Chiral Symmetry and Lattice Fermions: Lectures 1-2

David B. Kaplan

Institute for Nuclear Theory, University of Washington, Seattle, WA 98195-1550

Lectures delivered at the "Frontiers in Lattice QCD" Summer School
Peking University, June 24-July 12, 2019

Acknowledgements

This work is supported in part by U.S. DOE grant No. DE-FG02-00ER41132.

Contents

1

Chiral symmetry and anomalies in d=1+1.

1.1 Introduction

Chiral symmetries play an important role in the spectrum and phenomenology of both the standard model and various theories for physics beyond the standard model. In many cases chiral symmetry is associated with nonperturbative physics which can only be quantitatively explored in full on a lattice. It is therefore important to implement chiral symmetry on the lattice, which turns out to be less than straightforward. In these lectures I discuss what chiral symmetry is, why it is important, how it is broken, and ways to implement it on the lattice. There have been many hundreds of papers on the subject and this is not an exhaustive review; the limited choice of topics I cover reflects on the scope of my own understanding and not the value of the omitted work.

1.2 Dirac fermions in 1+1 dimensions

1.2.1 γ-matrices

Consider a free, massive Dirac fermion in $d = 1 + 1$ dimensions, whose coordinates we will call $x^0 = t$, $x^1 = x$. The Lagrange density is

$$\mathcal{L} = \bar{\psi} \left(i\partial_\mu \gamma^\mu - m \right) \psi \ , \tag{1.1}$$

where ψ is a 2-component spinor, and the γ-matrices satisfy

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \ , \qquad \eta^{\mu\nu} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \ . \tag{1.2}$$

A convenient representation for the γ-matrices in terms of the Pauli matrices is

$$\gamma^0 = \sigma_1 \ , \qquad \gamma^1 = -i\sigma_2 \ . \tag{1.3}$$

Lorentz transformation of ψ takes the form

$$\psi(x) \to e^{i\frac{1}{2}\omega_{\mu\nu}\sigma^{\mu\nu}} \psi(\Lambda^{-1}x) \ , \qquad \sigma^{\mu\nu} = \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right] \ , \tag{1.4}$$

where $\Lambda_\mu^{\ \nu}(\omega)$ is the corresponding Lorentz transformation matrix for a 2-vector. This is a bit heavy-handed: life in $d = 1 + 1$ dimensions is life on a wire, and the only Lorentz transformations are boosts in the x direction; for

$$\omega_{01} = -\omega_{10} = \theta \tag{1.5}$$

we have

$$\Lambda_\mu^{\ \nu} = \begin{pmatrix} \cosh\theta \ \sinh\theta \\ \sinh\theta \ \cosh\theta \end{pmatrix} , \qquad \sigma^{01} = i\frac{\sigma_3}{2} . \tag{1.6}$$

Note that we can define a third γ-matrix I will denote Γ which is Hermitian and which anticommutes with both γ^μ, the analogue of γ_5 in $d = 3+1$:

$$\Gamma = \Gamma^\dagger , \qquad \Gamma^2 = 1 , \qquad \{\Gamma, \gamma^\mu\} = 0 \quad \Longrightarrow \quad [\Gamma, \sigma^{\mu\nu}] = 0 . \tag{1.7}$$

In the above basis, we can take $\Gamma = \sigma_3$. Since Γ commutes with Lorentz transformations, we conclude that we have a *reducible* representation of the Lorentz group. We can define projection operators

$$P_\pm = \frac{1 \pm \Gamma}{2} , \qquad P_+^2 = P_+ , \qquad P_-^2 = P_- , \qquad P_+ + P_- = 1 . \tag{1.8}$$

(where "1" means the 2×2 unit matrix). Then we define

$$\psi_R = P_+\psi , \quad \psi_L = P_-\psi , \quad \bar\psi_L = \psi_L^\dagger\gamma^0 = \bar\psi P_+ , \quad \bar\psi_R = \psi_R^\dagger\gamma^0 = \bar\psi P_- , \tag{1.9}$$

and we know that $\psi_{L,R}$ will not mix under Lorentz transformations. With the above definitions, writing $\psi = \psi_L + \psi_R$ and plugging back into our Lagrange density in eqn. (1.1) we find we can rewrite it as

$$\mathcal{L} = \bar\psi_L i\partial\!\!\!/\psi_L + \bar\psi_R i\partial\!\!\!/\psi_R - m\left(\bar\psi_L\psi_R + \bar\psi_R\psi_L\right) = \bar\psi_L i\partial\!\!\!/\psi_L + \bar\psi_R i\partial\!\!\!/\psi_R - \left(m\bar\psi_L\psi_R + \text{h.c.}\right) \tag{1.10}$$

1.2.2 The massless case and chiral symmetry

Let's set the fermion mass to zero. We see that the above Lagrange density has two $U(1)$ symmetries: we can rotate the fermions with the two independent phases α and β as

$$\begin{aligned} \psi_L &\to e^{i\alpha}\psi_L , & \bar\psi_L &\to e^{-i\alpha}\bar\psi_L , \\ \psi_R &\to e^{i\beta}\psi_R , & \bar\psi_R &\to e^{-i\beta}\bar\psi_R , \end{aligned} \tag{1.11}$$

without affecting $\mathcal{L}$ in eqn. (1.10), provided that $m = 0$. Especially interesting is that the symmetry persists even if we add gauge interactions, replacing ∂_μ by a gauge covariant derivative D_μ. Therefore, even with gauge interactions, there are apparently two conserved currents

$$j_R^\mu = \bar\psi\gamma^\mu P_+\psi , \qquad j_L^\mu = \bar\psi\gamma^\mu P_-\psi . \tag{1.12}$$

Evidently these are both symmetry currents only for massless fermions, since the mass term in eqn. (1.10) couples ψ_L to ψ_R, and the Lagrange density is not invariant under independent phase rotations. It is useful, therefore, to consider two different linear combinations of these currents, referred to as the vector and axial currents,

$$j^\mu = \bar{\psi}\gamma^\mu\psi \ , \qquad j^\mu_A = \bar{\psi}\gamma^\mu\Gamma\psi \ . \tag{1.13}$$

These two currents correspond to the two independent transformations

$$\psi \to e^{i\theta}\psi \ , \qquad \psi \to e^{i\omega\Gamma}\psi \tag{1.14}$$

respectively. The first conserved quantity is just fermion number, the second, which counts right minus left number, is called axial charge. The fact that $U(1)_A$ axial transformations are a symmetry of the kinetic operator is a consequence of the property

$$\{\Gamma, \not{D}\} = 0 \ . \tag{1.15}$$

Later we will see that on the lattice, even for massless fermions it is not possible to define a kinetic operator analogous to $\not{D}$ which anti-commutes with Γ, but that the above equation will have to be modified.

The reason why both $j^\mu_{L,R}$ currents are conserved for massless fermions is easy to see if we look at the equation of motion for the free massless fermion:

$$0 = i\not{\partial}\psi = i\begin{pmatrix} 0 & \partial_t - \partial_x \\ \partial_t + \partial_x & 0 \end{pmatrix}\psi \ , \tag{1.16}$$

which has plane wave solutions

$$\psi_R = e^{-ik(t-x)}\begin{pmatrix}1\\0\end{pmatrix} \qquad \psi_L = e^{-ik(t+x)}\begin{pmatrix}0\\1\end{pmatrix} \ , \tag{1.17}$$

with $P_+\psi_R = \psi_R$ and $P_-\psi_L = \psi_L$. We see that ψ_R corresponds to fermions moving at the speed of light to the right (positive x-direction) and ψ_L corresponds to particles moving to the left. Clearly, Lorentz boosts cannot change the number of either, which is therefore conserved quantities. Thus we expect to be able to write

$$\partial_\mu j^\mu_{L,R} = \partial_\mu j^\mu = \partial_\mu j^\mu_A = 0 \ . \tag{1.18}$$

We will show below, however, that this is not the case, and that quantum effects spoil some of these conservation laws through "anomalies".

Because ψ_L and ψ_R transform under independent irreducible representations of the Lorentz group, we can consider a theory with just one of them, such as

$$\mathcal{L} = \bar{\psi}_R i\not{D}\psi_R \ , \tag{1.19}$$

which, if gauged as in the above example, would be called an example of a "chiral gauge theory". Note that ψ_R is a 2-component spinor, where the lower component equals zero. So we could have written this as a 1-component fermion, but it is convenient often to write it as a Dirac spinor, with one component projected out by P_+. This theory looks like it should make sense because the gauge field appears to be coupled to a conserved current, but again, anomalies will spoil this assumption.

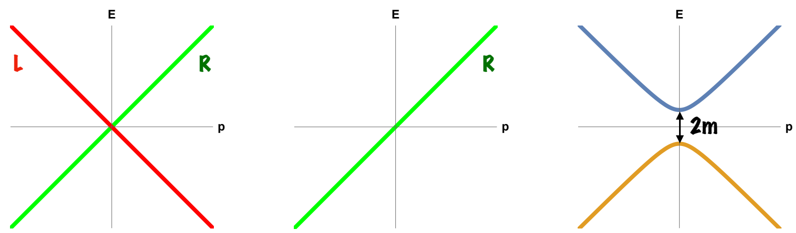

Fig. 1.1 *The spectrum of (i) a free massless Dirac fermion in $d = 1 + 1$, (ii) a free massless RH chiral fermion, (iii) a free massive Dirac fermion. In the first case both LH and RH fermion numbers are independently conserved — or equivalently, both fermion number and axial charge are conserved; in the second case there is a conserved RH fermion number; for the massive case there is only a conserved fermion number.*

1.2.3 The massive Dirac fermion

In contrast to the massless case, if the fermion is massive we can always boost between frames where a left-mover in one frame is a right-mover in the other, and so the number of either cannot be individually conserved. Figure 1.1 shows the spectrum for of free fermions for the cases we have discussed. If you take at the Lagrange density in eqn. (1.10) and perform the axial transformation in eqn. (1.14), you find

$$\mathcal{L} \to \bar{\psi}_L i\partial\!\!\!/\psi_L + \bar{\psi}_R i\partial\!\!\!/\psi_R - \left(me^{2i\omega}\bar{\psi}_L\psi_R + \text{h.c.}\right) \ , \tag{1.20}$$

which is equivalent to rotating the mass term by a phase, $m \to me^{2i\omega}$. (This shows that the phase of the fermion mass has no physical meaning in a theory where it is the only source of axial symmetry violation, since we can change the phase at will with a change of variables.) Noether's theorem then tells us that for the massive case we have

$$\partial_\mu j^\mu = 0, \qquad \partial_\mu j^\mu_A = 2im\bar{\psi}\Gamma\psi \ . \tag{1.21}$$

1.3 The $U(1)_A$ anomaly in $d = 1 + 1$

So far we have blithely assumed that symmetries of the Lagrange density imply symmetries of the quantum theory. However, one of the fascinating features of chiral symmetry is that sometimes it is not a symmetry of the quantum field theory even when it is a symmetry of the Lagrangian. In particular, Noether's theorem can be modified in a theory with an infinite number of degrees of freedom; the modification is called "an anomaly". Anomalies turn out to be very relevant both for phenomenology, and central for understanding the challenges for implementing chiral symmetry in lattice field theory. The reason anomalies affect chiral symmetries is that regularization requires a cut-off on the infinite number of modes above some mass scale, while chiral symmetry is incompatible with fermion masses[1].

[1]Dimensional regularization is not a loophole, since chiral symmetry cannot be analytically continued away from odd space dimensions.

A simple way to derive anomalies (and in some ways, overly simple) is to look at what happens to the ground state of a theory with a single flavor of massless Dirac fermion in $(1+1)$ dimensions in the presence of an electric field. Suppose one adiabatically turns on a constant positive electric field $E(t)$, then later turns it off; the equation of motion for the fermion is [2] $\frac{dp}{dt} = eE(t)$ and the total change in momentum is

$$\Delta p = e \int E(t)\, dt \ . \tag{1.22}$$

Thus the momenta of both left- and right-moving modes increase; if one starts in the ground state of the theory with filled Dirac sea, after the electric field has turned off, both the right-moving and left-moving sea levels have shifted to the right as in Fig. 1.2. The the final state differs from the original by the creation of particle- antiparticle pairs: right moving particles and left moving antiparticles. Thus while there is a fermion current in the final state, fermion number has not changed. This is what one would expect from conservation of the $U(1)$ current:

$$\partial_\mu j^\mu = 0 \ , \tag{1.23}$$

However, recall that right-moving and left-moving particles have positive and negative chirality respectively; therefore the final state in Fig. 1.2 has net axial charge, even though the initial state did not. This is peculiar, since the coupling of the electromagnetic field in the Lagrangian does not violate chirality. We can quantify the effect: if we place the system in a box of size L with periodic boundary conditions, momenta are quantized as $p_n = 2\pi n/L$. The change in axial charge is then

$$\Delta Q_A = 2\frac{\Delta p}{2\pi/L} = \frac{e}{\pi}\int d^2x\, E(t) = \frac{e}{2\pi}\int d^2x\, \epsilon_{\mu\nu}F^{\mu\nu} \ , \tag{1.24}$$

where I expressed the electric field in terms of the field strength F, where $F^{01} = -F^{10} = E$. This can be converted into the local equation using $\Delta Q_A = \int d^2x\, \partial_\mu j^\mu_A$, a modification of eqn. (1.18):

$$\partial_\mu j^\mu_A = \frac{e}{2\pi}\epsilon_{\mu\nu}F^{\mu\nu} \ , \tag{1.25}$$

where in the above equation I have included the classical violation due to a mass term as well. The second term is the axial anomaly in $1+1$ dimensions.

Exercise 1.1 Use the above arguments to derive the anomaly that results if one gauges the axial current instead of the vector current.

[2] While in much of these lectures I will normalize gauge fields so that $D_\mu = \partial_\mu + iA_\mu$, in this section I need to put the gauge coupling back in. If you want to return to the nicer normalization, rescale the gauge field by e so that there is no coupling constant in the covariant derivative and a $1/e^2$ factor appears in front of the gauge action.

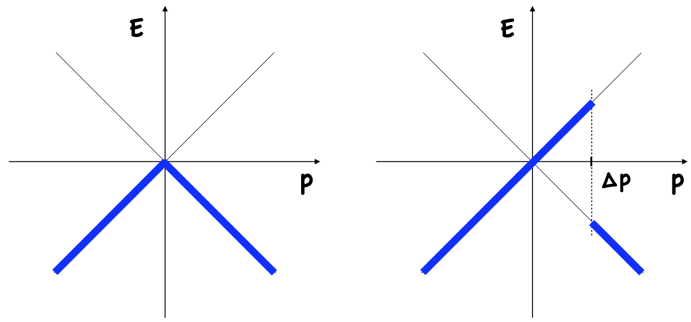

Fig. 1.2 *On the left: the ground state for a theory of a single massless Dirac fermion in* $(1 + 1)$ *dimensions; on the right: the theory after application of an adiabatic electric field with all states shifted to the right by* Δp*, given in eqn.* (1.22). *Filled states are indicated by the heavier blue lines.*

So how did an electric field end up violating chiral charge? Note that this analysis relied on the Dirac sea being infinitely deep. If there had been a finite number of negative energy states, then they would have shifted to higher momentum, but there would have been no change in the axial charge. With an infinite number of degrees of freedom, though, one can have a "Hilbert Hotel": the infinite hotel which can always accommodate another visitor, even when full, by moving each guest to the next room and thereby opening up a room for the newcomer. This should tell you that it will not be straightforward to represent chiral anomalies on the lattice: a lattice field theory approximates quantum field theory with a finite number of degrees of freedom — the lattice may be a big hotel, but it is quite conventional. In such a hotel there can be no anomaly, since there is no ambiguity about how many occupants it has.

This method of deriving the anomaly gives the correct answer, but is a bit too simplistic. For one thing, there is no need to assume that the gauge field must change adiabatically. For another, it doesn't help one figure out what happens in the case where there is a fermion mass and a gap, where the correct answer is that one just adds together the anomalous and explicit symmetry violation, modifying eqn. (1.21) to read

$$\partial_\mu j^\mu_A = 2im\overline{\psi}\Gamma\psi + \frac{e}{2\pi}\epsilon_{\mu\nu}F^{\mu\nu} \ , \tag{1.26}$$

We can derive the anomaly in other ways, such as by computing the anomaly diagram Fig. 1.3, or by following Fujikawa (Fujikawa, 1979; Fujikawa, 1980) and carefully accounting for the Jacobian from the measure of the path integral when performing a chiral transformation. It is particularly instructive for our later discussion of lattice fermions to compute the anomaly in perturbation theory using Pauli-Villars regulators of mass M. Consider the fermion determinant obtained from the path integral in Euclidian spacetime:

$$\det\left(\not{D} + m\right) \ . \tag{1.27}$$

Here we assume hermitian γ-matrices satisfying $\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}$ and so $\not{D}$ is an anti-hermitian operator with unbounded imaginary eigenvalues. The determinant is formally given by

$$\det\left(\not{D} + m\right) = \prod_i (i\lambda_i + m) \ , \tag{1.28}$$

but this is ill defined. To better define it, we consider instead a regulated version,

$$\lim_{M\to\infty} \frac{\det\left(\not{D} + m\right)}{\det\left(\not{D} + M\right)} = \lim_{M\to\infty} \prod_i \frac{(i\lambda_i + m)}{(i\lambda_i + M)} \ . \tag{1.29}$$

Note that at fixed M, for $\lambda_i \gg M \gg m$ the contributions to the regulated determinant all go to factors of 1, so the effect of the regulator is to cancel off contributions from those states. Of course, in the end we take $M \to \infty$ and recover the theory we are interested in. The Feynman rules for this regulated determinant are simple: we just add a heavy "Pauli-Villars" fermion Φ with a Dirac action with mass M, but instead of having a factor of -1 from each closed loop, we get a $+1$ in order to obtain the inverse determinant, as we would get from a boson field. It is important that the Φ have all the same couplings as the fermion, including to external sources. As a result, we should consider the divergence of the regulated axial current

$$j^\mu_{A,\mathrm{reg}} = \overline{\psi}\gamma^\mu\Gamma\psi + \overline{\Phi}\gamma^\mu\Gamma\Phi \ , \tag{1.30}$$

where it follows from Noether's theorem that

$$\partial_\mu j^\mu_{A,\mathrm{reg}} = 2im\overline{\psi}\Gamma\psi + 2iM\overline{\Phi}\Gamma\Phi \ . \tag{1.31}$$

note that Φ contributes a new contribution to axial symmetry breaking proportional to M. However, *this* current does not have any additional anomalous divergence, because we have essentially removed all high λ_i states from consideration and have a "conventional hotel". Therefore, if we are to recover the anaomaly, it must come the Pauli-Villars contribution somehow. As we are interested in matrix elements of $j^\mu_{A,\mathrm{reg}}$ in a background gauge field between states that do not contain any Pauli-Villars particles, we need to evaluate the expectation value $\langle 2iM\overline{\Phi}\Gamma\Phi\rangle$ in a background gauge field and take the limit $M \to \infty$, in order to see if $\partial_\mu j^\mu_{A,\mathrm{reg}}$ picks up any anomalous contributions that do not decouple as we remove the cutoff $M \to \infty$.

To compute $\langle 2iM\overline{\Phi}\Gamma\Phi\rangle$ we need to consider all Feynman diagrams with a Pauli-Villars loop, and insertion of the $\overline{\Phi}\Gamma\Phi$ operator, and any number of external $U(1)$ gauge fields. By gauge invariance, a graph with n external photon lines will contribute n powers of the field strength tensor $F^{\mu\nu}$. For power counting, it is convenient that we normalize the gauge field so that the covariant derivative is $D_\mu = (\partial_\mu + iA_\mu)$; then the gauge field has mass dimension 1, and $F^{\mu\nu}$ has dimension 2. In $(1+1)$ dimensions $\langle 2iM\overline{\Phi}\Gamma\Phi\rangle$ has dimension 2, and so simple dimensional analysis implies that the graph with n photon lines must make a contribution proportional to $(F^{\mu\nu})^n/M^{2(n-1)}$. Therefore only the graph in Fig. 1.3 with one photon insertion can make a contribution that survives the $M \to \infty$ limit (the graph with zero photons vanishes). Calculation of this diagram yields the same result for the divergence of the regulated axial current as we found in eqn. (1.26); to show this is an exercise.

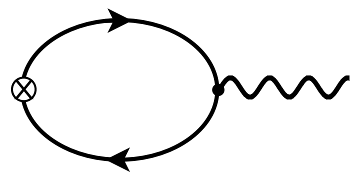

Fig. 1.3 *The anomaly diagram in 1+1 dimensions, with one Pauli-Villars loop and an insertion of $2iM\bar{\Phi}\Gamma\Phi$ at the X.*

Exercise 1.2 Compute the diagram in Fig. 1.3 using the conventional normalization of the gauge field $D_\mu = (\partial_\mu + ieA_\mu)$ and verify that $2iM\langle\bar{\Phi}\Gamma\Phi\rangle = \frac{e}{2\pi}\epsilon_{\mu\nu}F^{\mu\nu}$ when $M \to \infty$. Note that you are looking for a contribution proportional to $i\epsilon_{\mu\nu}k^\nu$, where k^ν is the momentum of the external gauge boson and μ is its polarization.

Note that in this description of the anomaly we (i) effectively rendered the number of degrees of freedom finite by introducing the regulator; (ii) the regulator explicitly broke the chiral symmetry; (iii) as the regulator was removed, the symmetry breaking effects of the regulator never decoupled, indicating that the anomaly arises when the two vertices in Fig. 1.3 sit at the same spacetime point. While we used a Pauli-Villars regulator here, the use of a lattice regulator will have qualitatively similar features, with the inverse lattice spacing playing the role of the Pauli-Villars mass, and we can turn these observations around: A lattice theory will not correctly reproduce anomalous symmetry currents in the continuum limit, unless that symmetry is broken explicitly by the lattice regulator. This means we would be foolish to expect that a continuum field theory with anomalies could ever be represented by a lattice theory with exact chiral symmetry.

1.4 A lattice Hamiltonian for d=1+1 fermions

1.4.1 Doubling of as chiral fermion

Let's reconsider the theory of a single gauged right-handed fermion, as in eqn. (1.19). We now know that the current will have an anomalous divergence, which means that the theory is not gauge invariant! Such theories are known to be sick, so it should not be possible to give them a definition on the lattice. To keep things simple, I will first consider a latticized version of the Hamiltonian for the free theory, which only involves discretizing space, not time. The continuum Hamiltonian in our γ-matrix basis dfor the free fermion is simply

$$H = -i\partial_x \tag{1.32}$$

with naive discretization

$$H = -i\frac{1}{2a}\sum_n c_n^\dagger\left(c_{n+1} - c_{n-1}\right) \tag{1.33}$$

where a is the lattice spacing, and the $c_n, c_n^\dagger$ are fermionic ladder operators at site n:

$$\{c_m, c_n\} = 0 \ , \qquad \{c_m, c_n^\dagger\} = \delta_{mn} \ . \tag{1.34}$$

This theory has an exact $U(1)$ symmetry, which is fermion number:

$$Q = \sum_n c_n^\dagger c_n \ , \qquad [Q, H] = 0 \ . \tag{1.35}$$

This is the symmetry we can gauge. The single-particle eigenstates of H are

$$|p\rangle = \sum_n e^{iapn} c_n^\dagger |0\rangle \tag{1.36}$$

with energy eigenvalue

$$H|p\rangle = E_p|p\rangle \ , \qquad E_p = \frac{\sin ap}{a} \ , \qquad -\frac{\pi}{a} \le p \le \frac{\pi}{a} \ . \tag{1.37}$$

Note from the construction of the state $|p\rangle$ that shifting $p \to p + 2\pi a$ gives back the same state, so p-space is a circle and so taking the above range for p (the Brillouin zone) accounts for all states.

What is the continuum limit of this theory? Naively, the continuum limit $ap \to 0$ gives $E_p = p$, the desired continuum result corresponding to a single right-mover, shown in Fig. 1.1. However, if we rewrite $p = \pi - k$, then the $ak \to 0$ limit gives $E_k = -k$, a left-mover! We see that the continuum theory describes a single massless Dirac fermion in the continuum, with both right and left modes, not a single right-mover. That is because the dispersion relation $E_p = \sin ap/a$ crosses the p-axis in two places, $p = 0$ and $p = \pm\pi$, so there will always be two low energy modes, even as $a \to 0$. Furthermore, the exact $U(1)$ symmetry of the lattice is just fermion number in the continuum theory, so if we gauge it, the result looks like QED in $d = 1 + 1$, a sensible theory with a conserved gauge current, unlike the chiral gauge theory in eqn. (1.19).

Can we add a local interaction that will gap the spectrum at $p = \pm\pi/a$, to get rid of the continuum left-mover? Obviously not: the function E_p will be a continuous function of p and therefore must be periodic; a periodic function of p cannot cross the p-axis an odd number of times.

And what about the anomalous $U(1)_A$ global symmetry in $d = 1 + 1$ QED? How does the lattice model realize the anomaly? The answer is that the lattice theory does not have a second $U(1)$ symmetry that we can identify with $U(1)_A$ in the continuum...that would require rotating states with $p \sim 0$ with the opposite phase from states near $p \sim \pm\pi/a$, which is not a symmetry of H. Consider an eigenstate of H at finite lattice spacing which we will call the vacuum when $a \to 0$, with every negative energy state occupied and every positive energy state empty, as shown on the left in Fig. 1.4. Now consider what happens when we turn on an electric field for some time in the x direction: all states will move to the right (increasing p) and we end up with the state shown on the right in that figure. In the continuum theory that corresponds to a state that still has no net fermion number, but a nonzero axial charge. The lattice correctly reproduces the axial anomaly by having no exact axial symmetry.

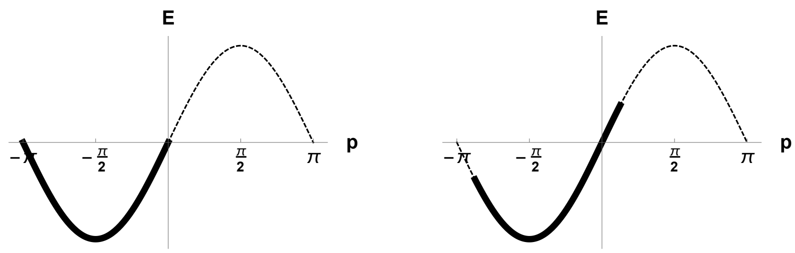

Fig. 1.4 *The ground state of a lattice Dirac fermion (left) and how it has evolved after application of an electric field in the x direction (right). The solid line denotes occupied 1-particle states, and the dashed line vacant states. In the continuum limit, it will appear as if an anomaly has violated axial charge, giving rise to right moving particles and left moving anti-particles.*

Note that it would have been wrong to consider the lattice theory to be similar to a continuum Dirac theory with a momentum cutoff at $p \sim 1/a$: such a theory would have an exact $U(1)_A$ symmetry (since a momentum cutoff is a regulator that does not violate axial rotations) but could not be gauged (since a momentum cutoff violates gauge symmetry). It behaves much more like a continuum theory with a Pauli-Villars regulator with mass $M \sim 1/a$, a regulator that preserves gauge symmetry while breaking axial symmetry.

It seems we should be happy that the lattice was "smart enough" to not give us a sick theory in the continuum, a gauge theory whose gauge symmetry was broken by an anomaly. However there are problems we can see even with very simple variants of this model. The first has to do with the role chiral symmetry plays in the Standard Model, protecting fermion masses from additive mass renormalization. The second has to do with creating a lattice regulator for chiral gauge theories that are not sick....like the Standard Model itself.

1.4.2 Problems for chiral gauge theories

It is possible to construct a theory with several chiral fermions that has an anomaly-free $U(1)$ symmetry that can be gauged. If the fermion representation is such that one cannot write down gauge-invariant mass terms for the fermions, then the theory is called a chiral gauge theory. From our discussion of the anomaly we see that an example of a chiral $U(1)$ gauge theory in $d = 1 + 1$ dimensions is the $3 - 4 - 5$ model which consists of three fermions, right-movers $\psi_{3,4}$ with electric charges 3 and 4, and a left-mover χ_5 with charge 5,

$$\mathcal{L} = \bar{\psi}_3 i\not{D} P_+ \psi_3 + \bar{\psi}_4 i\not{D} P_+ \psi_4 + \bar{\chi}_5 i\not{D} P_- \chi_5 \ . \tag{1.38}$$

Note that unlike in QED, it is impossible to write a gauge invariant fermion mass. That requires both a left- and a right-moving particle...however the two right movers we have in the theory have charges 3 and 4, while the only left-mover has charge 5, and neither $\bar{\chi}_5 P_+ \psi_3$ nor $\bar{\chi}_5 P_+ \psi_4$ operators are invariant under the gauge symmetry.

It appears that there are three conserved currents in this theory, one for each type of fermion number:

$$j_3^\mu = \bar\psi_3\gamma^\mu P_+\psi_3 \ , \qquad j_4^\mu = \bar\psi_4\gamma^\mu P_+\psi_4 \ , \qquad j_5^\mu = \bar\chi_5\gamma^\mu P_-\chi_5 \ , \tag{1.39}$$

Note that because $P_\pm = (1\pm\Gamma)/2$ these current can be written as a sum or difference of a vector and an axial current, with a factor of 1/2. We have seen that the vector currents are conserved in the presence of a background electric field, while the axial currents are anomalous, so that we get:

$$\partial_\mu j_3^\mu = 3\frac{e}{4\pi}\epsilon_{\mu\nu}F^{\mu\nu} \ , \qquad \partial_\mu j_4^\mu = 4\frac{e}{4\pi}\epsilon_{\mu\nu}F^{\mu\nu} \ , \qquad \partial_\mu j_5^\mu = -5\frac{e}{4\pi}\epsilon_{\mu\nu}F^{\mu\nu} \ . \tag{1.40}$$

However, the current that the gauge field couples to is divergenceless, by construction:

$$j^\mu = 3ej_3^\mu + 4ej_4^\mu + 5ej_5^\mu \ , \qquad \partial_\mu j^\mu = \frac{e^2}{4\pi}\left(3^2+4^2-5^2\right)\epsilon_{\mu\nu}F^{\mu\nu} = 0 \ . \tag{1.41}$$

Therefore we do not expect this to be a sick theory and would like to study it on a computer.

What happens when we try to construct this theory on the lattice, using a copy of H from eqn. (1.33) for each fermion and adding gauge fields with appropriate charges? We get a sensible theory in the continuum limit, but not the one we wanted: a theory of three Dirac fermions with Lagrangian

$$\mathcal{L} = \bar\psi_3 i\not{D}\psi_3 + \bar\psi_4 i\not{D}\psi_4 + \bar\chi_5 i\not{D}\chi_5 \ . \tag{1.42}$$

Note that this theory has no chiral projection operators in the kinetic terms, and that it is possible to write down gauge invariant Dirac mass terms for each field...this is not a chiral gauge theory. Obviously there are an infinite number of healthy chiral gauge theories and we do not seem to have a way to regulate them on the lattice. This is not just an academic problem because the Standard Model is a chiral gauge theory in $d = 3+1$. There have been ideas on how to construct lattice gauge theories which I will discuss later, but it remains an open problem.

1.5 Doubling of a Dirac fermion and the need for fine tuning

Some operators in a Lagrangian suffer from additive renormalizations, such as the unit operator (cosmological constant) and scalar mass terms, such as the Higgs mass in the Standard Model, $|H|^2$. Therefore, the mass scales associated with such operators will naturally be somewhere near the UV cutoff of the theory, unless the bare couplings of the theory are fine-tuned to cancel radiative corrections. Such fine tuning problems have obsessed particle theorists since the work of Wilson and 't Hooft on renormalization and naturalness in the 1970s. However, such intemperate behavior will not occur for operators which violate a symmetry respected by the rest of the theory: if the bare couplings for such operators were set to zero, the symmetry would ensure they could not be generated radiatively in perturbation theory. Fermion mass operators generally fall into this benign category.

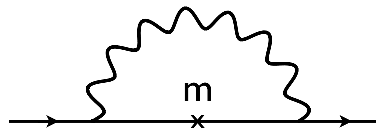

Fig. 1.5 *One-loop renormalization of the electron mass in QED due to photon exchange. A mass operator flips chirality, while gauge interactions do not. A contribution to the electron mass requires an odd number of chirality flips, and so there has to be at least one insertion of the electron mass in the diagram: the electron mass is multiplicatively renormalized. A scalar interaction flips chirality when the scalar is emitted, and flips it back when the scalar is absorbed, so replacing the photon with a scalar in the above graph again requires a fermion mass insertion to contribute to mass renormalization.*

Consider the following toy model: QED with a charge-neutral complex scalar field coupled to the electron:

$$\mathcal{L} = \overline{\psi}(i\not{D} - m)\psi + |\partial\phi|^2 - \mu^2|\phi|^2 - g|\phi|^4 + y\left(\overline{\psi}_R\phi\psi_L + \overline{\psi}_L\phi^*\psi_R\right) \ . \tag{1.43}$$

Note that in the limit $m \to 0$ this Lagrangian respects a chiral symmetry $\psi \to e^{i\alpha\gamma_5}\psi$, $\phi \to e^{-2i\alpha}\phi$. The symmetry ensures that if $m = 0$, a mass term for the fermion would not be generated radiatively in perturbation theory. With $m \neq 0$, this means that any renormalization of m must be proportional to m itself (i.e. m is "multiplicatively renormalized"). This is evident if one traces chirality through the Feynman diagrams; see Fig. 1.5. Multiplicative renormalization implies that the fermion mass can at most depend logarithmically on the cutoff (by dimensional analysis): $\delta m \sim (\alpha/4\pi)m \ln m/\Lambda$. Try plugging in some numbers here: with $\alpha = 1/137$ and $\Lambda = M_{\rm Planck} = 10^{19}$ GeV we get a radiative correction to the electron mass δm which is about 3% of the electron mass, not a shift that requires fine tuning.

In contrast, the scalar mass operator $|\phi|^2$ does not violate any symmetry and therefore suffers from additive renormalizations, such as through the graph in Fig. 1.6. By dimensional analysis, the scalar mass operator can have a coefficient that scales quadratically with the cutoff: $\delta\mu^2 \sim (y^2/16\pi^2)\Lambda^2$. This is called an additive renormalization, since $\delta\mu^2$ is not proportional to μ^2. It is only possible in general to have a scalar in the spectrum of this theory with mass much lighter than $y\Lambda/4\pi$ if the bare couplings are finely tuned to cause large radiative corrections to cancel. For $\Lambda = M_{\rm Planck}$, we require a bare mass term to cancel this one-loop radiative contribution to one part in 10^{30} to get a 100 GeV Higgs. When referring to the Higgs mass in the Standard Model, this is called the hierarchy problem.

Let's return to the problem of lattice Hamiltonians for fermions in $d = 1 + 1$. Suppose we want a lattice model to describe a massive Dirac fermion in $d = 1 + 1$. The continuum Hamiltonian is

$$H = \gamma^0(-i\gamma^1\partial_x + m) = -i\sigma_3\partial_x + m\sigma_1 \ , \tag{1.44}$$

and so we again naively write down a Hamiltonian for a free fermion, this time of the form

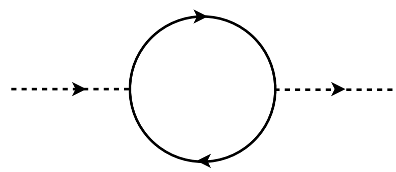

Fig. 1.6 *One-loop additive renormalization of the scalar mass due to a quadratically divergent fermion loop.*

$$H = \sum_n \psi_n^\dagger \left[-\frac{i}{2a}\sigma_3 \left(\psi_{n+1} - \psi_n\right) + m\sigma_1\psi_n \right] \tag{1.45}$$

where $\psi_{n,i}$ is a 2-component fermion ladder operator at site n with $\{\psi_{m,i}^\dagger, \psi_{n,j}\} = \delta_{mn}\delta_{ij}$, with all other anticommutators vanishing. The eigenvalues of H are the eigenvalues of the matrix

$$\left[\sigma_3 \frac{\sin ap}{a} + m\sigma_1\right] \ , \tag{1.46}$$

or

$$E_p = \pm\sqrt{\left(\frac{\sin ap}{a}\right)^2 + m^2} \tag{1.47}$$

Expanding about $a = 0$ we get $E_p \simeq \pm\sqrt{p^2+m^2}$ for $p = O(1)$, while writing $p = -\pi/a + k$ and expanding about $a = 0$ we get $E_k = \pm\sqrt{k^2+m^2}$... so we find *two* Dirac fermions in the continuum. This is the same doubling of the spectrum we saw when we tried to construct a lattice model for just a right moving 1-component fermion. In thaty case the doubling kep us from creating a sick theory....here it is just annoying, because a single massive Dirac fermion is a perfectly fine theory, and the one we want.

Can we make the mode near $p \sim \pi/a$ very heavy and get rid of it? Yes, now the spectrum never crosses the p axis and there is no reason a periodic function could exhibit a small gap $\sim m$ at $p = 0$ and a large gap $\sim 1/a$ at $p = \pi/a$. We can do that by adding to H a term of the form

$$-a\bar\psi\partial_x^2\psi \tag{1.48}$$

which looks like a mass term that will only affect modes with wavenumber $p \sim 1/a$ in the $a \to 0$ limit. The lattice realization of this operator is

$$H_w = -\frac{ra}{2}\sum_n \frac{1}{a^2}\bar\psi_n\left(\psi_{n+1} - 2\psi_n + \psi_{n-1}\right) \ , \tag{1.49}$$

where the $r/2$ factor is a parameter we can adjust. Now the energy is given by the eigenvalues of

$$\frac{\sin ap}{a}\sigma_3 + \left(m + r\frac{1-\cos ap}{a^2}\right)\sigma_1 \ , \tag{1.50}$$

or

$$E_p = \pm\sqrt{\left(\frac{\sin ap}{a}\right)^2 + \left(m + r\frac{1-\cos ap}{a^2}\right)^2} \tag{1.51}$$

Now if we expand about $ap = 0$ we get $E_p = \pm\sqrt{p^2+m^2}$ as before, the desired dispersion relation for a massive Dirac fermion in the continuum. However, when we set $p = -\pi/a + k$ and expand about $ak = 0$ we get $E_k = 2r/a + O(1)$. So we see that the unwanted mode at the edge of the Brillouin zone ($p = \pm\pi/a$) becomes infinitely heavy as $a \to 0$ and decouples from low energy physics. This is Wilson's solution for getting rid of the unwanted "doubler".

There has been a cost, however. Now we have two terms violating chiral symmetry in the Lagrangian: the $m\bar{\psi}\psi$ term, and the $a\bar{\psi}\nabla^2\psi$ term. Thus when we add interactions (e.g. by gauging fermion number) the Wilson interaction term $\bar{\psi}D^2\psi$ will renormalize the mass term through gauge boson loops, and we expect radiative corrections of size

$$\delta m \sim r\frac{\alpha}{a}\ , \tag{1.52}$$

which looks very similar to the Higgs mass fine tuning problem: the bare mass will have to be fine tuned to one part in $\sim ma/\alpha$ in order to obtain a physical mass m, which becomes harder to do the smaller a becomes. In getting rid of our doubler fermion we lost our approximate chiral symmetry, and have to fine tune parameters to recover it in the continuum limit. And the problem seemed to arise from the need for the lattice to properly account for anomalies.

Exercise 1.3 Compute the spectrum of the Wilson-Dirac Hamiltonian, $H + H_w$ and plot it for various values of m, r with $a = 1$.

1.6 What we have found

- Chiral symmetries appear in theories of massless fermions
- Chiral symmetries can be broken in the continuum by quantum effects called anomalies
- Anomalies cannot exist in a finite system; lattice theories of fermions break potentially anomalous symmetries explicitly
- ...or else double the fermions so that the exact lattice symmetries are vector symmetries in the continuum.
- Lattice doubling of the spectrum makes it problematic to construct sensible chiral gauge theories in the continuum limit
- Eliminating doubling in vector-like gauge theories eliminates global chiral symmetry and requires fine tuning in order to find light fermions in the continuum limit.

What we would like is a lattice fermion formulation which at the very least correctly accounts for anomalies, while protecting fermion masses from additive renormalization, like in the continuum. Better: we would like to know how to construct lattice models for chiral gauge theories.

2

Chiral symmetry, parity, and anomalies in higher dimensions.

2.1 Spinor representations of the Lorentz group

To understand chiral symmetry in $d = 3+1$ it is helpful to understand representations of the Lorentz group. Since we will be discussing fermions in various dimensions of spacetime, consider the generalization of the usual Lorentz group to d dimensions. The Lorentz group is defined by the real matrices Λ which preserve the form of the d-dimensional metric

$$\Lambda^T \eta \, \Lambda = \eta \ , \qquad \eta = \mathrm{diag}\,(1, -1, \ldots -1) \, . \tag{2.1}$$

With this definition, the inner product between two 4-vectors, $v^\mu \eta_{\mu\nu} w^\mu = v^T \eta w$, is preserved under the Lorentz transformations $v \to \Lambda v$ and $w \to \Lambda w$. This defines the group $SO(d-1,1)$, which — like $SO(d)$ — has $d(d-1)/2$ linearly independent generators, which may be written as $M^{\mu\nu} = -M^{\nu\mu}$, where the indices $\mu, \nu = 0, \ldots, (d-1)$ and

$$\Lambda = e^{i\theta_{\mu\nu} M^{\mu\nu}} \ , \tag{2.2}$$

with $\theta_{\mu\nu} = -\theta_{\nu\mu}$ being $d(d-1)/2$ real parameters. Note that μ, ν label the $d(d-1)/2$ generators, while in a representation R each M is a $d_R \times d_R$ matrix, where d_R is the dimension of R. By expanding eqn. (2.1) to order θ one sees that the generators M must satisfy

$$(M^{\mu\nu})^T \eta + \eta M^{\mu\nu} = 0. \tag{2.3}$$

It is really easy to find all the solutions to this equation in the defining representation (eg for the d-vector)! Given the form of η, the M^{ij} matrices (all spatial indices) have to be antisymmetric, while the M^{0i} matrices have to be symmetric. We can find a simple basis by just scattering factors of 1 and -1 in appropriate places, and then put in an overall factor of i to ensure that the transformation Λ is real. For example, in $d = 3+1$ we can find a representation for the generators of rotations and boosts in the z direction of the familiar form:

$$M^{12} = i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \qquad M^{03} = i \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{2.4}$$

and exponentiating them gives the familiar form for Λ,

$$e^{i\theta M^{12}} = \begin{pmatrix} 1 & & & \\ & \cos\theta & \sin\theta & \\ & -\sin\theta & \cos\theta & \\ & & & 1 \end{pmatrix} , \quad e^{i\theta M^{03}} = \begin{pmatrix} \cosh\theta & & & \sinh\theta \\ & 1 & & \\ & & 1 & \\ \sinh\theta & & & \cosh\theta \end{pmatrix} \tag{2.5}$$

Note that we find a Hermitian generator for rotations, but an anti-Hermitian generator for boosts. The transformation Λ is therefore not a unitary matrix. This is a hallmark of the Lorentz group being noncompact: if you keep boosting, you never return to your original frame, unlike if you keep rotating. In fact, there are no finite-dimensional unitary representations of the Lorentz group. (There are, however, infinite dimensional unitary representations, such as the ones that act on our Hilbert space and conserve probability!).

Once one knows the defining representation that solves eqn. (2.3) on can determine the commutation relations for the algebra,

$$\left[M^{\alpha\beta}, M^{\gamma\delta}\right] = i\left(\eta^{\beta\gamma}M^{\alpha\delta} - \eta^{\alpha\gamma}M^{\beta\delta} - \eta^{\beta\delta}M^{\alpha\gamma} + \eta^{\alpha\delta}M^{\beta\gamma}\right) . \tag{2.6}$$

By finding all representations of this algebra, one can construct all representations of the group by exponentiating the generators. A Dirac spinor representation can be constructed as

$$M^{\alpha\beta} \equiv \Sigma^{\alpha\beta} = \frac{i}{4}\left[\gamma^\alpha, \gamma^\beta\right] \tag{2.7}$$

where the gamma matrices satisfy the Clifford algebra:

$$\{\gamma^\alpha, \gamma^\beta\} = 2\eta^{\alpha\beta} \tag{2.8}$$

You can check explicitly that if you can find γ matrices satisfying eqn. (2.8), then Σ satisfies the commutation relations eqn. (2.6).

Solutions to the Clifford algebra are easy to find by making use of direct products of Pauli matrices. In a direct product space we can write a matrix as $M = a \otimes A$ where a and A are matrices of dimension d_a and d_A respectively, acting in different spaces; the matrix M then has dimension $(d_a \times d_A)$. Matrix multiplication is defined as $(a \otimes A)(b \otimes B) = (ab) \otimes (AB)$. It is usually much easier to construct a representation when you need one rather than to look one up and try to keep the conventions straight! One finds that solutions for the γ matrices in d Minkowski dimensions obey the following properties:

1. For both $d = 2k$ and $d = 2k + 1$, the γ-matrices are 2^k dimensional;
2. For even spacetime dimension $d = 2k$ (such as our own with $k = 2$) one can define a generalization of γ_5 to be

$$\Gamma = i^{k-1} \prod_{\mu=0}^{2k-1} \gamma^\mu \tag{2.9}$$

with the properties

$$\{\Gamma, \gamma^\mu\} = 0 , \quad \Gamma = \Gamma^\dagger = \Gamma^{-1} , \quad \mathrm{Tr}\left(\Gamma\gamma^{\alpha_1}\cdots\gamma^{\alpha_{2k}}\right) = 2^k i^{-1-k}\epsilon^{\alpha_1\ldots\alpha_{2k}} , \tag{2.10}$$

where $\epsilon_{012\ldots 2k-1} = +1 = -\epsilon^{012\ldots 2k-1}$.

3. In $d = 2k + 1$ dimensions one needs one more γ-matrix than in $d = 2k$, and one can take it to be $\gamma^{2k} = i\Gamma$.

Sometimes it is useful to work in a specific basis for the γ-matrices; a particularly useful choice is a "chiral basis", defined to be one where Γ is diagonal. For example, for $d = 2$ and $d = 4$ (Minkowski spacetime, i.e. $d = 1+1$ and $d = 3+1$) one can choose

$$d = 2: \quad \gamma^0 = \sigma_1 \ , \quad \gamma^1 = -i\sigma_2 \ , \quad \Gamma = \sigma_3 \tag{2.11}$$

$$d = 4: \quad \gamma^0 = -\sigma_1 \otimes 1 \ , \quad \gamma^i = i\sigma_2 \otimes \sigma_i \ , \quad \Gamma = \sigma_3 \otimes 1 \ . \tag{2.12}$$

It is easy to convert these direct product matrices into ordinary 4×4 matrices, for example one can write

$$\gamma^1 = i\sigma_2 \otimes \sigma_1 = i \begin{pmatrix} 0 & -i\sigma_1 \\ i\sigma_1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{2.13}$$

For spinors in $d = 3$ and $d = 5$ we can just take the above matrices for $d = 2$ and $d = 4$ respectively, and tack on $i\Gamma$ as the extra matrix:

$$d = 3: \quad \gamma^0 = \sigma_1 \ , \quad \gamma^1 = -i\sigma_2 \ , \quad \gamma^2 = i\sigma_3 \tag{2.14}$$

$$d = 5: \quad \gamma^0 = -\sigma_1 \otimes 1 \ , \quad \gamma^{i=1,2,3} = i\sigma_2 \otimes \sigma_i \ , \quad \gamma^5 = i\sigma_3 \otimes 1 \ . \tag{2.15}$$

It is evident that there does not exist a fourth 2×2 matrix in $d = 3$ or a sixth 4×4 matrix in $d = 5$ that anticommutes with all the other γ^μ matrices, and so there is no notion of chirality and Dirac fermions are irreducible for odd d.

2.1.1 γ-matrices in Euclidian spacetime

In Feynman diagrams one performs a Wick rotation $\int_{-\infty}^{\infty} dk_0 \to \int_{-i\infty}^{i\infty} dk_0 \equiv i \int_{-\infty}^{\infty} dk_0^E$, where the last step is a change of variables $k_0 = ik_0^E$. This implies that to go to Euclidian spacetime we should replace $\partial_0 \to i\partial_0^E$ and $x^0 \to -ix_0^E$. Therefore, to go to Euclidian spacetime with metric $\eta_E^{\mu\nu} = \delta_{\mu\nu}$, we take

$$\partial_0^M \to i\partial_0^E \ , \quad \partial_i^M \to \partial_i^E \tag{2.16}$$

and defines

$$\gamma_M^0 = \gamma_E^0 \ , \quad \gamma_M^i = i\gamma_E^i \ , \tag{2.17}$$

so that

$$(\gamma_E^\mu)^\dagger = \gamma_E^\mu \ , \qquad \{\gamma_E^\mu, \gamma_E^\nu\} = 2\delta_{\mu\nu} \tag{2.18}$$

and $\not{D}_M \to i\not{D}_E$, with

$$\not{D}_E = -\not{D}_E^\dagger \tag{2.19}$$

and the Euclidian Dirac operator is $(\not{D}_E + m)$. In the path integral

$$e^{iS_M} \to e^{-S_E}, \quad S_M = \int d^d x\, \bar{\psi}(i\not{D} - m)\psi, \quad S_E = \int d^4 x_E\, \bar{\psi}(\not{D}_E + m)\psi. \tag{2.20}$$

With the Euclidian metric there is no difference between upper and lower indices. (The annoying thing about working with the mostly minus metric is that one has to also flip the sign of scalar products, such as $x_\mu x^\mu \to -x_\mu^E x_\mu^E$...this can be avoided by working with the mostly plus metric, the drawback of that metric being that the non-relativistic limit is less convenient.) The matrix $\Gamma^{(2k)}$ in $2k$ dimensions is taken to equal γ_E^{2k} in $(2k+1)$ dimensions:

$$\Gamma_E^{(2k)} = \gamma_E^{2k} = \Gamma_M^{(2k)} , \qquad \mathrm{Tr}\,(\Gamma_E\, \gamma_E^{\alpha_1} \cdots \gamma_E^{\alpha_{2k}}) = -2^k i^k \epsilon^{\alpha_1 \ldots \alpha_{2k}} \tag{2.21}$$

where $\epsilon_{012\ldots 2k-1} = +1 = +\epsilon^{012\ldots 2k-1}$.

2.2 Chirality in even dimensions

As we saw in $d = 2$, the existence of Γ means that Dirac spinors are reducible representations of the Lorentz group, which in turn means we can have symmetries ("chiral symmetries") which transform different parts of Dirac spinors in different ways. To see this, define the projection operators

$$P_\pm = \frac{(1 \pm \Gamma)}{2} , \tag{2.22}$$

which have the properties

$$P_+ + P_- = \mathbf{1} , \qquad P_\pm^2 = P_\pm , \qquad P_+ P_- = 0 . \tag{2.23}$$

Since in odd spatial dimensions $\{\Gamma, \gamma^\mu\} = 0$ for all μ, it immediately follows that Γ commutes with the Lorentz generators $\Sigma^{\mu\nu}$ in eqn. (2.7): $[\Gamma, \Sigma^{\mu\nu}] = 0$. Therefore we can write $\Sigma^{\mu\nu} = \Sigma_+^{\mu\nu} + \Sigma_-^{\mu\nu}$ where

$$\Sigma_\pm^{\mu\nu} = P_\pm \Sigma^{\mu\nu} P_\pm , \tag{2.24}$$

Thus $\Sigma^{\mu\nu}$ is reducible: spinors $\psi_\pm$ which are eigenstates of Γ with eigenvalue ± 1 respectively transform independently under Lorentz transformations.

The word "chiral" comes from the Greek word for hand. We saw in $d = 1+1$ that massless $\Gamma = +1$ chirality fermions correspond to right movers, while $\Gamma = -1$ chirality fermions correspond to left movers. In $d = 3+1$ one finds that for massless fermions, chirality is equal to helicity, and again $\Gamma = \pm 1$ corresponds to RH and LH helicity respectively.

Exercise 2.1 You should perform the same exercise in $3+1$ dimensions and find that solutions $\psi_\pm$ to the massless Dirac equation satisfying $\Gamma\psi_\pm = \pm\psi_\pm$ must also satisfy $|\vec{p}| = E$ and $(2\vec{p} \cdot \vec{S}/E)\psi_\pm = \pm\psi_\pm$, where $S_i = \frac{1}{2}\epsilon_{0ijk}\Sigma^{jk}$ are the generators of rotations. Thus $\psi_\pm$ correspond to states with positive or negative helicity respectively, and are called right- and left-handed particles.

2.3 Chiral symmetry and fermion mass in four dimensions

In many ways, chiral symmetry is very similar in $d = 3+1$ dimensions to what we saw for $d = 1+1$. Consider the Lagrangian for a single flavor of Dirac fermion in 3+1 dimensions, coupled to a background gauge field

$$\mathcal{L} = \left(\overline{\psi}_L i\not{D}\psi_L + \overline{\psi}_R i\not{D}\psi_R\right) - \left(m\overline{\psi}_L\psi_R + \text{h.c.}\right) \tag{2.25}$$

where I have defined

$$\psi_L = P_-\psi\ , \quad \bar{\psi}_L = \psi_L^\dagger\gamma^0 = \bar{\psi}P_+\ , \quad \psi_R = P_+\psi\ , \quad \bar{\psi}_R = \bar{\psi}P_-\ . \tag{2.26}$$

For now I am assuming that $\psi_{L,R}$ are in the same complex representation of the gauge group, where D_μ is the gauge covariant derivative appropriate for that representation. It is important to note the the property $\{\gamma_5, \gamma^\mu\} = 0$ ensured that the kinetic terms in eqn. (2.26) do not couple left-handed and right-handed fermions; on the other hand, the mass terms do[1]. The above Lagrangian has an exact $U(1)$ symmetry, associated with fermion number, $\psi \to e^{i\alpha}\psi$. Under this symmetry, left-handed and right-handed components of ψ rotate with the same phase; this is often called a "vector symmetry". In the case where $m = 0$, it apparently has an additional symmetry where the left- and right-handed components rotate with the opposite phase, $\psi \to e^{i\alpha\gamma_5}\psi$; this is called an "axial symmetry", $U(1)_A$.

Symmetries are associated with Noether currents, and symmetry violation appears as a nonzero divergence for the current. Recall the Noether formula for a field ϕ and infinitesimal transformation $\phi \to \phi + \epsilon\delta\phi$:

$$j^\mu = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\delta\phi\ , \qquad \partial_\mu j^\mu = -\delta\mathcal{L}\ . \tag{2.27}$$

In the Dirac theory, the vector symmetry corresponds to $\delta\psi = i\psi$, and the axial symmetry transformation is $\delta\psi = i\gamma_5\psi$, so that the Noether formula yields the vector and axial currents:

$$U(1): \qquad j^\mu = \overline{\psi}\gamma^\mu\psi\ , \qquad \partial_\mu j^\mu = 0 \tag{2.28}$$

$$U(1)_A: \qquad j^\mu_A = \overline{\psi}\gamma^\mu\gamma_5\psi\ , \qquad \partial_\mu j^\mu_A = 2im\overline{\psi}\gamma_5\psi\ . \tag{2.29}$$

Some comments are in order:

- Eqn. (2.29) is not the whole story! As in $d = 1+1$, the axial current will also have an anomalous divergence from quantum effects.
- As in $d = 1+1$, the fact that the fermion mass explicitly breaks chiral symmetry means that fermion masses get multiplicatively renormalized, so that fermions can naturally be light.
- The variation of a general fermion bilinear $\overline{\psi}X\psi$ under chiral symmetry is

  $$\delta\overline{\psi}X\psi = i\overline{\psi}\{\gamma_5, X\}\psi\ . \tag{2.30}$$

  This will vanish if X can be written as the product of an odd number of γ^μ matrices. In any even dimension the chirally invariant bilinears include currents,

[1] I will use the familiar γ_5 in $3+1$ dimensions instead of Γ when there is no risk of ambiguity.

with $X = \gamma^\mu$ or $X = \gamma^\mu\Gamma$, and so gauge interactions are always invariant under chiral symmetry (up to anomalies). The bilinears which transform nontrivially under the chiral symmetry include not only mass terms, $X = \mathbf{1}, \Gamma$, but in $d = 4$ $X = \sigma^{\mu\nu}, \sigma^{\mu\nu}\Gamma$ as well. The latter operators when coupled to $F^{\mu\nu}$ give rise to contributions to magnetic and electric dipole moments, respectively.

The Lagrangian for N_f flavors of massive Dirac fermions in even d, coupled to some background gauge field may be written as

$$\mathcal{L} = \left(\overline{\psi}^a_L i\not{D}\psi^a_L + \overline{\psi}^a_R i\not{D}\psi^a_R\right) - \left(\overline{\psi}^a_L M_{ab}\psi^b_R + \overline{\psi}^a_R M^\dagger_{ab}\psi^b_L\right) \ . \tag{2.31}$$

The index on ψ denotes flavor, with $a, b = 1, \ldots N_f$, and M_{ab} is a general complex mass matrix (no distinction between upper and lower flavor indices). Again assuming the fermions to be in a complex representation of the gauge group, this theory is invariant under independent chiral transformations if the mass matrix vanishes:

$$\psi^a_R \to U_{ab}\psi^b_R \ , \quad \psi^a_L \to V_{ab}\psi^b_L \ , \quad U^\dagger U = V^\dagger V = \mathbf{1} \ . \tag{2.32}$$

where U and V are independent $U(N_f)$ matrices. Since $U(N_f) = SU(N_f) \times U(1)$, it is convenient to write

$$U = e^{i(\alpha+\beta)}R \ , \quad V = e^{i(\alpha-\beta)}L \ , \quad R^\dagger R = L^\dagger L = \mathbf{1} \ , \quad |R| = |L| = 1 \ , \tag{2.33}$$

so that the symmetry group is $SU(N_f)_L \times SU(N_f)_R \times U(1) \times U(1)_A$ with $L \in SU(N_f)_L$, $R \in SU(N_f)_R$.

If we turn on the mass matrix, the chiral symmetry is explicitly broken, since the mass matrix couples left- and right-handed fermions to each other. If $M_{ab} = m\delta_{ab}$ then the "diagonal" or "vector" symmetry $SU(N_f) \times U(1)$ remains unbroken, where $SU(N_f) \subset SU(N_f)_L \times SU(N_f)_R$ corresponding to the transformation eqn. (2.32), eqn. (2.33) with $L = R$. If M_{ab} is diagonal but with unequal eigenvalues, the symmetry may be broken down as far as $U(1)^{N_f}$, corresponding to independent phase transformations of the individual flavors. With additional flavor-dependent interactions, these symmetries may be broken as well.

2.4 Lorentz group in $d = 3 + 1$: $SU(2) \times SU(2)$ and Weyl fermions

We have seen that Dirac fermions in even dimensions form a reducible representation of the Lorentz group. Dirac notation is convenient when both LH and RH parts of the Dirac spinor transform as the same complex representation under a gauge group, and when there is a conserved fermion number. This sounds restrictive, but applies to QED and QCD. For other applications — such as chiral gauge theories (where LH and RH fermions carry different gauge charges, as under $SU(2) \times U(1)$), or when fermion number is violated (as is the case for neutrinos with a Majorana mass), or when fermions transform as a real representation of gauge group — then it is much more convenient to use irreducible fermion representations, called Weyl fermions.

The six generators of the Lorentz group may be chosen to be the three Hermitian generators of rotations J_i, and the three anti-Hermitian generators of boosts K_i, so that an arbitrary Lorentz transformation takes the form

$$\Lambda = e^{i(\theta_i J_i + \omega_i K_i)} \ . \tag{2.34}$$

In terms of the $M_{\mu\nu}$ generators in §2.1,

$$J_i = \tfrac{1}{2}\epsilon_{0i\mu\nu} M^{\mu\nu} \ , \qquad K_i = M^{0i} \ . \tag{2.35}$$

These generators have the commutation relations

$$[J_i, J_j] = i\epsilon_{ijk} J_k \ , \quad [J_i, K_j] = i\epsilon_{ijk} K_k \ , \quad [K_i, K_j] = -i\epsilon_{ijk} J_k \ . \tag{2.36}$$

It is convenient to define different linear combinations of generators

$$A_i = \frac{J_i + iK_i}{2} \ , \quad B_i = \frac{J_i - iK_i}{2} \ , \quad \Longrightarrow \quad J_i = \frac{A_i + B_i}{2} \ , \quad K_i = -i\frac{A_i - B_i}{2} \tag{2.37}$$

with

$$\Lambda = e^{i\left[(\vec{\theta}_i - i\vec{\omega})\cdot\vec{A} + (\vec{\theta}_i + i\vec{\omega})\cdot\vec{B})\right]} \ . \tag{2.38}$$

Life simplifies now since eqn. (2.36) implies that $\vec{A}$ and $\vec{B}$ are the six Hermitian generators of an $SU(2) \times SU(2)$ algebra:

$$[A_i, A_j] = i\epsilon_{ijk} A_k \ , \quad [B_i, B_j] = i\epsilon_{ijk} B_k, \quad [A_i, B_j] = 0 \ . \tag{2.39}$$

We already know all about representations of $SU(2)$! They are labeled by non-negative half-integer j. Thus Lorentz representations may be labelled with two $SU(2)$ spins, $j_{A,B}$ corresponding to the two $SU(2)$s: (j_A, j_B), transforming as

$$\Lambda(\vec{\theta}, \vec{\omega}) = D^{j_A}(\vec{\theta} - i\vec{\omega}) \times D^{j_B}(\vec{\theta} + i\vec{\omega}) \tag{2.40}$$

where the D^j is the usual $SU(2)$ rotation in the spin j representation; boosts appear as imaginary parts to the rotation angle; the D^{j_A} and D^{j_B} matrices act in different spaces and therefore commute. For example, under a general Lorentz transformation, a LH Weyl fermion $\psi = (\frac{1}{2}, 0)$ has $\vec{A} = \frac{1}{2}\vec{\sigma}$ and $\vec{B} = 0$, so that it transforms as

$$\psi \to e^{i(\vec{\theta} - i\vec{\omega})\cdot\vec{\sigma}/2}\psi \ . \tag{2.41}$$

Similarly, a RH Weyl fermion $\chi = (0, \frac{1}{2})$ transforms under Lorentz transformations as

$$\chi \to e^{i(\vec{\theta} + i\vec{\omega})\cdot\vec{\sigma}/2}\chi \ . \tag{2.42}$$

Evidently the two types of fermions transform the same way under rotations, but differently under boosts.

The dimension of the $(j_A,\, j_B)$ representation is $(2j_A+1)(2j_B+1)$. In this notation, the smaller irreducible Lorentz representations are labelled as:

$$\begin{aligned}
(0,0):&\quad \text{scalar}\\
(\tfrac{1}{2},0)\ ,\ (0,\tfrac{1}{2}):&\quad \text{LH and RH Weyl fermions}\\
(\tfrac{1}{2},\tfrac{1}{2}):&\quad \text{four-vector}\\
(1,0)\ ,\ (0,1):&\quad \text{self-dual and anti-self-dual antisymmetric tensors}
\end{aligned}$$

A Dirac fermion is the reducible representation $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ consisting of a LH and a RH Weyl fermion.

Parity takes $\vec{J}\to\vec{J}$ and $\vec{K}\to-\vec{K}$, and therefore interchanges $\vec{A}\leftrightarrow\vec{B}$, transforming a $(j_1,\, j_2)$ representation into $(j_2,\, j_1)$. So parity will change a LH Weyl fermion into a RH one, and vice versa, but takes a 4-vector into a 4-vector.

Charge conjugation taqkes a field to its complex conjugate, and therefore also effectively flips the sign of K_i in eqn. (2.40) due to the factor of i in front of ω, implying that if a field ϕ transforms as $(j_1,\, j_2)$, then $\phi^\dagger$ transforms as (j_2, j_1); for this reason, the combined symmetry CP does not alter the particle content of a chiral theory, so that CP violation must arise from complex coupling constants. This is in contrast to P violation, which will occur whenever RH and LH Weyl fermions do not have the same gauge charges, or when one of them is missing from the theory. Therefore a theory of N_L flavors of LH Weyl fermions ψ_i, and N_R flavors of RH Weyl fermions χ_a may be recast as a theory of (N_L+N_R) LH fermions by defining $\chi_a\equiv\omega_a^\dagger$. The fermion content of the theory can be described entirely in therms of LH Weyl fermions then, $\{\psi_i,\omega_a\}$; this often simplifies the discussion of parity violating theories, such as the Standard Model or Grand Unified Theories. Note that if the RH χ_a transformed under a gauge group as representation R, the conjugate fermions ω_a transform under the conjugate representation $\overline{R}$.

For example, QCD written in terms of Dirac fermions has the Lagrangian:

$$\mathcal{L}=\sum_{i=u,d,s\dots}\overline{\psi}_n(i\not{D}-m_n)\psi_n\ , \tag{2.43}$$

where D_μ is the $SU(3)_c$ covariant derivative, and the ψ_n fields (both LH and RH components) transform as a 3 of $SU(3)_c$. However, we could just as well write the theory in terms of the LH quark fields ψ_n and the LH anti-quark fields χ_n. Using the γ-matrix basis in eqn. (2.12), we write the Dirac spinor ψ in terms of two-component LH spinors ψ and χ as

$$\psi=\begin{pmatrix}-\sigma_2\chi^\dagger\\ \psi\end{pmatrix}\ . \tag{2.44}$$

Note that ψ transforms as a 3 of $SU(3)_c$, while χ transforms as a $\overline{3}$. Then the kinetic operator becomes (up to a total derivative)

$$\overline{\psi}i\not{D}\psi=\psi^\dagger iD_\mu\sigma^\mu\psi+\chi^\dagger iD_\mu\sigma^\mu\chi\ ,\qquad \sigma^\mu\equiv\{\mathbf{1},-\vec{\sigma}\}\ , \tag{2.45}$$

and the mass terms become

$$\overline{\psi}_R\psi_L=\chi\sigma_2\psi=\psi\sigma_2\chi$$

$$\overline{\psi}_L \psi_R = \psi^\dagger \sigma_2 \chi^\dagger = \chi^\dagger \sigma_2 \psi^\dagger \ , \tag{2.46}$$

where I used the fact that fermion fields anti-commute. Thus a Dirac mass in terms of Weyl fermions is just

$$m\overline{\psi}\psi = m(\psi\sigma_2\chi + h.c) \ , \tag{2.47}$$

and preserves a fermion number symmetry where ψ has charge $+1$ and χ has charge -1. On the other hand, one can also write down a Lorentz invariant mass term of the form

$$m(\psi\sigma_2\psi + h.c.) \tag{2.48}$$

which violates fermion number by two units; this is a Majorana mass, which is clumsy to write in Dirac notation. Experimentalists are trying to find out which form neutrino masses have — Dirac, or Majorana? If the latter, lepton number is violated by two units and could show up in neutrinoless double beta decay, where a nucleus decays by emitting two electrons and no anti-neutrinos.

The Standard Model is a relevant example of a chiral gauge theory. Written in terms of LH Weyl fermions, the quantum numbers of a single family under $SU(3) \times SU(2) \times U(1)$ are:

$$\begin{aligned} Q &= (3,2)_{+\frac{1}{6}} & L &= (1,2)_{-\frac{1}{2}} \\ U^c &= (\bar{3},1)_{-\frac{2}{3}} & E^c &= (1,1)_{+1} \\ D^c &= (\bar{3},1)_{+\frac{1}{3}} \ . \end{aligned} \tag{2.49}$$

Evidently this is a complex representation and chiral. If neutrino masses are found to be Dirac in nature (i.e. lepton number preserving) then a partner for the neutrino must be added to the theory, the "right handed neutrino", which can be described by a LH Weyl fermion which is neutral under all Standard Model gauge interactions, $N = (1,1)_0$.

If unfamiliar with two-component notation, you can find all the details in Appendix A of Wess and Bagger's classic book on supersymmetry (Wess and Bagger, 1992); the notation used here differs slightly as I use the metric and γ-matrix conventions of Itzykson and Zuber (Itzykson and Zuber, 1980), and write out the σ_2 matrices explicitly.

Exercise 2.2 Consider a theory of N_f flavors of Dirac fermions in a real or pseudo-real representation of some gauge group. (Real representations combine symmetrically to form an invariant, such as a triplet of $SU(2)$; pseudo-real representations combine anti-symmetrically, such as a doublet of $SU(2)$). Show that if the fermions are massless the action exhibits a $U(2N_f) = U(1) \times SU(2N_f)$ flavor symmetry at the classical level (the $U(1)$ subgroup being anomalous in the quantum theory). If the fermions condense as in QCD, what is the symmetry breaking pattern? How do the resultant Goldstone bosons transform under the unbroken subgroup of $SU(2N_f)$?

Exercise 2.3 To see how the $(\frac{1}{2}, \frac{1}{2})$ representation behaves like a four-vector, consider the 2×2 matrix $P = P_\mu \sigma^\mu$, where σ^μ is given in eqn. (2.45). Show that the transformation $P \to LPL^\dagger$ (with $\det L = 1$) preserves the Lorentz invariant inner product $P_\mu P^\mu = (P_0^2 - P_i P_i)$. Show that with L given by eqn. (2.41), P_μ transforms properly like a four-vector.

Exercise 2.4 Is it possible to write down an anomalous electric or magnetic moment operator in a theory of a single charge-neutral Weyl fermion?

2.5 Anomalies in 3+1 dimensions

2.5.1 The $U(1)_A$ anomaly

An analogous violation of the $U(1)_A$ current occurs in $3+1$ dimensions as well [2]. One might guess that the analogue of $\epsilon_{\mu\nu}F^{\mu\nu} = 2E$ in the anomalous divergence eqn. (1.26) would be the quantity $\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} = 8\vec{E}\cdot\vec{B}$, which has the right dimensions and properties under parity and time reversal. So we should consider the behavior a massless Dirac fermion in $(3+1)$ in parallel constant E and B fields. First turn on a B field pointing in the $\hat{z}$ direction: this gives rise to Landau levels, with energy levels E_n characterized by non-negative integers n as well as spin in the $\hat{z}$ direction S_z and momentum p_z, where

$$E_n^2 = p_z^2 + (2n+1)eB - 2eBS_z \ . \tag{2.50}$$

The number density of modes per unit transverse area is defined to be g_n, which can be derived by computing the zero-point energy in Landau modes and requiring that it yields the free fermion result as $B \to 0$. We have $g_n \to p_\perp dp_\perp/(2\pi)$ with $[(2n+1)eB - 2eBS_z] \to p_\perp^2$, implying that

$$g_n = eB/2\pi \ . \tag{2.51}$$

The dispersion relation looks like that of an infinite number of one-dimensional fermions of mass $m_{n,\pm}$, where

$$m_{n\pm}^2 = (2n+1)eB - 2eBS_z \ , \quad S_z = \pm\tfrac{1}{2} \ . \tag{2.52}$$

The state with $n=0$ and $S_z = +\frac{1}{2}$ is distinguished by having $m_{n,+} = 0$; it behaves like a *massless* one-dimensional Dirac fermion (with transverse density of states g_0) moving along the $\hat{z}$ axis with dispersion relation $E = |p_z|$. If we now turn on an electric field also pointing along the $\hat{z}$ direction we know what to expect from our analysis in $1+1$ dimensions: we find an anomalous divergence of the axial current equal to

$$g_0 eE/\pi = e^2 EB/2\pi^2 = \left(\frac{e^2}{16\pi^2}\right)\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} \ . \tag{2.53}$$

If we include an ordinary mass term in the $3+1$ dimensional theory, then we get

[2]Part of the content of this section comes directly from John Preskill's class notes on the strong interactions, available at his web page: http://www.theory.caltech.edu/∼preskill/notes.html.

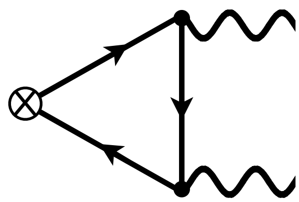

Fig. 2.1 *The $U(1)_A$ anomaly diagram in 3+1 dimensions, with one Pauli-Villars loop and an insertion of $2iM\overline{\Phi}\Gamma\Phi$.*

$$\partial_\mu j^\mu_A = 2im\overline{\psi}\Gamma\psi + \left(\frac{e^2}{16\pi^2}\right)\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} \ . \tag{2.54}$$

One can derive this result by computing $\langle M\overline{\Phi}i\Gamma\Phi\rangle$ for a Pauli-Villars regulator as in the $1+1$ dimensional example; now the relevant graph is the triangle diagram of Fig. 2.1.

If the external fields are nonabelian, the analogue of eqn. (2.54) is

$$\partial_\mu j^\mu_A = 2im\overline{\psi}\Gamma\psi + \left(\frac{g^2}{16\pi^2}\right)\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}_a F^{\rho\sigma}_b \,\mathrm{Tr}\, T_a T_b \ . \tag{2.55}$$

If the fermions transform in the defining representation of $SU(N)$, it is conventional to normalize the coupling g so that $\mathrm{Tr}\ T_a T_b = \frac{1}{2}\delta_{ab}$. This is still called an "Abelian anomaly", since j^μ_A generates a $U(1)$ symmetry.

2.5.2 Anomalies in Euclidian spacetime

Continuing to Euclidian spacetime by means of eqns. (2.16)-(2.21) changes the anomaly equations simply by eliminating the factor of i from in front of the fermion mass:

$$2d: \quad \partial_\mu j^\mu_A = 2m\overline{\psi}\Gamma\psi + \frac{e}{2\pi}\epsilon_{\mu\nu}F^{\mu\nu} \tag{2.56}$$

$$4d: \quad \partial_\mu j^\mu_A = 2m\overline{\psi}\Gamma\psi + \left(\frac{g^2}{16\pi^2}\right)\epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}_a F^{\rho\sigma}_b \,\mathrm{Tr}\, T_a T_b \ . \tag{2.57}$$

2.5.3 The index theorem in four dimensions

For nonabelian gauge theories the quantity on the far right of eqn. (2.57) is a topological charge density, with

$$\nu = \frac{g^2}{64\pi^2}\int d^4x_E\, \epsilon_{\mu\nu\rho\sigma}F^{\mu\nu}_a F^{\rho\sigma}_a \tag{2.58}$$

being the winding number associated with $\pi_3(G)$, the homotopy group of maps of S_3 (spacetime infinity) into the gauge group G. Instantons are specific gauge configurations with nontrivial winding number. You should recognize that topology is involved by the epsilon tensor. Recall that topology deals with how manifolds are connected, with no reference to a metric, the domain of differential geometry. The way you can see that the above operator is related to topology is that if we go to curved spacetime it does not depend on the metric. That is because in curved spacetime, the diffeomorphism invariant integration measure picks up a factor of $\sqrt{g}$, where g is the determinant of the metric. This term is required for diffeomorphism invariance because it cancels the Jacobean from a change of variables. On the other hand, the epsilon "tensor" is not really a tensor unless it is accompanied by a factor of $1/\sqrt{g}$, and the two factors cancel, and so no metric appears in the above intewgral, even in curved spacetime. In contrast, typical terms one encounters in the action of quantum field theories, such as mass terms, will depend on the metric and are not topological in nature.

It is striking that the above integrand looks just like the anomalous divergence in the axial current, and this has an interesting implication for fermions: that the Dirac operator in Euclidian spacetime must have exact zero eigenvalues in the presence of gauge fields with nonzero winding number. The exact connection between gauge field topology and eigenvalues of the Dirac operator is the index theorem.

Consider then continuing the anomaly equation eqn. (2.55) to Euclidian space and integrating over spacetime its vacuum expectation value in a background gauge field (assuming the fermions to be in the N-dimensional representation of $SU(N)$ so that $\mathrm{Tr}\ T_a T_b = \frac{1}{2}\delta_{ab}$). The integral of $\partial_\mu \langle j_A^\mu \rangle$ vanishes because it is a pure divergence, so we get

$$\int d^4x_E\, m\langle\overline{\psi}\Gamma\psi\rangle = -\nu\ . \tag{2.59}$$

The matrix element above on the right equals

$$\frac{\int [d\psi][d\overline{\psi}]\, e^{-S_E} \int d^4x_E\, m\, \overline{\psi}\Gamma\psi}{\int [d\psi][d\overline{\psi}]\, e^{-S_E}}\ . \tag{2.60}$$

where $S_E = \overline{\psi}(\not{D}_E + m)\psi$. We can expand ψ and $\overline{\psi}$ in terms of eigenstates of the anti-hermitian operator $\not{D}_E$, where

$$\not{D}_E\psi_n = i\lambda_n\psi_n\ , \qquad \int d^4x_E\, \psi_m^\dagger\psi_n = \delta_{mn}\ , \tag{2.61}$$

with

$$\psi = \sum c_n\psi_n\ , \qquad \overline{\psi} = \sum \bar{c}_n\psi_n^\dagger\ . \tag{2.62}$$

Then

$$S_E = \sum_n (i\lambda_n + m)\bar{c}_n c_n\ , \quad e^{-S_E} = \prod_n \left[1 - \bar{c}_n c_n(i\lambda_n + m)\right]\ , \tag{2.63}$$

and recalling the fermion integration is the same as differentiation (up to signs), we have

$$\int d^4x_E\, m\,\langle\overline{\psi}\Gamma\psi\rangle = \sum_n \int d^4x_E\, m\, \psi_n^\dagger\Gamma\psi_n \frac{\prod_{k\neq n} -(i\lambda_k + m)}{\prod_k -(i\lambda_k + m)}$$

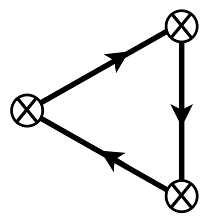

Fig. 2.2 *Anomalous three-point function of three currents.*

$$= -m \sum_n \frac{\int d^4x_E \psi_n^\dagger \Gamma \psi_n}{i\lambda_n + m} \ . \tag{2.64}$$

Recall that $\{\Gamma, \not{D}\} = 0$; thus

$$\not{D}\psi_n = i\lambda_n \psi_n \quad \text{implies} \quad \not{D}(\Gamma\psi_n) = -i\lambda_n(\Gamma\psi_n) \ . \tag{2.65}$$

Thus for $\lambda_n \neq 0$, the eigenstates ψ_n and $(\Gamma\psi_n)$ must be orthogonal to each other (they are both eigenstates of the anti-hermitian operator $\not{D}$ with different eigenvalues), and so $\psi_n^\dagger \Gamma \psi_n$ vanishes for $\lambda_n \neq 0$ and does not contribute to the sum in eqn. (2.64). In contrast, modes with $\lambda_n = 0$ can simultaneously be eigenstates of $\not{D}$ and of Γ; we refer to such solutions as "zeromodes". Let n_+, n_- be the number of RH ($\Gamma = +1$) and LH ($\Gamma = -1$) zeromodes respectively. The last integral in then just equals $(n_+ - n_-) = (n_R - n_L)$, and combining with eqn. (2.59) we arrive at the index equation

$$n_+ - n_- = \nu \ , \tag{2.66}$$

which states that the difference in the number of RH and LH zeromode solutions to the Euclidian Dirac equation in a background gauge field equals the winding number of the gauge field. With N_f flavors, the index equation is trivially modified to read

$$n_+ - n_- = N_f \nu \ . \tag{2.67}$$

This link between eigenvalues of the Dirac operator and the topological winding number of the gauge field provides a precise definition for the topological winding number of a gauge field on the lattice (where there is no obvious topology) — provided we have a definition of a lattice Dirac operator which exhibits exact zeromodes. We will see that the overlap operator is such an operator.

2.5.4 More general anomalies

Even more generally, one can consider the 3-point correlation function of three arbitrary currents as in Fig. 2.2,

$$\langle j_a^\alpha(k) j_b^\beta(p) j_c^\gamma(q) \rangle \ , \tag{2.68}$$

and show that the divergence with respect to any of the indices is proportional to a particular group theory factor

$$k_\mu \langle j_a^\mu(k) j_b^\alpha(p) j_c^\beta(q) \rangle \propto \mathrm{Tr}\ Q_a\{Q_b, Q_c\}\Big|_{R-L} \epsilon^{\alpha\beta\rho\sigma} k_\rho k_\sigma \ , \tag{2.69}$$

where the Qs are the generators associated with the three currents in the fermion representation, the symmetrized trace being computed as the difference between the contributions from RH and LH fermions in the theory. The anomaly $\mathcal{A}$ for the fermion representation is defined by the group theory factor

$$\mathrm{Tr}\ (Q_a\{Q_b, Q_c\})\Big|_{R-L} \equiv \mathcal{A}\, d_{abc} \ , \tag{2.70}$$

with d_{abc} being the totally symmetric invariant tensor of the symmetry group. For a simple group G (implying G is not $U(1)$ and has no factor subgoups), d_{abc} is only nonzero for $G = SU(N)$ with $N \geq 3$; even in the case of $SU(N)$, d_{abc} will vanish for real irreducible representations (for which $Q_a = -Q_a^*$), or for judiciously chosen reducible complex representations, such as $\bar{5} \oplus 10$ in $SU(5)$. For a semi-simple group $G_1 \times G_2$ (where G_1 and G_2 are themselves simple) there are no mixed anomalies since the generators are all traceless, implying that if $Q \in G_1$ and $\mathcal{Q} \in G_2$ then $\mathrm{Tr}\ (Q_a\{\mathcal{Q}_b, \mathcal{Q}_c\}) \propto \mathrm{Tr}\ Q_a = 0$. When considering groups with $U(1)$ factors there can be nonzero mixed anomalies of the form $U(1)G^2$ and $U(1)^3$ where G is simple; the $U(1)^3$ anomalies can involve different $U(1)$ groups. With a little group theory it is not difficult to compute the contribution to the anomaly of any particular group representation.

If a current with an anomalous divergence is gauged, then the theory does not make sense. That is because the divergenceless of the current is required for the unphysical modes in the gauge field A_μ to decouple; if they do not decouple, their propagator has a piece that goes as $k_\mu k_\nu / k^2$ which does not fall off at large momentum, and the theory is not renormalizable.

When global $U(1)$ currents have anomalous divergences, that is interesting. We have seen that the $U(1)_A$ current is anomalous, which explains the η' mass; the divergence of the axial isospin current explains the decay $\pi^0 \to \gamma\gamma$; the anomalous divergence of the baron number current in background $SU(2)$ in the Standard Model predicts baryon violation in the early universe and the possibility of weak-scale baryogenesis.

Exercise 2.5 Verify that all the gauge currents are anomaly-free in the standard model with the representation in eqn. (2.49). The only possible G^3 anomalies are for $G = SU(3)$ or $G = U(1)$; for the $SU(3)^3$ anomaly use the fact that a LH Weyl fermion contributes $+1$ to $\mathcal{A}$ if it transforms as a 3 of $SU(3)$, and contributes -1 to $\mathcal{A}$ if it is a $\bar{3}$. There are two mixed anomalies to check as well: $U(1)SU(2)^2$ and $U(1)SU(3)^2$.

This apparently miraculous cancellation is suggestive that each family of fermions may be unified into a spinor of $SO(10)$, since the vanishing of anomalies which happens automatically in $SO(10)$ is of course maintained when the symmetry is broken to a smaller subgroup, such as the Standard Model.

Exercise 2.6 Show that the global B (baryon number) and L (lepton number) currents are anomalous in the Standard Model eqn. (2.49), but that $B - L$ is not.

2.6 Parity and fermion mass in odd dimensions

Despite these lectures being about chiral fermions, it turns out that we will not only be interested in $d = 2, 4$ but also $d = 3, 5$! In these lectures I will be discussing fermions in $(2k + 1)$ dimensions with a spatially varying mass term which vanishes in some $2k$-dimensional region; in such cases we find chiral modes of a $2k$-dimensional effective theory bound to this mass defect. Such an example could arise dynamically when fermions have a Yukawa coupling to a real scalar ϕ which spontaneously breaks a discrete symmetry, where the surface with $\phi = 0$ forms a domain wall between two different phases; for this reason such fermions are called domain wall fermions, even though we will be putting the spatially dependent mass in by hand and not through spontaneous symmetry breaking.

In odd dimensions there is no analogue of Γ and therefore there is no such thing as chiral symmetry. Nevertheless, fermion masses still break a symmetry: parity. In a theory with parity symmetry one has extended the Lorentz group to include improper rotations: spatial rotations R for which the determinant of R is negative. Parity can be defined as a transformation where an odd number of the spatial coordinates flip sign. In even dimensions parity can be the transformation $\mathbf{x} \to -\mathbf{x}$ and

$$\psi(\mathbf{x}, t) \to \gamma^0 \psi(-\mathbf{x}, t) \qquad \text{(parity, } d \text{ even)} . \tag{2.71}$$

The role of the γ^0 is to transform the kinetic term correctly to realize $\vec{x} \to -\vec{x}$:

$$\gamma^0(\partial_0 \gamma^0)\gamma^0 = \partial_0 \gamma^0 \ , \qquad \gamma^0(\nabla_i \gamma^i)\gamma^0 = -\nabla_i \gamma^i \ . \tag{2.72}$$

Since $\{\gamma^0, \Gamma\} = 0$, ψ_L and ψ_R are exchanged under parity and a Dirac mass term is parity invariant.

However, in odd spacetime dimensions the transformation $\mathbf{x} \to -\mathbf{x}$ is just a proper rotation; instead we must define parity as the transformation which just flips the sign of one coordinate x^1 (or an odd number), and

$$\psi(\mathbf{x}, t) \to \gamma^1 \psi(\tilde{\mathbf{x}}, t) \ , \qquad \bar{\psi}(\mathbf{x}, t) \to -\bar{\psi}(\tilde{\mathbf{x}}, t)\gamma^1 \ , \qquad \tilde{\mathbf{x}} = (-x^1, x^2, \ldots, x^{2k}) \tag{2.73}$$

since

$$-\gamma^1(\partial_\mu \gamma^\mu)\gamma^1 = \begin{cases} +\partial_\mu \gamma^\mu & \mu \neq 1 \\ -\partial_\mu \gamma^\mu & \mu = 1 \end{cases} \qquad \text{(no sum on } \mu) \ . \tag{2.74}$$

Remarkably, a Dirac mass term flips sign under parity in this case; and since there is no chiral symmetry in odd d to rotate the phase of the mass matrix, the sign of the quark mass has physical meaning. Note that it is still possible to define a parity invariant theory of massive fermions, however, provided that they come in pairs with masses $\pm M$, and parity is defined to interchange the two, while flipping the sign of M.

2.7 The non-decoupling of parity violation in odd dimensions

We have seen that fermion masses break chiral symmetry for even d and that they can break parity for odd d. One might then that in a in odd d with a massless fermion,

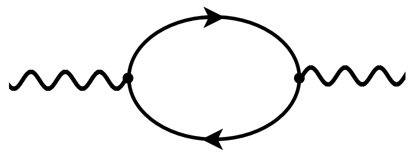

Fig. 2.3 *Integrating out a heavy fermion in three dimensions gives rise to the Chern Simons term in the effective action of eqn.* (2.77).

which is parity invariant, we might get anomalous parity violation from a massive regulator, such as a Pauli-Villars fermion with mass M. Indeed that occurs as can be seen both by power counting and explicit calculation. In $2k+1$ spacetime dimensions the Chern Simons form, which for an Abelian gauge field is proportional to

$$\epsilon^{\alpha_1\cdots\alpha_{2k+1}} A_{\alpha_1} F_{\alpha_2\alpha_3} \cdots F_{\alpha_{2k}\alpha_{2k+1}} \ . \tag{2.75}$$

Not that this operator violates parity and time reversal. This operator is not gauge invariant, but transforms into a total derivative under a gauge transformation, so that its inclusion in the Lagrangian does not spoil the gauge invariance of the action. The Chern Simons form for nonabelian gauge fields is more complicated, and requires a quantized coefficient in order for the action to be gauge invariant. let's count dimensions: with $D_\mu = (\partial_\mu + iA_\mu)$ we have the mass dimension of A_μ is 1, and therefore the mass dimension of the above operator is d. That means that its coefficient in the action must be dimensionless. On the other hand, its coefficient must be proportional to explicit parity symmetry violation. That suggests that on integrating out a Pauli-Villars fermion with mass M, this operator can and should be generated with a coefficient proportional to $M/|M|$, which flips sign when M flips sign. Such a coefficient can arise naturally from an integral such as

$$\int \frac{d^3k}{(2\pi)^3} \frac{M}{(k^2+M^2)^2} = \frac{1}{8\pi}\frac{M}{|M|} \ , \tag{2.76}$$

relevant for the $d=3$ case. Because the coefficient is proportional to $M/|M|$, it survives the limit $M\to\infty$. This effect is called a “parity anomaly” and arises because we have to break parity to regulate the theory.

For domain wall fermions we will be interested in a closely related but slightly different problem: the generation of a Chern Simons operator on integrating out a heavy fermion of mass m. In 2+1 dimensions with an Abelian gauge field one computes the graph in Fig. 2.3, which contributes to the low-energy Lagrangian

$$\mathcal{L}_{CS} = \frac{e^2}{8\pi}\frac{m}{|m|}\epsilon^{\alpha\beta\gamma} A_\alpha \partial_\beta A_\gamma \ . \tag{2.77}$$

It is striking how related the parity and chiral anomalies look. In particular, suppose one differentiates $\mathcal{L}_{CS}$ in eqn. (2.77) with respect to the gauge field to find the current this operator will contribute to Maxwell's equations. For example, in $d=2+1$ one finds:

$$j_\mu = \frac{1}{e}\frac{\partial \mathcal{L}_{CS}}{\partial A_\mu} = \frac{e}{8\pi}\frac{m}{|m|}\epsilon^{\mu\alpha\beta}F_{\alpha\beta} \tag{2.78}$$

where if I pick $\mu = 2$, what I get is related to the chiral anomaly we found in eqn. (1.26) for $1+1$ dimensions. This connection was made clear with a very physical model by Callan and Harvey, and is the basis for domain wall fermions on the lattice, the subject of the next lecture.

Exercise 2.7 Verify the coefficient in eqn. (2.77) by computing the diagram Fig. 2.3. By isolating the part that is proportional to $\epsilon_{\mu\nu\alpha}p^\alpha$ before performing the integral, one can make the diagram very easy to compute.

References

Fujikawa, Kazuo (1979). *Phys. Rev. Lett.*, **42**, 1195.

Fujikawa, Kazuo (1980). *Phys. Rev.*, **D21**, 2848.

Itzykson, C. and Zuber, J. B. (1980). *Quantum Field Theory.* New York, Usa: Mcgraw-hill, 705 P.(International Series In Pure and Applied Physics).

Wess, J. and Bagger, J. (1992). *Supersymmetry and supergravity.* Princeton, USA: Univ. Pr. (1992) 259 p.