Chiral Fermions on the Lattice: A Flatlander's Ascent into Five Dimensions

Urs Wenger[a]*

[a]John von Neumann Institut für Computing NIC, Platanenallee 6, D-15738 Zeuthen, Germany

We briefly review how exact chiral symmetry is realised on the lattice, in particular how chiral fermions are implemented in the context of lattice QCD. We discuss how the theory can be formulated in five dimensions and this provides a theoretical framework within which one can compare algorithmic alternatives for simulating light chiral fermions in QCD on the lattice.

1. IN THE FLATLAND

1.1. Lattice formulation of QCD

Quantumchromodynamics (QCD) is formally described by the Lagrange density

$$\mathcal{L}_{QCD} = \bar{\psi}(i\not{D} - m_q)\psi - \frac{1}{4}G_{\mu\nu}G^{\mu\nu}$$

where $\bar{\psi}, \psi$ describe the fermionic degrees of freedom of the quarks, m_q the bare quark mass, $\not{D}$ the Dirac operator and $G_{\mu\nu}$ is the gauge field tensor. At this point the quantum field theory associated with the Lagrange density is mathematically not well defined. To make it meaningful it has to be regularised and renormalised. One particular (non-perturbative) regularisation is provided by the lattice: Euclidean space-time is discretised on a hypercubic lattice with lattice spacing a. The fermionic degrees of freedom are defined only at the lattice sites, derivatives become finite differences and integrals become sums. The gauge potentials A_μ in $G_{\mu\nu}$ become link matrices $U_\mu \in$ SU(3) living on the links between the lattice sites. With this non-perturbative, gauge invariant regularisation it is possible to define the partition function

$$Z = \int \left(\mathcal{D}U\mathcal{D}\overline{\psi}\mathcal{D}\psi\right) \mathrm{e}^{-S[U;\overline{\psi},\psi]}$$

where $S[U;\overline{\psi},\psi]$ is now the Euclidean QCD lattice action associated with the Lagrangian above.

*This talk is based on work done in collaboration with B. Joó and A.D. Kennedy (Edinburgh University).

The partition function describes a mathematically well defined theory that can be used to study (numerically) the low energy physics of QCD. In order to describe continuum physics one has to renormalise the theory and remove the lattice regulator: this is achieved by taking the lattice spacing $a \to 0$ while keeping physical quantities fixed. Poincaré symmetries are restored automatically in the continuum limit, but the naive discretisation of the Dirac operator introduces additional unphysical fermion excitations, so-called doublers, which however can be eliminated at the expense of explicitly breaking chiral symmetry on the lattice at $m_q = 0$. As a consequence the restoration of chiral symmetry in the continuum requires a fine tuning of the bare quark mass to its critical value.

1.2. On-shell chiral symmetry

It is possible to have chiral symmetry on the lattice without doublers if one only insists that the symmetry holds on-shell. Such an on-shell chiral symmetry transformation should be of the form [1]

$$\psi \to e^{i\alpha\gamma_5(1-aD)}\psi; \quad \overline{\psi} \to \overline{\psi}e^{i\alpha(1-aD)\gamma_5}$$

and the Dirac operator D on the lattice must be invariant:

$$D \to e^{i\alpha(1-aD)\gamma_5} D e^{i\alpha\gamma_5(1-aD)} = D.$$

For an infinitesimal transformation this implies that

$$(1-aD)\gamma_5 D + D\gamma_5(1-aD) = 0$$

which is the Ginsparg-Wilson (GW) relation [2]

$$\gamma_5 D + D\gamma_5 = 2aD\gamma_5 D.$$

One can find a solution D_{GW} of the GW relation as follows. Let the lattice Dirac operator be γ_5-hermitian, $D_{GW}^\dagger = \gamma_5 D_{GW} \gamma_5$. Then the operator

$$aD_{GW} = \frac{1}{2}(1 + \gamma_5\hat{\gamma}_5)$$

with $\hat{\gamma}_5^\dagger = \hat{\gamma}_5$ satisfies the GW relation if $\hat{\gamma}_5^2 = 1$. Furthermore it must have the correct naive continuum limit $D_{GW} \to \not{\partial}$ when $a \to 0$, i.e.,

$$\hat{\gamma}_5 = \gamma_5(2a\not{\partial} - 1) + O(a^2).$$

Both conditions are satisfied if we define

$$\hat{\gamma}_5 = \frac{H_w}{\sqrt{H_w^\dagger H_w}} = \mathrm{sgn}\,[H_w]$$

where the kernel operator $H_w \equiv \gamma_5(D_w - 1)$ is the hermitian Wilson Dirac operator with its negative mass -1 at the lattice cut-off. The resulting massless overlap Dirac operator [3,4] eventually reads

$$D_{GW} = \frac{1}{2}\left(1 + \gamma_5\,\mathrm{sgn}\,[H_w]\right)$$

and a bare quark mass μ can be introduced by defining

$$D(\mu) = \left(1 - \frac{\mu}{2}\right) D_{GW} + \mu.$$

As a consequence of the exact lattice chiral symmetry the operator has exact zero modes with exact chirality and hence fulfils an exact index theorem [5]. There is no mixing of four-fermion operators in different chiral representations, no additive mass renormalisation and hence no tuning of the bare quark mass [6].

1.3. Approximations and representations

The sign-function above is a matrix function of the sparse, but huge kernel operator and hence it can not be computed directly. Instead one employs an approximation of the sign-function which in practice can be evaluated by iterative applications of the matrix.

The explicit construction of the overlap operator then allows for three different variations. Firstly, one can choose different kernel operators – here the only requirements are that the kernel has the correct continuum limit and enjoys γ_5-hermiticity. Obviously the Wilson Dirac operator used above is a straightforward choice but there are certainly other choices available. Secondly, one can use different approximations to the sign function, e.g. Chebyshev's polynomial approximations or the rational approximations by Zolotarev or Higham. And thirdly, one can choose a certain representation of the (rational) approximation, e.g. a continued fraction, a partial fraction or a (Euclidean) Cayley transform representation. For the overlap operator the choice of the representation is just a practical one and the partial fraction representation is in fact the most convenient choice. However, when we ascent into five dimensions, we shall see below that the choice of the representation has crucial consequences since the representation determines the explicit form and structure of the corresponding five-dimensional (5D) chiral fermion operator. More precisely, each of the above mentioned four-dimensional (4D) representations leads to a completely different class of 5D operators with completely different symmetry and transformation properties.

2. INTO FIVE DIMENSIONS

Let us consider a rational approximation to the sign-function of the hermitian operator H, $\mathrm{sgn}(H) \simeq R_{n,m}(H) = \frac{P_n(H)}{Q_m(H)}$, where n and m denote the order of the numerator and denominator polynomial, respectively. The inverse of H and hence the denominator polynomial is in general not explicitly available, but only through some iterative Krylov space procedure. The key to a practical evaluation of $R_{n,m}(H)$ is then provided by its partial fraction decomposition

$$R_{2n+1,2n}(H) = x\left(c_0 + \sum_{k=1}^{n} \frac{c_k}{H^2 + q_k}\right)$$

which can be obtained from the rational function by matching poles and residues. The partial fraction can now be evaluated by using multi-shift linear system solvers which calculate all the terms

$H^2 + q_k$ at one fell swoop. However, since all the physics requires the inverse of $D(\mu)$, e.g. for the calculation of propagators or forces in the course of a molecular dynamics evolution, this leads to a two level nested linear system solution, i.e. a complicated nested Krylov space procedure. It turns out that this can be avoided by introducing a tower of auxiliary fields living in a fictitious fifth dimension hence leading to a 5D representation of the sign-function. As a consequence the 4D nested Krylov space problem reduces to finding a solution in a single 5D Krylov space.

The key to our ascent into five dimensions is provided by the Schur complement which yields the detailed connection between the 4D representations and the corresponding 5D operators.

2.1. Schur complement

Let us consider the block matrix

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where the blocks A, B, C, D represent matrices acting in 4D and thus M can be regarded as a matrix acting in 5D. We may block diagonalise M by a LDU decomposition (Gaussian elimination) as

$$M = \begin{pmatrix} 1 & 0 \\ CA^{-1} & 1 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} 1 & A^{-1}B \\ 0 & 1 \end{pmatrix}$$

where the bottom right block of the block diagonal matrix is the Schur complement $S \equiv D - CA^{-1}B$.

The machinery of the Schur complement now paves the way for two applications. Firstly, from the LDU decomposition above it is clear that inverting the Schur complement S (in other words the 4D overlap operator) is equivalent to inverting directly the 5D matrix M. This comes about by noting that the inverse of L is simply

$$L^{-1} = \begin{pmatrix} 1 & 0 \\ -CA^{-1} & 1 \end{pmatrix}.$$

Then considering the solution of the linear system

$$M \begin{pmatrix} \phi \\ \psi \end{pmatrix} = \begin{pmatrix} 0 \\ \chi \end{pmatrix}$$

by using the LDU decomposition, we find that

$$L^{-1} \begin{pmatrix} 0 \\ \chi \end{pmatrix} = \begin{pmatrix} 0 \\ \chi \end{pmatrix}, \quad U \begin{pmatrix} \phi \\ \psi \end{pmatrix} = \begin{pmatrix} \tilde{\phi} \\ \psi \end{pmatrix}$$

and hence

$$M \begin{pmatrix} \phi \\ \psi \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} \tilde{\phi} \\ \psi \end{pmatrix} = \begin{pmatrix} 0 \\ \chi \end{pmatrix}.$$

So if we are interested in inverting the Schur complement S on some source χ, i.e. $\psi = S^{-1}\chi$, we can obtain the solution ψ also by inverting directly the 5D matrix M. The 4D solution is simply one particular 4D component of the 5D solution vector. Obviously, in this way the nested inversions in 4D are completely avoided at the expense of inverting a 5D matrix.

The second application to which the Schur complement leads the way is the construction of the effective 4D theory from the 5D one. This is accomplished by expressing the determinant of the 5D block matrix M in terms of the determinants of the 4D blocks. By noting that $\det U = \det L = 1$ we derive from the LDU decomposition

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A) \det(D - CA^{-1}B).$$

So on the level of the determinants, and hence of the effective fermionic action, the 5D operator M describes the same theory as the 4D operator represented by the Schur complement, i.e. the 4D overlap operator.

2.2. Continued fractions

Consider now a 5D matrix of the form

$$M = \begin{pmatrix} A_0 & 1 & 0 & 0 \\ 1 & A_1 & 1 & 0 \\ 0 & 1 & A_2 & 1 \\ 0 & 0 & 1 & A_3 \end{pmatrix}$$

and its LDU decomposition

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ S_0^{-1} & 1 & 0 & 0 \\ 0 & S_1^{-1} & 1 & 0 \\ 0 & 0 & S_2^{-1} & 1 \end{pmatrix}$$

$$D = \begin{pmatrix} S_0 & 0 & 0 & 0 \\ 0 & S_1 & 0 & 0 \\ 0 & 0 & S_2 & 0 \\ 0 & 0 & 0 & S_3 \end{pmatrix}$$

$$U = \begin{pmatrix} 1 & S_0^{-1} & 0 & 0 \\ 0 & 1 & S_1^{-1} & 0 \\ 0 & 0 & 1 & S_2^{-1} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

where $S_0 = A_0$; $S_n + \frac{1}{S_{n-1}} = A_n$. Then the Schur complement of the matrix is the continued fraction

$$S_3 = A_3 - \frac{1}{A_2 - \frac{1}{A_1 - \frac{1}{A_0}}}.$$

We may now use this representation to linearise the continued fraction representation of an approximation to the sign function,

$$\mathrm{sgn}(H) \simeq k_0 H + \frac{1}{k_1 H + \frac{1}{k_2 H + \ddots + \frac{1}{k_n H}}},$$

as the Schur complement of the 5D matrix

$$\begin{pmatrix} -k_n H & 1 & & & \\ 1 & \ddots & & & \\ & & k_2 H & 1 & \\ & & 1 & -k_1 H & 1 \\ & & & 1 & k_0 H \end{pmatrix}.$$

Analogously, one can define a 5D matrix such that its Schur complement is exactly the 4D overlap operator where the approximation of the sign-function is represented as a continued fraction. In fact, due to the invariance of continued fractions under equivalence transformations [7,8], we find that there is a whole class of 5D operators which all reduce to the same 4D overlap operator.

2.3. Partial fractions

Now consider a 5D matrix of the form

$$M = \begin{pmatrix} A_1 & 1 & 0 & 0 & 1 \\ 1 & -B_1 & 0 & 0 & 0 \\ 0 & 0 & A_2 & 1 & 1 \\ 0 & 0 & 1 & -B_2 & 0 \\ -1 & 0 & -1 & 0 & R \end{pmatrix}$$

where $A_i = \frac{x}{p_i}, B_i = \frac{p_i x}{q_i}$. As before we can compute its LDU decomposition and find its Schur complement

$$R + \frac{p_1 x}{x^2 + q_1} + \frac{p_2 x}{x^2 + q_2},$$

so we can use this representation to linearise the partial fraction approximation to the sgn-function

$$\mathrm{sgn}(H) \simeq H \sum_j \frac{p_j}{H^2 + q_j}.$$

So again, one can define a 5D matrix such that its Schur complement is exactly the 4D overlap operator where the approximation of the sign-function is represented as a partial fraction. In fact, due to the invariance of partial fractions under certain transformations, we find that there is a whole class of 5D operators which all reduce to the same 4D overlap operator.

2.4. Euclidean Cayley transform

Finally consider a 5D matrix of the (transfer matrix) form

$$M = \begin{pmatrix} 1 & -A_1 & 0 & 0 \\ 0 & 1 & -A_2 & 0 \\ 0 & 0 & 1 & -A_3 \\ -A_0 & 0 & 0 & C \end{pmatrix}$$

with its Schur complement $C - A_3 A_2 A_1 A_0$. We can then use this representation to linearise the (Euclidean) Cayley transform of the approximation to the sgn-function,

$$\mathrm{sgn}(H) \simeq \frac{1 - \prod_j T_j(H)}{1 + \prod_j T_j(H)},$$

where $T_j(H) = A_j(H)^{-1} = (\omega_j - H)/(\omega_j + H)$ and so this corresponds to the standard domain wall fermion formulation [9,10]. As before, due to the invariance of the Euclidean Cayley transform under certain transformations, we find that there is a whole class of 5D domain wall fermion operators which all reduce to the same 4D overlap operator.

3. THE VIEW FROM ABOVE

3.1. Panorama view

Let us stop here for a moment and contemplate what we see from above. We see that each representation of the rational approximation to the sign-function leads to a different 5D Dirac operator. They all have the same 4D, effective lattice fermion Dirac operator, namely the overlap

Dirac operator which satisfies the GW relation and hence guarantees an exact chiral symmetry on the lattice. We also see that each 5D operator has different symmetry properties which stem from the different properties of the individual representations. Due to this, each 5D operator has a very different spectrum and condition number and therefore a very different calculational behaviour, e.g. when calculating inverses, even though in 4D all the representations are completely equivalent.

The fact that there exist 5D formulations of lattice chiral fermions other than the standard domain wall formulation naturally raises the question about the physical significance of the fifth dimension. Is it simply a technical tool to handle a complicated 4D operator, or is it physically meaningful? Domain wall fermions point towards the latter being true, since they provide a beautiful and compelling mechanism for creating chiral fermions in 4D. Accepting this it is then natural to ask what the physical meaning of continued or partial fraction fermions is.

3.2. Chiral symmetry breaking

Any approximation of the overlap operator will destroy the exact lattice chiral symmetry. Such chiral symmetry breaking can be detected by examining the violation of the Ginsparg-Wilson relation in the chiral limit,

$$\gamma_5 D + D\gamma_5 - 2aD\gamma_5 D = \gamma_5 \Delta$$

where the r.h.s. is called the Ginsparg-Wilson defect. For an approximate overlap operator of the form $aD = \frac{1}{2}(1 + \gamma_5 R_n(H))$ the defect simply yields $a\Delta_n = \frac{1}{2}(1 - R_n(H)^2)$.

Any explicit chiral symmetry breaking manifests itself in terms of the defect Δ or, to be more precise, in terms of moments of Δ in expectation values of the operators under consideration. For example the residual quark mass m_{res} in the chiral limit is simply given by a first moment of Δ_n with respect to the pion propagator G,

$$m_{res} = \frac{\langle G^\dagger \Delta_n G\rangle}{\langle G^\dagger G\rangle}.$$

Here we only note that the residual quark mass can be calculated directly in four and five dimensions and provides a way to compare the various formulations and determine their efficiency in reducing the residual chiral symmetry breaking at a fixed extent of the fifth dimension [11]

Obviously, for other physical quantities higher moments might be important and in fact more relevant in order to quantify the effects from the explicit chiral symmetry breaking.

3.3. Summary and conclusions

We have a thorough understanding of various 5D formulations of chiral fermions on the lattice. In five dimensions there are more possibilities to formulate chiral fermions on the lattice than in four dimensions. While from a 4D view the formulations are all equivalent, from a 5D view they look very different. Computationally, better alternatives than the commonly used domain wall fermions seem to exist.

Finally, for dynamical simulations of QCD with light chiral sea quarks using Molecular Dynamics and Hybrid Monte Carlo, it is an open question whether it is preferable to follow a 5D or a 4D strategy.

REFERENCES

1. M. Luscher, Phys. Lett. B428 (1998) 342, hep-lat/9802011.
2. P.H. Ginsparg and K.G. Wilson, Phys. Rev. D25 (1982) 2649.
3. R. Narayanan and H. Neuberger, Nucl. Phys. B443 (1995) 305, hep-th/9411108.
4. H. Neuberger, Phys. Lett. B417 (1998) 141, hep-lat/9707022.
5. P. Hasenfratz, V. Laliena and F. Niedermayer, Phys. Lett. B427 (1998) 125, hep-lat/9801021.
6. P. Hasenfratz, Nucl. Phys. B525 (1998) 401, hep-lat/9802007.
7. A. Borici et al., Nucl. Phys. Proc. Suppl. 106 (2002) 757, hep-lat/0110070.
8. U. Wenger, Proc. 3rd QCDNA Workshop, Springer, Berlin, 2005, hep-lat/0403003.
9. D.B. Kaplan, Phys. Lett. B288 (1992) 342, hep-lat/9206013.
10. Y. Shamir, Nucl. Phys. B406 (1993) 90, hep-lat/9303005.
11. R.G. Edwards, B. Joó, A.D. Kennedy,

K. Orginos and U. Wenger, in preparation.