2. The Integer Quantum Hall Effect

In this section we discuss the integer quantum Hall effect. This phenomenon can be understood without taking into account the interactions between electrons. This means that we will assume that the quantum states for a single particle in a magnetic field that we described in Section 1.4 will remain the quantum states when there are many particles present. The only way that one particle knows about the presence of others is through the Pauli exclusion principle: they take up space. In contrast, when we come to discuss the fractional quantum Hall effect in Section 3, the interactions between electrons will play a key role.

2.1 Conductivity in Filled Landau Levels

Let's look at what we know. The experimental data for the Hall resistivity shows a number of plateaux labelled by an integer ν. Meanwhile, the energy spectrum forms Landau levels, also labelled by an integer. Each Landau level can accommodate a large, but finite number of electrons.

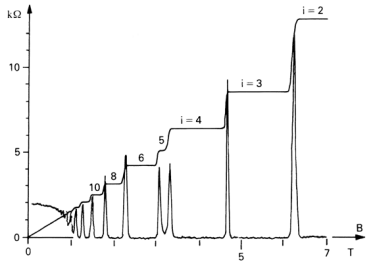


Figure 12: Integer quantum Hall effect

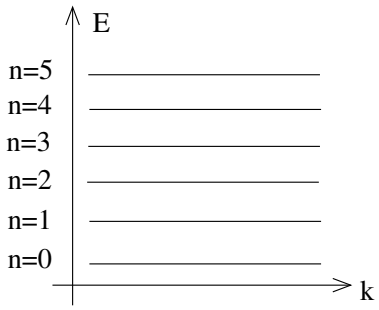


Figure 13: Landau levels

It's tempting to think that these integers are the same: $\rho_{xy} = 2\pi\hbar/e^2\nu$ and when precisely ν Landau levels are filled. And this is correct.

Let's first check that this simple guess works. If know that on a plateau, the Hall resistivity takes the value

$$\rho_{xy} = \frac{2\pi\hbar}{e^2}\frac{1}{\nu}$$

with $\nu \in \mathbf{Z}$. But, from our classical calculation in the Drude model, we have the expectation that the Hall conductivity should depend on the density of electrons, n

$$\rho_{xy} = \frac{B}{ne}$$

Comparing these two expressions, we see that the density needed to get the resistivity of the ν^{th} plateau is

$$n = \frac{B}{\Phi_0}\nu \tag{2.1}$$

with $\Phi_0 = 2\pi\hbar/e$. This is indeed the density of electrons required to fill ν Landau levels.

Further, when ν Landau levels are filled, there is a gap in the energy spectrum: to occupy the next state costs an energy $\hbar\omega_B$ where $\omega_B = eB/m$ is the cyclotron frequency. As long as we're at temperature $k_BT \ll \hbar\omega_B$, these states will remain empty. When we turn on a small electric field, there's nowhere for the electrons to move: they're stuck in place like in an insulator. This means that the scattering time $\tau \to \infty$ and we have $\rho_{xx} = 0$ as expected.

Conductivity in Quantum Mechanics: a Baby Version

The above calculation involved a curious mixture of quantum mechanics and the classical Drude mode. We can do better. Here we'll describe how to compute the conductivity for a single free particle. In section 2.2.3, we'll derive a more general formula that holds for any many-body quantum system.

We know that the velocity of the particle is given by

$$m\dot{\mathbf{x}} = \mathbf{p} + e\mathbf{A}$$

where p_i is the canonical momentum. The current is $\mathbf{I} = -e\dot{\mathbf{x}}$, which means that, in the quantum mechanical picture, the total current is given by

$$\mathbf{I} = -\frac{e}{m}\sum_{\text{filled states}} \langle\psi| - i\hbar\nabla + e\mathbf{A}|\psi\rangle$$

It's best to do these kind of calculations in Landau gauge, $\mathbf{A} = xB\hat{\mathbf{y}}$. We introduce an electric field E in the x-direction so the Hamiltonian is given by (1.23) and the states by (1.24). With the ν Landau levels filled, the current in the x-direction is

$$I_x = -\frac{e}{m}\sum_{n=1}^{\nu}\sum_k \langle\psi_{n,k}| - i\hbar\frac{\partial}{\partial x}|\psi_{n,k}\rangle = 0$$

This vanishes because it's computing the momentum expectation value of harmonic oscillator eigenstates. Meanwhile, the current in the y-direction is

$$I_y = -\frac{e}{m}\sum_{n=1}^{\nu}\sum_k \langle\psi_{n,k}| - i\hbar\frac{\partial}{\partial y} + exB|\psi_{n,k}\rangle = -\frac{e}{m}\sum_{n=1}^{\nu}\sum_k \langle\psi_{n,k}|\hbar k + eBx|\psi_{n,k}\rangle$$

The second term above is computing the position expectation value $\langle x \rangle$ of the eigenstates. But we know from (1.20) and (1.24) that these harmonic oscillator states are shifted from the origin, so that $\langle \psi_{n,k}|x|\psi_{n,k}\rangle = -\hbar k/eB - mE/eB^2$. The first of these terms cancels the explicit $\hbar k$ term in the expression for I_y. We're left with

$$I_y = e\nu \sum_k \frac{E}{B} \tag{2.2}$$

The sum over k just gives the number of electrons which we computed in (1.21) to be $N = AB/\Phi_0$. We divide through by the area to get the current density $\mathbf{J}$ instead of the current $\mathbf{I}$. The upshot of this is that

$$\mathbf{E} = \begin{pmatrix} E \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{J} = \begin{pmatrix} 0 \\ e\nu E/\Phi_0 \end{pmatrix}$$

Comparing to the definition of the conductivity tensor (1.6), we have

$$\sigma_{xx} = 0 \quad \text{and} \quad \sigma_{xy} = \frac{e\nu}{\Phi_0} \quad \Rightarrow \quad \rho_{xx} = 0 \quad \text{and} \quad \rho_{xy} = -\frac{\Phi_0}{e\nu} = -\frac{2\pi\hbar}{e^2\nu} \tag{2.3}$$

This is exactly the conductivity seen on the quantum Hall plateaux. Although the way we've set up our computation we get a negative Hall resistivity rather than positive; for a magnetic field in the opposite direction, you get the other sign.

2.1.1 Edge Modes

There are a couple of aspects of the story which the simple description above does not capture. One of these is the role played by disorder; we describe this in Section 2.2.1. The other is the special importance of modes at the edge of the system. Here we describe some basic facts about edge modes; we'll devote Section 6 to a more detailed discussion of edge modes in the fractional quantum Hall systems.

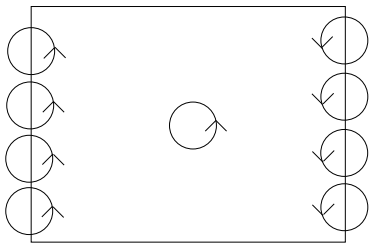


Figure 14:

The fact that something special happens along the edge of a quantum Hall system can be seen even classically. Consider particles moving in circles in a magnetic field. For a fixed magnetic field, all particle motion is in one direction, say anti-clockwise. Near the edge of the sample, the orbits must collide with the boundary. As all motion is anti-clockwise, the only option open to these particles is to bounce back. The result is a skipping motion in which the particles along the one-dimensional boundary move

only in a single direction, as shown in the figure. A particle restricted to move in a single direction along a line is said to be *chiral*. Particles move in one direction on one side of the sample, and in the other direction on the other side of the sample. We say that the particles have opposite *chirality* on the two sides. This ensures that the net current, in the absence of an electric field, vanishes.

We can also see how the edge modes appear in the quantum theory. The edge of the sample is modelled by a potential which rises steeply as shown in the figure. We'll work in Landau gauge and consider a rectangular geometry which is finite only in the x-direction, which we model by $V(x)$. The Hamiltonian is

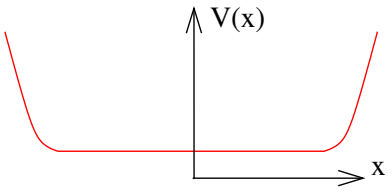


Figure 15:

$$H = \frac{1}{2m}\left(p_x^2 + (p_y + eBx)^2\right) + V(x)$$

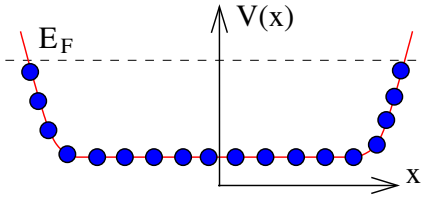
In the absence of the potential, we know that the wavefunctions are Gaussian of width l_B. If the potential is smooth over distance scales l_B, then, for each state, we can Taylor expand the potential around its location X. Each wavefunction then experiences the potential $V(x) \approx V(X) + (\partial V/\partial x)(x-X) + \ldots$. We drop quadratic terms and, of course, the constant term can be neglected. We're left with a linear potential which is exactly what we solved in Section 1.4.2 when we discussed Landau levels in a background electric field. The result is a drift velocity in the y-direction (1.26), now given by

$$v_y = -\frac{1}{eB}\frac{\partial V}{\partial x}$$

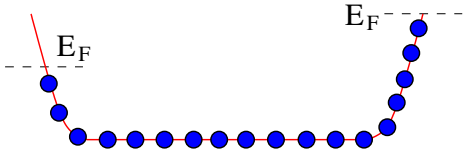
Each wavefunction, labelled by momentum k, sits at a different x position, $x = -kl_B^2$ and has a different drift velocity. In particular, the modes at each edge are both chiral, travelling in opposite directions: $v_y > 0$ on the left, and $v_y < 0$ on the right. This agrees with the classical result of skipping orbits.

Having a chiral mode is rather special. In fact, there's a theorem which says that you can't have charged chiral particles moving along a wire; there *has* to be particles which can move in the opposite direction as well. In the language of field theory, this follows from what's called the *chiral anomaly*. In the language of condensed matter physics, with particles moving on a lattice, it follows from the *Nielsen-Ninomiya theorem*. The reason that the simple example of a particle in a magnetic field avoids these theorems is because the chiral fermions live on the boundary of a two-dimensional system, rather than in a one-dimensional wire. This is part of a general story: there are physical phenomena which can only take place on the boundary of a system. This story plays a prominent role in the study of materials called *topological insulators*.

Let's now look at what happens when we fill the available states. We do this by introducing a chemical potential. The states are labelled by y-momentum $\hbar k$ but, as we've seen, this can equally well be thought of as the position of the state in the x-direction. This means that we're justified in drawing the filled states like this:



From our usual understanding of insulators and conductors, we would say that the bulk of the material is an insulator (because all the states in the band are filled) but the edge of the material is a metal. We can also think about currents in this language. We simply introduce a potential difference $\Delta\mu$ on the two sides of the sample. This means that we fill up more states on the right-hand edge than on the left-hand edge, like this:



To compute the resulting current we simply need to sum over all filled states. But, at the level of our approximation, this is the same as integrating over x

$$I_y = -e\int \frac{dk}{2\pi}\, v_y(k) = \frac{e}{2\pi l_B^2}\int dx\, \frac{1}{eB}\frac{\partial V}{\partial x} = \frac{e}{2\pi\hbar}\,\Delta\mu \tag{2.4}$$

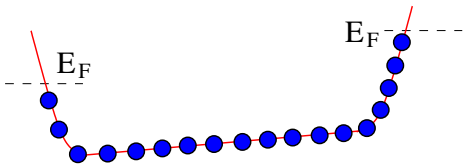
The Hall voltage is $eV_H = \Delta\mu$, giving us the Hall conductivity

$$\sigma_{xy} = \frac{I_y}{V_H} = \frac{e^2}{2\pi\hbar} \tag{2.5}$$

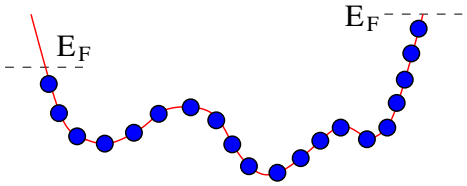
which is indeed the expected conductivity for a single Landau level.

The picture above suggests that the current is carried entirely by the edge states, since the bulk Landau level is flat so these states carry no current. Indeed, you can sometimes read this argument in the literature. But it's a little too fast: in fact, it's even in conflict with the computation that we did previously, where (2.2) shows that all states contribute equally to the current. That's because this calculation included the fact that the Landau levels are tilted by an electric field, so that the effective potential

and the filled states looked more like this:

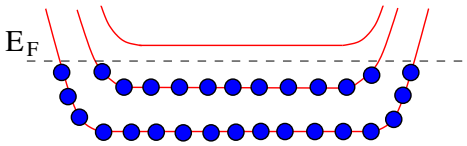


Now the current is shared among all of the states. However, the nice thing about the calculation (2.4) is that it doesn't matter what shape the potential V takes. As long as it is smooth enough, the resulting Hall conductivity remains quantised as (2.5). For example, you could consider the random potential like this



and you still get the quantised answer (2.4) as long as the random potential $V(x)$ doesn't extend above E_F. As we will describe in Section 2.2.1, these kinds of random potentials introduce another ingredient that is crucial in understanding the quantised Hall plateaux.

Everything we've described above holds for a single Landau level. It's easily generalised to multiple Landau levels. As long as the chemical potential E_F lies between Landau levels, we have n filled Landau levels, like this



Correspondingly, there are n types of chiral mode on each edge.

A second reason why chiral modes are special is that it's hard to disrupt them. If you add impurities to any system, they will scatter electrons. Typically such scattering makes the electrons bounce around in random directions and the net effect is often that the electrons don't get very far at all. But for chiral modes this isn't possible simply because all states move in the same direction. If you want to scatter a left-moving electron into a right-moving electron then it has to cross the entire sample. That's a long way for an electron and, correspondingly, such scattering is highly suppressed. It

means that currents carried by chiral modes are immune to impurities. However, as we will now see, the impurities do play an important role in the emergence of the Hall plateaux.

2.2 Robustness of the Hall State

The calculations above show that *if* an integer number of Landau levels are filled, then the longitudinal and Hall resistivities are those observed on the plateaux. But it doesn't explain why these plateaux exist in the first place, nor why there are sharp jumps between different plateaux.

To see the problem, suppose that we fix the electron density n. Then we only completely fill Landau levels when the magnetic field is exactly $B = n\Phi_0/\nu$ for some integer ν. But what happens the rest of the time when $B \neq n\Phi_0/\nu$? Now the final Landau level is only partially filled. Now when we apply a small electric field, there are accessible states for the electrons to scatter in to. The result is going to be some complicated, out-of-equilibrium distribution of electrons on this final Landau level. The longitudinal conductivity σ_{xx} will surely be non-zero, while the Hall conductivity will differ from the quantised value (2.3).

Yet the whole point of the quantum Hall effect is that the experiments reveal that the quantised values of the resistivity (2.3) persist over a range of magnetic field. How is this possible?

2.2.1 The Role of Disorder

It turns out that the plateaux owe their existence to one further bit of physics: disorder. This arises because experimental samples are inherently dirty. They contain impurities which can be modelled by adding a random potential $V(\mathbf{x})$ to the Hamiltonian. As we now explain, this random potential is ultimately responsible for the plateaux observed in the quantum Hall effect. There's a wonderful irony in this: the glorious precision with which these integers ν are measured is due to the dirty, crappy physics of impurities.

To see how this works, let's think about what disorder will likely do to the system. Our first expectation is that it will split the degenerate eigenstates that make up a Landau level. This follows on general grounds from quantum perturbation theory: any generic perturbation, which doesn't preserve a symmetry, will break degeneracies. We will further ask that the strength of disorder is small relative to the splitting of the Landau levels,

$$V \ll \hbar\omega_B \tag{2.6}$$

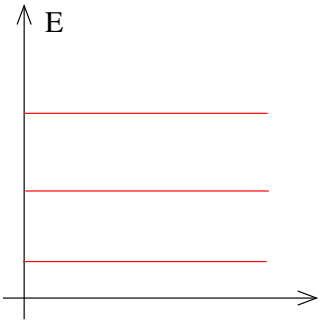


Figure 16: Density of states without disorder...

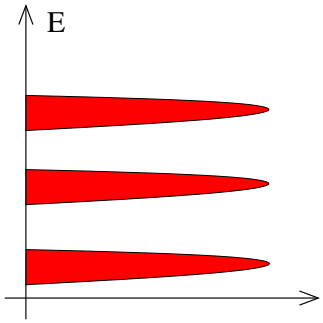


Figure 17: ...and with disorder.

In practice, this means that the samples which exhibit the quantum Hall effect actually have to be very clean. We need disorder, but not too much disorder! The energy spectrum in the presence of this weak disorder is the expected to change the quantised Landau levels from the familiar picture in the left-hand figure, to the more broad spectrum shown in the right-hand figure.

There is a second effect of disorder: it turns many of the quantum states from *extended* to *localised*. Here, an extended state is spread throughout the whole system. In contrast, a localised state is restricted to lie in some region of space. We can easily see the existence of these localised states in a semi-classical picture which holds if the potential, in addition to obeying (2.6), varies appreciably on distance scales much greater than the magnetic length l_B,

$$|\nabla V| \ll \frac{\hbar\omega_B}{l_B}$$

With this assumption, the cyclotron orbit of an electron takes place in a region of essentially constant potential. The centre of the orbit, $\mathbf{X}$ then drifts along equipotentials. To see this, recall that we can introduce quantum operators (X, Y) describing the centre of the orbit (1.33),

$$X = x - \frac{\pi_y}{m\omega_B} \quad \text{and} \quad Y = y + \frac{\pi_x}{m\omega_B}$$

with $\boldsymbol{\pi}$ the mechanical momentum (1.14). (Recall that, in contrast to the canonical momentum, $\boldsymbol{\pi}$ is gauge invariant). The time evolution of these operators is given by

$$\begin{aligned} i\hbar\dot{X} &= [X, H+V] = [X, V] = [X, Y]\frac{\partial V}{\partial Y} = il_B^2\frac{\partial V}{\partial Y} \\ i\hbar\dot{Y} &= [Y, H+V] = [Y, V] = [Y, X]\frac{\partial V}{\partial X} = -il_B^2\frac{\partial V}{\partial X} \end{aligned}$$

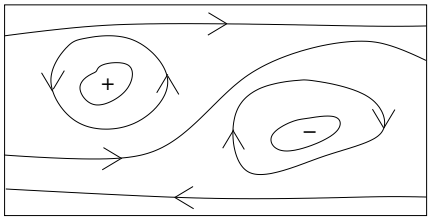


Figure 18: The localisation of states due to disorder.

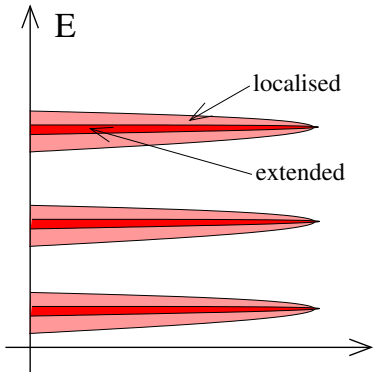


Figure 19: The resulting density of states.

where we used the fact (1.34) that, in the absence of a potential, $[X, H] = [Y, H] = 0$, together with the commutation relation $[X, Y] = il_B^2$ (1.35). This says that the centre of mass drifts in a direction $(\dot{X}, \dot{Y})$ which is perpendicular to ∇V; in other words, the motion is along equipotentials.

Now consider what this means in a random potential with various peaks and troughs. We've drawn some contour lines of such a potential in the left-hand figure, with + denoting the local maxima of the potential and − denoting the local minima. The particles move anti-clockwise around the maxima and clockwise around the minima. In both cases, the particles are trapped close to the extrema. They can't move throughout the sample. In fact, equipotentials which stretch from one side of a sample to another are relatively rare. One place that they're guaranteed to exist is on the edge of the sample.

The upshot of this is that the states at the far edge of a band — either of high or low energy — are localised. Only the states close to the centre of the band will be extended. This means that the density of states looks schematically something like the right-hand figure.

Conductivity Revisited

For conductivity, the distinction between localised and extended states is an important one. Only the extended states can transport charge from one side of the sample to the other. So only these states can contribute to the conductivity.

Let's now see what kind of behaviour we expect for the conductivity. Suppose that we've filled all the extended states in a given Landau level and consider what happens as we decrease B with fixed n. Each Landau level can accommodate fewer electrons.

But, rather than jumping up to the next Landau level, we now begin to populate the localised states. Since these states can't contribute to the current, the conductivity doesn't change. This leads to exactly the kind of plateaux that are observed, with constant conductivities over a range of magnetic field.

So the presence of disorder explains the presence of plateaux. But now we have to revisit our original argument of why the resistivities take the specific quantised values (2.3). These were computed assuming that all states in the Landau level contribute to the current. Now we know that many of these states are localised by impurities and don't transport charge. Surely we expect the value of the resistivity to be different. Right? Well, no. Remarkably, current carried by the extended states increases to compensate for the lack of current transported by localised states. This ensures that the resistivity remains quantised as (2.3) despite the presence of disorder. We now explain why.

2.2.2 The Role of Gauge Invariance

Instead of considering electrons moving in a rectangular sample, we'll instead consider electrons moving in the annulus shown in the figure. In this context, this is sometimes called a *Corbino ring*. We usually console ourselves by arguing that if the Hall conductivity is indeed quantised then it shouldn't depend on the geometry of the sample. (Of course, the flip side of this is that if we've really got the right argument, that shouldn't depend on the geometry of the sample either; unfortunately this argument does.)

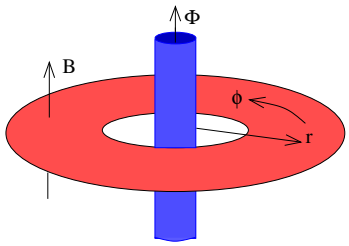


Figure 20:

The nice thing about the ring geometry is that it provides us with an extra handle[10]. In addition to the background magnetic field B which penetrates the sample, we can thread an additional flux Φ through the centre of the ring. Inside the ring, this Φ is locally pure gauge. Nonetheless, from our discussion in Section 1.5, we known that Φ can affect the quantum states of the electrons.

Let's first see what Φ has to do with the Hall conductivity. Suppose that we slowly increase Φ from 0 to $\Phi_0 = 2\pi\hbar/e$. Here "slowly" means that we take a time $T \gg 1/\omega_B$. This induces an emf around the ring, $\mathcal{E} = -\partial\Phi/\partial t = -\Phi_0/T$. Let's suppose that we

[10]This argument was first given by R. B. Laughlin in "*Quantized Hall Conductivity in Two Dimensions*", Phys. Rev, **B23** 5632 (1981). Elaborations on the role of edge states were given by B. I. Halperin in "*Quantized Hall conductance, current carrying edge states, and the existence of extended states in a two-dimensional disordered potential*," Phys. Rev. **B25** 2185 (1982).

can argue that n electrons are transferred from the inner circle to the outer circle in this time. This would result in a radial current $I_r = -ne/T$.

$$\rho_{xy} = \frac{\mathcal{E}}{I_r} = \frac{2\pi\hbar}{e^2}\frac{1}{n} \tag{2.7}$$

This is the result we want. Our task, therefore, is to argue that n electrons are indeed transferred across the ring as the flux is increased to Φ_0.

Spectral Flow in Landau Levels

The key idea that we need is that of *spectral flow*, introduced in Section 1.5.3. The spectrum of the Hamiltonian is the same whenever Φ is an integer multiple of Φ_0. However, if we start with a particular energy eigenstate when $\Phi = 0$, this will evolve into a different energy eigenstate with $\Phi = \Phi_0$. As the change is done suitably slowly, over a time $T \gg 1/\omega_B$, the adiabatic theorem ensures that the final energy eigenstate must lie in the same Landau level as the initial state.

To illustrate this, let's first look at the situation with no disorder. For the ring geometry, it is sensible to use symmetric gauge and radial coordinates, $z = x - iy = re^{i\phi}$. The wavefunctions in the lowest Landau level are (1.30),

$$\psi_m \sim z^m e^{-|z|^2/4l_B^2} = e^{im\phi} r^m e^{-r^2/4l_B^2}$$

The m^{th} wavefunction is strongly peaked at a radius $r \approx \sqrt{2ml_B^2}$ (where, of course, we must now chose $m \in \mathbf{Z}$ such that the wavefunction lies inside the annulus). From the discussion in Section 1.5.3, we see that if we increase the flux from $\Phi = 0$ to $\Phi = \Phi_0$, the wavefunctions shift from m to $m + 1$,

$$\psi_m(\Phi = 0) \;\longrightarrow\; \psi_m(\Phi = \Phi_0) = \psi_{m+1}(\Phi = 0)$$

This means that each state moves outwards, from radius $r = \sqrt{2ml_B^2}$ to $r = \sqrt{2(m+1)l_B^2}$. The net result is that, if all states in the Landau level are filled, a single electron is transferred from the inner ring to the outer ring as the flux is increased from $\Phi = 0$ to $\Phi = \Phi_0$. It is simple to check that the same result holds for higher Landau levels. If n Landau levels are filled, then n electrons are transferred from the inner to the outer ring and the Hall resistivity is given by (2.7) as required.

Spectral Flow in the Presence of Disorder

The discussion above merely reproduces what we already know. Let's now see how it changes in the presence of disorder. In polar coordinates, the Hamiltonian takes the

form

$$H_{\Phi=0} = \frac{1}{2m}\left[-\hbar^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \left(-\frac{i\hbar}{r}\frac{\partial}{\partial\phi} + \frac{eBr}{2}\right)^2\right] + V(r,\phi)$$

where $V(r,\phi)$ is the random potential capturing the effects of disorder. Note that this depends on ϕ, so angular momentum is no longer a good quantum number in this system. Adding the flux through the centre changes the Hamiltonian to

$$H_{\Phi} = \frac{1}{2m}\left[-\hbar^2\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \left(-i\frac{\hbar}{r}\frac{\partial}{\partial\phi} + \frac{eBr}{2} + \frac{e\Phi}{2\pi r}\right)^2\right] + V(r,\phi)$$

Importantly, the flux Φ affects only the extended states. It does not change the localised states. To see this, we attempt to undo the flux by a gauge transformation,

$$\psi(r,\phi) \to e^{-ie\Phi\phi/2\pi\hbar}\psi(r,\phi)$$

For the localised states, where ψ is non-zero only in some finite region, there's no problem in doing this. However for the extended states, which wrap around the annulus, we also have the requirement that the wavefunction ψ is single-valued as $\phi \to \phi + 2\pi$. We see that this is only true when Φ is an integer multiple of $\Phi_0 = 2\pi\hbar/e$.

The upshot of this argument is that the spectrum of the Hamiltonian is again left unchanged when Φ is an integer multiple of Φ_0. But, this time, as the flux is increased from 0 to Φ_0, the localised states don't change. Only the extended states undergo spectral flow; these alone must map onto themselves.

There are always at least two extended states: one near the inner ring and one near the outer ring. The spectral flow happens in the same heuristic manner as described above: an extended state localised at one radius is transformed into an extended state at the next available radius. The presence of disorder means that there are fewer extended states, but this doesn't change the overall conclusion: if all extended states in a given Landau level are filled, then the net effect of dialling the flux from $\Phi = 0$ to $\Phi = \Phi_0$ is to transport one electron from the inner to the outer edge. If n Landau levels are filled, we again get the result (2.7).

The arguments above involving gauge transformations start to give a hint of the topological nature of the quantum Hall effect. In fact, there are much deeper topological ideas underlying the quantisation of the Hall conductivity. We'll describe these in Section 2.2.4 and, in a slightly different context, in Section 2.3. However, before we proceed we first need a basic result which gives an expression for the conductivity in any quantum mechanical system.

2.2.3 An Aside: The Kubo Formula

Before we get to anything related to topology, we first need to lay some groundwork. Our task in this section will be to derive a formula for the Hall conductivity σ_{xy} in terms of quantum mechanical observables. The expression that we're looking for is called the *Kubo formula*; it is part of more general story that goes by the name of *linear response*[11].

We'll derive the Kubo formula for a general, multi-particle Hamiltonian H_0 where the subscript 0 means that this is the unperturbed Hamiltonian before we turn on an electric field. At this point, H_0 could be that of many non-interacting particles each, for example, obeying the single-particle Hamiltonian (1.13) that we saw previously, or it could be something more complicated. Later, we'll apply the Kubo formula both to Hamiltonians which describe particles moving in the continuum and to Hamiltonians that describe particles moving on a lattice. We denote the energy eigenstates of H_0 as $|m\rangle$, with $H_0|m\rangle = E_m|m\rangle$.

Now we add a background electric field. We work in the gauge with $A_t = 0$ so that the electric field can be written as $\mathbf{E} = -\partial_t \mathbf{A}$. The new Hamiltonian takes the form $H = H_0 + \Delta H$ with

$$\Delta H = -\mathbf{J} \cdot \mathbf{A} \tag{2.8}$$

where $\mathbf{J}$ is the quantum operator associated to the electric current. For the simple Hamiltonians that we considered in Section 1.4, $\mathbf{J}$ is equal (up to constants) to the mechanical momentum $\boldsymbol{\pi} = \mathbf{p} + e\mathbf{A} = m\dot{\mathbf{x}}$ which we defined in equation (1.14). However, we'll use more general definitions of $\mathbf{J}$ in what follows.

At this point, there's a couple of tricks that makes life simpler. First, we're ultimately interested in applying a constant, DC electric field. However, it turns out to be simpler to apply an AC electric field, $\mathbf{E}(t) = \mathbf{E}e^{-i\omega t}$ with frequency ω, and to then take the limit $\omega \to 0$. Second, it's also somewhat simpler if we work with a complexified $\mathbf{A}$. There's nothing deep in this: it's just easier to write $e^{-i\omega t}$ than, say, $\cos(\omega t)$. Because all our calculations will be to linear order only, you can take the real part at any time. We then have

$$\mathbf{A} = \frac{\mathbf{E}}{i\omega} e^{-i\omega t} \tag{2.9}$$

[11] You can read about this story in the lecture notes on *Kinetic Theory* where a slightly more sophisticated discussion of the Kubo formula can be found in Sections 4.3 and 4.4. In particular, there is often an extra term proportional to $\mathbf{A}^2$ in ΔH which contributes to σ_{xx} but not σ_{xy} so is ignored in the present discussion.

Our goal is to compute the current $\langle \mathbf{J} \rangle$ that flows due to the perturbation ΔH. We will assume that the electric field is small and proceed using standard perturbation theory.

We work in the interaction picture. This means that operators evolve as $\mathcal{O}(t) = V^{-1}\mathcal{O}V$ with $V = e^{-iH_0 t/\hbar}$. In particular $\mathbf{J}$, and hence $\Delta H(t)$ itself, both vary in time in this way. Meanwhile states $|\psi(t)\rangle$, evolve by

$$|\psi(t)\rangle_I = U(t, t_0)|\psi(t_0)\rangle_I$$

where the unitary operator $U(t, t_0)$ is defined as

$$U(t, t_0) = T \exp\left(-\frac{i}{\hbar}\int_{t_0}^{t} \Delta H(t')\, dt'\right) \tag{2.10}$$

Here T stands for time ordering; it ensures that U obeys the equation $i\hbar\, dU/dt = \Delta H\, U$.

We're interested in systems with lots of particles. Later we'll only consider non-interacting particles but, importantly, the Kubo formula is more general than this. We prepare the system at time $t \to -\infty$ in a specific many-body state $|0\rangle$. This is usually taken to be the many-body ground state, although it needn't necessarily be. Then, writing $U(t) = U(t, t_0 \to -\infty)$, the expectation value of the current is given by

$$\begin{aligned}
\langle \mathbf{J}(t) \rangle &= \langle 0(t)|\, \mathbf{J}(t) |0(t)\,\rangle \\
&= \langle 0|\, U^{-1}(t)\mathbf{J}(t)U(t)\, |0\rangle \\
&\approx \langle 0| \left(\mathbf{J}(t) + \frac{i}{\hbar}\int_{-\infty}^{t} dt'\, \left[\Delta H(t'), \mathbf{J}(t)\right]\right) |0\rangle
\end{aligned}$$

where, in the final line, we've expanded the unitary operator (2.10), keeping only the leading terms. The first term is the current in the absence of an electric field. We'll assume that this term vanishes. Using the expressions (2.8) and (2.9), the current due to the electric field is then

$$\langle J_i(t) \rangle = \frac{1}{\hbar\omega}\int_{-\infty}^{t} dt'\, \langle 0|[J_j(t'), J_i(t)]|0\rangle\, E_j\, e^{-i\omega t'}$$

Because the system is invariant under time translations, the correlation function above can only depend on $t'' = t - t'$. We can then rewrite the expression above as

$$\langle J_i(t) \rangle = \frac{1}{\hbar\omega}\left(\int_0^{\infty} dt''\, e^{i\omega t''}\, \langle 0|[J_j(0), J_i(t'')]|0\rangle\right) E_j e^{-i\omega t}$$

The only t dependence in the formula above sits outside as $e^{-i\omega t}$. This is telling us that if you apply an electric field at frequency ω, the current responds by oscillating at the

same frequency ω. This is the essence of *linear response*. The proportionality constant defines the frequency-dependent conductivity matrix $\sigma(\omega)$. The Hall conductivity is the off-diagonal part

$$\sigma_{xy}(\omega) = \frac{1}{\hbar\omega}\int_0^\infty dt\ e^{i\omega t}\,\langle 0|[J_y(0), J_x(t)]|0\rangle$$

This is the *Kubo formula* for the Hall conductivity.

We can massage the Kubo formula into a slightly more useful form. We use the fact that the current operator evolves as $\mathbf{J}(t) = V^{-1}\,\mathbf{J}(0)\,V$ with $V = e^{-iH_0t/\hbar}$. We then evaluate $\sigma_{xy}(\omega)$ by inserting complete basis of energy eigenstates of H_0,

$$\sigma_{xy}(\omega) = \frac{1}{\hbar\omega}\int_0^\infty dt\ e^{i\omega t}\sum_n \left[\langle 0|J_y|n\rangle\langle n|J_x|0\rangle e^{i(E_n-E_0)t/\hbar} - \langle 0|J_x|n\rangle\langle n|J_y|0\rangle e^{i(E_0-E_n)t/\hbar}\right]$$

We now perform the integral over $\int dt$. (There's a subtlety here: to ensure convergence, we should replace $\omega \to \omega + i\epsilon$, with ϵ infinitesimal. There is a story related to causality and where poles can appear in the complex ω plane which you can learn more about in the Kinetic Theory lecture notes.) Since the states with $|n\rangle = |0\rangle$ don't contribute to the sum, we get

$$\sigma_{xy}(\omega) = -\frac{i}{\omega}\sum_{n\neq 0}\left[\frac{\langle 0|J_y|n\rangle\langle n|J_x|0\rangle}{\hbar\omega + E_n - E_0} - \frac{\langle 0|J_x|n\rangle\langle n|J_y|0\rangle}{\hbar\omega + E_0 - E_n}\right] \tag{2.11}$$

Now, finally, we can look at the DC $\omega \to 0$ limit that we're interested in[12]. We expand the denominators as

$$\frac{1}{\hbar\omega + E_n - E_0} \approx \frac{1}{E_n - E_0} - \frac{\hbar\omega}{(E_n - E_0)^2} + \mathcal{O}(\omega^2)\ldots$$

and similar for the other term. The first term looks divergent. Indeed, such divergences do arise for longitudinal conductivities and tell us something physical, often that momentum is conserved due to translational invariance so there can be no DC resistivity. However, in the present case of the Hall conductivity, there is no divergence because this term vanishes. This can be shown on general grounds from gauge invariance or,

[12]There is yet another subtlety that appears when computing the conductivity using the Kubo formula. In general, one should do these calculations at finite temperature rather than at $T = 0$. Then one should take the limit $\omega \to 0$ *before* taking the $T \to 0$ limit. In the present context we get lucky and the two limits commute, so working at $T = 0$ throughout does not give the wrong answer. But, in other contexts (and in particular in systems with a gapless mode) this is not the case and one should be more careful. I'm grateful to Steve Kivelson for an explanation on this point.

equivalently, from the conservation of the current. Alternatively – although somewhat weaker – it can quickly seen by rotational invariance which ensures that the expression should be invariant under $x \to y$ and $y \to -x$. We're then left only with a finite contribution in the limit $\omega \to 0$ given by

$$\sigma_{xy} = i\hbar \sum_{n \neq 0} \frac{\langle 0|J_y|n\rangle\langle n|J_x|0\rangle - \langle 0|J_x|n\rangle\langle n|J_y|0\rangle}{(E_n - E_0)^2} \tag{2.12}$$

This is the Kubo formula for Hall conductivity.

Before we proceed, I should quickly apologise for being sloppy: the operator that we called $\mathbf{J}$ in (2.8) is actually the current rather than the current density. This means that the right-hand-side of (2.12) should, strictly speaking, be multiplied by the spatial area of the sample. It was simpler to omit this in the above derivation to avoid clutter.

2.2.4 The Role of Topology

In this section, we describe a set-up in which we can see the deep connections between topology and the Hall conductivity. The set-up is closely related to the gauge-invariance argument that we saw in Section 2.2.2. However, we will consider the Hall system on a spatial torus $\mathbf{T}^2$. This can be viewed as a rectangle with opposite edges identified. We'll take the lengths of the sides to be L_x and L_y.

We thread a uniform magnetic field B through the torus. The first result we need is that B obeys the Dirac quantisation condition,

$$BL_xL_y = \frac{2\pi\hbar}{e}\, n \qquad n \in \mathbf{Z} \tag{2.13}$$

This quantisation arises for the same reason that we saw in Section 1.5.2 when discussing the Berry phase. However, it's an important result so here we give an alternative derivation.

We consider wavefunctions over the torus and ask: what periodicity requirements should we put on the wavefunction? The first guess is that we should insist that wavefunctions obey $\psi(x, y) = \psi(x + L_x, y) = \psi(x, y + L_y)$. But this turns out to be too restrictive when there is a magnetic flux through the torus. Instead, one has to work in patches; on the overlap between two different patches, wavefunctions must be related by a gauge transformation.

Operationally, there is a slightly simpler way to implement this. We introduce the magnetic translation operators,

$$T(\mathbf{d}) = e^{-i\mathbf{d}\cdot\mathbf{p}/\hbar} = e^{-i\mathbf{d}\cdot(i\nabla + e\mathbf{A}/\hbar)}$$

These operators translate a state $\psi(x, y)$ by position vector $\mathbf{d}$. The appropriate boundary conditions will be that when a state is translated around a cycle of the torus, it comes back to itself. So $T_x\psi(x, y) = \psi(x, y)$ and $T_y\psi(x, y) = \psi(x, y)$ where $T_x = T(\mathbf{d} = (L_x, 0))$ and $T_y = T(\mathbf{d} = (0, L_y))$.

It is clear from the expression above that the translation operators are not gauge invariant: they depend on our choice of $\mathbf{A}$. We'll choose Landau gauge $A_x = 0$ and $A_y = Bx$. With this choice, translations in the x direction are the same as those in the absence of a magnetic field, while translations in the y direction pick up an extra phase. If we take a state $\psi(x, y)$, translated around a cycle of the torus, it becomes

$$\begin{aligned} T_x\psi(x, y) &= \psi(x + L_x, y) = \psi(x, y) \\ T_y\psi(x, y) &= e^{-ieBL_y x/\hbar}\,\psi(x, y + L_y) = \psi(x, y) \end{aligned}$$

Notice that we can see explicitly in the last of these equations that the wavefunction is not periodic in the naive sense in the y direction: $\psi(x, y + L_y) \neq \psi(x, y)$. Instead, the two wavefunctions agree only up to a gauge transformation.

However, these equations are not consistent for any choice of B. This follows by comparing what happens if we translate around the x-cycle, followed by the y-cycle, or if we do these in the opposite order. We have

$$T_y T_x = e^{-ieBL_x L_y/\hbar}\, T_x T_y \tag{2.14}$$

Since both are required to give us back the same state, we must have

$$\frac{eBL_x L_y}{\hbar} \in 2\pi\mathbf{Z}$$

This is the Dirac quantisation condition (2.13).

There is an interesting story about solving for the wavefunctions of a free particle on a torus in the presence of a magnetic field. They are given by theta functions. We won't discuss them here.

Adding Flux

Now we're going to perturb the system. We do this by threading two fluxes, Φ_x and Φ_y through the x and y-cycles of the torus respectively. This means that the gauge potential becomes

$$A_x = \frac{\Phi_x}{L_x} \quad \text{and} \quad A_y = \frac{\Phi_y}{L_y} + Bx$$

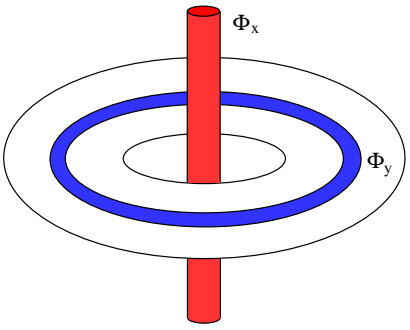


Figure 21:

This is the same kind of set-up that we discussed in Section 2.2.2; the only difference is that now the geometry allows us to introduce two fluxes instead of one. Just as in our previous discussion, the states of the quantum system are only sensitive to the non-integer part of Φ_i/Φ_0 where $\Phi_0 = 2\pi\hbar/e$ is the quantum of flux. In particular, if we increase either Φ_i from zero to Φ_0, then the spectrum of the quantum system must be invariant. However, just as in Section 2.2.2, the system can undergo spectral flow.

The addition of the fluxes adds an extra term to the Hamiltonian of the form (2.8),

$$\Delta H = -\sum_{i=x,y} \frac{J_i \Phi_i}{L_i}$$

We want to see how this affects the ground state of the system which we will denote as $|\psi_0\rangle$. (We called this $|0\rangle$ when deriving the Kubo formula, but we'll want to differentiate it soon and the expression $\frac{\partial 0}{\partial \Phi}$ just looks too odd!). We'll assume that the ground state is non-degenerate and that there is a gap to the first excited state. Then, to first order in perturbation theory, the ground state becomes

$$|\psi_0\rangle' = |\psi_0\rangle + \sum_{n \neq \psi_0} \frac{\langle n|\Delta H|\psi_0\rangle}{E_n - E_0} |n\rangle$$

Considering infinitesimal changes of Φ_i, we can write this as

$$|\frac{\partial \psi_0}{\partial \Phi_i}\rangle = -\frac{1}{L_i} \sum_{n \neq \psi_0} \frac{\langle n|J_i|\psi_0\rangle}{E_n - E_0} |n\rangle$$

But the right-hand-side is exactly the kind of expression that appeared in the Kubo formula (2.12). This means that, including the correct factors of the spatial area, we

can write the Hall conductivity as

$$\begin{aligned}\sigma_{xy} &= i\hbar L_x L_y \sum_{n\neq\psi_0} \frac{\langle\psi_0|J_y|n\rangle\langle n|J_x|\psi_0\rangle - \langle\psi_0|J_x|n\rangle\langle n|J_y|\psi_0\rangle}{(E_n - E_0)^2} \\ &= i\hbar\left[\langle\frac{\partial\psi_0}{\partial\Phi_y}|\frac{\partial\psi_0}{\partial\Phi_x}\rangle - \langle\frac{\partial\psi_0}{\partial\Phi_x}|\frac{\partial\psi_0}{\partial\Phi_y}\rangle\right] \\ &= i\hbar\left[\frac{\partial}{\partial\Phi_y}\langle\psi_0|\frac{\partial\psi_0}{\partial\Phi_x}\rangle - \frac{\partial}{\partial\Phi_x}\langle\psi_0|\frac{\partial\psi_0}{\partial\Phi_y}\rangle\right]\end{aligned}$$

As we now explain, this final way of writing the Hall conductivity provides a novel perspective on the integer quantum Hall effect.

Hall Conductivity and the Chern Number

The fluxes Φ_i appear as parameters in the perturbed Hamiltonian. As we discussed above, the spectrum of the Hamiltonian only depends on Φ_i mod Φ_0, which means that these parameters should be thought of as periodic: the space of the flux parameters is itself a torus, $\mathbf{T}^2_\Phi$, where the subscript is there to distinguish it from the spatial torus that we started with. We'll introduce dimensionless angular variables, θ_i to parameterise this torus,

$$\theta_i = \frac{2\pi\Phi_i}{\Phi_0} \qquad \text{with} \quad \theta_i \in [0, 2\pi)$$

As we discussed in Section 1.5, given a parameter space it is natural to consider the Berry phase that arises as the parameters are varied. This is described by the Berry connection which, in this case, lives over $\mathbf{T}^2_\Phi$. It is

$$\mathcal{A}_i(\Phi) = -i\langle\psi_0|\frac{\partial}{\partial\theta_i}|\psi_0\rangle$$

The field strength, or curvature, associated to the Berry connection is given by

$$\mathcal{F}_{xy} = \frac{\partial\mathcal{A}_x}{\partial\theta_y} - \frac{\partial\mathcal{A}_y}{\partial\theta_x} = -i\left[\frac{\partial}{\partial\theta_y}\langle\psi_0|\frac{\partial\psi_0}{\partial\theta_x}\rangle - \frac{\partial}{\partial\theta_x}\langle\psi_0|\frac{\partial\psi_0}{\partial\theta_y}\rangle\right]$$

This is precisely our expression for the Hall conductivity! We learn that, for the torus with fluxes, we can write

$$\sigma_{xy} = -\frac{e^2}{\hbar}\mathcal{F}_{xy}$$

This is a nice formula. But, so far, it doesn't explain why σ_{xy} is quantised. However, suppose that we average over all fluxes. In this case we integrate over the torus $\mathbf{T}^2_\Phi$ of

parameters to get

$$\sigma_{xy} = -\frac{e^2}{\hbar}\int_{\mathbf{T}^2_\Phi} \frac{d^2\theta}{(2\pi)^2}\ \mathcal{F}_{xy}$$

The integral of the curvature over $\mathbf{T}^2_\Phi$, is a number known as the *first Chern number*

$$C = \frac{1}{2\pi}\int_{\mathbf{T}^2_\Phi} d^2\theta\ \mathcal{F}_{xy}$$

Importantly, this is always an integer: $C \in \mathbf{Z}$. This follows from the same kind of argument that we made in Section 1.5 (or, alternatively, the kind of argument that we made at the beginning of this section on Dirac quantisation). The net result is that if we average over the fluxes, the Hall conductivity is necessarily quantised as

$$\sigma_{xy} = -\frac{e^2}{2\pi\hbar}\, C \tag{2.15}$$

This, of course, is the integer quantum Hall effect. The relationship between the Hall conductivity and the Chern number is usually referred to as the *TKNN invariant* (after Thouless, Kohomoto, Nightingale and den Nijs) although, strictly speaking, this name should be reserved for a very similar expression that we'll discuss in the next section.

2.3 Particles on a Lattice

We saw in the previous section that there is a deep relationship between the Hall conductivity and a certain topological quantity called the Chern number that is related to the Berry phase. Here we'll continue to explore this relationship but in the slightly different context of particles moving on a lattice. The kind of ideas that we will describe have had a resurgence in recent years when it was realised that they are the key to understanding the subject of *topological insulators*.

The advantage of looking at the particle on a lattice is that its momentum lies on a torus $\mathbf{T}^2$, known as the *Brillouin zone*. It is this torus that will allow us to find interesting topological features. Indeed, it will play a very similar role to the parameter space $\mathbf{T}^2_\Phi$ that we met in the previous section. We'll learn that one can define a Berry connection over the Brillouin zone and that the associated Chern number determines the Hall conductivity.

2.3.1 TKNN Invariants

We'll consider a particle moving on a rectangular lattice. The distance between lattice sites in the x-direction is a; the distance in the y-direction is b. Recall from earlier

courses that the energy spectrum of this system form bands. Within each band, states are labelled by lattice momentum which takes values in the Brillouin zone, parameterised by

$$-\frac{\pi}{a} < k_x \leq \frac{\pi}{a} \quad \text{and} \quad -\frac{\pi}{b} < k_y \leq \frac{\pi}{b} \tag{2.16}$$

The states with momenta at the edges of the Brillouin zone are identified. This means that the Brillouin zone is a torus $\mathbf{T}^2$ as promised. The wavefunctions in a given band can be written in Bloch form as

$$\psi_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}}\, u_{\mathbf{k}}(\mathbf{x}) \tag{2.17}$$

where $u_{\mathbf{k}}(\mathbf{x})$ is usually periodic on a unit cell so that $u_{\mathbf{k}}(\mathbf{x}+\mathbf{e}) = u_{\mathbf{k}}(\mathbf{x})$ with either $\mathbf{e} = (a, 0)$ or $\mathbf{e} = (0, b)$.

We're now in a position to describe the topology underlying the quantum Hall effect. The results below are very general: they don't rely on any specific Hamiltonian, but rather apply to any system that satisfies a few simple criteria.

- First, we will assume that the single particle spectrum decomposes into bands, with each band parameterised by a momentum label $\mathbf{k}$ which lives on a torus $\mathbf{T}^2$. This is obviously true for simple lattice models. As we explain in Section 2.3.3, it is also true (under certain assumptions) for particles moving in a lattice in the presence of a magnetic field where the torus in question is slightly different concept called a magnetic Brillouin zone. (In this case, the periodicity conditions on $u_{\mathbf{k}}$ are altered slightly but the formula we derive below still holds.)

- Second, we'll assume that the electrons are non-interacting. This means that we get the multi-particle spectrum simply by filling up the single-particle spectrum, subject to the Pauli exclusion principle.

- Finally, we'll assume that there is a gap between bands and that the Fermi energy E_F lies in one of these gaps. This means that all bands below E_F are completely filled while all bands above E_F are empty. In band theory, such a situation describes an *insulator*.

Whenever these three criteria are obeyed, one can assign an integer-valued topological invariant $C \in \mathbf{Z}$ to each band. The topology arises from the way the phase of the states winds as we move around the Brillouin zone $\mathbf{T}^2$. This is captured by a $U(1)$ Berry connection over $\mathbf{T}^2$, defined by

$$\mathcal{A}_i(\mathbf{k}) = -i\langle u_{\mathbf{k}}|\frac{\partial}{\partial k^i}|u_{\mathbf{k}}\rangle$$

There is one slight conceptual difference from the type of Berry connection we met previously. In Section 1.5, the connections lived on the space of parameters of the Hamiltonian; here the connection lives on the space of states itself. Nonetheless, it is simple to see that many of the basic properties that we met in Section 1.5 still hold. In particular, a change of phase of the states $|u_{\mathbf{k}}\rangle$ corresponds to a change of gauge of the Berry connection.

We can compute the field strength associated to $\mathcal{A}_i$. This is

$$\mathcal{F}_{xy} = \frac{\partial \mathcal{A}_x}{\partial k^y} - \frac{\partial \mathcal{A}_y}{\partial k^x} = -i\left\langle \frac{\partial u}{\partial k^y}\middle| \frac{\partial u}{\partial k^x}\right\rangle + i\left\langle \frac{\partial u}{\partial k^x}\middle| \frac{\partial u}{\partial k^y}\right\rangle \tag{2.18}$$

Once again, we can compute the first Chern number by integrating $\mathcal{F}$ over the Brillouin zone $\mathbf{T}^2$,

$$C = -\frac{1}{2\pi}\int_{\mathbf{T}^2} d^2k\ \mathcal{F}_{xy} \tag{2.19}$$

In the present context, it is usually referred to as the TKNN invariant[13]. As we've seen before, the Chern number is always an integer: $C \in \mathbf{Z}$. In this way, we can associate an integer C_α to each band α.

The Chern number once again has a beautiful physical manifestation: it is related to the Hall conductivity σ_{xy} of a non-interacting band insulator by

$$\sigma_{xy} = \frac{e^2}{2\pi\hbar}\sum_\alpha C_\alpha \tag{2.20}$$

where the sum is over all filled bands α and C_α is the Chern class associated to that band. This is the famous TKNN formula. It is, of course, the same formula (2.15) that we met previously, although the context here is rather different.

Let's now prove the TKNN formula. Our starting point is the Kubo formula (2.12). We previously wrote this in terms of multi-particle wavefunctions. If we're dealing with non-interacting particles, then these can be written as tensor products of single particle wavefunctions, each of which is labelled by the band α and momentum $\mathbf{k} \in \mathbf{T^2}$. The expression for the Hall conductivity becomes

$$\sigma_{xy} = i\hbar \sum_{E_\alpha < E_F < E_\beta} \int_{\mathbf{T}^2} \frac{d^2k}{(2\pi)^2}\ \frac{\langle u^\alpha_{\mathbf{k}}|J_y|u^\beta_{\mathbf{k}}\rangle\langle u^\beta_{\mathbf{k}}|J_x|u^\alpha_{\mathbf{k}}\rangle - \langle u^\alpha_{\mathbf{k}}|J_x|u^\beta_{\mathbf{k}}\rangle\langle u^\beta_{\mathbf{k}}|J_y|u^\alpha_{\mathbf{k}}\rangle}{(E_\beta(\mathbf{k}) - E_\alpha(\mathbf{k}))^2}$$

[13]As we mentioned in the previous section, the initials stand for Thouless, Kohomoto, Nightingale and den Nijs. The original paper is "*Quantized Hall Conductance in a Two-Dimensional Periodic Potential*", Phys. Rev. Lett. 49, 405 (1982).

where the index α runs over the filled bands and β runs over the unfilled bands. We note that this notation is a little lazy; there are really separate momentum integrals for each band and no reason that the states in the expression have the same momentum $\mathbf{k}$. Our lazy notation will save us from adding yet more annoying indices and not affect the result below.

To make progress, we need to define what we mean by the current $\mathbf{J}$. For a single, free particle in the continuum, the current carried by the particle was simply $\mathbf{J} = e\dot{\mathbf{x}}$ where the velocity operator is $\dot{\mathbf{x}} = (\mathbf{p} + e\mathbf{A})/m$. Here we'll use a more general definition. We first look at the Schrödinger equation acting on single-particle wavefunctions of Bloch form (2.17),

$$\begin{aligned} H|\psi_{\mathbf{k}}\rangle = E_{\mathbf{k}}|\psi_{\mathbf{k}}\rangle \quad &\Rightarrow \quad (e^{-i\mathbf{k}\cdot\mathbf{x}}He^{i\mathbf{k}\cdot\mathbf{x}})|u_{\mathbf{k}}\rangle = E_{\mathbf{k}}|u_{\mathbf{k}}\rangle \\ &\Rightarrow \quad \tilde{H}(\mathbf{k})|u_{\mathbf{k}}\rangle = E_{\mathbf{k}}|u_{\mathbf{k}}\rangle \quad \text{with} \quad \tilde{H}(\mathbf{k}) = e^{-i\mathbf{k}\cdot\mathbf{x}}He^{i\mathbf{k}\cdot\mathbf{x}} \end{aligned}$$

We then define the current in terms of the group velocity of the wavepackets,

$$\mathbf{J} = \frac{e}{\hbar}\frac{\partial \tilde{H}}{\partial \mathbf{k}}$$

Before proceeding, it's worth checking that coincides with our previous definition. In the continuum, the Hamiltonian was simply $H = (\mathbf{p} + e\mathbf{A})^2/2m$, which gives $\tilde{H} = (\mathbf{p} + \hbar\mathbf{k} + e\mathbf{A})^2/2m$ and the current due to a single particle is $\mathbf{J} = e\dot{\mathbf{x}}$ as expected.

From now on it's merely a question of doing the algebra. The Kubo formula becomes

$$\sigma_{xy} = \frac{ie^2}{\hbar}\sum_{E_\alpha<E_F<E_\beta}\int_{\mathbf{T}^2}\frac{d^2k}{(2\pi)^2}\,\frac{\langle u^\alpha_{\mathbf{k}}|\partial_y\tilde{H}|u^\beta_{\mathbf{k}}\rangle\langle u^\beta_{\mathbf{k}}|\partial_x\tilde{H}|u^\alpha_{\mathbf{k}}\rangle - \langle u^\alpha_{\mathbf{k}}|\partial_x\tilde{H}|u^\beta_{\mathbf{k}}\rangle\langle u^\beta_{\mathbf{k}}|\partial_y\tilde{H}|u^\alpha_{\mathbf{k}}\rangle}{(E_\beta(\mathbf{k}) - E_\alpha(\mathbf{k}))^2}$$

where ∂_x and ∂_y in this expression are derivatives with respect to momenta k_x and k_y respectively. We can then write

$$\begin{aligned} \langle u^\alpha_{\mathbf{k}}|\partial_i\tilde{H}|u^\beta_{\mathbf{k}}\rangle &= \langle u^\alpha_{\mathbf{k}}|\partial_i\left(\tilde{H}|u^\beta_{\mathbf{k}}\rangle\right) - \langle u^\alpha_{\mathbf{k}}|\tilde{H}|\partial_i u^\beta_{\mathbf{k}}\rangle \\ &= (E_\beta(\mathbf{k}) - E_\alpha(\mathbf{k}))\langle u^\alpha_{\mathbf{k}}|\partial_i u^\beta_{\mathbf{k}}\rangle \\ &= -(E_\beta(\mathbf{k}) - E_\alpha(\mathbf{k}))\langle \partial_i u^\alpha_{\mathbf{k}}|u^\beta_{\mathbf{k}}\rangle \end{aligned}$$

The missing term, proportional to $\partial_i E_\beta$, doesn't appear because α and β are necessarily distinct bands. Substituting this into the Kubo formula gives

$$\sigma_{xy} = \frac{ie^2}{\hbar}\sum_{E_\alpha<E_F<E_\beta}\int_{\mathbf{T}^2}\frac{d^2k}{(2\pi)^2}\,\langle\partial_y u^\alpha_{\mathbf{k}}|u^\beta_{\mathbf{k}}\rangle\langle u^\beta_{\mathbf{k}}|\partial_x u^\alpha_{\mathbf{k}}\rangle - \langle\partial_x u^\alpha_{\mathbf{k}}|u^\beta_{\mathbf{k}}\rangle\langle u^\beta_{\mathbf{k}}|\partial_y u^\alpha_{\mathbf{k}}\rangle$$

But now we can think of the sum over the unfilled bands as $\sum_\beta |u^\beta_{\mathbf{k}}\rangle\langle u^\beta_{\mathbf{k}}| = \mathbf{1} - \sum_\alpha |u^\alpha_{\mathbf{k}}\rangle\langle u^\alpha_{\mathbf{k}}|$. The second term vanishes by symmetry, so we're left with

$$\sigma_{xy} = \frac{ie^2}{\hbar}\sum_\alpha \int_{\mathbf{T}^2} \frac{d^2k}{(2\pi)^2}\ \langle \partial_y u^\alpha_{\mathbf{k}} | \partial_x u^\alpha_{\mathbf{k}}\rangle - \langle \partial_x u^\alpha_{\mathbf{k}} | \partial_y u^\alpha_{\mathbf{k}}\rangle$$

where now the sum is only over the filled bands α. Comparing to (2.18), we see that the Hall conductivity is indeed given by the sum of the Chern numbers of filled bands as promised,

$$\sigma_{xy} = -\frac{e^2}{2\pi\hbar}\sum_\alpha C_\alpha$$

The TKNN formula is the statement that the Hall conductivity is a topological invariant of the system. It's important because it goes some way to explaining the robustness of the integer quantum Hall effect. An integer, such as the Chern number C, can't change continuously. This means that if we deform our system in some way then, as long as we retain the assumptions that went into the derivation above, the Hall conductivity can't change: it's pinned at the integer value.

The existence of the TKNN formula is somewhat surprising. The right-hand side is simple and pure. In contrast, conductivities are usually thought of as something complicated and messy, depending on all the intricate details of a system. The essence of the TKNN formula, and indeed the quantum Hall effect itself, is that this is not the case: the Hall conductivity is topological.

2.3.2 The Chern Insulator

Let's look at an example. Perhaps surprisingly, the simplest examples of lattice models with non-vanishing Chern numbers don't involve any magnetic fields at all. Such lattice models with filled bands are sometimes called *Chern insulators*, to highlight the fact that they do something interesting — like give a Hall response — even though they are insulating states.

The simplest class of Chern insulators involve just two bands. The single-particle Hamiltonian written, written in momentum space, takes the general form

$$\tilde{H}(\mathbf{k}) = \vec{E}(\mathbf{k})\cdot\vec{\sigma} + \epsilon(\mathbf{k})\mathbf{1}$$

where $\mathbf{k}\in\mathbf{T}^2$ and $\vec{\sigma} = (\sigma_1,\sigma_2,\sigma_3)$ are the three Pauli matrices. The energies of the two states with momentum $\mathbf{k}$ are $\epsilon(\mathbf{k}) \pm |\vec{E}(\mathbf{k})|$. An insulator requires a gap between the upper and lower bands; we then fill the states of the lower band. An insulator can only occur when $\vec{E}(\mathbf{k}) \neq 0$ for all $\mathbf{k}$.

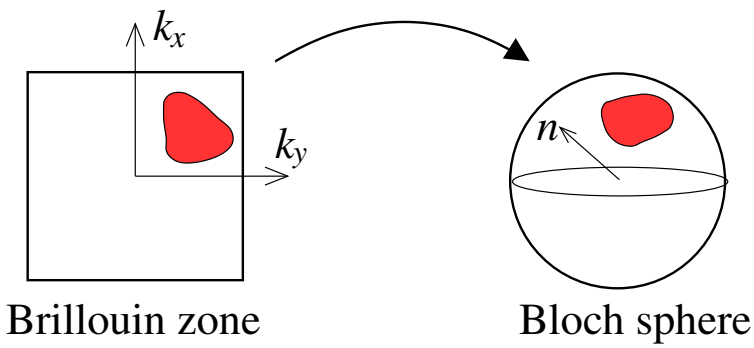


Figure 22: The map from Brillouin zone to Bloch sphere

For any such model, we can introduce a unit three-vector,

$$\vec{n}(\mathbf{k}) = \frac{\vec{E}(\mathbf{k})}{|\vec{E}(\mathbf{k})|}$$

Clearly $\vec{n}$ describes a point on a two-dimensional sphere $\mathbf{S}^2$. This is the *Bloch sphere*. As we move in the Brillouin zone, $\vec{n}(\mathbf{k})$ gives a map from $\mathbf{T}^2 \to \mathbf{S}^2$ as shown in the figure. This Chern number (2.19) for this system can be written in terms of $\vec{n}$ as

$$C = \frac{1}{4\pi}\int_{\mathbf{T}^2} d^2k\ \vec{n}\cdot\left(\frac{\partial\vec{n}}{\partial k_x}\times\frac{\partial\vec{n}}{\partial k_y}\right)$$

There is a particularly nice interpretation of this formula: it measures the area of the unit sphere (counted with sign) swept out as we vary $\mathbf{k}$ over $\mathbf{T}^2$. In other words, it counts how many times $\mathbf{T}^2$ wraps around $\mathbf{S}^2$.

Perhaps the simplest lattice model with a non-trivial Chern number arises on a square lattice, with the Hamiltonian in momentum space given by[14].

$$\tilde{H}(\mathbf{k}) = (\sin k_x)\sigma_1 + (\sin k_y)\sigma_2 + (m + \cos k_x + \cos k_y)\sigma_3$$

In the continuum limit, this becomes the Hamiltonian for a 2-component Dirac fermion in $d = 2+1$ dimensions. For this reason, this model is sometimes referred to as a *Dirac-Chern* insulator.

[14]This model was first constructed in Xiao-Liang Qi, Yong-Shi Wu and Shou-Cheng Zhang, "*Topological quantization of the spin Hall effect in two-dimensional paramagnetic semiconductors*", cond-mat/0505308. An earlier model of a quantum Hall effect without a magnetic field, involving a honeycomb lattice, was described by Duncan Haldane, "*Model for a Quantum Hall Effect without Landau Levels: Condensed Matter Realisation of the Parity Anomaly*", Phys. Rev. Lett. 61, 2015 (1988).

For general values of m, the system is an insulator with a gap between the bands. There are three exceptions: the gap closes and the two bands touch at $m = 0$ and at $m = \pm 2$. As m varies, the Chern number — and hence the Hall conductivity — remains constant as long as the gap doesn't close. A direct computation gives

$$C = \begin{cases} -1 & -2 < m < 0 \\ 1 & 0 < m < 2 \\ 0 & |m| > 2 \end{cases}$$

2.3.3 Particles on a Lattice in a Magnetic Field

So far, we've discussed the integer quantum Hall effect in lattice models but, perhaps surprisingly, we haven't explicitly introduced magnetic fields. In this section, we describe what happens when particles hop on a lattice in the presence of a magnetic field. As we will see, the physics is remarkably rich.

To orient ourselves, first consider a particle hopping on two-dimensional square lattice in the absence of a magnetic field. We'll denote the distance between adjacent lattice sites as a. We'll work in the tight-binding approximation, which means that the position eigenstates $|\mathbf{x}\rangle$ are restricted to the lattice sites $\mathbf{x} = a(m, n)$ with $m, n \in \mathbf{Z}$. The Hamiltonian is given by

$$H = -t \sum_{\mathbf{x}} \sum_{j=1,2} |\mathbf{x}\rangle\langle \mathbf{x} + \mathbf{e}_j| + \text{h.c.} \tag{2.21}$$

where $\mathbf{e}_1 = (a, 0)$ and $\mathbf{e}_2 = (0, a)$ are the basis vectors of the lattice and t is the hopping parameter. (Note: t is the standard name for this parameter; it's not to be confused with time!) The lattice momenta $\mathbf{k}$ lie in the Brillouin zone $\mathbf{T}^2$, parameterised by

$$-\frac{\pi}{a} < k_x \leq \frac{\pi}{a} \quad \text{and} \quad -\frac{\pi}{a} < k_y \leq \frac{\pi}{a} \tag{2.22}$$

Suppose that we further make the lattice finite in spatial extent, with size $L_1 \times L_2$. The momenta k_i are now quantised in units of $2\pi/L_i$. The total number of states in the Brillouin zone is then $(\frac{2\pi}{a}/\frac{2\pi}{L_1}) \times (\frac{2\pi}{a}/\frac{2\pi}{L_1}) = L_1 L_2/a^2$. This is the number of sites in the lattice which is indeed the expected number of states in the Hilbert space.

Let's now add a background magnetic field to the story. The first thing we need to do is alter the Hamiltonian. The way to do this is to introduce a gauge field $A_j(\mathbf{x})$ which lives on the links between the lattice sites. We take $A_1(\mathbf{x})$ to be the gauge field on the link to the right of point $\mathbf{x}$, and $A_2(\mathbf{x})$ to be the gauge field on the link above

point $\mathbf{x}$, as shown in the figure. The Hamiltonian is then given by

$$H = -t \sum_{\mathbf{x}} \sum_{j=1,2} |\mathbf{x}\rangle e^{-iaeA_j(\mathbf{x})/\hbar} \langle \mathbf{x} + \mathbf{e}_j| + \text{h.c.} \tag{2.23}$$

It might not be obvious that this is the correct way to incorporate a magnetic field. To gain some intuition, consider a particle which moves anti-clockwise around a plaquette. To leading order in t, it will pick up a phase $e^{-i\gamma}$, where

$$\begin{aligned} \gamma &= \frac{ea}{\hbar} \left(A_1(\mathbf{x}) + A_2(\mathbf{x} + \mathbf{e}_1) - A_1(\mathbf{x} + \mathbf{e}_2) - A_2(\mathbf{x}) \right) \\ &\approx \frac{ea^2}{\hbar} \left(\frac{\partial A_2}{\partial x^1} - \frac{\partial A_1}{\partial x^2} \right) = \frac{ea^2 B}{\hbar} \end{aligned}$$

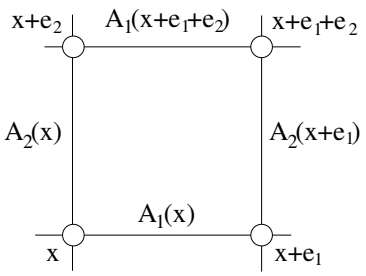


Figure 23:

where B is the magnetic field which passes through the plaquette. This expression is the same as the Aharonov-Bohm phase (1.49) for a particle moving around a flux $\Phi = Ba^2$.

Let's now restrict to a constant magnetic field. We can again work in Landau gauge,

$$A_1 = 0 \quad \text{and} \quad A_2 = Bx_1 \tag{2.24}$$

We want to understand the spectrum of the Hamiltonian (2.23) in this case and, in particular, what becomes of the Brillouin zone.

Magnetic Brillouin Zone

We saw above that the key to finding topology in a lattice system was the presence of the Brillouin zone $\mathbf{T}^2$. Yet it's not immediately obvious that the Brilliouin zone survives in the presence of the magnetic field. The existence of lattice momenta $\mathbf{k}$ are a consequence of the discrete translational invariance of the lattice. But, as usual, the choice of gauge breaks the explicit translational invariance of the Hamiltonian, even though we expect the underlying physics to be translational invariant.

In fact, we'll see that the interplay between lattice effects and magnetic effects leads to some rather surprising physics that is extraordinarily sensitive to the flux $\Phi = Ba^2$ that threads each plaquette. In particular, we can define a magnetic version of the Brillouin zone whenever Φ is a rational multiple of $\Phi_0 = 2\pi\hbar/e$,

$$\Phi = \frac{p}{q} \Phi_0 \tag{2.25}$$

with p and q integers which share no common divisor. We will see that in this situation the spectrum splits up into q different bands. Meanwhile, if Φ/Φ_0 is irrational, there are no distinct bands in the spectrum: instead it takes the form of a Cantor set! Nonetheless, as we vary Φ/Φ_0, the spectrum changes continuously. Needless to say, all of this is rather odd!

We start by defining the gauge invariant translation operators

$$T_j = \sum_{\mathbf{x}} |\mathbf{x}\rangle e^{-iea A_j(\mathbf{x})/\hbar} \langle \mathbf{x} + \mathbf{e}_j|$$

This shifts each state by one lattice site; T_1 moves us to the left and $T_1^\dagger$ to the right, while T_2 moves us down and $T_2^\dagger$ up, each time picking up the appropriate phase from the gauge field. Clearly we can write the Hamiltonian as

$$H = -t\left(\sum_{j=1,2} T_j + T_j^\dagger\right)$$

These translation operators do not commute. Instead it's simple to check that they obey the nice algebra

$$T_2\, T_1 = e^{ie\Phi/\hbar} T_1\, T_2 \tag{2.26}$$

This is the discrete version of the magnetic translation algebra (2.14). In the present context it means that $[T_i, H] \neq 0$ so, in the presence of a magnetic field, we don't get to label states by the naive lattice momenta which would be related to eigenvalues of T_i. This shouldn't be too surprising: the algebra (2.26) is a reflection of the fact that the gauge invariant momenta don't commute in a magnetic field, as we saw in (1.15).

However, we can construct closely related operators that do commute with T_j and, hence, with the Hamiltonian. These are defined by

$$\tilde{T}_j = \sum_{\mathbf{x}} |\mathbf{x}\rangle e^{-iea \tilde{A}_j(\mathbf{x})/\hbar} \langle \mathbf{x} + \mathbf{e}_j|$$

where the new gauge field $\tilde{A}_j$ is constructed to obey $\partial_k \tilde{A}_j = \partial_j A_k$. In Landau gauge, this means that we should take

$$\tilde{A}_1 = Bx_2 \quad \text{and} \quad \tilde{A}_2 = 0$$

When this holds, we have

$$[T_j, \tilde{T}_k] = [T_j^\dagger, \tilde{T}_k] = 0 \quad \Rightarrow \quad [H, \tilde{T}_j] = 0$$

These operators commute with the Hamiltonian, but do not themselves commute. Instead, they too obey the algebra (2.26).

$$\tilde{T}_2\,\tilde{T}_1 = e^{ie\Phi/\hbar}\tilde{T}_1\,\tilde{T}_2 \tag{2.27}$$

This means that we could label states of the Hamiltonian by eigenvalues of, say, $\tilde{T}_2$ but not simultaneously by eigenvalues of $\tilde{T}_1$. This isn't enough to construct a Brillouin zone.

At this point, we can see that something special happens when the flux is a rational multiple of Φ_0, as in (2.25). We can now build commuting operators by

$$[\tilde{T}_1^{n_1}, \tilde{T}_2^{n_2}] = 0 \quad \text{whenever} \quad \frac{p}{q} n_1 n_2 \in \mathbf{Z}$$

This means in particular that we can label energy eigenstates by their eigenvalue under $\tilde{T}_2$ and, simultaneously, their eigenvalue under $\tilde{T}_1^q$. We call these states $|\mathbf{k}\rangle$ with $\mathbf{k} = (k_1, k_2)$. They are Bloch-like eigenstates, satisfying

$$H|\mathbf{k}\rangle = E(\mathbf{k})|\mathbf{k}\rangle \quad \text{with} \quad T_1^q|\mathbf{k}\rangle = e^{iqk_1 a}|\mathbf{k}\rangle \quad \text{and} \quad T_2|\mathbf{k}\rangle = e^{ik_2 a}|\mathbf{k}\rangle$$

Note that the momenta k_i are again periodic, but now with the range

$$-\frac{\pi}{qa} < k_1 \leq \frac{\pi}{qa} \quad \text{and} \quad -\frac{\pi}{a} < k_2 \leq \frac{\pi}{a} \tag{2.28}$$

The momenta k_i parameterise the *magnetic Brillouin zone*. It is again a torus $\mathbf{T}^2$, but q times smaller than the original Brillouin zone (2.22). Correspondingly, if the lattice has size $L_1 \times L_2$, the number of states in each magnetic Brillouin zone is $L_1 L_2/qa^2$. This suggests that the spectrum decomposes into q bands, each with a different range of energies. For generic values of p and q, this is correct.

The algebraic structure above also tells us that any energy eigenvalue in a given band is q-fold degenerate. To see this, consider the state $\tilde{T}_1|\mathbf{k}\rangle$. Since $[H, \tilde{T}_1] = 0$, we know that this state has the same energy as $|\mathbf{k}\rangle$: $H\tilde{T}_1|\mathbf{k}\rangle = E(\mathbf{k})\tilde{T}_1|\mathbf{k}\rangle$. But, using (2.27), the k_y eigenvalue of this state is

$$\tilde{T}_2(\tilde{T}_1|\mathbf{k}\rangle) = e^{ie\Phi/\hbar}\tilde{T}_1\tilde{T}_2|\mathbf{k}\rangle = e^{i(2\pi p/q + k_2 a)}\tilde{T}_1|\mathbf{k}\rangle$$

We learn that $|\mathbf{k}\rangle$ has the same energy as $\tilde{T}_1|\mathbf{k}\rangle \sim |(k_1, k_2 + 2\pi p/qa)\rangle$.

The existence of a Brillouin zone (2.28) is the main result we need to discuss Hall conductivities in this model. However, given that we've come so far it seems silly not to carry on and describe what the spectrum of the Hamiltonian (2.23) looks like. Be warned, however, that the following subsection is a slight detour from our main goal.

Hofstadter Butterfly

To further understand the spectrum of the Hamiltonian (2.23), we'll have to roll up our sleeves and work directly with the Schrödinger equation. Let's first look in position space. We can write the most general wavefunction as a linear combination of the position eigenstates $|\mathbf{x}\rangle$,

$$|\psi\rangle = \sum_{\mathbf{x}} \psi(\mathbf{x})|\mathbf{x}\rangle$$

The Schrödinger equation $H|\psi\rangle = E|\psi\rangle$ then becomes an infinite system of coupled, discrete equations

$$\Big[\psi(\mathbf{x}+\mathbf{e}_1) + \psi(\mathbf{x}-\mathbf{e}_1) + e^{-i2\pi px^1/qa}\psi(\mathbf{x}+\mathbf{e}_2) + e^{+i2\pi px^1/qa}\psi(\mathbf{x}-\mathbf{e}_2)\Big] = -\frac{E}{t}\psi(\mathbf{x})$$

We want to find the possible energy eigenvalues E.

The way we usually solve these kinds of problems is by doing a Fourier transform of the wavefunction to work in momentum space, with

$$\tilde{\psi}(\mathbf{k}) = \sum_{\mathbf{x}} e^{-i\mathbf{k}\cdot\mathbf{x}}\,\psi(\mathbf{x}) \tag{2.29}$$

where, since $\mathbf{x}$ takes values on a discrete lattice, $\mathbf{k}$ takes values in the original Brillouin zone (2.22). In the absence of a magnetic field, modes with different momenta $\mathbf{k}$ decouple from each other. However, if you try the same thing in the presence of a magnetic field, you'll find that the modes with momentum $\mathbf{k} = (k_1, k_2)$ couple to modes with momentum $(k_1 + 2\pi p/qa, k_2)$. The reflects the fact that, as we have seen, the magnetic Brillouin zone (2.28) is q times smaller. For this reason, we instead split the wavefunction (2.29) into q different wavefunctions $\tilde{\psi}_r(\mathbf{k})$, with $r = 1, \ldots, q$ as

$$\tilde{\psi}_r(\mathbf{k}) = \sum_{\mathbf{x}} e^{-i(k_1+2\pi pr/qa,k_2)\cdot\mathbf{x}}\,\psi(\mathbf{x})$$

These contain the same information as (2.29), but now the argument $\mathbf{k}$ ranges over the magnetic Brillouin zone (2.28). Given the wavefunctions $\tilde{\psi}_r$, we can always reconstruct $\psi(x)$ by the inverse Fourier transform,

$$\psi(\mathbf{x}) = \sum_{r=1}^{q} \int_{-\pi/qa}^{+\pi/qa} \frac{dk_1}{2\pi} \int_{-\pi/a}^{+\pi/a} \frac{dk_2}{2\pi}\, e^{i\mathbf{k}\cdot\mathbf{x}}\,\tilde{\psi}_r(\mathbf{k})$$

In this way, we see that we have a q-component vector of wavefunctions, $\tilde{\psi}_r(\mathbf{k})$ living on the magnetic Brillouin zone.

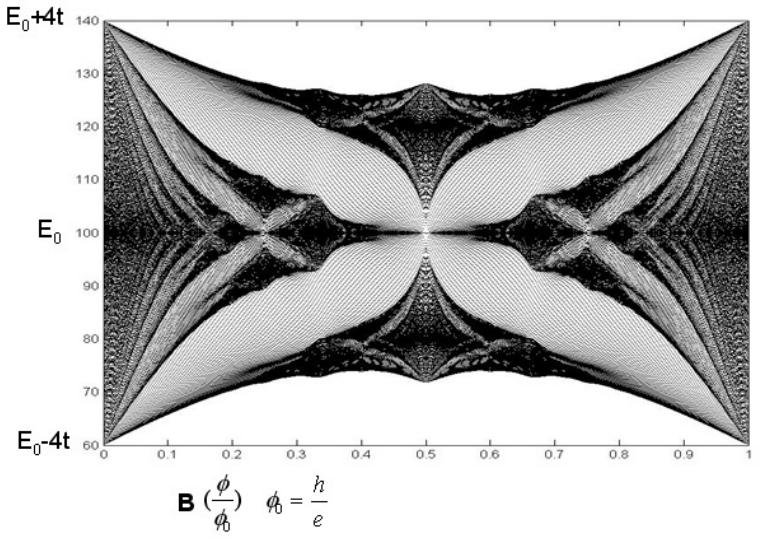


Figure 24: The Hofstadter butterfly

Taking the Fourier transform of the discrete Schrödinger equation in position space yields the following equation

$$2\cos\left(k_1 a + \frac{2\pi p r}{q}\right)\tilde{\psi}_r(\mathbf{k}) + e^{ik_2 a}\tilde{\psi}_{r+1}(\mathbf{k}) + e^{-ik_2 a}\tilde{\psi}_{r-1}(\mathbf{k}) = -\frac{E(\mathbf{k})}{t}\,\tilde{\psi}_r(\mathbf{k})$$

This is known as the *Harper equation*.

The Harper equation can be solved numerically. The resulting spectrum is quite wonderful. For rational values, $\Phi/\Phi_0 = p/q$, the spectrum indeed decomposes into q bands with gaps between them, as we anticipated above. Yet the spectrum also varies smoothly as we change Φ. Obviously if we change Φ/Φ_0 continuously it will pass through irrational values; when this happens the spectrum forms something like a Cantor set. The result is a beautiful fractal structure called the *Hofstadter butterfly*[15] shown in Figure 24. Here, a point is drawn in black if there is a state with that energy. Otherwise it is white. To get a sense of the structure, you could look at the specific values $\Phi/\Phi_0 = 1/q$, above which you should see q vertical bands of black.

[15]The spectrum was first solved numerically by Douglas Hofstadter in "*Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields*", Phys. Rev. B14, 2239 (1976). The picture of the butterfly was taken from Or Cohen's webpage http://phelafel.technion.ac.il/∼orcohen/butterfly.html where you can find a nice description of the techniques used to generate it.

TKNN Invariants for Particles on a Lattice in a Magnetic Field

Finally we reach our main goal: to compute the Hall conductivity of the lattice model for a particle in a background magnetic field. We can only do this for rational fluxes $\Phi = p\Phi_0/q$ for which there exists a magnetic Brillouin zone. In this case, we can use the TKNN formula (2.20), but with the Chern number, which used to be defined by integrating over the Brillouin zone, now arising by integrating over the magnetic Brillouin zone.

The computation of the Chern numbers is not so straightforward. (You can find it in the original paper of TKNN or, in more detail, in the book by Fradkin). Here we just state the answer. Even this is not totally straightforward.

First consider the r^{th} of the q bands. Then, to compute the Chern number, you are invited to solve the linear Diophantine equation

$$r = qs_r + pt_r$$

with $|t_r| \leq q/2$. The Chern number of the r^{th} band is given by

$$C_r = t_r - t_{r-1}$$

where $t_0 \equiv 0$. If the first r bands are filled, so that $E_r < E_F < E_{r+1}$, then the Hall conductivity is given by

$$\sigma_{xy} = \frac{e^2}{2\pi\hbar}\, t_r$$

It's helpful to look at some examples. First, when $\Phi = p\Phi_0$, there is only a single band and the Hall conductivity vanishes. A more complicated, illustrative example is given by $p/q = 11/7$. Here the solutions to the Diophantine equation are $(s_r, t_r) = (-3, 2), (5, -3), (2, -1), (-1, 1), (-4, 3), (4, -2), (1, 0)$. As we fill consecutive bands, the second number t_r in these pairs determines the Hall conductivity. We see that the Hall conductivity varies in an interesting way, sometimes negative and sometimes positive.