Zero modes of fermions in flux backgrounds

Ethan Lake

Preface:

These notes (pulled from previous physics diary entries) contain a few calculations relating to free fermions placed in various flux backgrounds.

Contents

1 Zero modes of $i\not{D}_A$ on the sphere

In this section we're going to a calculation whose results I've heard mentioned in a few papers, but have never actually seen worked out anywhere (I likely just did not search hard enough). We will explicitly construct the zero modes of the Dirac operator on a sphere, in the presence of an arbitrary $U(1)$ magnetic flux.

◇ ◇ ◇

Spherical coordinates

In what follows, we will be using veilbeins, since that's the only method we have for dealing with fermions on curved spaces.[1] Recall that the veilbeins are found by fractionalizing the metric:

$$g_{\mu\nu} = e^a_\mu e^a_\nu, \qquad e^a_\mu e^b_\nu g^{\mu\nu} = \delta^{ab}. \tag{1}$$

[1] This is because fermion actions need γ^a matrices to be defined, which represent Clifford algebras. We want to represent a Clifford algebra with the relation $\{\gamma^a, \gamma^b\} = 2\delta^{ab}$ (or maybe η^{ab}), and definitely don't want to have the anticommutator be equal to $g^{ab}(x)$; this would be a mess. Thus we need veilbeins to switch between spacetime and a frame in which the Clifford generators can be defined.

Since we don't want to constantly be phantoming when writing stuff out, our convention will be that, when viewing the veilbeins as a matrix, the greek (spacetime) letter will always denote the row index of the matrix, and the roman (internal space) letter will always denote the column index. When we break apart the metric like this, we pick up a gauge redundancy, since the transformation $e^a_\mu \mapsto [O]^a_b e^b_\mu$ for $O \in O(s,t)$ leaves the splitting $g_{\mu\nu} = e^a_\mu e^a_\nu$ invariant (in what follows we will only be concerned with 2+0 dimensions, so that the relevant "gauge group" is $O(2)$).

The Dirac operator is (roman indices can be raised / lowered with impunity)

$$\not{D}_A = \gamma_a e^{\mu a}(\partial_\mu + i(\omega_\mu + A_\mu)), \tag{2}$$

where ω, A are the spin and gauge connections, with

$$\omega_\mu{}^a{}_b = e^a_\nu \partial_\mu e^\nu_b + e^a_\nu \Gamma^\nu_{\mu\lambda} e^\lambda_b. \tag{3}$$

We've tried to take a sign convention that is maximally simple; ours differs from the conventions in many other places though so be careful. The spin connection is needed to ensure that $\not{D}_A\psi$ transforms covariantly under local $O(2)$ gauge rotations of the coordinate frames.[2]

Since the generators of Spin(d) are $-i[\gamma_a,\gamma_b]/4$, the spin connection is, quite generally,

$$\omega_\mu = \frac{1}{2}\omega^{ab}_\mu \Sigma_{ab}, \qquad \Sigma_{ab} = \frac{-i}{4}[\gamma_a,\gamma_b]. \tag{8}$$

[2] It is also needed to ensure that the (real-time) action is Hermitian! The added spin-connection term is actually not Hermitian, and this compensates for the non-Hermiticity of $\bar\psi(i\not{\partial})\psi$ when working on a curved manifold. Indeed, under Hermitian conjugation,

$$\dagger : i\bar\psi e^{a\mu}\gamma_a\partial_\mu\psi \mapsto i\bar\psi e^{a\mu}\gamma_a\partial_\mu\psi + i\bar\psi(\partial_\mu e^{a\mu})\gamma_a\psi. \tag{4}$$

Now let's look at the spin connection part. For simplicity, we will work in Riemann normal coordinates around a certain point p, where the Christoffel symbols (but not their derivatives) can be chosen to vanish. The spin connection part of the Lagrangian density at this point is then

$$\begin{aligned}\mathcal{L} \ni -\frac{1}{2}\sum_{a,b\neq c}\psi^\dagger\gamma^0\gamma_a e^{a\mu}\omega^{bc}_\mu\Sigma_{bc}\psi &= \frac{i}{2}\sum_{a,b\neq c}\psi^\dagger\gamma^0\gamma_a e^{a\mu}(e^b_\nu\partial_\mu e^{c\nu})\gamma_b\gamma_c\psi \\ = \sum_{a\neq b\neq c}\frac{i}{4}\bar\psi e^{a\mu}e^b_\nu\partial_\mu e^{\nu c}\gamma_a\gamma_b\gamma_c\psi &+ \sum_{a\neq c}\frac{i}{4}\bar\psi e^{a\mu}(e^a_\nu\partial_\mu e^{\nu c} - e^c_\nu\partial_\mu e^{\nu a})\gamma_c\psi \\ = \sum_{a\neq b\neq c}\frac{i}{4}\bar\psi e^{a\mu}e^b_\nu\partial_\mu e^{\nu c}\gamma_a\gamma_b\gamma_c\psi &+ \frac{i}{2}\sum_c\bar\psi\partial_\nu e^{\nu c}\gamma_c\psi\end{aligned} \tag{5}$$

The first term here is Hermitian: using $\gamma_a^\dagger\gamma_0^\dagger = \gamma_0\gamma_a$ (we are in $\mathbb{R}$ time, remember), we have

$$\dagger : \frac{i}{4}\bar\psi e^{a\mu}e^b_\nu\partial_\mu e^{\nu c}\psi \mapsto -\frac{i}{4}\psi^\dagger e^{a\mu}e^b_\nu\partial_\mu e^{\nu c}(\gamma_c^\dagger\gamma_b^\dagger\gamma_a^\dagger)\gamma_0^\dagger = \frac{i}{4}\bar\psi e^{a\mu}e^b_\nu\partial_\mu e^{\nu c}\gamma_c\gamma_b\gamma_a = \frac{i}{4}\bar\psi e^{a\mu}e^b_\nu\partial_\mu e^{\nu c}\psi, \tag{6}$$

since here $a \neq b \neq c$.

However, the second term at the end of (5) is actually anti-Hermitian:

$$\dagger : \frac{i}{2}\bar\psi\partial_\nu e^{\nu c}\gamma_c\psi \mapsto -\frac{i}{2}\bar\psi\partial_\nu e^{\nu c}\gamma_c\psi. \tag{7}$$

However, when we add in the $+i\bar\psi(\partial_\mu e^{a\mu})\gamma_a\psi$ from (4), we see that it combines with the RHS of the above equation to yield the LHS, giving an action that is Hermitian. Thus the second term at the end of (5) is a counterterm that ensures that the full action is Hermitian.

The veilbeins for spherical coordinates on the unit S^2 are easy to write down:

$$e^a_\mu = \begin{pmatrix} 1 & 0 \\ 0 & \sin\theta \end{pmatrix}^a_{\ \mu}, \qquad e^{\mu a} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^{-1}\theta \end{pmatrix}^{\mu a}. \tag{9}$$

Here the fact that the tetrads are the "square root of the metric" is made manifest. Of course, there are infinitely many other choices, related by local $O(2)$ transformations.

To get the spin connection, we will need to know that the nonzero Christoffel symbols on the sphere are

$$\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta, \qquad \Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta. \tag{10}$$

Then we calculate

$$\omega^{ab}_\theta = 0, \qquad \omega^{ab}_\phi = \cos\theta J^{ab}. \tag{11}$$

Then, if we adopt the gamma matrices $\gamma^1 = X, \gamma^2 = Y$ (this is the best choice since it makes the splitting $S = S_+ \oplus S_-$ manifest), we get

$$\omega_\theta = 0, \qquad \omega_\phi = \frac{1}{2}\cos\theta J^{ab}\Sigma_{ab} = -\frac{\cos\theta}{2}Z. \tag{12}$$

Finally, we need an expression for A_μ. We will make the usual choice for a monopole on S^2, namely (note how similar the forms of the gauge and spin connections are!)

$$A^{N/S} = n\frac{\pm 1 - \cos\theta}{2}d\phi, \tag{13}$$

which gives $\int_{S^2} F = 2\pi n$. The covariant derivatives are

$$\nabla_\theta = \partial_\theta, \qquad \nabla_\phi = \partial_\phi - \frac{iZ}{2}\cos\theta + in\frac{\pm 1 - \cos\theta}{2}. \tag{14}$$

We can now finally write down the expression for $i\not{D}_A\psi = 0$, which is

$$\not{D}_A\psi^{(N/S)} = \left[X\left(\partial_\theta + \frac{\cot\theta}{2}\right) + Y\csc\theta\left(\partial_\phi + in\left(\frac{\pm 1 - \cos\theta}{2}\right)\right)\right]\psi^{(N/S)} = 0 \tag{15}$$

or written out, (note to self: may need to come back and sort out some minus signs)

$$\begin{aligned} \left(\partial_\theta + \frac{\cot\theta}{2} - i\csc\theta\partial_\phi - n\csc\theta\left(\frac{\pm 1 - \cos\theta}{2}\right)\right)\psi_R^{(N/S)} &= 0 \\ \left(\partial_\theta + \frac{\cot\theta}{2} + i\csc\theta\partial_\phi + n\csc\theta\left(\frac{\pm 1 - \cos\theta}{2}\right)\right)\psi_L^{(N/S)} &= 0 \end{aligned} \tag{16}$$

These equations are actually very easy to solve: the ϕ dependence is $e^{il\phi}$ by symmetry, while the θ dependence is figured out by the common factor of $1/\sin\theta$ in the equations.

For $n = 0$, a solution with $l = 0$ is $\psi_{R/L} = 1/\sqrt{\sin\theta}$. However, while normalizable, this is not differentible, and therefore is not an allowed solution.[3] So there are no zero modes when $n = 0$. This follows from the fact that there are no zero modes on spaces where the curvature scalar is nowhere negative; more on this in a separate diary entry.

Now consider $n = 1$. From the index theorem we know there should be one more L zero mode than R zero mode. And indeed, there is one L zero mode, and no R zero modes. The L zero mode is just

$$\psi_L^N(1, 1/2) = e^{i\phi/2}, \qquad \psi_L^S(1, 1/2) = e^{-i\phi/2}, \tag{18}$$

where we have adopted the notation $\psi_{L/R}^{N/S}(n, l)$. Note that the gluing transition function on the equator correctly recovers the $n = 1$ flux. Conversely, if $n = -1$ then we have an R zero mode and no L zero mode: the R zero mode is

$$\psi_R^N(-1, -1/2) = e^{-i\phi/2}, \qquad \psi_R^S(-1, -1/2) = e^{i\phi/2}. \tag{19}$$

Now take $n = \pm 2$. For $n = 2$ we expect two more L zero modes than R zero modes, and indeed, the two L zero modes are (there are no R zero modes)

$$\psi_L^N(2, 3/2) = \sin(\theta/2) e^{i3\phi/2}, \qquad \psi_L^S(2, -1/2) = \sin(\theta/2) e^{-i\phi/2} \tag{20}$$

and

$$\psi_L^N(2, 1/2) = \cos(\theta/2) e^{i\phi/2}, \qquad \psi_L^S(2, -3/2) = \cos(\theta/2) e^{-3i\phi/2}. \tag{21}$$

The situation is reversed for $n = -2$: the two zero modes are

$$\psi_R^N(-2, -3/2) = \sin(\theta/2) e^{-i3\phi/2}, \qquad \psi_R^S(-2, 3/2) = \sin(\theta/2) e^{i3\phi/2} \tag{22}$$

and

$$\psi_R^N(-2, -1/2) = \cos(\theta/2) e^{-i\phi/2}, \qquad \psi_R^S(-2, 3/2) = \cos(\theta/2) e^{3i\phi/2}. \tag{23}$$

In general, for flux n, there are n L zero modes and no R zero modes if $n > 0$, while there are n R zero modes and no L zero modes if $n < 0$. The n modes come in a series $\psi_{L/R}^N(n, l)$ with $l = n - 1/2, n - 3/2, \dots, \pm 1/2$, with the $+$ for L and the $-$ for R. The corresponding functions on the southern patch are related by $l \mapsto l - n$, which ensures that the transition

[3] We are looking only for solutions in the domain of $i\not{D}_A$, which by definition are C^∞ sections of the bundle $S \otimes E$, where S is the spinor bundle and E is the gauge bundle. Indeed, the Laplacian and the Dirac operator are only defined on infinitely smooth functions, and stuff can go wrong if our functions are not infinitely differentible. For example, consider $\psi_R(\theta, \phi) = 1/\sqrt{\sin(\theta)}$. This looks like a totally fine zero mode solution for the zero-flux case $n = 0$, since $i\not{D}_0\psi = 0$ and also $\int \psi_R^\dagger \psi_R \sin\theta = 2\pi^2$ is finite, so that ψ_R is L_2 on the S^2. However, weird pathologies come up due to the non-differentibility at $0, \pi$. For example, while $\not{D}_0\psi = 0$,

$$\not{D}_0^2\psi = \left(-\csc\theta\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{4} + \frac{1}{4\sin^2\theta} \right)\psi = \sin^{-5/2}\theta(1 - \cos^2(\theta)/4) - \sin^{-1/2}(\theta)/4 \neq 0. \tag{17}$$

The problem here is that $\not{D}(\not{D}\psi) \neq (\not{D}^2)\psi$ and so the positive-definiteness arguments we made don't apply to ψ (this is why we don't want to let functions like ψ be part of the domain of definition of the Dirac operator).

Anyway, I'm still not really sure about the physicality of such a restriction. Should I really be bothered if ψ is non-differentible, provided that $\sqrt{|g|}|\psi(x)|^2$ is finite everywhere? Not sure.

function on the equator is a large gauge transformation with winding $2\pi n$. To find out what the spins (alias angular momentum) of these zero modes are, we need to calculate the L^2 angular momentum operator, which is modified by the spin and gauge connections; we will do this in a little bit.

We've been working with a uniform field strength, but of course (by the index theorem), we know that the zero modes must persist if we take an arbitrary field configuration with the same value of $\int F$. For example, we might add the vector potential $\widetilde{A} = g(\theta)d\phi$ to the existing monopole potential, where $\widetilde{A}$ is globally defined on the sphere. Then we just have to modify our zero mode solution by $\psi_{L/R} \mapsto f_{L/R}(\theta)\psi_{L/R}$, where $f_{L/R}(\theta)$ satisfies

$$(\partial_\theta \pm \csc(\theta)g(\theta))f(\theta) = 0 \implies \psi_{L/R} = \exp\left(\mp\int_0^\theta d\theta'\ \csc(\theta')g(\theta')\right)\psi^0_{L/R}, \tag{24}$$

where $\psi^0_{L/R}$ is the solution for the homogeneous field. From this expression we see that we get a well-defined answer only if $g(\theta) \to \theta$ as $\theta \to 0, \pi$; this is the condition that $\widetilde{A}$ go to zero at the poles in this coordinate system, so that $\widetilde{A}$ is topologically trivial. For example, consider the L zero mode, and let $g(\theta) = \sin 2\theta$. Then we see that the zero-mode solution is modified by a factor of $f_L(\theta) = e^{-2\sin\theta}$, so that the weight of the wavefunction becomes concentrated near the poles, where the normalized field strength $F/\sin\theta$ becomes largest; this is because as we have seen, the L zero modes "like" positive field strength. More generally, for an arbitrary asymmetric $\widetilde{A}$, we just have to multiply our symmetric zero mode solution by a factor $f(\theta,\phi)$, where

Anyway, now some more comments on the symmetric (uniform field strength) case. we've seen that the zero mode states fall into half-odd-integer representations of $SU(2)$. This is no surprise given the symmetry of the problem (in fact, the whole spectrum of $\not{D}$ falls into $SU(2)$ representations, not just the zero modes).

To find the generators of the $SU(2)$, it is not simply enough to covariantize by making the replacement $\mathbf{L} = -i\mathbf{n}\times\partial \mapsto -i\mathbf{n}\times\nabla$, with $\mathbf{n}$ the unit vector on the sphere. Indeed, doing this leads to generators that fail to satisfy the correct $SU(2)$ commutation relations, since in the presence of background field strengths the covariant momenta $-i\nabla$ fail to commute (their commutator measures the field strength). The commutation relations are ruined not just by the gauge background field, but also by the field strength of the spin connection (the geometric curvature of the sphere). Now viewing the S^2 as living in three dimensions, the Hodge duals of the field strengths of both the gauge field and the spin connection are oriented along $\mathbf{n}$ (the magnetic field for both spin and gauge connections is radial), and so we can write $(d[\omega + A])_{\mu\nu} = \epsilon_{\mu\nu\lambda}n^\lambda B$, where in the present case $B = \mathbf{1}n/2 + Z/2$. One can check that in this case,

$$[(-i\mathbf{n}\times\nabla)^\mu, (-i\mathbf{n}\times\nabla)^\nu] = i\epsilon^{\mu\nu\lambda}((-i\mathbf{n}\times\nabla)_\lambda - Bn_\lambda). \tag{25}$$

This prompts us to take the ansatz

$$L_\mu = \epsilon_{\mu\nu\lambda}n^\nu\nabla^\lambda + Bn^\mu \tag{26}$$

for the angular momentum generators. Indeed, one can check (see the next diary for the calculation) that with this choice, the L_μ satisfy the usual $SU(2)$ commutation relations.

We can now write down the angular momentum generators explicitly. We have

$$\begin{aligned} L_z &= -i\nabla_\phi + \cos\theta B \\ L_\pm &= e^{\pm i\phi}\left(\pm\nabla_\theta + i\cot\theta\nabla_\phi + B\sin\theta\right). \end{aligned} \tag{27}$$

When simplifying this, we will use the covariant derivative $\nabla_\phi = \partial_\phi - i(Z\cos\theta + n\cos\theta)/2$, which differs from the one written above by the term $\pm in/2$, which doesn't affect the field strength and hence can be dropped without affecting the angular momentum commutation relations. However, one must use caution with this convention, since changing the convention for the gauge field *does* change the expressions for the eigenstates of $\not{D}_A$. This means that the zero mode eigenstates obtained above will *not* be appropriate eigenfunctions of the L_z and L^2 operators obtained below! Maybe someday I'll come back and redo this so that the conventions are the same. Anyway, plugging in and simplifying, we find

$$\begin{aligned} L_z &= -i\partial_\phi \\ L_\pm &= e^{\pm i\phi}\left(\pm\partial_\theta + i\cot\theta\partial_\phi + \frac{1}{2\sin\theta}(Z+n)\right). \end{aligned} \tag{28}$$

I've done a check in mathematica of the commutation relations, and they work! Yay! Note that if we had chosen to keep the factor of $\pm d\phi/2$ in the gauge connection, the eigenvalues of L_z would be shifted by $\pm 1/2$: therefore the eigenvalues of L_z are not a gauge-invariant thing to calculate, and they do not tell you about the spin of the zero mode.

As a reminder, one should not confuse the total angular momentum operator $L^2 = L_z^2 + \frac{1}{2}(L_+L_- + L_-L_+)$ with the (negative of, depending on conventions) covariant Laplacian $-\nabla_\mu\nabla^\mu$. Indeed, the Laplacian is (we are working in conventions where the Laplacian is positive-definite)

$$\begin{aligned} -\Delta = \nabla_\mu\nabla^\mu &= \nabla_\theta^2 + \cot\theta\,\nabla_\theta + \csc^2\theta\,\nabla_\phi^2 \\ &= \partial_\theta^2 + \cot\theta\,\partial_\theta + \csc^2\theta\left(\partial_\phi^2 - i(Z+n)\cos\theta\partial_\phi - \frac{\cos^2\theta}{4}(Z+n)^2\right). \end{aligned} \tag{29}$$

In contrast, the angular momentum operator is

$$L^2 = \Delta + \frac{1}{4}(Z+n)^2. \tag{30}$$

We would like to use this result to figure out the angular momentum of the zero modes obtained previously. In our current conventions, the zero mode equation is, on the N hemisphere

$$\begin{aligned} \left(\partial_\theta + \frac{\cot\theta}{2} + i\csc\theta\partial_\phi + \frac{n}{2}\cot\theta\right)\psi_R = 0 \\ \left(\partial_\theta + \frac{\cot\theta}{2} + i\csc\theta\partial_\phi - \frac{n}{2}\cot\theta\right)\psi_L = 0 \end{aligned} \tag{31}$$

This choice for the gauge field makes finding the zero modes slightly easier. For example, if $n = 1$ we see that we get a solution where $\psi_R = 1/\sqrt{4\pi}, \psi_L = 0$. The fact that the zero mode is a constant suggests that it has spin zero: to check, we act on it with L^2:

$$L^2(n=1)\psi_R = \Delta(n=1)\psi_R = 0, \tag{32}$$

as expected.

The last thing we square a lot is the Dirac operator: more on this and its relation to Δ in a future diary entry.

Stereographic projection

Now we will go through the problem again using stereographic coordinates. These coordinates are nicer since they are less singular than spherical coordinates. We usually cover the sphere in two hemispherical patches, but doing it this way means that both patches will contain a coordinate singularity for ϕ, which is no good (although for us it's okay since the zero mode solutions vanish at the poles so we can still use two patches to construct a single-valued zero mode solution). Getting a good covering of S^2 requires 4 patches, and in order to escape coordinate singularities, the patches have to have different definitions of ϕ, θ such that the $\theta = 0, \pi$ points in a given patch's coordinate system don't occur within that patch itself. Gluing together zero mode solutions like this really is a hopeless mess.

By contrast, stereographic coordinates are great! We only need two patches to cover the S^2 (not goodly, but that's okay), and within each patch we can use a metric which is perfectly singularity-free, and in fact is conformally equivalent to flat Euclidean space. Indeed, recall from a few diary entries ago that in stereographic projection of S^2 onto the plane, the metric assumes the conformally flat form (assuming the sphere has radius 1 for simplicity)

$$ds^2 = \frac{4}{(1+r^2)^2}(dx^2 + dy^2). \tag{33}$$

The tetrads are simple is this coordinate system:

$$e^a_\mu = \frac{2}{1+r^2}\mathbf{1}^a_\mu, \qquad e^{\mu a} = \frac{1+r^2}{2}\mathbf{1}^{\mu a}. \tag{34}$$

The Christoffel symbols are easily calculated to be (I won't write out the algebra)

$$\Gamma^\mu_{\nu\lambda} = \frac{2}{1+r^2}(x^\mu \delta_{\nu\lambda} - x_\nu \delta^\mu_\lambda - x_\lambda \delta^\mu_\nu). \tag{35}$$

We can then calculate

$$\omega^{ab}_\mu = \frac{2x_\mu}{1+r^2}\delta^{ab} + \delta^a_\nu \delta^{b\lambda}\Gamma^\nu_{\mu\lambda}. \tag{36}$$

Only the off-diagonal part is non-zero:

$$\omega^{12}_\mu = -\omega^{21}_\mu = \frac{2}{1+r^2}(x\delta_{\mu y} - y\delta_{x\mu}), \tag{37}$$

and so the full spin connection is

$$\omega_\mu = \frac{1}{2}\omega^{ab}_\mu \Sigma_{ab} = -\frac{i}{2}\sum_{a<b}\omega^{ab}_\mu \gamma_a\gamma_b = \frac{Z}{1+r^2}(x\delta_{\mu y} - y\delta_{\mu x}). \tag{38}$$

Therefore the equation $\not{D}_A\psi = 0$ reads

$$\left[X\left(\partial_x + iA_x - iZ\frac{y}{1+r^2}\right) + Y\left(\partial_y + iA_y + iZ\frac{x}{1+r^2}\right)\right]\psi = 0 \tag{39}$$

Decomposing this with $\psi = (\psi_L, \psi_R)^T$,

$$\begin{aligned} \left(2\partial + 2iA - \frac{\bar{z}}{1+|z|^2}\right)\psi_R = 0 \\ \left(2\bar{\partial} + 2i\bar{A} - \frac{z}{1+|z|^2}\right)\psi_L = 0, \end{aligned} \tag{40}$$

where $z = x + iy$, $A = (A_x - iA_y)/2$, and $\partial = (\partial_x - i\partial_y)/2$.

Now we need to get an expression for A. We want the field strength to be proportional to the volume form, which is $4/(1+r^2)^2$. In complex coordinates,

$$\frac{4}{(1+r^2)^2}(dx^2 + dy^2) = \frac{4}{(1+|z|^2)^2}dzd\bar{z}. \tag{41}$$

Thus if we take

$$A = in\frac{\bar{z}}{2(1+|z|^2)}, \quad \bar{A} = -in\frac{z}{2(1+|z|^2)}, \tag{42}$$

then we have

$$F_{z\bar{z}} = \partial\bar{A} - \bar{\partial}A = i\frac{n}{1+|z|^2}\left(1 - \frac{|z|^2}{1+|z|^2}\right) = i\frac{n}{(1+|z|^2)^2}, \tag{43}$$

which gives

$$\int_{\mathbb{R}^2} F_{z\bar{z}}\, dz \wedge d\bar{z} = \int_{\mathbb{R}^2} \frac{n}{(1+r^2)^2} i(-2idx \wedge dy) = 2\pi n \int_0^\infty dr\, \frac{2r}{(1+r^2)^2} = 2\pi n \tag{44}$$

as desired. Thus the equations for the zero modes are (I think I missed a factor of 2 somehwere, which I am re-instating below in a very ad hoc manner: I think it is needed to get the right zero mode solutions)

$$\begin{aligned} \left(2\partial - (1/2+n)\frac{\bar{z}}{1+|z|^2}\right)\psi_R = 0 \\ \left(2\bar{\partial} - (1/2-n)\frac{z}{1+|z|^2}\right)\psi_L = 0. \end{aligned} \tag{45}$$

The solutions are then

$$\begin{aligned} \psi_R(z,\bar{z}) = f_R(\bar{z})(1+|z|^2)^{(1/2+n)/2} \\ \psi_L(z,\bar{z}) = f_L(z)(1+|z|^2)^{(1/2-n)/2}, \end{aligned} \tag{46}$$

where the Laurent series for $f_R(\bar{z})$ and $f_L(z)$ only involve terms of non-negative degree since we require $\psi_{R/L}$ to be finite at the origin. Now, requiring that the zero mode solutions be finite at ∞ tells us that no L zero modes exist if $n \leq 0$, while no R zero modes exist if $n \geq 0$, in agreement with what we found before. Wolog, take $n > 0$ and look at the $\psi_L(z,\bar{z})$ solutions. We can take them to be eigenstates of L_z, which in complex coordinates is

$$L_z = -i(X\partial_Y - Y\partial_X) = -i(x\partial_y - y\partial_x) = \frac{z+\bar{z}}{2}(\partial - \bar{\partial}) + \frac{z-\bar{z}}{2}(\partial + \bar{\partial}) = z\partial - \bar{z}\bar{\partial}, \tag{47}$$

where X, Y are coordinates in 3-space on the S^2 (note that the contribution from the connection to the covariant derivative has canceled with the modification of the angular momentum generators required in a magnetic field; see a subsequent diary entry for details). Note that functions only of $|z|$ have no angular momentum, as required. This means that z^α is an eigenfunction of L_z with eigenvalue $+\alpha$, while $\bar{z}^\alpha$ is an eigenfunction with eigenvalue $-\alpha$. Therefore in a basis in which L_z is diagonalized with eigenvalue $l \in \frac{1}{2}\mathbb{Z}$, the $f_L(z), f_R(\bar{z})$ will be proportional to z^l and $\bar{z}^l$, respectively. So, the zero mode eigenfunctions of L_z for $n > 0$ are of the form $z^l(1+|z|^2)^{1/4-n/2}$. At $r \to \infty$ this goes as $r^{l-n+1/2}$, so we require that $l \leq n - 1/2$ (having a constant is okay since constants have finite integrals due to the conformal factor in the metric). This recovers the situation where we have n L zero modes of angular momentum $l = n-1/2, n-3/2, \dots, 1/2$ if $n > 0$, and n R zero modes with $l = n+1/2, \dots, -1/2$ if $n < 0$.

Another way to derive the parity anomaly

Finally we mention a zero-mode-focused way of deriving the parity anomaly that I learned about from one of Seiberg's lectures at the Jerusalem winter school.

Consider the case of a 2+1D theory on $S^2 \times \mathbb{R}$, with a unit of flux through the spatial S^2. Then the results above tell us that the Hamiltonian (which here is just the spatial part of the Lagrangian) has a single zero mode; when we quantize we thus get two states $|0\rangle$ and $|1\rangle = \chi^\dagger|0\rangle$.

Now the charges of $|1\rangle$ and $|0\rangle$ must satisfy $q_0 = q_1 - 1$: this is just a consequence of $Q\chi^\dagger = \chi^\dagger(Q+1)$. Now, using CT symmetry,[4] we have (taking $CT|0\rangle = |1\rangle$; a possible phase factor here doesn't contribute to the discussion)

$$CTe^{iQ}|0\rangle = CTe^{iq_0}|0\rangle = e^{-iq_0}|1\rangle \tag{48}$$

but also

$$CTe^{iQ}|0\rangle = e^{iQ}CT|0\rangle = e^{q_1}|1\rangle, \tag{49}$$

so that $q_1 = -q_0$. Thus if CT really is a symmetry, we have $q_0 = -1/2, q_1 = +1/2$. This however means that both $|0\rangle$ and $|1\rangle$ are not gauge-invariant; a contradiction. Hence CT must actually be broken. Of course, this is the parity anomaly, and the way that CT gets broken is by T getting broken.

◇ ◇ ◇

2 Zero modes for Dirac fermions on the torus

This section is another simple exercise, but which I hadn't seen worked out in textbooks: finding the spectrum of Dirac fermions on a torus in the presence of nonzero net magnetic flux.

[4] The presence of $\int dxdy\, F_{xy} = 2\pi$ means that both C and T are broken (using the historical definition of T under which magnetic fields are odd), while CT is preserved. Thus it makes sense to ask about how CT acts within the subspace of the zero modes, but not C or T individually.

◇ ◇ ◇

We will work on a square torus with both side lengths set to 1 for simplicity. Suppose that over the torus, the integral of the field strength is $B \equiv \int F = 2\pi n$. We will split the torus up into two cylindrical coordinate patches. The first will be $U_1 = \{0 \le x \le 1/2, 0 \le y \le 1\}$, and the second will be $U_2 = \{1/2 \le x \le 1, 1/2 \le y \le 1\}$, with $1 \sim 0$ in both limits. The two patches overlap along the line $x = 1/2$ and along the line $x = 0 \sim 1$. We will take the transition function to be trivial on the former overlap region, and to be the exponential of a function that winds by $2\pi n$ around the y direction, viz. $g_{12} = e^{iyB}$. The natural choice for the gauge field is then $A = (0, Bx)$. This gives the correct field strength and the way it is glued between the two patches is determined correctly as

$$A^2_\mu(1, y) = g^{-1}_{12}(A^1_\mu(0, y) - i\partial_\mu)g_{12}. \tag{50}$$

The Hamiltonian $H = -i(X[\partial_x - iA_x] + Y[\partial_y - iA_y])$ is then, in this gauge,

$$H = -i \begin{pmatrix} 0 & \partial_x - i\partial_y - Bx \\ \partial_x + i\partial_y + Bx & 0 \end{pmatrix}. \tag{51}$$

Define the operator

$$\gamma \equiv \frac{1}{\sqrt{2|B|}}(\partial_x + i\partial_y + Bx), \qquad H = -i\sqrt{2|B|}\begin{pmatrix} 0 & -\gamma^\dagger \\ \gamma & 0 \end{pmatrix}. \tag{52}$$

Then

$$[\gamma, \gamma^\dagger] = \mathrm{sgn}(B). \tag{53}$$

In what follows we will assume $B > 0$, and so $\gamma, \gamma^\dagger$ obey the usual harmonic oscillator algebra.

Now if we square $H\psi = E\psi$, we get

$$2B \begin{pmatrix} \gamma^\dagger\gamma & 0 \\ 0 & \gamma^\dagger\gamma + 1 \end{pmatrix} \psi = E^2\psi. \tag{54}$$

Therefore the energy levels are

$$E_n = \pm\sqrt{2Bm}, \qquad m \in \mathbb{N}. \tag{55}$$

Now since $\gamma^\dagger\gamma$ has only non-negative eigenvalues, we see that for $\psi = (\psi_L,\, \psi_R)^T$, we can have ψ_L zero modes, but cannot have ψ_R zero modes. If we were to change the sign of the flux $B > 0$ by $B \mapsto -B$, then in order to maintain the right commutation relations, we would need to interchange γ and $\gamma^\dagger$. This would then give $H = 2|B|(\gamma^\dagger\gamma + 1) \oplus \gamma^\dagger\gamma$, which is the same as for $B > 0$, but with left and right components switched. Therefore for $B > 0$ we can have only L zero modes, while for $B < 0$ we can only have R zero modes. This is in agreement with the index theorem.

Anyway, to solve for the eigenspectrum, we need to find $|0\rangle$ such that $\gamma|0\rangle = 0$; we can then build up the spectrum by acting on this with creation operators. Since y doesn't appear in γ we can give the $|0\rangle$ wavefunction a y dependence of e^{iyk_y}. We then have $(\partial_x - k_y + Bx)\psi_0(x,y) = 0$, where $\langle x,y|0\rangle = \psi_0(x,y)$. This gives

$$\psi_0(x,y) = e^{iyk_y}e^{-(x-k_y/B)^2B/2}. \tag{56}$$

If we set PBC in the y direction, then we need $k_y \in 2\pi\mathbb{Z}$. For the x direction, setting $x \sim x+1$ means that $k_y \sim k_y + B$, Therefore $B = 2\pi n$ means that we have n different options for k_y, and so the degeneracy of the zero energy states is n. The zero modes are then $\psi = (\psi_0, 0)^T$. Note that the zero modes survive the introduction of a non-uniform perturbing flux, so long as that flux integrates to zero over the torus. For example, adding on $\widetilde{A} = \widetilde{A}_y(x)dy$ to A modifies the zero mode solution by a factor of $\exp(-\int_0^x dx'\,\widetilde{A}_y(x))$, which preserves the boundary conditions on ψ_0 provided that $\int d\widetilde{A} = 0$ (for example, we could take $\widetilde{A}_y(x) = \sin(2\pi x)$).

Excited states are constructed by acting with $(\gamma^\dagger)^n$ on $|0\rangle$ in the usual way. For example, $\gamma^\dagger|0\rangle = |1\rangle = \sqrt{2B}(x-k_y/B)\psi_0$. The state with energy $\pm\sqrt{2Bm}$, $m>0$ is formed by taking $|m\rangle$ for the left component and $\mp i|m-1\rangle$ for the right component:

$$H\psi_{\pm m} = H\begin{pmatrix} |m\rangle \\ \mp i|m-1\rangle \end{pmatrix} = \sqrt{2B}\begin{pmatrix} 0 & i\gamma^\dagger \\ -i\gamma & 0 \end{pmatrix}\begin{pmatrix} |m\rangle \\ \mp i|m-1\rangle \end{pmatrix} = \pm\sqrt{2Bn}\begin{pmatrix} |m\rangle \\ \mp i|m-1\rangle \end{pmatrix}. \tag{57}$$

Suppose we require that the system be CT symmetric. Now in this basis, charge conjugation $C : \psi \mapsto C\psi^\dagger$ needs to satisfy $[C^\dagger(X,Y)C]^T = -(X,Y)$, and so we can take $C = Y$. If T acts as $T : \psi \mapsto Y\psi$, $i \mapsto -i$, then CT acts just as $\psi \mapsto \psi^\dagger, i \mapsto -i$. Note that $CT : \psi_{\pm m} \mapsto \psi_{\mp m}$, as expected for particle-hole symmetry.

Let $|+\rangle$ be the many-body state with the zero mode filled, and $|-\rangle$ be the many-body state with the zero mode unfilled. Since the charge operator $e^{iQ} = e^{i\int dx\,dy\,\psi^\dagger\psi}$ commutes with CT, by acting on $|\pm\rangle$ with CTe^{iQ} in two different ways (directly, or by moving the e^{iQ} to the left first), we get $e^{iq_+} = e^{-iq_-}$. Since $q_+ = q_- + 1$ (because $C : |\pm \leftrightarrow |\mp\rangle$ and because $e^{iQ}C = Ce^{-iQ}$), we then conclude that CT symmetry implies $q_\pm = \pm 1/2$. If we break C symmetry with e.g. $\delta H \propto \gamma^\dagger\gamma\mathbf{1}$, then since the charge is quantized, the charge assignment of the zero modes will remain unchanged. The fact that the preservation of T symmetry implies a charge of $\pm1/2$ for the monopole can also be inuited from fact that T symmetry can be preserved in this model by adding "$\frac{1}{8\pi}\int AdA$" to the Lagrangian.

◇ ◇ ◇

3 $SO(3)$ monopoles and zero modes

Since $\pi_1(SO(N)) = \mathbb{Z}_2$, $SO(N)$ gauge theories have $\mathbb{Z}_2$ monopoles. In this section we will consider an $SO(3)$ gauge theory coupled to a Dirac fermion in the spin 1 representation on various spatial manifolds of different topologies (S^2, T^2, etc). We will discuss whether or not the Hamiltonian generically has zero modes when a nontrivial $\mathbb{Z}_2$ flux configuration is turned on. The motivation for thinking about this is to compare with the case where the monopoles are characterized by an integer topological invariant, like in the case with a $U(1)$

gauge field. In this case the larger the flux the larger the average field strength passing through the spatial manifold, and since the solutions of the Dirac equation depend on the total field strength of the background fields (both geometric and gauge), the existence and number of zero modes depends on the flux. It is not a priori obvious that a similar story should hold when the flux is discrete however, and we will try to argue that for generic flux configurations there are no zero modes.

◇ ◇ ◇

We can basically answer this question by using what we know about the index of $\not{D}_A$ to discuss the expected number of zero modes. In two dimensions (for us these are spatial dimensions since we are interested in the zero modes of the Hamiltonian, which for Dirac fermions is just the spatial part of the spacetime Dirac operator), the Dirac operator $\not{D}$: $\Gamma(S^{\pm} \otimes E \;\to) \to \Gamma(S^{\mp} \otimes E)$ has a pseudoreal structure if the associated gauge bundle $E = P_G \times_\rho \mathbb{C}^{\dim \rho}$ is such that the representation ρ is either real or pseudoreal (here P_G is a principal G-bundle, with G the gauge group).

The reason for this is as follows: if ρ is real, then we can choose a connection such that $\not{A}^a T^a \mathcal{K} = -\mathcal{K} \not{A}^a T^a$, where A is the gauge connection (we are working in physicist conventions where the T^a are Hermitian). Then choosing the γ matrices to be X, Y, we see that the operator $\mathcal{J} = \mathcal{K}(J \otimes \mathbf{1})$ (notation is spin $\otimes$ gauge) is such that

$$i\not{D}_A \mathcal{J} = \mathcal{J}(i\not{D}_A), \qquad \mathcal{J}^2 = -\mathbf{1}. \tag{58}$$

If ρ is pseudoreal, then we can find a unitary U_G such that

$$\not{A}^a T^a \mathcal{K} U_G = -\mathcal{K} U_G \not{A}^a T^a, \qquad (\mathcal{K} U_G)^2 = -\mathbf{1}. \tag{59}$$

Then we can choose the γ matrices to be the real matrices X, Z, which tells us that the operator $\mathcal{J}' = \mathcal{K}(\mathbf{1} \otimes U_G)$ satisfies

$$i\not{D}_A \mathcal{J}' = -\mathcal{J}'(i\not{D}_A), \qquad (\mathcal{J}')^2 = -\mathbf{1}. \tag{60}$$

Therefore if ρ is not complex, we can find an antilinear operator $\mathcal{J}$ that squares to $-\mathbf{1}$ and either commutes or anticommutes with $i\not{D}_A$. This means that if $\not{D}_A \psi = 0$, then $\mathcal{J}\psi$ is also a zero mode. Since $\mathcal{J}\psi$ has opposite chirality to ψ,[5] the index of $\not{D}_A$ must vanish. Therefore if the fermions transform in a representation of the gauge group that is not complex, ind $\not{D}_A = 0$, and there is nothing that protects zero modes, if they do exist, from being lifted.[6] Therefore in a generic situation, we expect no zero modes. The remainder of the diary entry is just an attempt to confirm this and to make sure we aren't missing any other symmetry that might protect the zero modes from being lifted.

[5] In the case where ρ is real, this is clear since the J tensor factor in $\mathcal{J}$ is off-diagonal, and in this basis $\bar{\gamma} = Z$. For the pseudoreal case, the choice of γ matrices means $\bar{\gamma} = Y$, so that eigenspinors of $\pm$ chirality look like $(1, \pm i)^T$. The complex conjugation in $\mathcal{J}$ exchanges these, and hence $\mathcal{J}$ anticommutes with $\bar{\gamma}$.

[6] Of course, another way to derive this would just have been to say that since in two dimensions the only gauge-invariant 2-forms for the gauge curvature that are non-vanishing are those from $U(1)$ groups, so that gauge groups like $SO(N)$ can make no contribute to ind $\not{D}_A$, by the index theorem. However the argument given in the main text doesn't depend on taking the index theorem for granted, which I think is nice.

We should also point out that this conclusion is rather special to two dimensions (or to be pedantic, for dimensions d where $d = 2 \mod 8$). This is because the γ matrices in two dimensions admit both a real and a pseudoreal structure, which meant that we could get a pseudoreal structure for the full connection on $S \otimes E$ with either a real or pseudoreal gauge connection. We also used the fact that the pseudoreal structure $\mathcal{J}$ anticommutes with $\bar{\gamma}$; in four dimensions this is not true (more on this later).

On the plane / torus

We will now specialize to $SO(3)$ gauge theory with 1 Dirac fermion in the spin 1 representation. On the plane / torus, we will work in Landau gauge, where $A_x = 0$ and A_y is a function of x only. Then $\not{D}_A\psi = 0$ reads

$$(\partial_x - k_y + A_y)\psi_L = 0, \qquad (\partial_x + k_y - A_y)\psi_R = 0, \tag{61}$$

where k_y is the y component of the momentum. For a uniform $U(1)$ flux, we would take $A_y = Bx$, where B is the flux density. For $SO(3)$, we let $A_y = BxT^3$.[7] After diagonalizing T^3 from $Y \oplus 0$ to $T^3 = Z \oplus 0$ and writing the spinors in flavor space as (f_1, f_2, f_3) with each f_i a two-component spinor, this gives a left-handed zero mode $(f_L, 0, 0)$ and a right-handed zero mode $(0, f_R, 0)$ (again handedness here means with respect to space; since we are in 2+1D there is no notion of chirality in spacetime).

We would like to know whether these zero modes still exist after we perturb with some field strength that does not point uniformly in one direction in flavor space.[8] To do this, consider as an example adding the connection $\widetilde{A} = \widetilde{A}_y dy$, where

$$\widetilde{A}_y = \widetilde{A}_y^2 T^2 = \epsilon \cos(x) e^{-x^2/2} T^2. \tag{62}$$

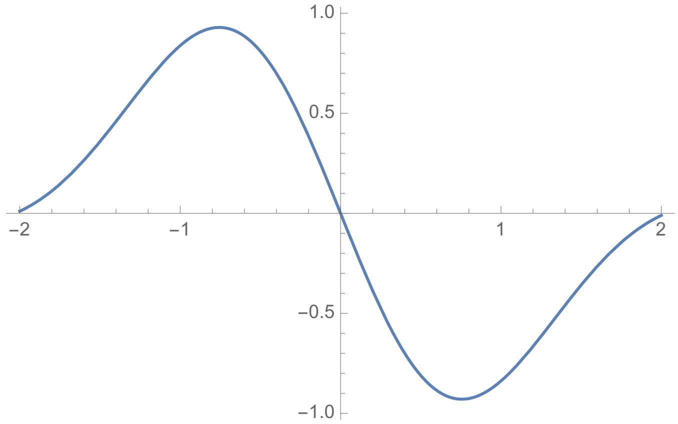
This has a field strength which has zero integral over the plane[9], and so it is topologically

[7] For other gauge groups G with $\pi_1(G) \neq 0$, we can just take $A_y = A_y^a T^a = BxT^a$ for some particular generator T^a.

[8] This wording isn't very precise, since the fact that F transforms adjointly under gauge transformations means that we can perform a gauge transformation to take our uniform flux field to one in which $F_{\theta\phi}(\theta, \phi)$ has constant $\mathrm{Tr}[F \wedge \star F]$, but which has a direction in flavor space that is an arbitrary function of θ, ϕ. So we are really interested in making a perturbation that changes $\mathrm{Tr}[F \wedge \star F]$ to be something non-uniform.

[9] Or torus. If we are on the torus, we take it to be big enough that the usual $e^{-(x-k_y/B)^2 B/2}$ zero mode wavefunctions have support only within width δx which is small compared to the size of the torus, so that the fact that the above wavefunctions are technically speaking not smooth over the torus doesn't really matter.

trivial. The field strength as a function of x looks like



(63)

After diagonalizing T^3, the matrix T^2 in flavor space becomes

$$T^2 = \begin{pmatrix} & i \\ -i & \end{pmatrix} \to \frac{1}{\sqrt{2}} \begin{pmatrix} & & i \\ & & i \\ -i & -i & \end{pmatrix}. \tag{64}$$

Now since the fermions are not in a complex representation, we have ind $\not{D}_A = 0$ and we know that there will always be as many left zero modes as right zero modes.[10] Therefore to see whether adding the $\widetilde{A}_y$ term to the connection does anything to $\ker i\not{D}_A$, we can focus wolog on a certain chirality, which we will take to be L for definiteness. Therefore we are interested in whether we can find normalizable solutions to the following equations (setting $k_y = 0$ for simplicity)

$$\begin{aligned} (\partial_x + Bx) f_1 + i\frac{\widetilde{A}_y}{\sqrt{2}} f_3 &= 0 \\ (\partial_x - Bx) f_2 + i\frac{\widetilde{A}_y}{\sqrt{2}} f_3 &= 0 \\ \partial_x f_3 - i\frac{\widetilde{A}_y}{\sqrt{2}}(f_1 + f_2) &= 0. \end{aligned} \tag{66}$$

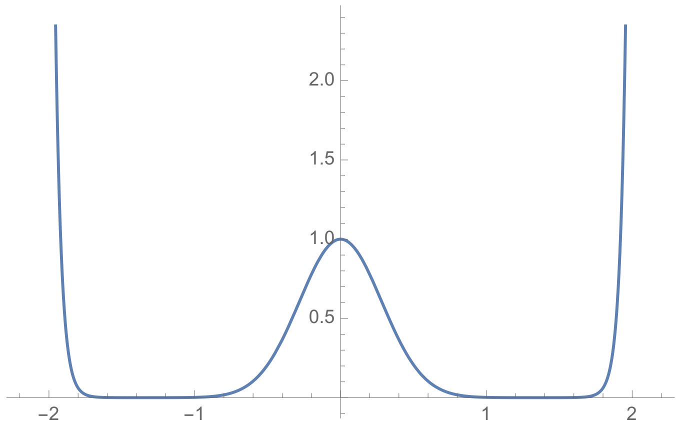
If $\widetilde{A}_y = 0$ then we just take $f_2 = f_3 = 0$, and let f_1 be the usual harmonic oscillator solution. However, if $\widetilde{A}_y \neq 0$, this is not possible: the last equation means that either f_2 or f_3 must be nonzero if f_1 is nonzero, and the second equation then ensures that in fact $f_2 \neq 0$. Now f_2 is the mode that doesn't have a normalizable solution when $\widetilde{A}_y = 0$, and so we might expect that the $\widetilde{A}_y$ coupling ruins the normalizability of the solution. Indeed, this is what

[10] An operator that provides the (pseudo)real structure here is $\mathcal{K}(Y \otimes \mathbf{1})$, where the first tensor factor is for the spin indices and the second is for the gauge indices. Indeed, as in the previous section, using X and Y as the γ matrices, and working in a basis where the gauge generator matrices are purely imaginary and antisymmetric, we have

$$[i\not{D}_A, \mathcal{K}(Y \otimes \mathbf{1})] = 0, \tag{65}$$

and so $\mathcal{K}(Y \otimes \mathbf{1})$ provides a way to take a zero mode of a certain chirality and construct another zero mode with opposite chirality.

appears to happen: using the form of $\widetilde{A}_y$ above with $B = 2\pi$ and $\epsilon = \sqrt{2}/10$, a plot of the magnitude $\sum_i f_i^* f_i$ as a function of x shows a divergence:



(67)

There of course may be something I've missed, or some tricky choice of initial conditions (the above plot was for $f_1(0) = 1$, $f_2(0) = f_3(0) = 0$; modifying the latter two to be nonzero makes the divergence worse) that allow this divergence to be avoided, but for now it seems to be a generic consequence of taking $\widetilde{A} \neq 0$.

On the sphere

We first need to choose a gauge connection. For a $U(1)$ monopole of flux n, the standard choice is

$$A^{N/S} = n\frac{\pm 1 - \cos\theta}{2}d\phi, \tag{68}$$

which gives $\int_{S^2} F = 2\pi n$. For a gauge group with $\pi_1(G) \neq 0$, the simplest choice for a monopole field is the above but with a T^a tacked on, where T^a is a particular (Hermitian) generator of $\mathfrak{g}$. Using results from our earlier entry on zero modes, the covariant derivatives are

$$\nabla_\theta = \partial_\theta, \qquad \nabla_\phi = \partial_\phi - \frac{iZ\otimes\mathbf{1}}{2}\cos\theta + in\frac{\pm 1 - \cos\theta}{2}\mathbf{1}\otimes T^a, \tag{69}$$

where the first tensor factor is the spin indices and the second is the gauge indices (we won't bother to explicitly write the $\otimes$ in what follows). The expression $i\not{D}_A\psi = 0$, is then, for our uniform monopole field,

$$\not{D}_A\psi^{(N/S)} = \left[X\left(\partial_\theta + \frac{\cot\theta}{2}\right) + Y\csc\theta\left(\partial_\phi + in\left(\frac{\pm 1 - \cos\theta}{2}\right)T^a\right)\right]\psi^{(N/S)} = 0 \tag{70}$$

or written out in chiral components,

$$\begin{aligned}\left(\partial_\theta + \frac{\cot\theta}{2} - i\csc\theta\partial_\phi + n\csc\theta\left(\frac{\pm 1 - \cos\theta}{2}\right)T^a\right)\psi_R^{(N/S)} &= 0\\ \left(\partial_\theta + \frac{\cot\theta}{2} + i\csc\theta\partial_\phi - n\csc\theta\left(\frac{\pm 1 - \cos\theta}{2}\right)T^a\right)\psi_L^{(N/S)} &= 0\end{aligned} \tag{71}$$

In what follows, we will take $n = 1$ for concreteness. Then in the $U(1)$ case, we see that we get a single R zero mode, $\psi_R = e^{-i\phi/2}$ (the reason we get an R zero mode and not an L one is because of our sign conventions for the covariant derivative). Also note that this zero mode actually has spin zero; to see this one needs to properly calculate the angular momentum generators, which I won't go into here.

Now for $SO(3)$. If we let the field strength point in the T^3 direction, we see that we get a single L and a single R zero mode, as expected. Do these zero modes survive when a perturbation is added? Let us add the potential

$$\widetilde{A} = \widetilde{A}^2_\phi T^2\, d\phi = \epsilon \sin(2\theta) T^2\, d\phi. \tag{72}$$

This is well-defined on the sphere since $\widetilde{A}(\theta = 0, \pi) = 0$, and it is topologically trivial since $\int_{S^2} d\widetilde{A} = 0$. As mentioned before, since the zero modes for real gauge groups always come in left-right pairs, we can focus on a single handedness (we will look at R) wolog. We therefore want to find normalizable solutions to (working on the N coordinate patch)

$$\begin{aligned} \left(\partial_\theta - i\csc\theta\partial_\phi + \frac{1}{2}\csc\theta\right) f_1 + i\epsilon\frac{\csc(\theta)\sin(2\theta)}{\sqrt{2}} f_3 = 0 \\ \left(\partial_\theta + \cot\theta - i\csc\theta\partial_\phi - \frac{1}{2}\csc\theta\right) f_2 + i\epsilon\frac{\csc(\theta)\sin(2\theta)}{\sqrt{2}} f_3 = 0 \\ \left(\partial_\theta + \frac{1}{2}\cot\theta - i\csc\theta\partial_\phi\right) f_3 - i\epsilon\frac{\csc(\theta)\sin(2\theta)}{\sqrt{2}}(f_1 + f_2) = 0 \end{aligned} \tag{73}$$

When $\epsilon = 0$ we just take $f_1 = e^{-i\phi/2}, f_2 = f_3 = 0$.[11] When $\epsilon \neq 0$ the coupling between the different modes kicks in, and as in the planar case we seem to run into normalizability problems caused by the troublesome modes f_2, f_3 being forced to be nonzero. For example, set $\epsilon = \sqrt{2}$. The natural choices for the ϕ dependence of the three modes is $f_1 \propto e^{-i\phi/2}$, $f_2 \propto e^{i\phi/2}$, and with f_3 having no ϕ dependence. With these assignments of ϕ dependence, the

[11] Here the f_3 mode has no normalizable solution when $\epsilon = 0$, since taking the ϕ dependence to be trivial means

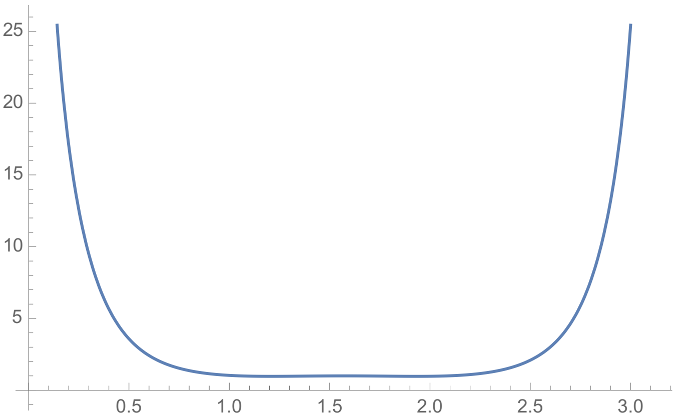
$$f_3 \propto \frac{1}{\sqrt{\sin\theta}}, \tag{74}$$

which is not acceptable (it integrates to something finite because of the $\sin\theta$ in the measure, but it is not differentible). Similarly the f_2 mode has to be zero, since otherwise we have

$$f_2 \propto e^{i\phi/2}\csc(\theta), \tag{75}$$

which is also no good.

volume-element-normalized magnitude $\sum_i f_i^* f_i \sin\theta$ as a function of θ looks like



(76)

So, it blows up at the poles, and we don't get a legit zero mode solution. This seems to be the generic behavior for any $\epsilon \neq 0$.

◇ ◇ ◇

4 Number of zero modes allowed by magnetic fluxes in two dimensions

Consider massless Dirac fermions on a spatial $\mathbb{R}^2$ (2+1 dimensions) with a background field A. The Hamiltonian is

$$H = \gamma^0 \gamma^j (i\partial_j + A_j), \tag{77}$$

where we will take $\gamma^0 = Z, \gamma^x = X, \gamma^y = Y$. Our goal today will be to show that the if the total magnetic flux though the plane is $\int F = \Phi > 0$, then H generically supports

$$\left\lceil \frac{\Phi}{2\pi} \right\rceil - 1 \tag{78}$$

zero modes. If the magnetic flux is oriented positively (negatively) with respect to the plane, we will see that all the zero modes are generically right (left) handed.

◇ ◇ ◇

The proof is actually pretty quick once one figures out the trick. As is common with these types of questions, the idea is to find the functional form of the wavefunctions in the kernel of H, and then to see what subspace of the kernel is picked out by requiring normalizability.

Letting the candidate zero mode be $\psi = (\psi_L, \psi_R)^T$, $H\psi = 0$ reads

$$(\partial - iA_-)\psi_L = 0, \qquad (\bar\partial - iA_+)\psi_R = 0, \tag{79}$$

with $\partial = \partial_z = \frac{1}{2}(\partial_x - i\partial_y)$, $\bar\partial = \partial_{\bar z} = \frac{1}{2}(\partial_x - i\partial_y)$, and $A_\pm = \frac{1}{2}(A_x \pm iA_y)$. These parallel transport equations can be solved in the usual way by attaching Wilson lines to the fermions, but this gives us yucky non-local expressions that are not well-suited for our present needs. Better is to choose the gauge $d^\dagger A = 0$, allowing us to write $A = d^\dagger \phi$ since we're on $\mathbb{R}^2$ (this is really Coulomb gauge $\partial_j A^j = 0$ not $\partial_\mu A^\mu = 0$ since we are only working in space. From now on, all d's and $\star$'s and so on will take place in $\mathbb{R}^2$). Since we don't want to work with 2-forms, we will actually write $A = -\star d\lambda$, where $\lambda = -\star\phi$ is a zero-form (the minus sign is just for convenience). This means that

$$A_- = -\partial_y \lambda - i\partial_x \lambda, \qquad A_+ = -\partial_y \lambda + i\partial_x \lambda, \tag{80}$$

so that we need to solve

$$\left[\partial_x - \sigma(\partial_x \lambda) - i\partial_y + i\sigma(\partial_y \lambda)\right]\psi_\sigma = 0, \tag{81}$$

where $\sigma = +1$ for the L fermions and $\sigma = -1$ for the R fermions. So evidently we need

$$\psi_L = f(\bar{z})e^{\lambda}, \qquad \psi_R = f(z)e^{-\lambda}, \tag{82}$$

where $f(\bar{z})$ and $f(z)$ are polynomial functions in $\bar{z}$ and z, respectively. We've included these functions since they get killed by ∂ and $\bar{\partial}$ in turn and so don't affect the equations for the ψ_σ's (they need to be polynomials since we don't want any singularities and since we need the fields to be single-valued).

Now we need to make sure that these candidate solutions are normalizeable and have appropriate fall-off behavior at infinity. To examine this, we need to know what happens to λ at $|z| = \infty$. Now $dA = F$ reads $-d\star d\lambda = F$, which is $\partial_j \partial^j \lambda = F$ (hence the weird sign choice). Thus

$$\int d^2r' G(r, r')F(r') = \lambda(r), \tag{83}$$

where

$$G(r, r') = -\frac{1}{2\pi}\ln\left(\frac{|r - r'|}{a}\right) \tag{84}$$

is the propagator for free scalars in two dimensions (a is some short-distance cutoff). Let us now suppose that $F(r') \to 0$ as $r' \to \infty$, and take $r \to \infty$ in the above expression. Then we write

$$G(r, r') = -\frac{1}{2\pi}\left(\ln(r/a) + \frac{1}{2}\ln(1 + r'^2/r^2 - 2(r\cdot r')/r^2)\right). \tag{85}$$

The second term can be dropped in the integral: it only contributes at very large r', and we have assumed that $F(r')$ dies off at large r'. Thus as $r \to \infty$,

$$\lambda(r) \to -\frac{1}{2\pi}\int d^2r' \ln(r/a)F(r') = \frac{\Phi}{2\pi}\ln(a/r), \tag{86}$$

where $\Phi = \int F$ is the total flux through the plane. This means that our candidate fermion fields behave at $|z| \to \infty$ like

$$\psi_L(z, \bar{z}) = f(\bar{z})\left(\frac{a}{|z|}\right)^{\Phi/2\pi}, \qquad \psi_R(z, \bar{z}) = f(z)\left(\frac{a}{|z|}\right)^{-\Phi/2\pi}. \tag{87}$$

Thus for the fields to be finite at infinity, we need $\Phi > 0$ for the L fermions and $\Phi < 0$ for the R fermions: the handedness of the magnetic flux means that only fermions whose handedness matches that of the flux can be zero modes (or whose handedness is opposite to that of the flux; maybe our labelling of L/R was unfortunate and should be swapped).

Let us take $\Phi > 0$ wolog, and suppose that $f(\bar{z})$ is a polynomial of degree $l > 0$ in $\bar{z}$. Then in order for $\int |\psi_L|^2$ to be finite, we count powers of $|z|$ and see that the degree of the polynomial has to satisfy

$$l < \frac{\Phi}{2\pi} - 1. \tag{88}$$

Since a degree l polynomial gives us $l+1$ linearly independent functions to choose from, the number of degenerate zero modes is evidently

$$N = \left\lceil \frac{\Phi}{2\pi} \right\rceil - 1, \tag{89}$$

as promised. If the flux Φ is negative, we get the same number of zero modes, and in that case

$$N = \left\lfloor \frac{\Phi}{2\pi} \right\rfloor + 1. \tag{90}$$

◇ ◇ ◇