Chiral anomalies and Wilson fermions

Michael Creutz

Senior Physicist Emeritus

Physics Department

Brookhaven National Laboratory

Upton, NY 11973, USA *

(Dated: November 28, 2025)

Abstract

The Wilson formulation of fermions in lattice gauge theory provides a unified description of the chiral anomalies in the standard model. The discrete Dirac operator diagonalizes into a series of two by two blocks. In each block the possible eigenvalues either form a complex pair or separate into two real eigenvalues that have specific chirality. The collision of these pairs of eigenvalues occurs outside the perturbative region and provides a path between topological sectors.

I. INTRODUCTION

The importance of chiral anomalies to the standard model has a long history, going back to the calculation of the decay of the neutral pion to two photons by Steinberger [1]. Naive arguments suggest that this should be suppressed when quark masses are small. This issue was elucidated in the classic papers of Adler, Bell, and Jackiw [2–4].

In the context of the strong interactions, anomalies associated with the axial current emerged as crucial to understanding why the eta prime meson does not behave as a pseudo-Goldstone boson, unlike the pions and kaons. Also, a portion of the proton mass arises via a similar mechanism.

Perhaps the most unexpected consequence of chiral anomalies is the prediction by 't Hooft [5] that proton decay should occur in the standard model of the weak interactions. This rate is extremely small, perhaps unmeasurable, but its existence is a necessary consequence of the theory.

All these phenomena are tied to an anomaly with a simple change of fermionic variables in path integral

$$\begin{aligned} \psi &\longrightarrow e^{i\phi\gamma_5}\psi \\ \overline{\psi} &\longrightarrow \overline{\psi}e^{i\phi\gamma_5} \end{aligned} . \tag{1}$$

This is a symmetry of the classical standard model Lagrangian when the fermion masses vanish. However all renormalization schemes must violate this symmetry. In particular left handed fields, $\psi_L = (1-\gamma_5)\psi/2$, ultimately must mix with right handed ones, $\psi_R = (1+\gamma_5)\psi/2$, although the mechanism for this depends on the details of the regulator.

The Wilson lattice approach provides a specific regularization making the Feynman path integral mathematically well defined. As such the approach must allow for the effects of

* mike@latticeguy.net

anomalies. A summary of the long history of this problem appears in Ref. [6]. The Smit-Swift model [7, 8] when extended to include an entire generation of fermions provides a specific lattice formulation for the standard model [9], based on three commuting gauge groups $SU(3) \otimes SU(2). \otimes U(1)$. Here we expand on that presentation with more details on the connection between anomalies and colliding eigenvalues of the Dirac matrix. Among various interesting issues is where does the baryon decay predicted by 't Hooft [5] arise? We will see that it comes from multiple fermion combinations simultaneously coupling in a gauge invariant way with complex configurations of Higgs and gauge fields.

A typical lattice gauge simulation generates configurations for the gauge and Higgs fields and then explores Dirac operator on these configurations. In something like the hybrid Monte Carlo method [10] for a dynamical simulation, feedback is included from the fermion loops. Here we will concentrate on the Dirac operator itself and not be concerned with the details of generating the relevant gauge and Higgs configurations.

II. DOUBLERS AND THE WILSON TRICK

In lattice gauge theory, the free fermion action starts with the derivative term replaced by nearest neighbor differences

$$i\partial_\mu \gamma_\mu \psi = \!\!\longrightarrow \sum_\mu \frac{i\gamma_\mu}{a}(\psi_{i+e_\mu} - \psi_i) \tag{2}$$

The lattice sites are labeled i, space-time directions μ, and lattice spacing a. Before including bosonic fields, this difference takes a particularly simple form in momentum space:

$$p_\mu \longrightarrow \frac{1}{a}\sin(ap_\mu) = p_\mu + O(a) \tag{3}$$

The natural range of each component of lattice momentum is. $-\frac{\pi}{a} < p_\mu \le \frac{\pi}{a}$ Unfortunately within this range there is an extra pole in the propagator at $p_\mu \sim \frac{\pi}{a}$. The Wilson [11] solution to this issue removes this pole by giving the doubler a large mass

$$i\gamma_\mu p_\mu \longrightarrow \frac{1}{a}(i\gamma_\mu \sin(ap_\mu) + 1 - \cos(ap_\mu)) \tag{4}$$

Effectively this adds a $O(ap^2)$ correction to the naive result, a formally irrelevant operator $\sim \overline{\psi}\partial^2\psi$.

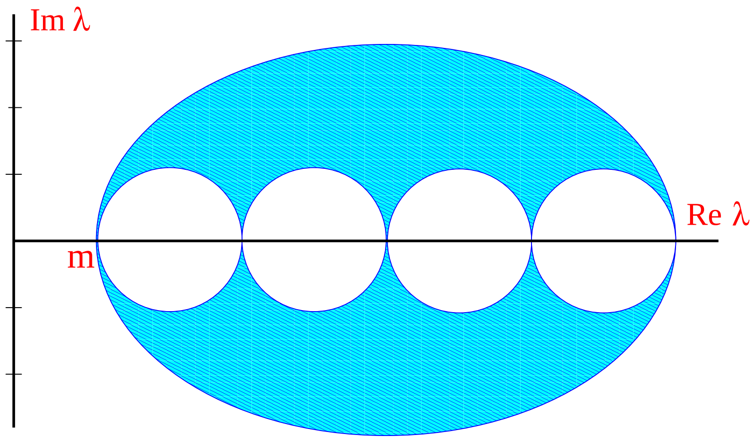

FIG. 1. The spectrum of the free four dimensional Dirac operator. The crossings of the real axis correspond to how many doubler momentum components are at π/a. The multiplicities at these points follow the pattern 1, 4, 6, 4, 1.

The added heavy states explicitly break chiral symmetry and are ultimately the source of anomalies. One might think of them as an analogy of the heavy states in a Pauli-Villars [12] approach.

In the usual continuum theory, $i\not{D} + m$ is an anti-hermitian kinetic term plus a real constant mass term, As such, all eigenvalues lie on a line at fixed real part m. On the lattice the Wilson term smears the mass term at high momentum. For the free lattice theory the sin and cosine factors in the above equations give rise to a set of nested circles as shown in Fig. 1.

When interactions turn on, the eigenvalues move around and momentum is no longer a useful parameter. Indeed, the Dirac operator is a non-normal operator with $[D, D^\dagger] \neq 0$. This comes about because of the complex gauge factors as the fermions hop around. Indeed, different eigenvectors of D need not be orthogonal. Also the right and left eigenvectors are not generally the same.

III. GAMMA-FIVE HERMITICITY

Although the lattice Dirac operator is neither Hermitian or normal, it does satisfy what is called gamma-5 hermiticity

$$\gamma_5 D = D^\dagger \gamma_5. \tag{5}$$

The kinetic or "hopping" terms involving gauge fields are anti-Hermitian and contain a gamma matrix factor while the mass and Wilson term are Hermitian and commute with γ_5.

This "symmetry" has important consequences for the eigenvalue spectrum of the Dirac operator. Consider some particular complex right eigenvalue $D|\psi\rangle = \lambda|\psi_1\rangle$. Playing with the characteristic equation tells us

$$0 = |D - \lambda| = |D^\dagger - \lambda^*| = |D - \lambda^*| = |D^\dagger - \lambda| \tag{6}$$

Thus the complex conjugate of λ is also an eigenvalue, and both D and $D^\dagger$ have the same eigenvalues. Of course since our matrix is not normal, the various eigenvectors are not in general the same or even orthogonal.

To proceed it is convenient to consider a two dimensional space spanned by our eigenvector $|\psi_1\rangle$ and $|\psi_2\rangle \equiv \gamma_5|\psi_1\rangle$. Since $\gamma_5^2 = 1$, these each have the same norm. By definition these satisfy

$$\begin{array}{c} D|\psi_1\rangle = \lambda|\psi_1\rangle \\ D^\dagger|\psi_2\rangle = \lambda|\psi_2\rangle \end{array}. \tag{7}$$

Conjugation gives

$$\begin{array}{c} \langle\psi_1^\dagger|D^\dagger = \lambda^*\langle\psi_1^\dagger| \\ \langle\psi_2^\dagger|D = \lambda^*\langle\psi_2^\dagger| \end{array} \tag{8}$$

From these, when λ is complex,

$$\langle\psi_2^\dagger|D|\psi_1\rangle = \lambda\langle\psi_2^\dagger|\psi_1\rangle = \lambda^*\langle\psi_2^\dagger|\psi_1\rangle = 0 \tag{9}$$

and our basis is in fact orthogonal. On this two dimensional space

$$\begin{array}{c} D = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^* \end{pmatrix} \\ D^\dagger = \begin{pmatrix} \lambda^* & 0 \\ 0 & \lambda \end{pmatrix} \end{array}. \tag{10}$$

The Dirac matrix is thus block diagonalized into 2 by 2 blocks.

A complication occurs when the basis becomes singular with $|\psi_1\rangle = |\psi_2\rangle$. To treat this situation it is useful to mix the modes into a "chiral" basis

$$\begin{array}{c} |\psi_+\rangle = (|\psi_1\rangle + |\psi_2\rangle)/2 = \frac{1+\gamma_5}{2}|\psi_1\rangle \\ |\psi_-\rangle = (|\psi_1\rangle - |\psi_2\rangle)/2 = \frac{1-\gamma_5}{2}|\psi_1\rangle \end{array}. \tag{11}$$

In this basis γ_5 reduces to the Pauli matrix σ_3.

We now write our sub-matrix of D as the most general 2 by 2 matrix satisfying $D^\dagger = \sigma_3 D \sigma_3$

$$D \to a_0 + ia_1\sigma_1 + ia_2\sigma_2 + a_3\sigma_3 \tag{12}$$

where the coefficients a_i are real. This is easily diagonalized with eigenvalues

$$\lambda = a_0 \pm i\sqrt{a_1^2 + a_2^2 - a_3^2} \tag{13}$$

The contribution to fermion determinant is the expected

$$|D| = a_0^2 + a_1^2 + a_2^2 - a_3^2 = \lambda\lambda^*. \tag{14}$$

The gamma-five hermeticity also implies

$$H = \gamma_5 D \tag{15}$$

is a hermitian quantity and always has real eigenvalues. This is also contained in our 2 by 2 blocks, with eigenvalues

$$\lambda_H = a_3 \pm \sqrt{a_0^2 + a_1^2 + a_2^2} \tag{16}$$

The real parts of the eigenvalues of D are constrained to lie between these.

The parameter space of the four a's separates nicely into two domains. When $a_1^2 + a_2^2 > a_3^2$ we have two complex eigenvalues. However when $a_1^2 + a_2^2 < a_3^2$ the eigenvalues become real [19]

$$\lambda = a_0 \pm \sqrt{a_3^2 - a_1^2 - a_2^2}. \tag{17}$$

As the gauge and Higgs fields evolve in a simulation, we can transit between these situations. In the process the two eigenvalues collide and move off in opposite directions on the real axis,as sketched in Fig. 2.

Large values of the parameter a_3 give rise to real modes of D. If we now cool our configuration down to classical fields, these become the topological objects commonly called "instantons" or "sphalerons." Indeed, these are the famous zero modes of the index theorem. Note that γ_5 commutes with set of all real modes. Thus we can simultaneously diagonalize this set with γ_5. The modes are then chiral with $\gamma_5|\psi\rangle = \pm|\psi\rangle$. The modes inherently mix left and right fields and are at the heart of chiral anomalies.

In our two dimensional sub-space, the role of γ_5 is played by σ_3. The free theory contains no terms directly involving γ_5, so this term can only be generated from fermion loops in the determinant of D. Indeed a loop that involves hoppings in all four dimensions can generate such a term. Ref. [14] used the loops circumnavigating hyper-cubes to define a local topological susceptibility. See also Ref. [15]. These loops necessarily mix kinetic hoppings with hoppings involving the Wilson term.

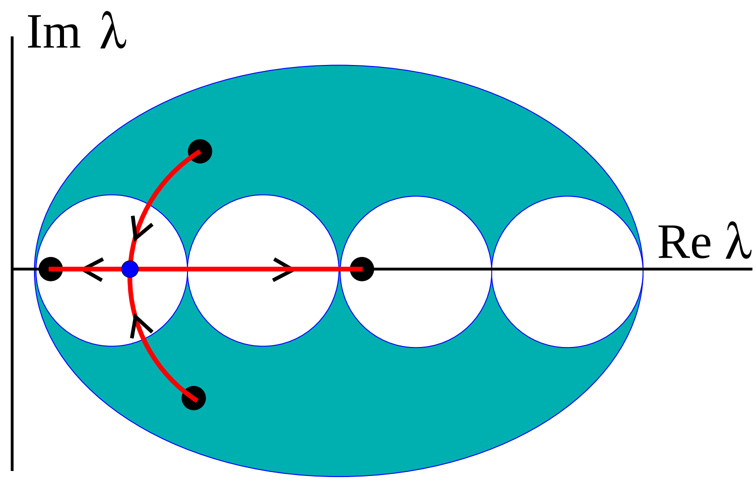

FIG. 2. As gauge fields evolve, two complex eigenvalues can collide and move off on the real axis.

Note that the eigenvalue collision does not occur in perturbation theory. The effect does not occur until chiral effects are sufficiently large. The anomalies do appear immediately after the real modes separate. It is not required to go to the zero mode limit. Classical instantons and the index theorem do prove chiral modes must exist.

If one were to construct Neuberger's overlap operator [16] by projecting all eigenvalues onto a circle, singularities will appear as either of the modes pass through the projection point [17]. In general these singularities do not occur when the eigenvalues collide, but when either mode crosses the circle center.

Small eigenvalues are formally suppressed in the path integral

$$Z = \int (dA)(d\overline{\psi} d\psi)\; e^{-S_g + \overline{\psi} D \psi} = \int (dA)\; e^{-S_g(A)}\; \prod \lambda_i \tag{18}$$

Naively that would suggest that zero modes might be irrelevant. However 't Hooft [5] showed how their effects could be enhanced in physical observables. This can be intuitively seen by adding source terms to the path integral

$$\begin{aligned} Z(\eta, \overline{\eta}) &= \int (dA)\; (d\overline{\psi} d\psi)\; e^{-S_g + \overline{\psi} D \psi + \overline{\psi}\eta + \overline{\eta}\psi} \\ &= \int (dA)\; e^{-S_g + \overline{\eta}\; D^{-1}\; \eta/4}\; \prod \lambda_i \end{aligned} \tag{19}$$

The presence of D^{-1} is how the zero modes remain relevant to anomalies even as the quark masses go to zero.

IV. PHYSICAL CONSEQUENCES

In a simulation D is gauge gauge dependent; so, for physical results we must combine the fields into gauge invariant observables. As usual in lattice gauge theory, we do not consider fixing a gauge. As such the various fields are highly fluctuating. To avoid these fluctuations we need to look at gauge invariant combinations of the fields.

Here we concentrate in observables that are sensitive to the chiral modes under discussion. We briefly discuss effects with each of the three commuting standard model groups $SU(3) \otimes SU(2) \otimes U(1)$.

A. Electrodynamics

First we discuss electrodynamics $U(1)$ where it is well known that one cannot conserve both the vector and axial current. In this case the above modes immediately break the axial symmetry of Eq. 1. In this case it is perhaps simplest to think of the doublers as playing the role of a Pauli-Villars field [12].

B. QCD

Here we consider only one generation of fermions. Thus we have u and d quarks which represent two SU(3) triplets. When they are degenerate the couple to exactly the same gauge and Higgs fields and thus both see the same eigenvalues of D. A chiral mode will flip both spins together. This generates an effective mass term for the eta prime meson as sketched in Fig. 3.

The chiral modes can influence the masses of other gauge singlet particles including the proton. The effective $\overline{\psi}\psi$ operator affects each quark color, generating an effective proton mass even if the quarks are mass-less.

The magnitude of these effects is controlled by the renormalization scale of the strong interactions Λ_{qcd}

$$m_{\eta'}, m_p \propto \Lambda_{qcd} \propto \frac{1}{a}\ g^{-\beta_1/\beta_0^2}\ \exp\left(\frac{-1}{2\beta_0 g^2}\right). \tag{20}$$

This scale is not from classical instantons alone with $\exp(-S_I) = \exp(\frac{-8\pi^2}{g^2})$, but rather is enhanced by fermionic loops and quantum fluctuations. The precise numberical values can

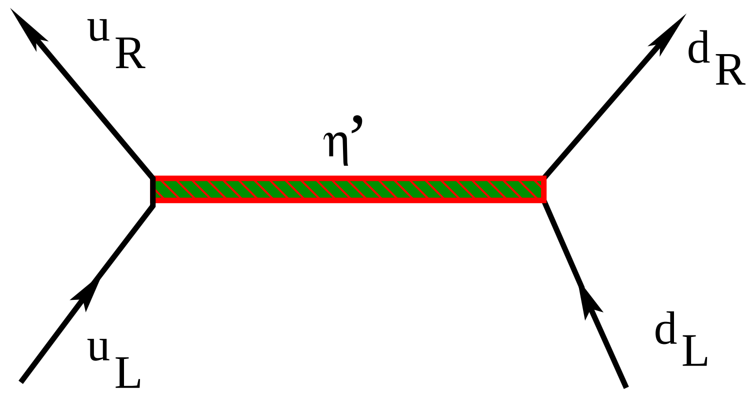

FIG. 3. As the up and down quarks flip spin from the chiral eigenmode, the eta prime acquires an effective mass.

in principle be determined via numerical simulations.

C. Weak $SU(2)$

Now we turn to the group of the weak interactions. Here 't Hooft [5] has shown there is a small contribution to baryon decay. Because the basic electroweak coupling is small, the effects are probably too small to observe.

The first generation standard model is built on four $SU(2)$ left handed doublets

$$\begin{pmatrix} u^r \\ d^r \end{pmatrix}_L \begin{pmatrix} u^g \\ d^g \end{pmatrix}_L \begin{pmatrix} u^b \\ d^b \end{pmatrix}_L \begin{pmatrix} \nu \\ e^- \end{pmatrix}_L \tag{21}$$

where the superscripts r, g, b refer to the strong interaction "colors." Their antiparticles form four right handed doublets $\psi_R^c = \tau_2\gamma_2\psi_L^*$. Explicitly these are

$$\begin{pmatrix} d^{r,c} \\ r^{r,c} \end{pmatrix}_R \begin{pmatrix} d^{g,c} \\ r^{g,c} \end{pmatrix}_R \begin{pmatrix} d^{b,c} \\ r^{b,c} \end{pmatrix}_R \begin{pmatrix} e^+ \\ \nu^c \end{pmatrix}_R \tag{22}$$

All of the above fermion variables have conjugate fields which are independent Grassmann variables in the path integral.

When the gauge and Higgs fields of a configuration form the chiral modes as discussed earlier, all of these doublets can be sensitive to them. To form gauge invariant observables, combine the left handed doublets with the right handed ones to define a four by four matrix of $SU(2)$ singlet combinations

$$T_{ij} = \overline{\psi}_i^R \psi_j^L. \tag{23}$$

To proceed we need to obtain a quantity that is also singlet under the strong and electromagnetic groups. This is accomplished in the determinant of this matrix

$$\det(T_{ij}). \tag{24}$$

This antisymmetrizes over $SU(3)$ colors, making an $SU(3)$ singlet. The quark and lepton charges also cancel making it electrically neutral. So this combination is is a fully gauge invariant operator sensitive the chiral modes. This is the determinant that 't Hooft [5] found to summarize the weak anomalies in the standard model.

This determinant includes terms that couple two quarks to an antiquark and a lepton. [20] This allows for baryon-lepton mixing while preserving all gauged symmetries. This includes $p \leftrightarrow e^+$ and $n \leftrightarrow \overline{\nu}$.

V. SUMMARY

Given a configuration of gauge and Higgs fields in the path integral, the corresponding Dirac operator can be reduced to 2 by 2 blocks, each representing two possible eigenvalues contributing to the determinant of D. As the configuration evolves in a simulation, these modes can collide and transition between complex pairs and real chiral modes. This provides a common mechanism for chiral anomalies in all three gauge symmetries of the standard model. These real modes occur outside the perturbative limit. Cooling procedures to smooth fields drives such modes to known topological configurations, although the anomalies do not require this limit.

An interesting question concerns the counting of non-perturbative topological parameters. With Wilson fermions these are related to relative phases between the fermion masses and the Wilson terms [18]. In QCD flavored chiral rotations can move this phase Theta between different quark species. For example it can be placed entirely in the phase of a single quark, i.e. m_u or even m_t. How many such parameters are in the full standard model? Do both $SU(3)$ of QCD and $SU(2)$ of the weak interactions have independent Theta parameters? And how does QED fit into this picture?

ACKNOWLEDGMENT

This manuscript has been authored under contract number DE-AC02-98CH10886 with the U.S. Department of Energy. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

---

[1] J. Steinberger. On the Use of subtraction fields and the lifetimes of some types of meson decay. *Phys. Rev.*, 76:1180–1186, 1949.

[2] Stephen L. Adler. Axial vector vertex in spinor electrodynamics. *Phys.Rev.*, 177:2426–2438, 1969.

[3] Stephen L. Adler and William A. Bardeen. Absence of higher order corrections in the anomalous axial vector divergence equation. *Phys.Rev.*, 182:1517–1536, 1969.

[4] J.S. Bell and R. Jackiw. A PCAC puzzle: pi0 to gamma gamma in the sigma model. *Nuovo Cim.*, A60:47–61, 1969.

[5] Gerard 't Hooft. Computation of the quantum effects due to a four-dimensional pseudoparticle. *Phys.Rev.*, D14:3432–3450, 1976.

[6] Jan Smit. A confederacy of anomalies. *Eur. Phys. J. H*, 50(1):5, 2025.

[7] J. Smit. Fermions on a Lattice. *Acta Phys. Polon. B*, 17:531, 1986.

[8] P. V. D. Swift. The Electroweak Theory on the Lattice. *Phys. Lett. B*, 145:256–260, 1984.

[9] Michael Creutz. Standard model and the lattice. *Phys. Rev. D*, 109(3):034514, 2024.

[10] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys.Lett.*, B195:216–222, 1987.

[11] Kenneth G. Wilson. Confinement of quarks. *Phys. Rev.*, D10:2445–2459, 1974.

[12] W. Pauli and F. Villars. On the Invariant regularization in relativistic quantum theory. *Rev. Mod. Phys.*, 21:434–444, 1949.

[13] Notice an analogy with the connection between the Lorentz group and the group $SL(2)_C$, where vectors also divide into two classes: time-like and space-like.

[14] Michael Creutz. Anomalies, gauge field topology, and the lattice. *Annals Phys.*, 326:911–925, 2011.

[15] M. Teper. Instantons in the quantized su(2) vacuum: A lattice monte carlo investigation. *Phys.Lett.*, B162:357, 1985.

[16] Herbert Neuberger. Exactly massless quarks on the lattice. *Phys.Lett.*, B417:141–144, 1998.

[17] Michael Creutz. Transiting topological sectors with the overlap. *Nucl. Phys. Proc. Suppl.*, 119:837–839, 2003.

[18] E. Seiler and I.O. Stamatescu. Lattice fermions and Theta vacua. *Phys.Rev.*, D25:2177, 1982.

[19] Notice an analogy with the connection between the Lorentz group and the group $SL(2)_C$, where vectors also divide into two classes: time-like and space-like.

[20] This is consistent with the strong gauge group since in $SU(3)$ the product of two triplet representations contains an antitriplet, i.e. $3 \otimes 3 = 6 \oplus \overline{3}$.