably be traced back to the heavy modes of such a hamiltonian system (the spectrum is continuous, because the extent of space-time in the extra dimension is taken to infinity).

Appendix A

A.1 Indices and Dirac matrices

Lorentz indices $\mu, \nu, \ldots$ are taken from the middle of the Greek alphabet and run from 0 to 3. Spatial vectors are printed in bold face and have components labelled by Latin indices $k, l, \ldots$ ranging from 1 to 3. For the Dirac indices Greek letters from the beginning of the alphabet are used. Unless stated otherwise the Einstein summation convention is applied and scalar products are taken with euclidean metric.

The symbol $\epsilon_{\mu\nu\rho\sigma}$ denotes the totally antisymmetric tensor with $\epsilon_{0123} = 1$ and the conventions for the Dirac matrices are

$$(\gamma_\mu)^\dagger = \gamma_\mu, \qquad \{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}, \qquad \gamma_5 = \gamma_0\gamma_1\gamma_2\gamma_3. \tag{A.1}$$

In particular, γ_5 is hermitian and $(\gamma_5)^2 = 1$.

A.2 Lattice derivatives

The forward and backward difference operators ∂_μ and ∂_μ^* act on lattice functions according to

$$\partial_\mu f(x) = \frac{1}{a}\{f(x + a\hat{\mu}) - f(x)\}, \tag{A.2}$$

$$\partial_\mu^* f(x) = \frac{1}{a}\{f(x) - f(x - a\hat{\mu})\}, \tag{A.3}$$

where $\hat{\mu}$ denotes the unit vector in direction μ. The eigenfunctions of these operators are the plane waves, and with the standard abbreviations

$$\mathring{p}_\mu = (1/a)\sin(ap_\mu), \qquad \hat{p}_\mu = (2/a)\sin(ap_\mu/2), \tag{A.4}$$

the derivatives that occur in the Wilson-Dirac operator are given by

$$\tfrac{1}{2}(\partial_\mu^* + \partial_\mu)\,\mathrm{e}^{ipx} = i\mathring{p}_\mu \mathrm{e}^{ipx}, \qquad \partial_\mu^*\partial_\mu\,\mathrm{e}^{ipx} = -\hat{p}^2\mathrm{e}^{ipx}. \tag{A.5}$$

Using this result it is straightforward to obtain the momentum space representation (2.1) of the propagator.

References

[1] P. H. Ginsparg, K. G. Wilson, Phys. Rev. D25 (1982) 2649

[2] P. Hasenfratz, Nucl. Phys. B (Proc. Suppl.) 63A-C (1998) 53; Nucl. Phys. B525 (1998) 401

[3] H. Neuberger, Phys. Lett. B417 (1998) 141; *ibid.* B427 (1998) 353

[4] M. Lüscher, Phys. Lett. B428 (1998) 342

[5] F. Niedermayer, Nucl. Phys. (Proc. Suppl.) 73 (1999) 105