Fermions And Topological Phases, II

Edward Witten

PSCCMP/PiTP Lectures

Last time we talked about Weyl fermions in condensed matter and the boundary-localized modes that they produce. It is also possible to get boundary-localized modes from Dirac fermions, and since this is important in understanding topological insulators, I will say a little about it.

In the absence of discrete symmetries, and without tuning any parameters, it is not natural in condensed matter physics to get a massless Dirac (as opposed to Weyl) fermion. That is simply because for a four-component fermion field ψ with both chiralities, mass terms are possible, in fact there are two such terms:

$$\left(i \sum_\mu \gamma^\mu \partial_\mu - m - m' \gamma_5 \right) \psi = 0.$$

To get a massless Dirac fermion, we need a reason for $m = m' = 0$. It turns out that it is actually possible to achieve this in the context of a crystal with suitable discrete symmetries, but unfortunately there will not be time to explain this.

The only symmetry that we will assume is time-reversal symmetry, and this is enough to set one of the two parameters to 0 but not both. The Dirac equation becomes

$$\left(i \sum_\mu \gamma^\mu \partial_\mu - m \right) \psi = 0.$$

Generically m is not 0 but of course if we adjust one parameter (for example the chemical composition of an alloy) we can hope to pass through a point with $m = 0$. It turns out that this is the phase transition between an ordinary insulator and a topological one. For now, we consider a sample with boundary and look for a mode localized near the boundary.

We take the boundary at $x_1 = 0$ and the natural Dirac boundary condition $i\gamma_1\psi| = \psi|$. (The sign does not really matter and there is no reason for the boundary condition to jump when m passes through 0.) Let $x_\parallel$ be the coordinates along the boundary and

$$i\gamma \cdot \partial^\parallel = i\sum_{\mu \neq 1} \gamma^\mu \partial_\mu$$

the $2+1$-dimensional Dirac operator along the boundary. We can obey the $3+1$-dimensional Dirac equation with

$$\psi = \exp(mx_1)\psi_\parallel(x_\parallel)$$

where

$$i\gamma_1\psi_\parallel = \psi_\parallel, \quad \gamma \cdot \partial^\parallel \psi_\parallel(x_\parallel) = 0.$$

For $m > 0$, this solution is highly unnormalizable. But for $m < 0$, it is plane-wave normalizable and localized along the boundary. We have found one of the main characteristics of the topological insulator phase: the boundary supports a gapless fermion mode.

To be more precise, since $\psi_\parallel$ was constrained to obey the massless $2+1$ dimensional Dirac equation

$$\gamma \cdot \partial^\parallel \psi_\parallel(x_\parallel) = 0,$$

we get a $2+1$ dimensional massless Dirac fermion. (Because of the additional constraint $i\gamma_1\psi_\parallel = \psi_\parallel$ which removes half the modes, the original 4-component fermion in $3+1$ dimensions gives a single 2-component massless Dirac fermion in $2+1$ dimensions.)

Generically in the context of condensed matter physics the fermi energy ε_F does not pass through the Dirac point in the boundary theory
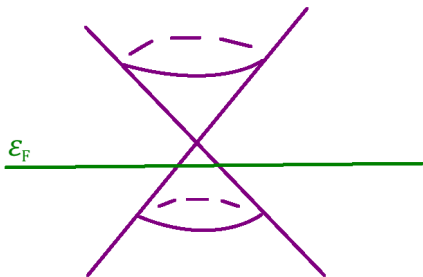
so the boundary of a topological insulator is more like an ordinary metal than some of the Weyl semimetals that we talked about yesterday.

To summarize the last point, we have learned that in a generic 1-parameter family of T-conserving band theories, there can be a phase transition between a phase with no localized boundary modes and a phase with a massless 2d fermion on the boundary. At the phase transition point, there is a massless 3d Dirac fermion.

As I have already remarked, the other way to get massless 3d Dirac fermions in condensed matter is to take advantage of certain discrete crystal symmetries. This has actually been part of the background for some of the experimental talks that we have heard.

However, I want to go in a different direction involving an introduction to some aspects of the integer quantum Hall effect. First I just want to explain from the point of view of effective field theory why there is an integer quantum Hall effect in the first place. We consider a material that not only is an insulator, but more than that has no relevant degrees of freedom – not even topological ones – in the sense that its interaction with an electromagnetic field can be described by an effective action for the $U(1)$ gauge field A only, without any additional degrees of freedom. (This would certainly not be true in a conductor, whose interaction with an electromagnetic field cannot be described without including the charge carriers in the description, along with A. But more subtly, as we will discuss, it is not true in a fractional quantum Hall system, whose effective field theory requires topological degrees of freedom coupled to A.)

In a $3+1$-dimensional material with no relevant degrees of freedom, the effective action for the electromagnetic field can have all sorts of terms associated to various familiar effects. For example, ferromagnetism and ferroelectricity correspond to terms in the effective action that are linear in $\vec{E}$ or $\vec{B}$

$$I' = \int_{W_3 \times \mathbb{R}} \left(\vec{a} \cdot \vec{E} + \vec{b} \cdot \vec{B} \right).$$

(Here W_3 is the spatial volume of the material and $\mathbb{R}$ parametrizes the time, so the "world-volume" of the material is $M_4 = W_3 \times \mathbb{R}$.) Similarly, electric and magnetic susceptibilities correspond to terms bilinear in $\vec{E}$ or $\vec{B}$:

$$I'' = \int_{W_3 \times \mathbb{R}} \left(\alpha_{ij} E_i E_j + \beta_{ij} B_i B_j \right).$$

And so on.

All these terms are manifestly gauge-invariant in the sense that they are integrals of gauge-invariant functions – the integrands are constructed only from $\vec{E}$ and $\vec{B}$ (and possibly their derivatives). In $2 + 1$ dimensions, there is a unique term that is gauge-invariant but does *not* have this property. This is the Chern-Simons coupling

$$I_{\mathrm{CS}} = \frac{1}{4\pi} \int_{\mathcal{M}_2 \times \mathbb{R}} \mathrm{d}^3 x \epsilon^{ijk} A_i \partial_j A_k.$$

The density $\epsilon^{ijk} A_i \partial_j A_k$ that is being integrated is definitely not gauge-invariant, but the integral is gauge-invariant up to a total derivative. In fact, under

$$A_i \to A_i + \partial_i \phi,$$

we have

$$\epsilon^{ijk} A_i \partial_j A_k \to \epsilon^{ijk} A_i \partial_j A_k + \partial_i \left(\epsilon^{ijk} \phi \partial_j A_k \right).$$

Roughly speaking, this shows that I_{CS} is gauge-invariant, but we have to be more careful because electric charge quantization, with a field of charge 1 transforming as

$$\psi \to e^{i\phi}\psi$$

means that we should consider ϕ to be defined only modulo 2π:

$$\phi \cong \phi + 2\pi.$$

Given this fact, the previous proof of gauge-invariance of I_{CS} is not quite correct and we will be more careful in a moment.

Before I go on, though, I want to point out that logically, one could consider a theory in which one is only allowed to make a gauge transformation $A_i \rightarrow A_i + \partial_i \phi$ with a single-valued ϕ. But that theory is not the real world. Dirac showed that the Schrodinger equation of electrons, protons, and neutrons can be consistently coupled with magnetic monopoles, and that this consistency is only possible because the Schrodinger equation is invariant under gauge transformations in which $e^{i\phi}$ is single-valued although ϕ is not. This is needed to make the Dirac string unobservable:

Anyway our microscopic knowledge that the Schrodinger equation is invariant under any gauge transformation such that $e^{i\phi}$ is single-valued (even if ϕ is not single-valued) implies constraints on the effective action that we would not have without that knowledge. We want to understand those constraints.

To do this, we will consider the following situation: we take our two-dimensional material to be a closed two-manifold, for instance S^2, and we will take "time" to be a circle S^1 of circumference β. (For example, we might be computing $\mathrm{Tr}\, e^{-\beta H}$.) Thus we consider a material whose "worldvolume" is $M_3 = S^2 \times S^1$:
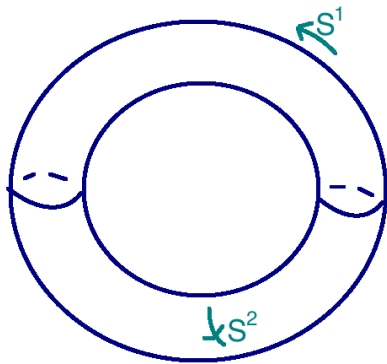
One might not be able to engineer this situation in the real world but it is clear that the Schdrodinger equation makes sense in this situation. So we can consider it in deducing constraints on the effective action that can arise from the Schrodinger equation.

The gauge field that we want to consider on $M_3 = S^2 \times S^1$ is characterized by the following: We place a unit of Dirac magnetic flux on S^2

$$\int_{S^2} \mathrm{d}x_1 \mathrm{d}x_2 \; \frac{F}{2\pi} = 1.$$

(This is the right quantum of flux if the covariant derivative of the electron is $D_i\psi = (\partial_i - iA_i)\psi$, meaning that I am writing A for what is often called eA. This lets us avoid factors of e in many formulas.) And we take a constant gauge field in the time direction:

$$A_0 = \frac{s}{\beta}$$

with constant s. (Remember the time direction is a circle of circumference β.) For this gauge field, one can calculate

$$I_{\mathrm{CS}} = \frac{1}{4\pi} \int_{M_3 = S^2 \times S^1} \mathrm{d}^3 x \, \epsilon^{ijk} A_i \partial_j A_k = s.$$

(This is actually a slightly tricky calculation.)

Note that the holonomy of A around the "time" circle is

$$\exp\left(i\int_0^\beta A_0 \mathrm{d}t\right) = \exp\left(i\int_0^\beta (s/\beta)\mathrm{d}t\right) = \exp(is).$$

The gauge transformation

$$\phi = \frac{2\pi t}{\beta},$$

which was chosen to make $e^{i\phi}$ periodic, acts by

$$s \to s + 2\pi$$

and so leaves the holonomy invariant. (This must be true, because with my normalization of A, this holonomy is the phase factor when an electron is parallel-transported around the circle and so is physically meaningful.)

So we have in this example $I_{\mathrm{CS}} = s$, and a gauge transformation can act by $s \to s + 2\pi$. So I_{CS} is not quite gauge-invariant. Here we must remember what is essentially the same fact that was exploited by Dirac in his theory of the magnetic monopole. The classical action I enters quantum mechanics only via a factor $\exp(iI)$ in the Feynman path integral (or $\exp(iI/\hbar)$ if one restores $\hbar$), so it is enough if I is well-defined and gauge-invariant mod $2\pi\mathbb{Z}$. Since I_{CS} is actually gauge-invariant mod $2\pi\mathbb{Z}$ (we showed this in an example but it is actually true in general), it can appear in the effective action with an integer coefficient:

$$I_{\mathrm{eff}} = kI_{\mathrm{CS}} + \dots.$$

The point of this explanation has been to explain why k has to be an integer – sometimes called the "level." The fact that k is an integer gives a macroscopic explanation of the quantization of the Hall current. Indeed for any material whose interaction with an electromagnetic potential A is governed by an effective action I_{eff}, the induced current in the material is

$$J_i = -\frac{\delta I_{\mathrm{eff}}}{\delta A_i}.$$

We are interested in the case that

$$I_{\mathrm{eff}} = k I_{\mathrm{CS}} = \frac{k}{4\pi} \int_{M_3} \mathrm{d}^3 x\, \epsilon^{ijk} A_i \partial_j A_k.$$

Let us consider a material sitting at rest at $x_3 = 0$ and thus parametrized by x_1, x_2.
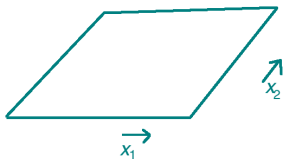
The current in the x_2 direction is

$$J_2 = -\frac{\delta I_{\text{eff}}}{\delta A_1} = \frac{k F_{01}}{2\pi} = \frac{k E_1}{2\pi}.$$

This is called a Hall current: an electric field in the x_1 direction has produced a current in the x_2 direction. The Hall current has a quantized coefficient $k/2\pi$ (usually called $ke^2/2\pi$, recall that my A is usually called eA; also I set $\hbar = 1$ so $h = 2\pi$), where the quantization follows from the fact that I_{CS} is not quite gauge-invariant.

One may wonder "How then can one have a fractional quantum Hall effect?" I will give a short answer for now, postponing for later a deeper study of the fractional quantum Hall effect. One cannot get an integer quantum Hall effect in a description in which A is the only relevant degree of freedom. However, from a macroscopic point of view, this can happen in a material that generates an additional "emergent" $U(1)$ gauge field a that only propagates in the material. We will write $f_{ij} = \partial_i a_j - \partial_j a_i$ for the field strength of a. An example (Wen, Wilczek, Zee, ...) of a gauge-invariant effective action that leads to a fractional quantum Hall effect is

$$I_{\text{eff}} = \frac{1}{2\pi} \int_{M_3} \mathrm{d}^3 x \; \epsilon^{ijk} A_i \partial_j a_k - \frac{r}{4\pi} \int_{M_3} \mathrm{d}^3 x \; \epsilon^{ijk} a_i \partial_j a_k.$$

An oversimplified explanation of why this gives a fractional quantum Hall effect is the following. One argues that as a appears only quadratically in the effective action, one can integrate it out using its equation of motion. This equation is

$$f = \frac{1}{r}F,$$

implying that up to a gauge transformation $a = A/r$. Substituting this in I_{eff}, we get an effective action for A only that describes a fractional quantum Hall effect:

$$I'_{\text{eff}} = \frac{1}{r}I_{\text{CS}}(A) = \frac{1}{r}\frac{1}{4\pi}\int_{M_3} \mathrm{d}^3x \; \epsilon^{ijk}A_i\partial_j A_k.$$

Here $1/r$ appears where k usually does, and this suggests that the Hall conductivity in this model is $1/r$. That is correct. But there clearly is something wrong with the derivation because the claimed answer for the effective action $I'_{\text{eff}} = (1/r)I_{\text{CS}}(A)$ does not make sense as it violates gauge invariance. The mistake is that in general, as F may have a flux quantum of 2π, and f has the same allowed flux quantum (otherwise the action we assumed would not be gauge-invariant), for a given A it is not possible to solve the equation

$$f = \frac{1}{r}F$$

for a. Thus, it is not possible to eliminate a from this system and give a description in terms of A only. The reason that "integrating out a" gives the right answer for the Hall current is that this procedure is valid locally and this is enough to determine the Hall current. The system has more subtle properties (fractionally charged quasiparticles and topological degeneracies) that can only be properly understood in the description with a as well as A.

Going back to a theory that can be described in terms of A only, we have then an integer k in the macroscopic description. But there is also an integer in the microscopic description of a band insulator (Thouless-Kohmoto-Nightingale-den Nijs 1982). It arises as follows. We consider a crystal with N bands, of which n are filled. We assume the system is completely gapped.
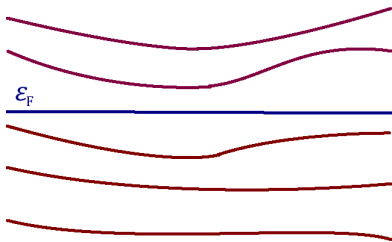

say with n filled bands fand $N - n$ empty bands. As we learned yesterday, in a 2d system it is generic to have no band crossings.

We are in the same situation as in our discussion yesterday of Weyl semimetals, except that there are no bad points, so we work over the whole Brillouin zone B, without removing anything. As we are in two-dimensions, B is a two-torus. At momentum p, let $\mathcal{H}_p \cong \mathbb{C}^N$ be the full space of all states, and $\mathcal{H}'_p$ the subspace of filled levels. We can regard $\mathcal{H}_p$ as a rank (or dimension) N "trivial bundle" over B and $\mathcal{H}'_p$ as a "subbundle" of rank n. The integer we want, which we will call k', is the first Chern class $c_1(\mathcal{H}'_p)$, integrated over B. In terms of the Berry connection $\mathcal{A}$ on the filled bands that we discussed yesterday, whose curvature we call $\mathcal{F}$, this integer is

$$k' = c_1(\mathcal{H}'_p) = \int_B \frac{\operatorname{Tr} \mathcal{F}}{2\pi}.$$

The basic claim of TKNN is that k', the flux of the Berry connection, is the same as k, the coefficient of the quantum Hall current. The original proof was based on literally just calculating the current from first principles in terms of a matrix element in the fermion ground state – which is written as an integral of single particle matrix elements over the Brillouin zone. I want to explain a different viewpoint that will emphasize that k' is not just a band concept but can be defined in the full many-body theory. (I hope to describe tomorrow another approach essentially due to Haldane to the relation $k = k'$.)

We consider a finite sample, say on an $L_1 \times L_2$ lattice
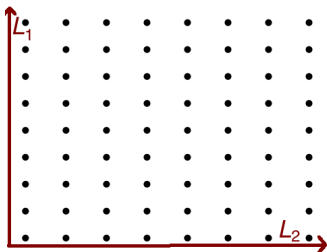
for very large L_1, L_2, where I will take lattice constants a_1, a_2 in the two directions. We assume periodic boundary conditions, maintaining the lattice translation symmetries. However, for the finite system, the momenta take discrete values

$$p_1 = \frac{2\pi s_1}{a_1 L_1}, \quad 0 \leq s_1 \leq L_1 - 1$$

$$p_2 = \frac{2\pi s_2}{a_2 L_2}, \quad 0 \leq s_2 \leq L_2 - 1.$$

The ground state of the finite system is of course obtained by filling all of the states in the first n bands with these values of the momenta.

Now, however, we turn on a background electromagnetic vector potential that is chosen such that the magnetic field vanishes, but an electron going all the way around the x_1 direction or the x_2 direction picks up a phase:

$$A_1 = \frac{\alpha_1}{a_1 L_1}, \quad A_2 = \frac{\alpha_2}{a_2 L_2}.$$

The phase picked up by an electron going around the x_1 (or x_2) direction is $\exp(i\alpha_1)$ (or $\exp(i\alpha_2)$) and up to a gauge transformation the range of these parameters is

$$0 \leq \alpha_1, \ \alpha_2 \leq 2\pi.$$

From the point of view of band theory, the effect of turning on the parameters α_1, α_2 is just to shift the momenta of the electrons, which become

$$p_1 = \frac{2\pi s_1 + \alpha_1}{a_1 L_1}, \quad 0 \le s_1 \le L_1 - 1$$

$$p_2 = \frac{2\pi s_2 + \alpha_2}{a_2 L_2}, \quad 0 \le s_2 \le L_2 - 1.$$

This actually shows that the spectrum is invariant under a 2π shift of α_1 or of α_2 (up to an integer shift of s_1 or s_2). For any α_1, α_2, from the point of view of band theory, the ground state is found by filling all states in the first n bands with these shifted values of the momenta.

Now we think of the parameters α_1, α_2 as parameters that are going to vary adiabatically. Since they are each defined mod 2π, they parametrize a torus that I will call $\widehat{B}$. ($\widehat{B}$ can be viewed as a sort of rescaled version of the Brillouin zone B.) Since Berry's construction is universal for adiabatic variation of parameters, we can construct a Berry connection $\widehat{\mathcal{A}}$ over $\widehat{B}$, with curvature $\widehat{\mathcal{F}}$. $\widehat{\mathcal{A}}$ is a connection that can be used to transport the ground state as the parameters α_1, α_2 are varied. All we need to know to define it is that the ground state is always nondegenerate as α_1, α_2 are varied. We do not need to assume a single-particle picture (i.e. band theory). But I should say that for the conclusions we draw to be useful, at least in the form I will state, we need the gap from the ground state to be independent of L_1, L_2 as they become large. (Otherwise in practice our measurements in the lab may not be adiabatic. The stated assumption is not true for a fractional quantum Hall system, as I hope to explain tomorrow.)

Using the Berry connection over $\widehat{B}$, we can define an integer:

$$\widehat{k}' = \int_{\widehat{B}} \mathrm{d}\alpha_1 \, \mathrm{d}\alpha_2 \; \frac{\widehat{\mathcal{F}}}{2\pi}.$$
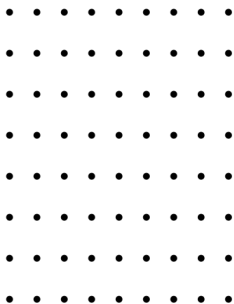
But I claim that this is the same as the integer k' defined in band theory:

$$k' = \widehat{k}'.$$

The reason that this is useful is that the definition of $\widehat{k}'$ is more general. To define k', we assume band theory – that is, a single-particle description based on free electrons. The definition of $\widehat{k}'$ assumes much less.

To understand why $k' = \widehat{k}'$, I have drawn

the discrete points in the Brillouin zone that obey the finite volume condition

$$p_1 = \frac{2\pi s_1}{a_1 L_1}, \quad 0 \le s_1 \le L_1 - 1$$

$$p_2 = \frac{2\pi s_2}{a_2 L_2}, \quad 0 \le s_2 \le L_2 - 1.$$

at $\alpha_1 = \alpha_2 = 0$. The parameters α_1, α_2 parametrize one of the little rectangles in the picture, say the one at the lower left.

Turning on α_1, α_2 shifts the allowed momenta as shown.
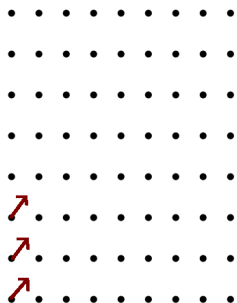
To compute k', we integrate over B, the full Brillouin zone. To compute $\widehat{k}'$, we integrate over the little rectangle, but for each point in the little rectangle, we sum over the corresponding shifted momenta
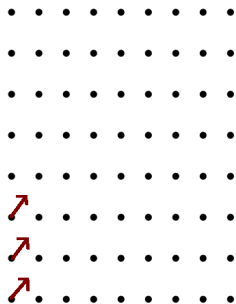
These are two different ways to organize the same calculation, so $\widehat{k}' = k'$.

So instead of proving the original TKNN formula $k = k'$, we can prove $k = \widehat{k}'$. This has the following advantage: $\widehat{k}'$ is defined in terms of the response of the system to a changing electromagnetic vector potential A, so we can determine $\widehat{k}'$ just from a knowledge of the effective action for A.

As practice, before determining the Berry connection for A, I am going to determine the Berry connection for an arbitrary dynamical system with dynamical variables $x^i(t)$. You can think of $x^i(t)$, $i = 1, \ldots, 3$ as representing the position coordinates of a particle, but they really could be anything else (for example $x^i(t)$ could have $3N$ components representing the positions of N particles). Regardless, we assume an action

$$I = \frac{1}{2} \int \mathrm{d}t \, g_{ij}(x) \frac{\mathrm{d}x^i}{\mathrm{d}t} \frac{\mathrm{d}x^j}{\mathrm{d}t} + \int \mathrm{d}t \mathbf{A}_i(x) \frac{\mathrm{d}x^i}{\mathrm{d}t} - \int \mathrm{d}t V(x) + \ldots.$$

(There might be higher order terms but it will be clear in a moment that they are not important.) We shall compute the Berry connection in the space of semiclassical states of zero energy, a condition that we satisfy by imposing the condition $V(x) = 0$. (This semiclassical approximation is valid in our problem because we do not need to treat the electromagnetic vector potential A quantum mechanically. We can view it as a given external field.)

Setting $V = 0$ means that we will evaluate the Berry phase not for all values of x but only for values of x that ensure $V(x) = 0$.
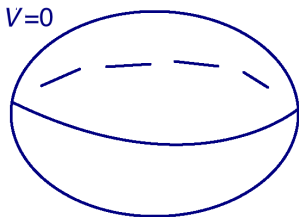
So we drop the $V(x)$ term from the action, and only carry out transport in the subspace of the configuration space with $V = 0$. In adiabatic transport, we can also ignore the term

$$I_{\mathrm{kin}} = \frac{1}{2} \int \mathrm{d}t \, g_{ij}(x) \frac{\mathrm{d}x^i}{\mathrm{d}t} \frac{\mathrm{d}x^j}{\mathrm{d}t}$$

in the action, and any other term with two or more time derivatives. That is because if we transport from a starting point p to an ending point p' in time T, the derivative $\mathrm{d}x^i/\mathrm{d}t$ is of order $1/T$, and $I_{\mathrm{kin}} \sim 1/T$.

So the only term in the action that we need to keep is the term with precisely one time derivative:

$$I' = \int \mathrm{d}t \, \mathbf{A}_i(x) \frac{\mathrm{d}x^i}{\mathrm{d}t} = \int_p^{p'} \mathbf{A}_i(x)\mathrm{d}x^i.$$

As I have indicated, this term depends only on the path followed from p to p', and not on how it is parametrized. Now remember that the phase that a quantum particle acquires in propagating from p to p' along a given trajectory is $e^{iI/\hbar}$, where I is the action for that trajectory.
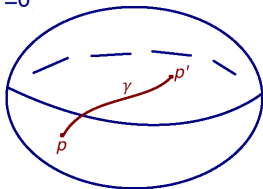
For us this phase is just $\exp\left(i \int_\gamma \mathbf{A}_i \mathrm{d}x^i\right)$.

But the connection which on parallel transport along a path γ gives a phase $\exp\left(i\int_\gamma \mathbf{A}_i \mathrm{d}x^i\right)$ is just $\mathbf{A}$. What we have learned, in other words, is that for a system in which a quantum ground state can be considered to be equivalent to a classical ground state, the Berry connection is just the classical connection $\mathbf{A}$ that can be read off from the classical action.

For the electromagnetic field in our problem, the action is

$$I = \frac{1}{2e^2} \int_{\mathbb{R}^{3,1}} \mathrm{d}^3x\mathrm{d}t \left(\vec{E}^2 - \vec{B}^2 \right) + \frac{k}{4\pi} \int_{W_3} \mathrm{d}^2x\mathrm{d}t \, \epsilon^{ijk} A_i \partial_j A_k + \dots$$

We assume, for example, periodic boundary conditions with very long periods L_1, L_2 in the two directions that are filled by our quantum Hall sample. (It doesn't matter if we assume periodic boundary conditions in the third direction.) A classical state of zero energy is labeled by the two angles α_1, α_2 that were introduced before:

$$A_1 = \frac{\alpha_1}{a_1 L_1}, \quad A_2 = \frac{\alpha_2}{a_2 L_2}.$$

To compute the Berry phase, we are supposed to substitute this formula in the action and keep only the part of the action that has precisely 1 time derivative. This comes only from the Chern-Simons term.

After integration over x_1 and x_2, the relevant part of the action is just

$$I' = -\frac{k}{2\pi} \int \mathrm{d}t\, \alpha_1 \frac{\mathrm{d}\alpha_2}{\mathrm{d}t}.$$

From this we read off the Berry connection

$$\nabla \equiv \left(\frac{D}{D\alpha_1}, \frac{D}{D\alpha_2} \right) = \left(\frac{\partial}{\partial \alpha_1}, \frac{\partial}{\partial \alpha_2} + i\frac{k\alpha_1}{2\pi} \right)$$

and hence the Berry curvature

$$\widehat{\mathcal{F}}_{\alpha_1 \alpha_2} = -i \left[\frac{D}{D\alpha_1}, \frac{D}{D\alpha_2} \right] = \frac{k}{2\pi}.$$

(If we add to I' a total derivative term $\int \mathrm{d}t\, \partial_t f(\alpha_1, \alpha_2)$, this will change the formula for ∇ but it will not change $\widehat{\mathcal{F}}$.)

We remember that the integer $\widehat{k}'$ is supposed to be the integral of $\widehat{\mathcal{F}}/2\pi$ over the Brillouin zone. We can now compute

$$\widehat{k}' = \int_0^{2\pi} \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \, \frac{\widehat{\mathcal{F}}}{2\pi} = \int_0^{2\pi} \mathrm{d}\alpha_1 \mathrm{d}\alpha_2 \, \frac{k}{(2\pi)^2} = k.$$

Thus we arrive at a version of the famous formula of TKNN: the coefficient k of the quantum Hall current can be computed as a flux integral of the Berry connection. (A similar explanation can be given for the result about the polarization of a 1-dimensional system that Charlie Kane described in his first lecture. Also, let me say again that the original proof that $k = k'$ was based on a direct evaluation of the Kubo formula for the conductivity in the context of band theory.)

Yesterday, we explained why a purely 1d quantum electron gas cannot have an imbalance between left-moving and right-moving electron excitations. As a reminder, the reason was that in a periodic orbit, "what goes up must come down"
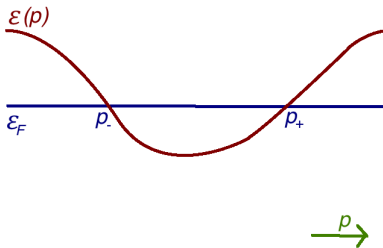
From a field theory point of view, this is needed because right-moving gapless fermions without left-moving ones cannot be quantized in a gauge-invariant fashion. There is an "Adler-Bell-Jackiw anomaly."

However, one of the hallmarks of a quantum Hall system is that on its boundary it has precisely such an imbalance. The reason that this must happen is that when we verified the invariance of the Chern-Simons action

$$kI_{\mathrm{CS}} = \frac{k}{4\pi} \int_{\mathcal{M}_2 \times \mathbb{R}} \mathrm{d}^3 x \epsilon^{ijk} A_i \partial_j A_k$$

under a gauge transformation $A_i \to A_i + \partial_i \phi$, we had to integrate by parts. This integration by parts produces a surface term on the surface of our material – that is on $\partial M_2 \times \mathbb{R}$. There is not any way to cancel this failure of gauge invariance by adding to the action a surface term supported on $\partial M_2 \times \mathbb{R}$. You can try to replace I_{CS} by

$$I_{\mathrm{CS}} + \int_{\partial M_2 \times \mathbb{R}} \mathrm{d}t \mathrm{d}x \, (?????)$$

where ????? is some polynomial in A and its derivatives, but whatever you try will not work. (I recommend this exercise.).

To cancel the "anomaly," that is the failure of gauge invariance of I_{CS} along the boundary, requires the existence on the boundary of modes that are (1) gapless, so they cannot be integrated out to produce a local effective action for A only, and (2) "anomalous," that is they are not possible in a purely 1-dimensional system. What fills the bill is precisely what we found does not exist in a purely 1-dimensional system: "chiral fermions," that is right-moving gapless modes not accompanied by left-moving ones.

Since the failure of kI_{CS} is proportional to k, the "chiral asymmetry" that is needed to cancel it is also proportional to k. In fact, the hallmark of an integer quantum Hall system with a Hall conductivity of k is precisely that

$$n_+ - n_- = k$$

where n_+ and n_- are the numbers of "right-moving" and "left-moving" gapless edge modes. Instead of giving a technical analysis of field theory "anomalies" to explain how this works, I will give a couple of possibly more physical explanations – one today and hopefully one tomorrow.

Let us think of a quantum Hall system in the form of a long cylinder:

In fact for starters, think of an infinite cylinder.

We introduce the same sort of "twist parameter" α as before. We can imagine that there is a magnetic flux α through a solenoid inside the cylinder such that the magnetic field is 0 (or at least independent of α) in the cylinder itself but

$$\oint_\gamma A \cdot \mathrm{d}\ell = \alpha.$$

Just as before, the parameter α is only gauge-invariant mod 2π.

We adiabatically increase α from 0 to 2π, with the scalar potential assumed to be 0. Since the electric field is then

$$\vec{E} = \frac{\partial \vec{A}}{\partial t},$$

increasing α turns on an electric field that goes "around" the cylinder. But in the case of a quantum Hall system, this drives a current that is perpendicular to $\vec{E}$, in other words the current flows "along" the cylinder.
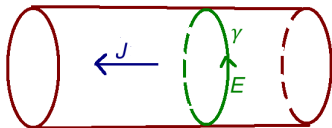
The electrons therefore are pushed to the left (or right, depending on the sign of k).

An early explanation of the integer quantum Hall effect by Laughlin was the following. We assume that when $\alpha = 2\pi$, the system returns to the same state that it was in at $\alpha = 0$. (This assumption is not valid for fractional quantum Hall systems, as discussed later.) However, in the process, each electron may move k steps to the left, for some integer k. Notice that since the cylinder has a finite circumference S, the number of electrons per unit length is finite and thus it makes sense to say that each one moves k steps to the left, for some k. This was interpreted as the basic integrality of the integer quantum Hall effect. It does lead to the value $k/2\pi$ for the Hall conductivity.

Now let us consider a cylinder that is only *semi*-infinite, with a boundary at let us say the left end:

The same parameter α as before makes sense, and we can still adiabatically increase it by 2π. Since a quantum Hall system is gapped, if we make a measurement far from the boundary, we will still see the same flux of valence electrons to the left as before, assuming that only valence bands (states below the fermi energy) are filled.

But what happens to the electrons when they arrive at the left boundary? A partial answer is that there are edge states, and electrons go from the valence bands to the edge states. But this is not enough: since the boundary has finite length, only finitely many electrons can go into edge states (of reasonable energy). The only interpretation is that as electrons flow in to the left from the valence bands (the bands below the usual ε_F in the bulk) they must eventually flow back out to the right in the conduction bands (the bands above the usual ε_F). Moreover, all this is happening continuously in energy so it must be possible for an electron to evolve *continuously* from the valence bands in the bulk, to the conduction bands in the bulk, somehow passing through edge states.

The spectrum must therefore look something like this (drawn for $k = 2$):
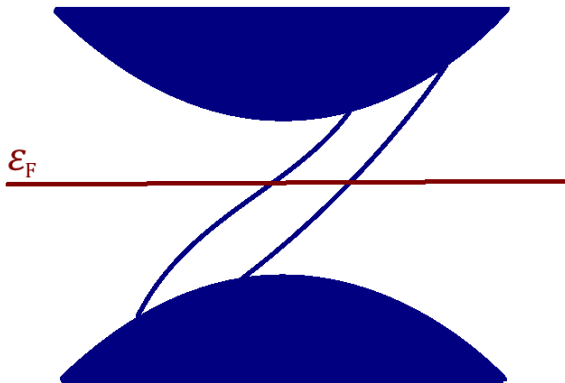
Actually the edge states are a continuum, as drawn in the last picture, only in the limit that the circumference S of the cylinder goes to infinity. For a finite S, the spectrum of edge states is discrete, as shown here:
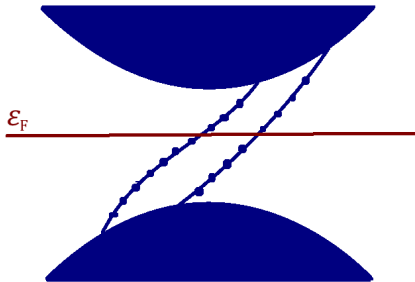
The little beads indicate "allowed states" of the edge modes, for a given S and a given value of the angle α. As we adiabatically increase α, each little bead moves up along the curve and under $\alpha \to \alpha + 2\pi$, each bead is shifted in position to the next one. So under $\alpha \to \alpha + 2\pi$, there is a net charge flow of 1 from the valence bands to the conduction bands for each right-moving edge mode.

Recall that as we discussed yesterday, a 1d mode is rightmoving if $d\varepsilon/dp > 0$ at $\varepsilon = \varepsilon_F$. A left-moving mode has $d\varepsilon/dp < 0$ at $\varepsilon = \varepsilon_F$, and under $\alpha \to \alpha + 2\pi$ produces a net charge flow of -1 from the valence band to the conduction band. Thus with n_+ and n_- as the numbers of right- and left-moving modes, the net charge flow under $\alpha \to \alpha + 2\pi$ is

$$k = n_+ - n_-.$$

To tie up some loose ends: The 1d edge modes cannot be defined on the whole 1d Brillouin zone (which is a circle) because then we would be stuck with the fact that in a periodic orbit "what goes up must come down," leading to $n_+ = n_-$. The asymmetry comes from branches of edge mode that exist in only a finite range of momenta $p_- \leq p \leq p_+$. What happens at the endpoints? The answer is the same as it was in a somewhat similar example that we looked at yesterday. The way that a family of edge-localized states can cease to exist at some momentum p_- is by ceasing to be normalizable. This happens when the edge state becomes indistinguishable from a bulk state.

That is part of what makes it possible to have adiabatic transport from the valence bands (the states normally filled) to the conduction bands (the states normally empty), through the edge states. At the endpoint of the edge state spectrum, an edge state is indistinguishable from a bulk state:
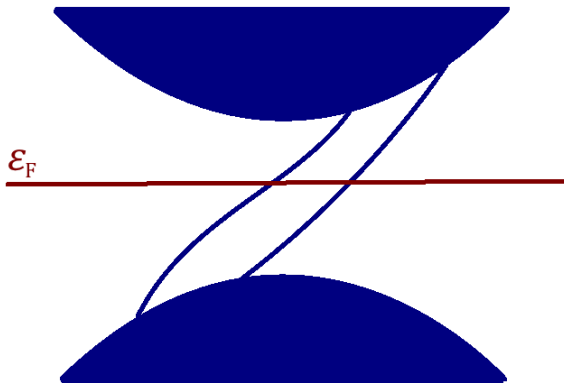
Also, for all this to make sense

the total Hall conductivity of the empty (conduction) bands must
be minus the Hall conductivity of the filled (valence) bands. That
is actually a property of the Berry connection. Let $\mathcal{A}$ be the usual
Berry connection for the filled bands and $\mathcal{F}$ the corresponding
curvature; and similarly let $\mathcal{A}'$ and $\mathcal{F}'$ be the Berry connection and
curvature of the empty bands. Then $\operatorname{Tr} \mathcal{F} + \operatorname{Tr} \mathcal{F}' = 0$, basically
because for all bands together there is no Berry curvature.

The Hall conductivities of filled and empty bands are respectively

$$\int_B \frac{\operatorname{Tr} \mathcal{F}}{2\pi}, \quad \int_B \frac{\operatorname{Tr} \mathcal{F}'}{2\pi}.$$

So the relation $\operatorname{Tr} \mathcal{F} + \operatorname{Tr} \mathcal{F}' = 0$ means that they have opposite Hall conductivities, and in our thought experiment, the flow of "filled" states (i.e. states that would be filled in the ground state on an infinite cylinder) to the left equals the flow of "empty" states to the right.

I would like to explain the assertion that a fractional quantum Hall system does not return to its previous state under $\alpha \to \alpha + 2\pi$. We will use the same macroscopic model of a fractional quantum Hall system that we used before in terms of the electromagnetic vector potential A and an emergent $U(1)$ gauge field a that only exists inside the material:

$$I_{\text{eff}} = \frac{1}{2\pi} \int_{M_3} \mathrm{d}^3x \; \epsilon^{ijk} A_i \partial_j a_k - \frac{r}{4\pi} \int_{M_3} \mathrm{d}^3x \; \epsilon^{ijk} a_i \partial_j a_k.$$

First let us discuss how to characterize the state of the system for a given α
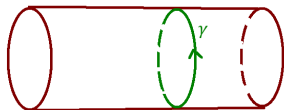
We recall that α was defined as $\oint_\gamma A$.

The parameter $\alpha = \oint_\gamma A$ can be controlled, in principle, by varying the magnetic flux threaded by the cylinder. But there is an analogous parameter

$$\widehat{\alpha} = \oint_\gamma a$$

that cannot be controlled in that way. Just like α, $\widehat{\alpha}$ is gauge-invariant mod 2π.

What can we say about $\widehat{\alpha}$? Recalling that $F = \mathrm{d}A$, $f = \mathrm{d}a$ are the ordinary electromagnetic field strength and its analog for a, the classical field equation for this system is

$$rf = F.$$

In the limit of an infinite cylinder, a can be treated classically. (We postpone the more interesting case of a finite cylinder until tomorrow.) In the gauge $A_0 = a_0 = 0$, the equation $rf_{0i} = F_{0i}$ becomes
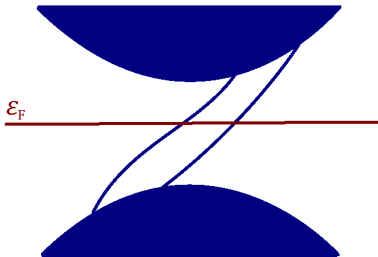
$$r\frac{\mathrm{d}a_i}{\mathrm{d}t} = \frac{\mathrm{d}A_i}{\mathrm{d}t},$$

and therefore

$$r\frac{\mathrm{d}\widehat{\alpha}}{\mathrm{d}t} = \frac{\mathrm{d}\alpha}{\mathrm{d}t}.$$

Hence when we adiabatically increase α by 2π, $\widehat{\alpha}$ increases adiabatically by $2\pi/r$. Since $\widehat{\alpha}$ is gauge-invariant mod 2π, the shift $\widehat{\alpha} \to \widehat{\alpha} + 2\pi/r$ does not return the system to its original state. We need to take $\alpha \to \alpha + 2\pi r$, and therefore the Hall conductivity can be smaller than its usual "quantum" by a factor of r.

The picture

still has some sort of analog, but the edge states cannot be free electron states: They have to be capable of transporting a fractional current under $\alpha \to \alpha + 2\pi$, and returning to their original state only under $\alpha \to \alpha + 2\pi r$.

There is actually one more topic that I would like to tidy up for today. Yesterday, we studied Weyl semimetals and by explicitly solving the Dirac equation, we learned that there must be surface states – Fermi arcs. We considered the Hamiltonian

$$H = -i\vec{\sigma} \cdot \frac{\partial}{\partial \vec{x}}$$

on a half-space $x_1 \geq 0$ with the boundary condition

$$\sigma_2 \psi| = \psi.$$

We found surface localized states at zero energy

$$\psi = \exp(ikx_3 - kx_1)\psi_0, \quad \sigma_2\psi_0 = \psi_0, \quad k > 0.$$

I should have pointed out that there are analogous surface-localized states of energy ε, for any ε:

$$\psi = \exp(i\varepsilon x_2)\exp(ikx_3 - kx_1)\psi_0, \quad \sigma_2\psi_0 = \psi_0, \quad k > 0.$$

Pictures I drew that only showed surface-localized states of $\varepsilon = 0$ were a little misleading.

Actually, the original paper predicting these states (Wan, Turner, Vishwanath, and Savrasov, 2011) did not proceed by explicitly picking a boundary condition and solving the Dirac equation. Rather, they deduced the result from some things that I have just explained. I want to explain how this goes.

First we recall the basic setup. Weyl points arise at special points in the Brillouin zone at which valence and conduction bands meet.
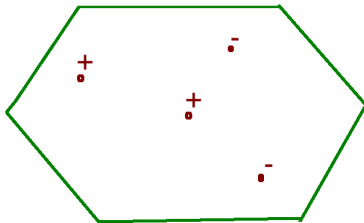
Near a boundary of a finite sample, only two of the three components of momentum are conserved. So it is useful to project the Brillouin zone and the bad points in it to two dimensions, "forgetting" the component of momentum that is not conserved:
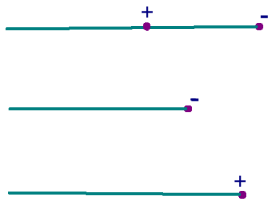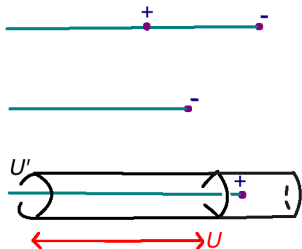

It is important to remember that in a crystal, the momentum components, including the component that is being "forgotten", are periodic, and in particular the horizontal direction in the picture represents a circle $U \cong S^1$, though it is hard to draw this.

Now draw a little circle U' around the projection of one of the bad points:

The product $U \times U'$ is a two-torus

We define an integer k^* as the Berry flux through $U \times U'$:

$$k^* = \int_{U \times U'} \mathrm{d}^2 p \, \frac{\mathrm{Tr}\,\mathcal{F}}{2\pi}.$$

It receives a contribution of 1 or -1 for each positive or negative Weyl point enclosed by $U \times U'$. So in the example drawn, $k^* = 1$, but we would get $k^* = 0$ or $k^* = -1$ if we take U' to encircle one of the other two special points in the projection.

We have arranged so that the two-torus $U \times U'$ does not intersect any of the Weyl points. So the restriction to $U \times U'$ of the original 3d band Hamiltonian on the 3d Brillouin zone B is a gapped Hamiltonian H^* on a two-torus $U \times U'$. We can intepret H^* as the band Hamiltonian of some 2d lattice system that has a Hall conductivity of k^*. So as we have learned, H^* has edge modes, equal in number to k^*, that "bridge the gap" in energy between the filled and empty bands
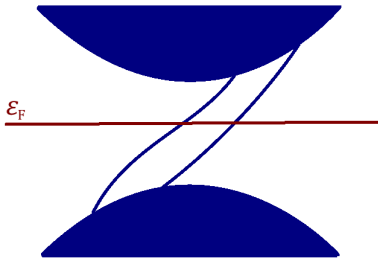
So there have to be edge states that intersect U' (the edge states are not labeled by U since U parametrizes the component of momentum that is not relevant to edge states). Since we had a lot of freedom in the choice of U', the spectrum of edge-localized points has to consist of arcs that link the appropriate boundary projections. In our example, this means edge states as shown
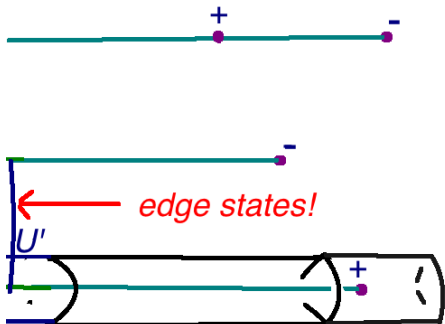
The auxiliary 2d quantum Hall system that was used in this argument does not have any simple relation to the 3d Weyl sem-metal that we were studying, as far as I know.

For tomorrow, there are three topics I hope to describe:

(1) More on the fractional quantum Hall effect.

(2) Another explanation of edge modes in the integral quantum Hall effect.

(3) Haldane's model of quantum Hall physics without an applied magnetic field.