

Palestinian Advanced Physics School

Condensed Matter Physics

David Tong

Department of Applied Mathematics and Theoretical Physics

University of Cambridge

U.K.

d.tong@damtp.cam.ac.uk

ABSTRACT: These lectures provide an introduction to solid state physics.

Contents

0. Introduction	1
1. Electrons Moving in One Dimension	3
1.1 The Tight-Binding Model	3
1.2 Nearly Free Electrons	7
1.3 Bloch's Theorem in One Dimension	14
2. Lattices	19
2.1 Bravais Lattices	19
2.2 The Reciprocal Lattice	25
2.3 The Brillouin Zone	26
2.4 Band Structure	28
2.5 Bloch's Theorem	29
2.6 Nearly Free Electrons in Three Dimensions	30
3. Fermi Surfaces	36
3.1 Fermi Surfaces	36
3.2 Metals vs Insulators	37
4. Graphene	43
5. Dynamics of Bloch Electrons	48
5.1 Velocity	48
5.2 The Effective Mass	50
5.3 Semi-Classical Equation of Motion	51
5.4 Holes	53
5.5 Drude Model Again	55

0. Introduction

Condensed matter physics is the study of “stuff”, of how the wonderfully diverse properties of solids and liquids and other forms of matter can emerge from the simple laws that govern the underlying electrons and atoms. It provides answers to the most basic, childlike questions that a scientist can ask. Why is water wet? Why are metals shiny? How can we use non-Abelian quasi-particles to store quantum information? (Ok...that last one wasn't very childlike.) But all of these questions have their answers in the beautiful world of quantum mechanics.

As a physicist, if you want to understand how a something works then it's very tempting to think about the underlying constituents. This is the approach of reductionism. For any material, we know exactly what the underlying constituents are: atoms. Often in a material, some of the electrons get dislodged from the atom, so we should really be talking about electrons and ions. All materials in the world are described by some number N_e of electrons and some number N_i of ions. We understand perfectly the equation which describes these constituents. It is the Hamiltonian

$$H = \sum_{j=1}^{N_e} \frac{\mathbf{p}_j^2}{2m} + \sum_{\alpha=1}^{N_i} \frac{\mathbf{P}_\alpha^2}{2M_\alpha} + \frac{e^2}{4\pi\epsilon_0} \left[\sum_{j \neq k} \frac{1}{|\mathbf{r}_j - \mathbf{r}_k|} - \sum_{j,\alpha} \frac{Z_\alpha}{|\mathbf{r}_j - \mathbf{R}_\alpha|} + \sum_{\alpha \neq \beta} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} \right]$$

where

- Each electron has mass m and charge $-e$. The j^{th} electron has position \mathbf{r}_j and momentum \mathbf{p}_j with $j = 1, \dots, N_e$.
- The α^{th} ion has mass M_α , charge $+Z_\alpha e$, position \mathbf{R}_α and momentum \mathbf{P}_α with $\alpha = 1, \dots, N_i$ is the number of ions.

For completeness, we should also include the spin of the electron. But that's it. The equation above includes everything that we need to understand the world around us. All we need to do is solve it and we can go home.

That turns out to be hard. We can solve the equation above when $N_e = N_i = 1$. This is the hydrogen atom. We can solve it numerically when $N_e, N_i \sim 10$ or so. But to understand solids, we need to solve this when

$$N_e, N_i \sim 10^{24}$$

We have no hope at all of doing this! The naive reductionist approach is not very useful when thinking about how stuff works.

So how to proceed? We shouldn't just forget about atoms – or electrons and ions – completely. After all, they are the building blocks of everything. But we need to be more creative when thinking about our starting point. We need to come up with new ways to think about what's happening. We often do this by writing down *toy models*. These are starting points which capture some of the basic, obvious physics but where we still have some work to do to understand the less intuitive behaviour of materials. This is the approach that we will take in this course.

Condensed matter physics was originally called “solid state” physics, as it was focussed on how solids work. But as the subject expanded to include more and more complicated forms of matter its name evolved. In these lectures we will talk about some of the foundational ideas in solid state physics. In particular, we will be interested in a very basic question: how do electrons move inside a material? This explains why some materials conduct electricity very well, while others are insulators. But, along the way, we will touch upon some of the most important ideas in condensed matter physics.

1. Electrons Moving in One Dimension

In this chapter, we start our journey into the world of condensed matter physics. This is the study of the properties of “stuff”. Here, our interest lies in a particular and familiar kind of stuff: solids.

Solids are collections of tightly bound atoms. For most solids, these atoms arrange themselves in regular patterns on an underlying crystalline lattice. Some of the electrons of the atom then disassociate themselves from their parent atom and wander through the lattice environment. The properties of these electrons determine many of the properties of the solid, not least its ability to conduct electricity.

One might imagine that the electrons in a solid move in a fairly random fashion, as they bounce from one lattice site to another, like a ball in a pinball machine. However, as we will see, this is not at all the case: the more fluid nature of quantum particles allows them to glide through a regular lattice, almost unimpeded, with a distorted energy spectrum the only memory of the underlying lattice.

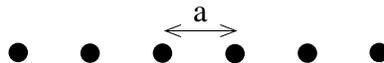
In this chapter, we will focus on understanding how the energy of an electron depends on its momentum when it moves in a lattice environment. The usual formula for kinetic energy, $E = \frac{1}{2}mv^2 = p^2/2m$, is one of the first things we learn in theoretical physics as children. As we will see, a lattice changes this in interesting ways, the consequences of which we will explore in chapter 3.

We begin with some particularly simple toy models which capture much of the relevant physics. These toy models describe an electron moving in a one-dimensional lattice. We’ll take what lessons we can from this before moving onto more realistic descriptions of electrons moving in higher dimensions.

1.1 The Tight-Binding Model

The tight-binding model is a caricature of electron motion in solid in which space is made discrete. The electron can sit only on the locations of atoms in the solid and has some small probability to hop to a neighbouring site due to quantum tunnelling.

To start with our “solid” consists of a one-dimensional lattice of atoms. This is described by N points arranged along a line, each separated by distance a .



Consider a single electron moving on this lattice. We will assume that the electron can only sit on a given lattice point; it's not allowed to roam between lattice points. This is supposed to mimic the idea that electrons are bound to the atoms in a lattice and goes by the name of the *tight-binding approximation*. (We'll see exactly what we're neglecting in this approximation later.)

When the electron sits on the n^{th} atom, we denote the quantum state as $|n\rangle$. These states are considered orthogonal to each other, so

$$\langle n|m\rangle = \delta_{nm}$$

Clearly the total Hilbert space has dimension N , and is spanned by $|n\rangle$ with $n = 1, \dots, N$.

What kind of Hamiltonian will govern the dynamics of this electron? If the electron just remains on a given atom, an appropriate Hamiltonian would be

$$H_0 = E_0 \sum_n |n\rangle\langle n|$$

Each of the position states $|n\rangle$ is an energy eigenstate of H_0 with energy E_0 . The electrons governed by this Hamiltonian don't move. This Hamiltonian is boring.

To make things more interesting, we need to include the possibility that the electron can tunnel from one site to another. How to do this? Well, the Hamiltonian governs time evolution. In some small time increment of time Δt , a state evolves as

$$|\psi\rangle \mapsto |\psi\rangle - \frac{i\Delta t}{\hbar} H |\psi\rangle + \mathcal{O}(\Delta t^2)$$

This means that if we want the possibility for the electron to hop from one site to another, we should include in the Hamiltonian a term of the form $|m\rangle\langle n|$ which takes an electron at site n and moves it to an electron at site m .

There is one last ingredient that we want to feed into our model: locality. We don't want electrons to disappear and reappear many thousands of lattice spacings down the line. We want our model to describe electrons hopping from one atom to neighbouring atoms. This motivates our final form of the Hamiltonian,

$$H = E_0 \sum_n |n\rangle\langle n| - t \sum_n \left(|n\rangle\langle n+1| + |n+1\rangle\langle n| \right) \quad (1.1)$$

First a comment on notation: the parameter t is called the *hopping parameter*. It is not time; it is simply a number which determines the probability that a particle will hop

to a neighbouring site. (More precisely, the ratio t^2/E_0^2 will determine the probability of hopping.) It's annoying notation, but unfortunately t is the canonical name for this hopping parameter so it's best we get used to it now.

Now back to the physics encoded in H . We've chosen a Hamiltonian that only includes hopping terms between neighbouring sites. This is the simplest choice; we will describe more general choices later. Moreover, the probability of hopping to the left is the same as the probability of hopping to the right. This is required because H must be a Hermitian operator.

There's one final issue that we have to address before solving for the spectrum of H : what happens at the edges? Again, there are a number of different possibilities but none of the choices affect the physics that we're interested in here. The simplest option is simply to declare that the lattice is periodic. This is best achieved by introducing a new state $|N + 1\rangle$, which sits to the right of $|N\rangle$, and is identified with $|N + 1\rangle \equiv |1\rangle$.

Solving the Tight-Binding Model

Let's now solve for the energy eigenstates of the Hamiltonian (1.1). A general state can be expanded as

$$|\psi\rangle = \sum_m \psi_m |m\rangle$$

with $\psi_n \in \mathbf{C}$. Substituting this into the Schrödinger equation gives

$$H|\psi\rangle = E|\psi\rangle \quad \Rightarrow \quad E_0 \sum_m \psi_m |m\rangle - t \left(\sum_m \psi_{m+1} |m\rangle + \psi_m |m + 1\rangle \right) = E \sum_n \psi_n |n\rangle$$

If we now take the overlap with a given state $\langle n|$, we get the set of linear equations for the coefficients ψ_n

$$\langle n|H|\psi\rangle = E\langle n|\psi\rangle \quad \Rightarrow \quad E_0\psi_n - t(\psi_{n+1} + \psi_{n-1}) = E\psi_n \quad (1.2)$$

These kind of equations arise fairly often in physics. They are solved by the ansatz

$$\psi_n = e^{ikna} \quad (1.3)$$

Or, if we want to ensure that the wavefunction is normalised, $\psi_n = e^{ikna}/\sqrt{N}$. The exponent k is called the *wavenumber*. The quantity $p = \hbar k$ plays a role similar to momentum in our discrete model; we will discuss the ways in which it is like momentum in Section 1.3. We'll also often be lazy and refer to k as *momentum*.

The wavenumber has a number of properties. First, the set of solutions remain the same if we shift $k \rightarrow k + 2\pi/a$ so the wavenumber takes values in

$$k \in \left[-\frac{\pi}{a}, +\frac{\pi}{a}\right) \quad (1.4)$$

This range of k is given the fancy name *Brillouin zone*. We'll see why this is a useful concept that deserves its own name in Section 2.

There is also a condition on the allowed values of k coming from the requirement of periodicity. We want $\psi_{N+1} = \psi_1$, which means that $e^{ikNa} = 1$. This requires that k is quantised in units of $2\pi/aN$. In other words, within the Brillouin zone (1.4) there are exactly N quantum states of the form (1.3). But that's what we expect as it's the dimension of our Hilbert space; the states (1.3) form a different basis.

States of the form (1.3) have the property that

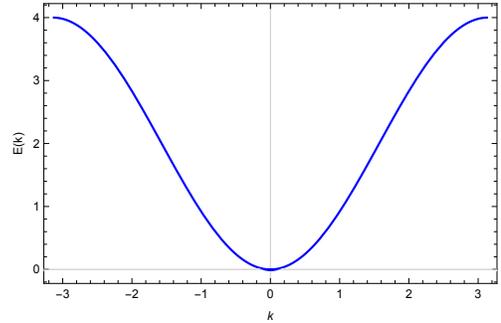
$$\psi_{n\pm 1} = e^{\pm ika} \psi_n$$

This immediately ensures that equation (1.2) is solved for any value of k , with the energy eigenvalue

$$E(k) = E_0 - 2t \cos(ka) \quad (1.5)$$

The spectrum is shown in the figure for $t > 0$.

(The plot was made with $a = t = 1$ and $E_0 = 2$.) The states with $k > 0$ describe electrons which move to the right; those with $k < 0$ describe electrons moving to the left.



There is a wealth of physics hiding in this simple result, and much of the following sections will be fleshing out these ideas. Here we highlight a few pertinent points

- The electrons do not like to sit still. The eigenstates $|n\rangle$ of the original Hamiltonian H_0 were localised in space. One might naively think that adding a tiny hopping parameter t would result in eigenstates that were spread over a few sites. But this is wrong. Instead, all energy eigenstates are spread throughout the whole lattice. Arbitrarily small local interactions result in completely delocalised energy eigenstates.
- The energy eigenstates of H_0 were completely degenerate. Adding the hopping term lifts this degeneracy. Instead, the eigenstates are labelled by the wavevector k and have energies (1.5) that lie in a range

$$E_0 - 2t \leq E(k) \leq E_0 + 2t$$

This range of energies is referred to a *band* and the difference between the maximum and minimum energy (which is $4t$ in this case) is called the *band width*. In our simple model, we have just a single energy band. In subsequent models, we will see multiple bands emerging.

- For suitably small momentum, $k \ll \pi/a$, we can Taylor expand the energy (1.5) as

$$E(k) \approx (E_0 - 2t) + ta^2k^2$$

Up to a constant, this takes the same form as a free particle moving in the continuum,

$$E_{\text{free}} = \frac{\hbar^2 k^2}{2m^*} \tag{1.6}$$

This is telling us that low energy, low momentum particles are unaware that they are moving on an underlying lattice. Instead, they act as if they are moving along a continuous line with *effective mass* $m^* = \hbar^2/2ta^2$. Notice that in this model the effective mass has nothing to do with the physical mass of the electron; it is inherited from properties of the lattice.

- There is a cute reciprocity between the properties of momentum and position. We know from our first course on quantum mechanics that if space is made finite — for example, a particle in a box, or a particle moving on a circle — then momentum becomes discrete. We also saw this above as the periodic boundary conditions enforced the wavenumber to be quantised in units of $2\pi/Na$.

However, our tight-binding model also exhibits the converse phenomenon: when we make space discrete, momentum becomes periodic: it has to lie in the Brillouin zone (1.4). More generally, discreteness is the Fourier transform of compactness.

1.2 Nearly Free Electrons

The tight-binding model is an extreme cartoon of the real physics in which space is discrete; electrons are stuck on atomic sites with a non-vanishing probability to hop to a neighbouring site. In this section we present another cartoon that is designed to capture the opposite extreme.

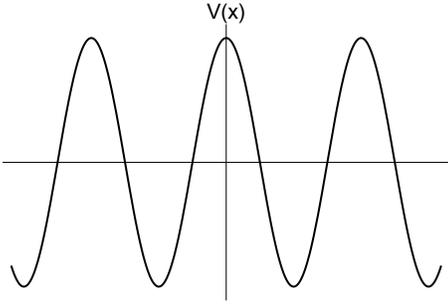


Figure 2: A periodic sine wave.

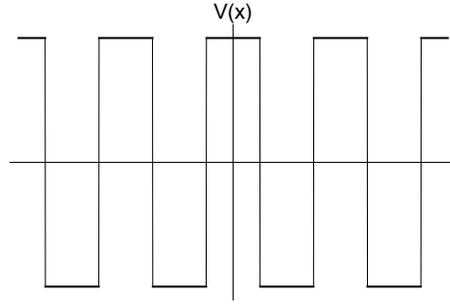


Figure 3: A periodic square wave.

We will assume that our electron is free to move anywhere along the line, parameterised by the position x . To mimic the underlying lattice, we add a weak, periodic potential $V(x)$. This means that we consider the Hamiltonian

$$H = \frac{p^2}{2m} + V(x)$$

where $p = -i\hbar d/dx$ is the usual momentum operator. The periodicity of the potential means that it satisfies

$$V(x + a) = V(x) \tag{1.7}$$

For example, the potential could take the form of a sine wave, or a square wave as shown in the figure, or it could be an infinite series of delta functions. For much of our discussion we won't need the exact form of the potential.

To avoid discussing edge effects, it's again useful to consider the particle moving on a circle \mathbf{S}^1 of length (circumference) L . This is compatible with the periodicity requirement (1.7) only if

$$\frac{L}{a} = N \in \mathbf{Z}$$

The integer N plays the role of the number of atoms in the lattice.

In the absence of the potential, the eigenstates are the familiar plane waves $|k\rangle$, labelled by the momentum $p = \hbar k$. Because we are on a circle, the wavenumber of k is quantised in units of $2\pi/L$. The associated wavefunctions are

$$\psi_k(x) = \langle x|k\rangle = \frac{1}{\sqrt{L}} e^{ikx} \tag{1.8}$$

These states are orthonormal, with

$$\langle k|k'\rangle = \frac{1}{L} \int dx e^{i(k'-k)x} = \delta_{k,k'} \tag{1.9}$$

(Recall that we are living on a circle, so the momenta k are discrete and the Kronecker delta is the appropriate thing to put on the right-hand side.) Meanwhile, the energy of a free particle is given by

$$E_0(k) = \frac{\hbar^2 k^2}{2m} \quad (1.10)$$

Our goal is to understand how the presence of the potential $V(x)$ affects this energy spectrum. To do this, we work perturbatively. However, perturbation theory in the present situation is a little more subtle than usual. Let's see why.

Perturbation Theory

Recall that the first thing we usually do in perturbation theory is decide whether we have non-degenerate or degenerate energy eigenstates. Which do we have in the present case? Well, all states are trivially degenerate because the energy of a free particle moving to the right is the same as the energy of a free particle moving to the left: $E_0(k) = E_0(-k)$. But the fact that the two states $|k\rangle$ and $|-k\rangle$ have the same energy does not necessarily mean that we have to use degenerate perturbation theory. This is only true if the perturbation causes the two states to mix.

To see what happens we will need to compute matrix elements $\langle k|V|k'\rangle$. The key bit of physics is the statement that the potential is periodic (1.7). This ensures that it can be Fourier expanded

$$V(x) = \sum_{n \in \mathbf{Z}} V_n e^{2\pi i n x / a} \quad \text{with} \quad V_n = V_{-n}^*$$

where the Fourier coefficients follow from the inverse transformation

$$V_n = \frac{1}{a} \int_0^a dx V(x) e^{-2\pi i n x / a}$$

The matrix elements are then given by

$$\langle k|V|k'\rangle = \frac{1}{L} \int dx \sum_{n \in \mathbf{Z}} V_n e^{i(k' - k + 2\pi n / a)x} = \sum_{n \in \mathbf{Z}} V_n \delta_{k - k', 2\pi n / a} \quad (1.11)$$

We see that we get mixing only when

$$k = k' + \frac{2\pi n}{a}$$

for some integer n . In particular, we get mixing between degenerate states $|k\rangle$ and $|-k\rangle$ only when

$$k = \frac{\pi n}{a}$$

for some n . The first time that this happens is when $k = \pi/a$. But we've seen this value of momentum before: it is the edge of the Brillouin zone (1.4). This is the first hint that the tight-binding model and nearly free electron model share some common features.

With this background, let's now try to sketch the basic features of the energy spectrum as a function of k .

Low Momentum: With low momentum $|k| \ll \pi/a$, there is no mixing between states at leading order in perturbation theory (and very little mixing at higher order). In this regime we can use our standard results from non-degenerate perturbation theory. Expanding the energy to second order, we have

$$E(k) = \frac{\hbar^2 k^2}{2m} + \langle k|V|k\rangle + \sum_{k' \neq k} \frac{|\langle k|V|k'\rangle|^2}{E_0(k) - E_0(k')} + \dots \quad (1.12)$$

From (1.11), we know that the first order correction is $\langle k|V|k\rangle = V_0$, and so just gives a constant shift to the energy, independent of k . Meanwhile, the second order term only gets contributions from $|k'\rangle = |k + 2\pi n/a\rangle$ for some n . When $|k| \ll \pi/a$, these corrections are small. We learn that, for small momenta, the particle moves as if unaffected by the potential. Intuitively, the de Broglie wavelength $2\pi/k$ of the particle much greater than the wavelength a of the potential, and the particle just glides over it unimpeded.

The formula (1.12) only holds for low momenta. Indeed, we see that the formula knows about its own failings, because if we attempt to use it when $k = n\pi/a$ for some n , the the numerator $\langle k|V|-k\rangle$ is finite while the denominator becomes zero. Whenever perturbation theory diverges in this manner it's because we're doing something wrong. In this case it's because we should be working with degenerate perturbation theory.

At the Edge of the Brillouin Zone: Let's consider the momentum eigenstates which sit right at the edge of the Brillouin zone, $k = \pi/a$, or at integer multiples

$$k = \frac{n\pi}{a}$$

As we've seen, these are the values which mix due to the potential perturbation and we must work with degenerate perturbation theory.

Let's recall the basics of degenerate perturbation theory. We focus on the subsector of the Hilbert space formed by the two degenerate states, in our case $|k\rangle$ and $|k'\rangle = |-k\rangle$. To leading order in perturbation theory, the new energy eigenstates will be some linear

combination of these original states $\alpha|k\rangle + \beta|k'\rangle$. We would like to figure out what choice of α and β will diagonalise the new Hamiltonian. There will be two such choices since there must, at the end of the day, remain two energy eigenstates. To determine the correct choice of these coefficients, we write the Schrödinger equation, restricted to this subsector, in matrix form

$$\begin{pmatrix} \langle k|H|k\rangle & \langle k|H|k'\rangle \\ \langle k'|H|k\rangle & \langle k'|H|k'\rangle \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (1.13)$$

We've computed the individual matrix elements above: using the fact that the states $|k\rangle$ are orthonormal (1.9), the unperturbed energy (1.10) and the potential matrix elements (1.11), our eigenvalue equation becomes

$$\begin{pmatrix} E_0(k) + V_0 & V_n \\ V_n^* & E_0(k') + V_0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (1.14)$$

where, for the value $k = -k' = n\pi/a$ of interest, $E_0(k) = E_0(k') = n^2\hbar^2\pi^2/2ma^2$. It's simple to determine the eigenvalues E of this matrix: they are given by the roots of the quadratic equation

$$(E_0(k) + V_0 - E)^2 - |V_n|^2 = 0 \quad \Rightarrow \quad E = \frac{\hbar^2 n^2\pi^2}{2m a^2} + V_0 \pm |V_n| \quad (1.15)$$

This is important. We see that a gap opens up in the spectrum at the values $k = \pm n\pi/a$. The size of the gap is proportional to $2|V_n|$.

It's simple to understand what's going on here. Consider the simple potential

$$V = 2V_1 \cos\left(\frac{2\pi x}{a}\right)$$

which gives rise to a gap only at $k = \pm\pi/a$. The eigenvectors of the matrix are $(\alpha, \beta) = (1, -1)$ and $(\alpha, \beta) = (-1, 1)$, corresponding to the wavefunctions

$$\begin{aligned} \psi_+(x) &= \langle x|(|k\rangle + |-k\rangle) \sim \cos\left(\frac{\pi x}{a}\right) \\ \psi_-(x) &= \langle x|(|k\rangle - |-k\rangle) \sim \sin\left(\frac{\pi x}{a}\right) \end{aligned}$$

The density of electrons is proportional to $|\psi_{\pm}|^2$. Plotting these densities on top of the potential, we see that ψ_+ describes electrons that are gathered around the peaks of the potential, while ψ_- describes electrons gathered around the minima. It is no surprise that the energy of ψ_+ is higher than that of ψ_- .

Close to the Edge of the Brillouin Zone: It's more subtle to look at the electron close to the edge of the Brillouin zone. Consider an electron with momentum

$$k = \frac{n\pi}{a} + \delta$$

for some small δ . As we've seen, the potential causes plane wave states to mix only if their wavenumbers differ by some multiple of $2\pi/a$. This means that $|k\rangle = |n\pi/a + \delta\rangle$ will mix with $|k'\rangle = |-n\pi/a + \delta\rangle$. These states don't quite have the same kinetic energy, but they have very *nearly* the same kinetic energy. And, as we will see, the perturbation due to the potential V will mean that these states still mix strongly.

To see this mixing, we need once again to solve the eigenvalue equation (1.13) or, equivalently, (1.14). The eigenvalues are given by solutions to the quadratic equation

$$\left(E_0(k) + V_0 - E\right)\left(E_0(k') + V_0 - E\right) - |V_n|^2 = 0 \quad (1.16)$$

The only difference from our previous discussion is that $E(k)$ and $E(k')$ are now given by

$$E(k) = \frac{\hbar^2}{2m} \left(\frac{n\pi}{a} + \delta\right)^2 \quad \text{and} \quad E(k') = \frac{\hbar^2}{2m} \left(\frac{n\pi}{a} - \delta\right)^2$$

and the quadratic equation (1.16) becomes

$$\left(\frac{\hbar^2}{2m} \left(\frac{n^2\pi^2}{a^2} + \delta^2\right) + V_0 - E\right)^2 - \left(\frac{\hbar^2}{2m} \frac{2n\pi\delta}{a}\right)^2 - |V_n|^2 = 0$$

This equation has two solutions, $E = E_{\pm}$, given by

$$E_{\pm} = \frac{\hbar^2}{2m} \left(\frac{n^2\pi^2}{a^2} + \delta^2\right) + V_0 \pm \sqrt{|V_n|^2 + \left(\frac{\hbar^2}{2m} \frac{2n\pi\delta}{a}\right)^2}$$

We're ultimately interested in this expression when δ is small, where we anticipate that the effect of mixing will be important. But, as a sanity check, let's first expand it in the opposite regime, when we're far from the edge of the Brillouin zone and δ is large compared to the gap V_n . In this case, a little bit of algebra shows that the eigenvalues can be written as

$$E_{\pm} = E_0(n\pi/a \pm \delta) + V_0 \pm \frac{|V_n|^2}{E_0(n\pi/a + \delta) - E_0(n\pi/a - \delta)}$$

But this coincides with the the expression that we got from second-order, non-degenerate perturbation theory (1.12). (Or, more precisely, because we have kept just a single mixing term in our discussion above we get just a single term in the sum in (1.12); for some choice of potentials, keeping further terms may be important.)

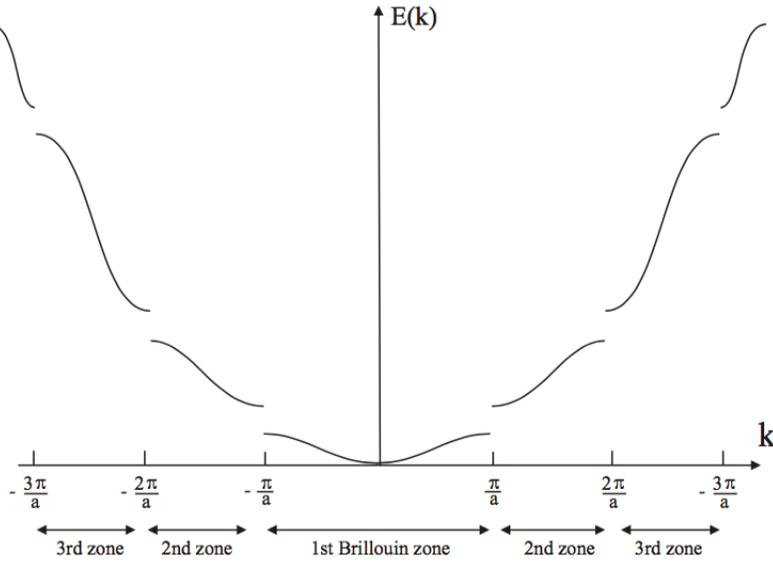


Figure 4: Energy dispersion for the free electron model.

Our real interest is what happens close to the edge of the Brillouin zone when δ is small compared to the gap V_n . In this case we can expand the square-root to give

$$E_{\pm} \approx \frac{\hbar^2}{2m} \frac{n^2 \pi^2}{a^2} + V_0 \pm |V_n| + \frac{\hbar^2}{2m} \left(1 \pm \frac{1}{|V_n|} \frac{n^2 \hbar^2 \pi^2}{ma^2} \right) \delta^2$$

The first collection of terms coincide with the energy at the edge of the Brillouin zone (1.15), as indeed it must. For us, the important new point is in the second term which tells us that as we approach the gaps, the energy is quadratic in the momentum δ .

Band Structure

We now have all we need to sketch the rough form of the energy spectrum $E(k)$. The original quadratic spectrum is deformed with a number of striking features:

- For small momenta, $k \ll \pi/a$, the spectrum remains roughly unchanged.
- The energy spectrum splits into distinct bands, with gaps arising at $k = n\pi/a$ with $n \in \mathbf{Z}$. The size of these gaps is given by $2|V_n|$, where V_n is the appropriate Fourier mode of the potential.

The region of momentum space corresponding to the n^{th} energy band is called the n^{th} Brillouin zone. However, we usually call the 1st Brillouin zone simply *the* Brillouin zone.

- As we approach the edge of a band, the spectrum is quadratic. In particular, $dE/dk \rightarrow 0$ at the end of a band.

The relationship $E(k)$ between energy and momentum is usually called the *dispersion relation*. In the present case, it is best summarised in a figure.

Note that the spectrum within the first Brillouin zone $|k| \leq \pi/a$, looks very similar to what we saw in the tight-binding model. In particular, both models have a gap at the edge of the Brillouin zone $k = \pm\pi/a$ which is a feature that will persist as we explore more complicated situations. The qualitative differences in the two models show arises because the tight-binding model has a finite number of states, all contained in the first Brillouin zone, while the nearly-free electron model has an infinite number of states which continue for $|k| > \pi/a$.

1.3 Bloch's Theorem in One Dimension

In both models described above, we ended up labelling states by momentum $\hbar k$. It's worth pausing to ask: why did we do this? And how should we think of k ?

Before we get to this, let's back up and ask an even more basic question: why do we label the states of a free particle by momentum? Here, the answer is because momentum is conserved. In the quantum theory, this means that the momentum operator commutes with the Hamiltonian: $[p, H] = 0$, so that we can simultaneously label states by both energy and momentum. Ultimately, Noether's theorem tells us that this conservation law arises because of translational invariance of the system.

Now let's look at our system with a lattice. We no longer have translational invariance. Correspondingly, in the nearly-free electron model, $[p, H] \neq 0$. Hopefully this now makes our original question sharper: why do we get to label states by k !?

While we don't have full, continuous translational invariance, both the models that we discussed do have a discrete version of translational invariance

$$x \rightarrow x + a$$

As we now show, this is sufficient to ensure that we can label states by something very similar to "momentum". However, the values of this momentum are restricted. This result is known as *Bloch's Theorem*. Here we prove the theorem for our one-dimensional system; we will revisit it in Section 2.5 in higher dimensions.

The Translation Operator

For concreteness, let's work with continuous space where states are described by a wavefunction $\psi(x)$. (There is a simple generalisation to discrete situations such as the tight-binding model that we describe below.) We introduce the translation operator T_l as

$$T_l \psi(x) = \psi(x + l)$$

First note that T_l is a unitary operator. To see this, we just need to look at the overlap

$$\begin{aligned} \langle \phi | T_l | \psi \rangle &= \int dx \phi(x)^* T_l \psi(x) = \int dx \phi(x)^* \psi(x + l) \\ &= \int dx \phi(x - l)^* \psi(x) = \int dx [T_{-l} \phi(x)]^* \psi(x) \end{aligned}$$

where, in the step to the second line, we've simply shifted the origin. This tells us that $T_l^\dagger = T_{-l}$. But clearly $T_l^{-1} = T_{-l}$ as well, so $T_l^\dagger = T_l^{-1}$ and the translation operator is unitary as claimed.

Next note that the set of translation operators form an Abelian group,

$$T_{l_1} T_{l_2} = T_{l_1 + l_2} \tag{1.17}$$

with $[T_{l_1}, T_{l_2}] = 0$.

The translation operator is a close cousin of the familiar momentum operator

$$p = -i\hbar \frac{d}{dx}$$

The relationship between the two is as follows: the unitary translation operator is the exponentiation of the Hermitian momentum operator

$$T_l = e^{ilp/\hbar}$$

To see this, we expand the exponent and observe that $T_l \psi(x) = \psi(x + l)$ is just a compact way of expressing the Taylor expansion of a function

$$\begin{aligned} T_l \psi(x) &= \left(1 + \frac{ilp}{\hbar} + \frac{1}{2} \left(\frac{ilp}{\hbar} \right)^2 + \dots \right) \psi(x) \\ &= \left(1 + l \frac{d}{dx} + \frac{l^2}{2} \frac{d^2}{dx^2} + \dots \right) \psi(x) = \psi(x + l) \end{aligned}$$

We say that the momentum operator is the “generator” of infinitesimal translations.

A quantum system is said to be invariant under translations by l if

$$[H, T_l] = 0 \tag{1.18}$$

Phrased in this way, we can describe both continuous translational symmetry and discrete translational symmetry. A system has continuous translational invariance if (1.18) holds for all l . In this case, we may equivalently say that $[p, H] = 0$. Alternatively, a system may have discrete translational invariance if (1.18) holds only when l is an integer multiple of the lattice spacing a . Now p does not commute with H .

Let's look at the case of discrete symmetry. Now we can't simultaneously diagonalise p and H , but we can simultaneously diagonalise T_a and H . In other words, energy eigenstates can be labelled by the eigenvalues of T_a . But T_a is a unitary operator and its eigenvalues are simply a phase, $e^{i\theta}$ for some θ . Moreover, we want the eigenvalues to respect the group structure (1.17). This is achieved if the eigenstates are labelled by some k , such that

$$T_a \psi_k(x) = \psi_k(x + a) = e^{ika} \psi_k(x)$$

Now comes the rub. Because the eigenvalue is a phase, there is an arbitrariness in this labelling: states labelled by k have the same eigenvalue under T_a as states labelled by $k + 2\pi/a$. To remedy this, we will simply require that k lies in the range

$$k \in \left[-\frac{\pi}{a}, \frac{\pi}{a} \right) \tag{1.19}$$

We recognise this as the first Brillouin zone.

This, then, is the essence of physics on a lattice. We can still label states by k , but it now lies in a finite range. Note that we can approximate a system with continuous translational symmetry by taking a arbitrarily small; in this limit we get the usual result $k \in \mathbf{R}$.

This discussion leads us directly to:

Bloch's Theorem in One Dimension: In a periodic potential, $V(x) = V(x + a)$, all energy eigenstates can be written as

$$\psi_k(x) = e^{ikx} u_k(x)$$

where $u_k(x) = u_k(x + a)$ is a periodic function and k lies in the Brillouin zone (1.19).

Proof: We take ψ_k to be an eigenstate of the translation operator T_a , so that $\psi_k(x + a) = e^{ika} \psi_k(x)$. Then $u_k(x + a) = e^{-ik(x+a)} \psi_k(x + a) = e^{-ikx} \psi_k(x) = u_k(x)$. \square

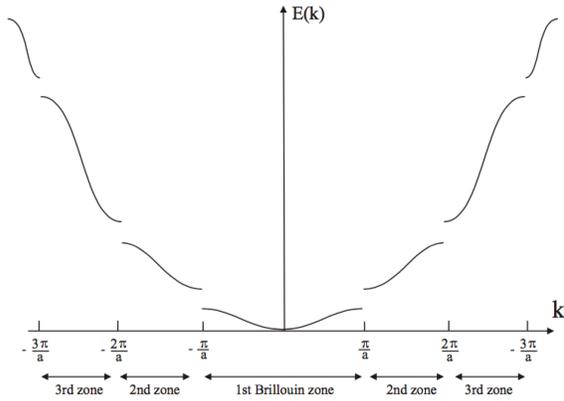


Figure 5: The extended zone scheme.

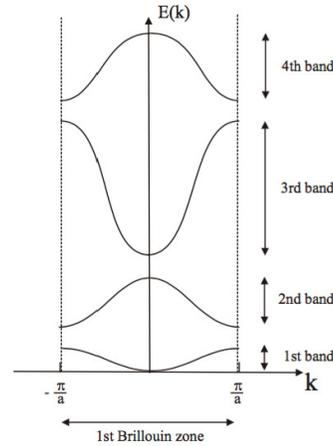


Figure 6: The reduced zone scheme.

Bloch’s theorem is rather surprising. One might think that the presence of a periodic potential would dramatically alter the energy eigenstates, perhaps localising them in some region of space. Bloch’s theorem is telling us that this doesn’t happen: instead the plane wave states e^{ikx} are altered only by a periodic function $u(x)$, sometimes referred to as a *Bloch function*, and the fact that the wavenumber is restricted to the first Brillouin zone.

Finally, note that we’ve couched the above discussion in terms of wavefunctions $\psi(x)$, but everything works equally well for the tight-binding model with the translation operator defined by $T_a|n\rangle = |n + 1\rangle$.

Crystal Momentum

The quantity $p = \hbar k$ is the quantity that replaces momentum in the presence of a lattice. It is called the *crystal momentum*. Note, however, that it doesn’t have the simple interpretation of “mass \times velocity”. (We will describe how to compute the velocity of a particle in terms of the crystal momentum in Section 5.1.)

Crystal momentum is conserved. This becomes particularly important when we consider multiple particles moving in a lattice and their interactions. This, of course, sounds the same as the usual story of momentum. Except there’s a twist: crystal momentum is conserved only mod $2\pi/a$. It is perfectly possible for two particles to collide in a lattice environment and their final crystal momentum to differ from their initial crystal momentum by some multiple of $2\pi/a$. Roughly speaking, the lattice absorbs the excess momentum.

This motivates us to re-think how we draw the energy spectrum. Those parts of the spectrum that lie outside the first Brillouin zone should really be viewed as having the same crystal momentum. To show this, we draw the energy spectrum as a multi-valued function of $k \in [-\pi/a, \pi/a)$. The spectrum that we previously saw in Figure 4 then looks like

The original way of drawing the spectrum is known as the *extended zone scheme*. The new way is known as the *reduced zone scheme*. Both have their uses. Note that edges of the Brillouin zone are identified: $k = \pi/a$ is the same as $k = -\pi/a$. In other words, the Brillouin zone is topologically a circle.

In the reduced zone scheme, states are labelled by both $k \in [-\pi/a, \pi/a)$ and an integer $n = 1, 2, \dots$ which tells us which band we are talking about.

2. Lattices

The ideas that we described above all go over to higher dimensions. The key difference is that lattices in higher dimensions are somewhat more complicated than a row of points. In this section, we introduce the terminology needed to describe different kinds of lattices. In Section 2.4, we'll return to look at what happens to electrons moving in these lattice environments.

2.1 Bravais Lattices

The simplest kind of lattice is called a *Bravais lattice*. This is a periodic array of points defined by integer sums of linearly independent basis vectors \mathbf{a}_i . In two-dimensions, a Bravais lattice Λ is defined by

$$\Lambda = \{\mathbf{r} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 \text{ , } n_i \in \mathbf{Z}\}$$

An obvious example is the square lattice shown to the right. We will see further examples shortly.

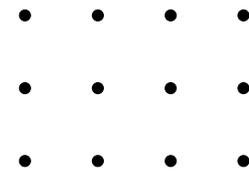


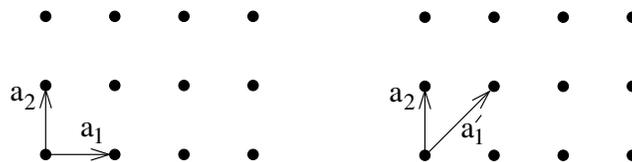
Figure 7:

In three dimensions, a Bravais lattice is defined by

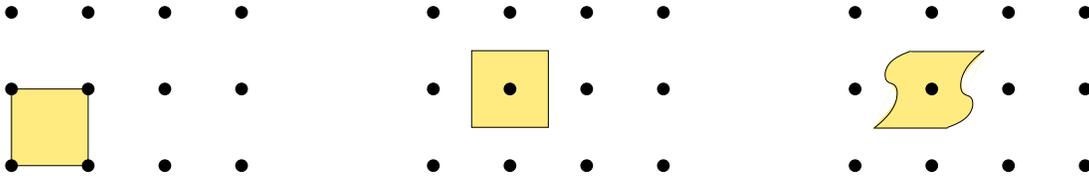
$$\Lambda = \{\mathbf{r} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \text{ , } n_i \in \mathbf{Z}\}$$

These lattices have the property that any point looks just the same as any other point. In mathematics, such an object would simply be called a lattice. Here we add the word *Bravais* to distinguish these from more general kinds of lattices that we will meet shortly.

The basis vectors \mathbf{a}_i are called *primitive lattice vectors*. They are not unique. As an example, look at the 2-dimensional square lattice below. We could choose basis vectors $(\mathbf{a}_1, \mathbf{a}_2)$ or $(\mathbf{a}'_1, \mathbf{a}_2)$. Both will do the job.



A *primitive unit cell* is a region of space which, when translated by the primitive lattice vectors \mathbf{a}_i , tessellates the space. This means that the cells fit together, without overlapping and without leaving any gaps. These primitive unit cells are not unique. As an example, let's look again at the 2-dimensional square lattice. Each of the three possibilities shown below is a good unit cell.



Each primitive unit cell contains a single lattice point. This is obvious in the second and third examples above. In the first example, there are four lattice points associated to the corners of the primitive unit cell, but each is shared by four other cells. Counting these as a 1/4 each, we see that there is again just a single lattice point in the primitive unit cell.

Although the primitive unit cells are not unique, each has the same volume. It is given by

$$V = |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)| \quad (2.1)$$

Because each primitive unit cell is associated to a single lattice point, $V = 1/n$ where n is the density of lattice points.

Note finally that the primitive unit cell need not have the full symmetry of the lattice. For example, the third possible unit cell shown above for the square lattice is not invariant under 90° rotations.

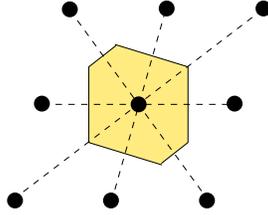
For any lattice, there is a canonical choice of primitive unit cell that does inherit the symmetry of the underlying lattice. This is called the *Wigner-Seitz cell*, Γ . (It sometimes goes by the name of the *Voronoi cell*.) Pick a lattice point which we choose to be at the origin. The Wigner-Seitz cell is defined to be the region of space around such that the origin is the closest lattice point. In equations,

$$\Gamma = \{ \mathbf{x} : |\mathbf{x}| < |\mathbf{x} - \mathbf{r}| \quad \forall \mathbf{r} \in \Lambda \text{ s.t. } \mathbf{r} \neq 0 \}$$

The Wigner-Seitz cells for square and triangular lattices are given by



There is a simple way to construct the Wigner-Seitz cell. Draw lines from the origin to all other lattice points. For each of these lines, construct the perpendicular bi-sectors; these are lines in 2d and planes in 3d. The Wigner-Seitz cell is the inner area bounded by these bi-sectors. Here's another example.



Examples of Bravais Lattices in 2d

Let's look at some examples. In two dimensions, a Bravais lattice is defined by two non-parallel vectors \mathbf{a}_1 and \mathbf{a}_2 , with angle $\theta \neq 0$ between them. However, some of these lattices are more special than others. For example, when $|\mathbf{a}_1| = |\mathbf{a}_2|$ and $\theta = \pi/2$, the lattice is square and enjoys an extra rotational symmetry.

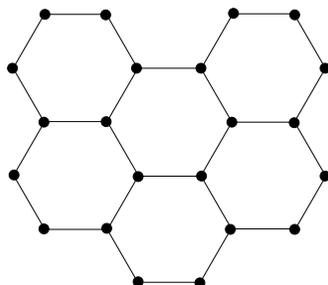
We will consider two Bravais lattices to be equivalent if they share the same symmetry group. With this definition, there are five possible Bravais lattices in two dimensions. They are

- **Square:** $|\mathbf{a}_1| = |\mathbf{a}_2|$ and $\theta = \pi/2$. It has four-fold rotation symmetry and reflection symmetry.
- **Triangular:** $|\mathbf{a}_1| = |\mathbf{a}_2|$ and $\theta = \pi/3$ or $\theta = 2\pi/3$. This is also sometimes called a hexagonal lattice. It has six-fold rotation symmetry.
- **Rectangular:** $|\mathbf{a}_1| \neq |\mathbf{a}_2|$ and $\theta = \pi/3$ or $\theta = 2\pi/3$. This has reflection symmetry.
- **Centred Rectangular:** $|\mathbf{a}_1| \neq |\mathbf{a}_2|$ and $\theta \neq \pi/2$, but the primitive basis vectors should obey $(2\mathbf{a}_2 - \mathbf{a}_1) \cdot \mathbf{a}_1 = 0$. This means that the lattice looks like a rectangle with an extra point in the middle.
- **Oblique:** $|\mathbf{a}_1| \neq |\mathbf{a}_2|$ and nothing special. This contains all other cases.

The square, triangular and oblique lattices were shown on the previous page where we also drew their Wigner-Seitz cells.

Not all Lattices are Bravais

Not all lattices of interest are Bravais lattices. One particularly important lattice in two dimensions has the shape of a honeycomb and is shown below.



This lattice describes a material called *graphene* that we will describe in more detail in Section 4. The lattice is not Bravais because not all points are the same. To see this, consider a single hexagon from the lattice as drawn below.

Each of the red points is the same: each has a neighbour directly to the left of them, and two neighbours diagonally to the right. But the white points are different. Each of them has a neighbour directly to the right, and two neighbours diagonally to the left.

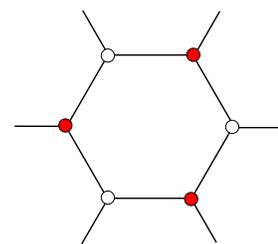


Figure 8:

Lattices like this are best thought by decomposing them into groups of atoms, where some element of each group sits on the vertices of a Bravais lattice. For the honeycomb lattice, we can consider the group of atoms \circ — \bullet . The red vertices form a triangular lattice, with primitive lattice vectors

$$\mathbf{a}_1 = \frac{\sqrt{3}a}{2}(\sqrt{3}, 1) \quad , \quad \mathbf{a}_2 = \frac{\sqrt{3}a}{2}(\sqrt{3}, -1)$$

Meanwhile, each red vertex is accompanied by a white vertex which is displaced by

$$\mathbf{d} = (-a, 0)$$

This way we build our honeycomb lattice.

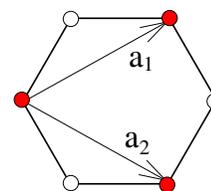


Figure 9:

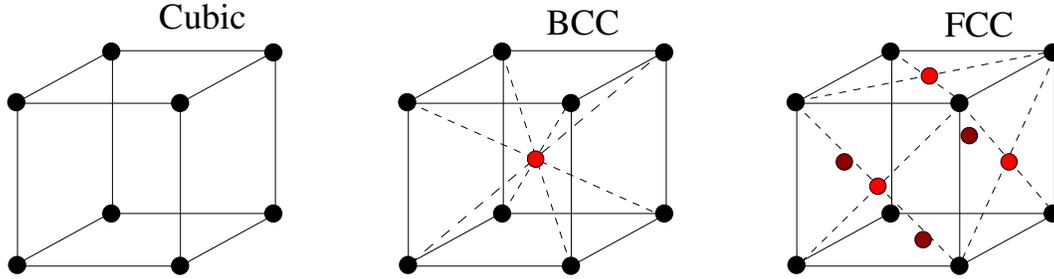


Figure 10: Three Bravais lattices. The different coloured atoms are there in an attempt to make the diagrams less confusing; they do not denote different types of atoms.

This kind of construction generalises. We can describe any lattice as a repeating group of atoms, where each group sits on an underlying Bravais lattice Λ . Each atom in the group is displaced from the vertex of the Bravais lattice by a vector \mathbf{d}_i . Each group of atoms, labelled by their positions \mathbf{d}_i is called the *basis*. For example, for the honeycomb lattice we chose the basis $\mathbf{d}_1 = 0$ for red atoms and $\mathbf{d}_2 = \mathbf{d}$ for white atoms, since the red atoms sat at the positions of the underlying triangular lattice. In general there's no requirement that any atom sits on the vertex of the underlying Bravais lattice. The whole lattice is then described by the union of the Bravais lattice and the basis, $\cup_i \{\Lambda + \mathbf{d}_i\}$.

Examples of Bravais Lattices in 3d

It turns out that there are 14 different Bravais lattices in three dimensions. Fortunately we won't need all of them. In fact, we will describe only the three that arise most frequently in Nature. These are:

- **Cubic:** This is the simplest lattice. The primitive lattice vectors are aligned with the Euclidean axes

$$\mathbf{a}_1 = a\hat{\mathbf{x}} \quad , \quad \mathbf{a}_2 = a\hat{\mathbf{y}} \quad , \quad \mathbf{a}_3 = a\hat{\mathbf{z}}$$

And the primitive cell looks has volume $V = a^3$. The Wigner-Seitz cell is also a cube, centered around one of the lattice points.

- **Body Centered Cubic (BCC):** This is a cubic lattice, with an extra point placed at the centre of each cube. We could take the primitive lattice vectors to be

$$\mathbf{a}_1 = a\hat{\mathbf{x}} \quad , \quad \mathbf{a}_2 = a\hat{\mathbf{y}} \quad , \quad \mathbf{a}_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})$$

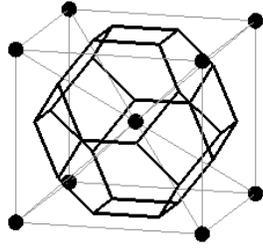


Figure 11: Wigner-Seitz cell for BCC

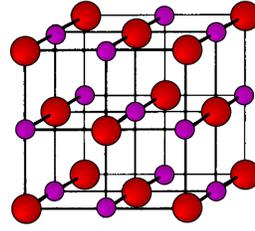


Figure 12: Salt.

However, a more symmetric choice is

$$\mathbf{a}_1 = \frac{a}{2}(-\hat{x} + \hat{y} + \hat{z}) \quad , \quad \mathbf{a}_2 = \frac{a}{2}(\hat{x} - \hat{y} + \hat{z}) \quad , \quad \mathbf{a}_3 = \frac{a}{2}(\hat{x} + \hat{y} - \hat{z})$$

The primitive unit cell has volume $V = a^3/2$.

The BCC lattice can also be thought of as a cubic lattice, with a basis of two atoms with $\mathbf{d}_1 = 0$ and $\mathbf{d}_2 = \frac{a}{2}(\hat{x} + \hat{y} + \hat{z})$. However, this doesn't affect the fact that the BCC lattice is itself Bravais.

The Alkali metals (*Li, Na, K, Rb, Cs*) all have a BCC structure, as do the Vanadium group (*V, Nb, Ta*) and Chromium group (*Cr, Mo, W*) and Iron (*Fe*). In each case, the lattice constant is roughly $a \approx 3$ to $6 \times 10^{-10} m$.

- **Face Centered Cubic (FCC):** This is again built from the cubic lattice, now with an extra point added to the centre of each face. The primitive lattice vectors are

$$\mathbf{a}_1 = \frac{a}{2}(\hat{y} + \hat{z}) \quad , \quad \mathbf{a}_2 = \frac{a}{2}(\hat{x} + \hat{z}) \quad , \quad \mathbf{a}_3 = \frac{a}{2}(\hat{x} + \hat{y})$$

The primitive unit cell has volume $V = a^3/4$.

The FCC lattice can also be thought of as a cubic lattice, now with a basis of four atoms sitting at $\mathbf{d}_1 = 0$, $\mathbf{d}_2 = \frac{a}{2}(\hat{x} + \hat{y})$, $\mathbf{d}_3 = \frac{a}{2}(\hat{x} + \hat{z})$ and $\mathbf{d}_4 = \frac{a}{2}(\hat{y} + \hat{z})$. Nonetheless, it is also a Bravais lattice in its own right.

Examples of FCC structures include several of the Alkaline earth metals (*Be, Ca, Sr*), many of the transition metals (*Sc, Ni, Pd, Pt, Rh, Ir, Cu, Ag, Au*) and the Noble gases (*Ne, Ar, Kr, Xe*) when in solid form, again with $a \approx 3$ to $6 \times 10^{-10} m$ in each case.

The Wigner-Seitz cells for the BCC and FCC lattices are polyhedra, sitting inside the cube. For example, the Wigner-Seitz cell for the BCC lattice is shown in the right-hand figure.

Examples of non-Bravais Lattices in 3d

As in the 2d examples above, we can describe non-Bravais crystals in terms of a basis of atoms sitting on an underlying Bravais lattice. Here are two particularly simple examples.

Diamond is made up of two, interlaced FCC lattices, with carbon atoms sitting at the basis points $\mathbf{d}_1 = 0$ and $\mathbf{d}_2 = \frac{a}{4}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})$. Silicon and germanium also adopt this structure.

Another example is salt ($NaCl$). Here, the basic structure is a cubic lattice, but with Na and Cl atoms sitting at alternate sites. It's best to think of this as two, interlaced FCC lattices, but shifted differently from diamond. The basis consists of a Na atom at $\mathbf{d} = 0$ and a Cl atom at $\mathbf{d}_2 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})$. This basis then sits on top of an FCC lattice.

2.2 The Reciprocal Lattice

Given a Bravais lattice Λ , defined by primitive vectors \mathbf{a}_i , the *reciprocal lattice* Λ^* is defined by the set of points

$$\Lambda^* = \{\mathbf{k} = \sum_i n_i \mathbf{b}_i, n_i \in \mathbf{Z}\}$$

where the new primitive vectors \mathbf{b}_i obey

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi \delta_{ij} \tag{2.2}$$

Λ^* is sometimes referred to as the *dual lattice*. In three dimensions, we can simply construct the lattice vectors \mathbf{b}_i by

$$\mathbf{b}_i = \frac{2\pi}{V} \frac{1}{2} \epsilon_{ijk} \mathbf{a}_j \times \mathbf{a}_k$$

where V is the volume of unit cell of Λ (2.1). We can also invert this relation to get

$$\mathbf{a}_i = \frac{2\pi}{V^*} \frac{1}{2} \epsilon_{ijk} \mathbf{b}_j \times \mathbf{b}_k$$

where $V^* = |\mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3)| = (2\pi)^3/V$ is the volume of Γ^* , the unit cell of Λ^* . Note that this shows that the reciprocal of the reciprocal lattice gives you back the original.

The condition (2.2) can also be stated as the requirement that

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 1 \quad \forall \mathbf{r} \in \Lambda, \mathbf{k} \in \Lambda^* \quad (2.3)$$

which provides an alternative definition of the reciprocal lattice.

Here are some examples:

- The cubic lattice has $\mathbf{a}_1 = a\hat{\mathbf{x}}$, $\mathbf{a}_2 = a\hat{\mathbf{y}}$ and $\mathbf{a}_3 = a\hat{\mathbf{z}}$. The reciprocal lattice is also cubic, with primitive vectors $\mathbf{b}_1 = (2\pi/a)\hat{\mathbf{x}}$, $\mathbf{b}_2 = (2\pi/a)\hat{\mathbf{y}}$ and $\mathbf{b}_3 = (2\pi/a)\hat{\mathbf{z}}$
- The BCC lattice has $\mathbf{a}_1 = \frac{a}{2}(-\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}})$, $\mathbf{a}_2 = \frac{a}{2}(\hat{\mathbf{x}} - \hat{\mathbf{y}} + \hat{\mathbf{z}})$ and $\mathbf{a}_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}})$. The reciprocal lattice vectors are $\mathbf{b}_1 = (2\pi/a)(\hat{\mathbf{y}} + \hat{\mathbf{z}})$, $\mathbf{b}_2 = (2\pi/a)(\hat{\mathbf{x}} + \hat{\mathbf{z}})$ and $\mathbf{b}_3 = (2\pi/a)(\hat{\mathbf{x}} + \hat{\mathbf{y}})$. But we've seen these before: they are the lattice vectors for a FCC lattice with the sides of the cubic cell of length $4\pi/a$.

We see that the reciprocal of a BCC lattice is an FCC lattice and vice versa.

2.3 The Brillouin Zone

The Wigner-Seitz cell of the reciprocal lattice is called the *Brillouin zone*.

We already saw the concept of the Brillouin zone in our one-dimensional lattice. Let's check that this coincides with the definition given above. The one-dimensional lattice is defined by a single number, a , which determines the lattice spacing. The Wigner-Seitz cell is defined as those points which lie closer to the origin than any other lattice point, namely $r \in [-a/2, a/2)$. The reciprocal lattice is defined by (2.2) which, in this context, gives the lattice spacing $b = 2\pi/a$. The Wigner-Seitz cell of this reciprocal lattice consists of those points which lie between $[-b/2, b/2) = [-\pi/a, \pi/a)$. This coincides with what we called the Brillouin zone in Section 1.

The Brillouin zone is also called the *first Brillouin zone*. As it is the Wigner-Seitz cell, it is defined as all points in reciprocal space that are closest to a given lattice point, say the origin. The n^{th} Brillouin zone is defined as all points in reciprocal space that are n^{th} closest to the origin. All these higher Brillouin zones have the same volume as the first.

We can construct the Brillouin zone boundaries by drawing the perpendicular bisectors between the origin and each other point in Λ^* . The region enclosing the origin is the first Brillouin zone. The region you can reach by crossing just a single bisector is the second Brillouin zone, and so on. In fact, this definition generalises the Brillouin zone beyond the simple Bravais lattices.

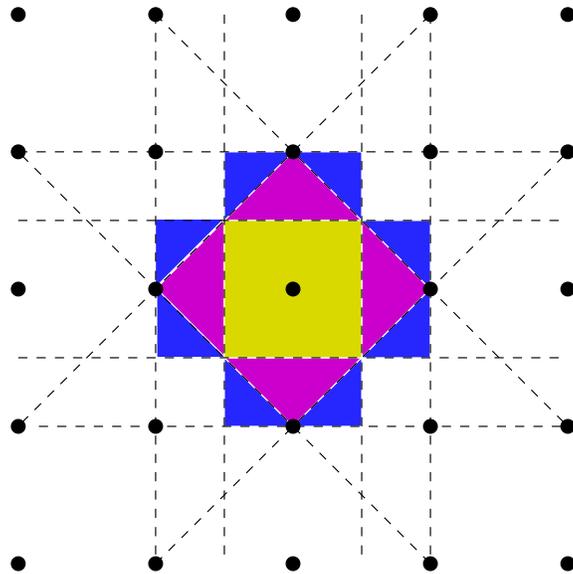


Figure 13: The Brillouin zones for a 2d square lattice. The first is shown in yellow, the second in pink, the third in blue.

As an example, consider the square lattice in 2d. The reciprocal lattice is also square. The first few Brillouin zones on this square lattice are shown in Figure 13.

For the one-dimensional lattice that we looked at in Section 1, we saw that the conserved momentum lies within the first Brillouin zone. This will also be true in higher dimensions. This motivates us to work in the *reduced zone scheme*, in which these higher Brillouin zones are mapped back into the first. This is achieved by translating them by some lattice vector. The higher Brillouin zones of the square lattice in the reduced zone scheme are shown in Figure 14.

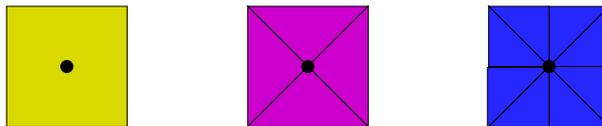


Figure 14: The first three Brillouin zones for a square lattice in the reduced zone scheme.

Finally, note that the edges of the Brillouin zone should be identified; they label the same momentum state \mathbf{k} . For one-dimensional lattices, this results in the Brillouin zone having the topology of a circle. For d -dimensional lattices, the Brillouin zone is topologically a torus \mathbf{T}^d .

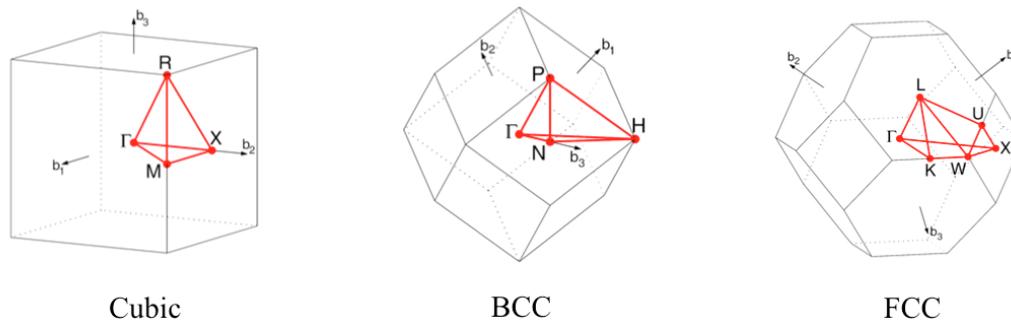


Figure 15: The labels for various special points on the Brillouin zone.

Crystallographic Notation

The Brillouin zone of real materials is a three-dimensional space. We often want to describe how certain quantities – such as the energy of the electrons – vary as we move around the Brillouin zone. To display this information graphically, we need to find a way to depict the underlying Brillouin zone as a two-dimensional, or even one-dimensional space. Crystallographers have developed a notation for this. Certain, highly symmetric points in the Brillouin zone are labelled by letters. From the letter, you’re also supposed to remember what underlying lattice we’re talking about.

For example, all Brillouin zones have an origin. The concept of an “origin” occurs in many different parts of maths and physics and almost everyone has agreed to label it as “0”. Almost everyone. But not our crystallographer friends. Instead, they call the origin Γ .

From hereon, it gets more bewildering although if you stare at enough of these you get used to it. For example, for a cubic lattice, the centre of each face is called X , the centre of each edge is M while each corner is R . Various labels for BCC and FCC lattices are shown in Figure 15

2.4 Band Structure

“When I started to think about it, I felt that the main problem was to explain how the electrons could sneak by all the ions in a metal... I found to my delight that the wave differed from a plane wave of free electron only by a periodic modulation.

This was so simple that I didn’t think it could be much of a discovery, but when I showed it to Heisenberg he said right away, ‘That’s it.’.”

Now that we've developed the language to describe lattices in higher dimensions, it's time to understand how electrons behave when they move in the background of a fixed lattice. We already saw many of the main ideas in the context of a one-dimensional lattice in Section 1. Here we will describe the generalisation to higher dimensions.

2.5 Bloch's Theorem

Consider an electron moving in a potential $V(\mathbf{x})$ which has the periodicity of a Bravais lattice Λ ,

$$V(\mathbf{x} + \mathbf{r}) = V(\mathbf{x}) \quad \text{for all } \mathbf{r} \in \Lambda$$

Bloch's theorem states that the energy eigenstates take the form

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

where $u_{\mathbf{k}}(\mathbf{x})$ has the same periodicity as the lattice, $u_{\mathbf{k}}(\mathbf{x} + \mathbf{r}) = u_{\mathbf{k}}(\mathbf{x})$ for all $\mathbf{r} \in \Lambda$.

There are different ways to prove Bloch's theorem. Here we will give a simple proof using the ideas of translation operators, analogous to the one-dimensional proof that we saw in Section 1.3. Later, in Section 2.6, we will provide a more direct proof by decomposing the Schrödinger equation into Fourier modes.

Our starting point is that the Hamiltonian is invariant under discrete translations by the lattice vectors $\mathbf{r} \in \Lambda$. As we explained in Section 1.3, these translations are implemented by unitary operators $T_{\mathbf{r}}$. These operators form an Abelian group,

$$T_{\mathbf{r}}T_{\mathbf{r}'} = T_{\mathbf{r}+\mathbf{r}'} \tag{2.4}$$

and commute with the Hamiltonian: $[H, T_{\mathbf{r}}] = 0$. This means that we can simultaneously diagonalise H and $T_{\mathbf{r}}$, so that energy eigenstates are also labelled by the eigenvalue of each $T_{\mathbf{r}}$. Because $T_{\mathbf{r}}$ is unitary, this is simply a phase. To respect the group structure (2.4), we should take this phase to be labelled by some vector \mathbf{k} ,

$$T_{\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{x}) = \psi_{\mathbf{k}}(\mathbf{x} + \mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{x})$$

Now we simply need to look at the function $u_{\mathbf{k}}(\mathbf{x}) = e^{-i\mathbf{k}\cdot\mathbf{x}}\psi_{\mathbf{k}}(\mathbf{x})$. The statement of Bloch's theorem is that $u_{\mathbf{k}}(\mathbf{x})$ has the periodicity of Λ which is indeed true, since $u_{\mathbf{k}}(\mathbf{x} + \mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{x}}e^{-i\mathbf{k}\cdot\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{x} + \mathbf{r}) = e^{-i\mathbf{k}\cdot\mathbf{x}}\psi_{\mathbf{k}}(\mathbf{x}) = u_{\mathbf{k}}(\mathbf{x})$.

Crystal Momentum

The energy eigenstates are labelled by the wavevector \mathbf{k} , called the *crystal momentum*. There is an ambiguity in the definition of this crystal momentum. This is not the same as the true momentum. The energy eigenstates do not have a well defined momentum because they are not eigenstates of the momentum operator $\mathbf{p} = -i\hbar\nabla$ unless $u_{\mathbf{k}}(\mathbf{x})$ is constant. Nonetheless, we will see as we go along that the crystal momentum plays a role similar to the true momentum. For this reason, we will often refer to \mathbf{k} simply as “momentum”.

There is an ambiguity in the definition of the crystal momentum. Consider a state with a crystal momentum $\mathbf{k}' = \mathbf{k} + \mathbf{q}$, with $\mathbf{q} \in \Lambda^*$ a reciprocal lattice vector. Then

$$\psi_{\mathbf{k}'}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{q}\cdot\mathbf{x}} u_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{u}_{\mathbf{k}}(\mathbf{x})$$

where $\tilde{u}_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{q}\cdot\mathbf{x}} u_{\mathbf{k}}(\mathbf{x})$ also has the periodicity of Λ by virtue of the definition of the reciprocal lattice (2.3).

As in the one-dimensional example, we have different options. We could choose to label states by \mathbf{k} which lie in the first Brillouin zone. In this case, there will typically be many states with the same \mathbf{k} and different energies. This is the *reduced zone scheme*. In this case, the energy eigenstates are labelled by two indices, $\psi_{\mathbf{k},n}$, where \mathbf{k} is the crystal momentum and n is referred to as the *band index*. (We will see examples shortly.)

Alternatively, we can label states by taking any $\mathbf{k} \in \mathbf{R}^d$ where d is the dimension of the problem. This is the *extended zone scheme*. In this case that states labelled by \mathbf{k} which differ by Λ^* have the same crystal momenta.

2.6 Nearly Free Electrons in Three Dimensions

Consider an electron moving in \mathbf{R}^3 in the presence of a weak potential $V(\mathbf{x})$. We'll assume that this potential has the periodicity of a Bravais lattice Λ , so

$$V(\mathbf{x}) = V(\mathbf{x} + \mathbf{r}) \quad \text{for all } \mathbf{r} \in \Lambda$$

We treat this potential as a perturbation on the free electron. This means that we start with plane wave states $|\mathbf{k}\rangle$ with wavefunctions

$$\langle \mathbf{x} | \mathbf{k} \rangle \sim e^{i\mathbf{k}\cdot\mathbf{x}}$$

with energy $E_0(\mathbf{k}) = \hbar^2 k^2 / 2m$. We want to see how these states and their energy levels are affected by the presence of the potential. The discussion will follow closely the one-dimensional case that we saw in Section 1.2 and we only highlight the differences.

When performing perturbation theory, we're going to have to consider the potential $V(\mathbf{x})$ sandwiched between plane-wave states,

$$\langle \mathbf{k} | V(\mathbf{x}) | \mathbf{k}' \rangle = \frac{1}{\text{Volume}} \int d^3x e^{i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{x}} V(\mathbf{x})$$

However, the Fourier transform of a periodic function can be written as a sum over wavevectors that lie in the reciprocal lattice Λ^* ,

$$V(\mathbf{x}) = \sum_{\mathbf{q} \in \Lambda^*} e^{i\mathbf{q} \cdot \mathbf{x}} V_{\mathbf{q}}$$

(Note: here $V_{\mathbf{q}}$ is the Fourier component of the potential and should not be confused with the volumes of unit cells which were denoted as V and V^* in Section 2.) This means that $\langle \mathbf{k} | V(\mathbf{x}) | \mathbf{k}' \rangle$ is non-vanishing only when the two momenta differ by

$$\mathbf{k} - \mathbf{k}' = \mathbf{q} \quad \mathbf{q} \in \Lambda^*$$

This has a simple physical interpretation: a plane wave state $|\mathbf{k}\rangle$ can scatter into another plane wave state $|\mathbf{k}'\rangle$ only if they differ by a reciprocal lattice vector. In other words, only momenta \mathbf{q} , with $\mathbf{q} \in \Lambda^*$, can be absorbed by the lattice.

Another Perspective on Bloch's Theorem

The fact that that a plane wave state $|\mathbf{k}\rangle$ can only scatter into states $|\mathbf{k} + \mathbf{q}\rangle$, with $\mathbf{q} \in \Lambda^*$, provides a simple viewpoint on Bloch's theorem, one that reconciles the quantum state with the naive picture of the particle bouncing off lattice sites like a ball in a pinball machine. Suppose that the particle starts in some state $|\mathbf{k}\rangle$. After scattering, we might expect it to be some superposition of all the possible scattering states $|\mathbf{k} + \mathbf{q}\rangle$. In other words,

$$\psi_{\mathbf{k}}(\mathbf{x}) = \sum_{\mathbf{q} \in \Lambda^*} e^{i(\mathbf{k} + \mathbf{q}) \cdot \mathbf{x}} c_{\mathbf{k} - \mathbf{q}}$$

for some coefficients $c_{\mathbf{k} - \mathbf{q}}$. We can write this as

$$\psi_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k} \cdot \mathbf{x}} \sum_{\mathbf{q} \in \Lambda^*} e^{i\mathbf{q} \cdot \mathbf{x}} c_{\mathbf{k} - \mathbf{q}} = e^{i\mathbf{k} \cdot \mathbf{x}} u_{\mathbf{k}}(\mathbf{x})$$

where, by construction, $u_{\mathbf{k}}(\mathbf{x} + \mathbf{r}) = u_{\mathbf{k}}(\mathbf{x})$ for all $\mathbf{r} \in \Lambda$. But this is precisely the form guaranteed by Bloch's theorem.

Although the discussion here holds at first order in perturbation theory, it is not hard to extend this argument to give an alternative proof of Bloch's theorem, which essentially comes down to analysing the different Fourier modes of the Schrödinger equation.

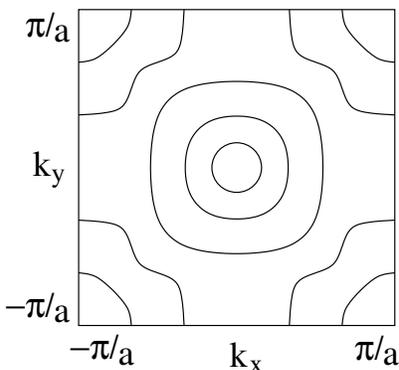


Figure 16: Energy contours for nearly-free electrons in the first Brillouin zone.

Band Structure

Let's now look at what becomes of the energy levels after we include the perturbation. We will see that, as in the 1d example, they form bands. The resulting eigenstates $\psi_{\mathbf{k},n}(\mathbf{x})$ and their associated energy levels $E_n(\mathbf{k})$ are referred to as the *band structure* of the system.

Low Momentum: Far from the edge of the Brillouin zone, the states $|\mathbf{k}\rangle$ can only scatter into states $|\mathbf{k} + \mathbf{q}\rangle$ with greatly different energy. In this case, we can work with non-degenerate perturbation theory to compute the corrections to the energy levels.

On the Boundary of the Brillouin zone: Things get more interesting when we have to use degenerate perturbation theory. This occurs whenever the state $|\mathbf{k}\rangle$ has the same energy as another state $|\mathbf{k} + \mathbf{q}\rangle$ with $\mathbf{q} \in \Lambda^*$,

$$E_0(\mathbf{k}) = E_0(\mathbf{k} + \mathbf{q}) \quad \Rightarrow \quad k^2 = (\mathbf{k} + \mathbf{q})^2 \quad \Rightarrow \quad 2\mathbf{k} \cdot \mathbf{q} + q^2 = 0$$

This condition is satisfied whenever we can write

$$\mathbf{k} = -\frac{1}{2}\mathbf{q} + \mathbf{k}_\perp$$

where $\mathbf{q} \cdot \mathbf{k}_\perp = 0$. This is the condition that we sit on the perpendicular bisector of the origin and the lattice point $-\mathbf{q} \in \Lambda^*$. But, as we explained in Section 2.3, these bisectors form the boundaries of the Brillouin zones. We learn something important: momentum states are degenerate only when they lie on the boundary of a Brillouin zone. This agrees with what we found in our one-dimensional example in Section 1.2.

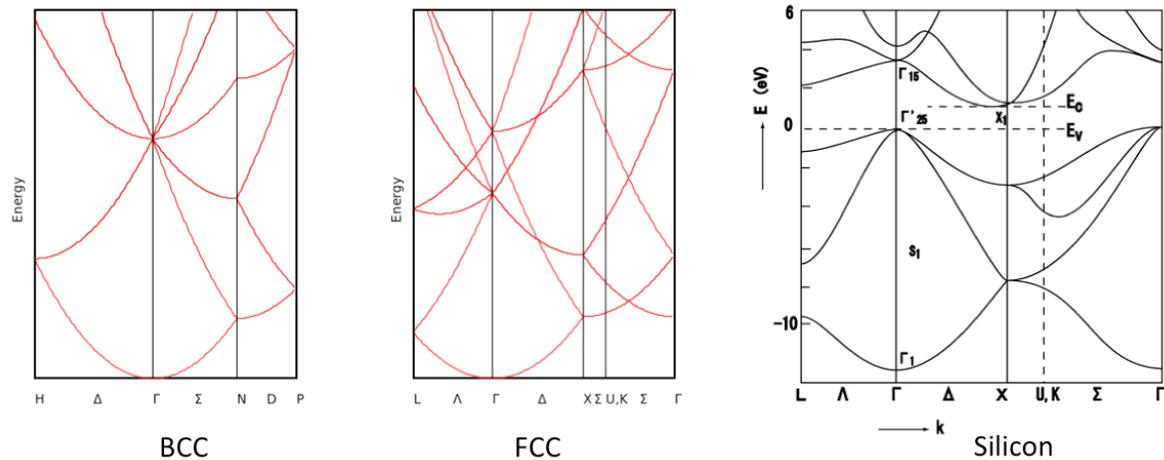


Figure 17: Free band structure (in red) for BCC and FCC, together with the band structure for silicon, exhibiting a gap.

We know from experience what the effect of the perturbation $V(\mathbf{x})$ will be: it will lift the degeneracy. This means that a gap opens at the boundary of the Brillouin zone. For example, the energy of states just inside the first Brillouin zone will be pushed down, while the energy of those states just outside the first Brillouin zone will be pushed up. Note that the size of this gap will vary as we move around the boundary..

There is one further subtlety that we should mention. At a generic point on the boundary of the Brillouin zone, the degeneracy will usually be two-fold. However, at special points — such as edges, or corners — it often higher. In this case, we must work with all degenerate states when computing the gap.

All of this is well illustrated with an example. However, it's illustrated even better if you do the example yourself! The problem of nearly free electrons in a two-dimensional square lattice is on the problem sheet. The resulting energy contours are shown in Figure 16.

Plotting Band Structures in Three Dimensions

For three-dimensional lattice, we run into the problem of depicting the bands. For this, we need the crystallographer's notation we described previously. The spectrum of free particles (i.e. with no lattice) is plotted in the Brillouin zone of BCC and FCC lattices in Figure 17¹.

¹Images plotted by Jan-Rens Reitsma, from Wikimedia commons.

We can then compare this to the band structure of real materials. The dispersion relation for silicon is also shown in Figure 17. This has a diamond lattice structure, which is plotted as FCC. Note that you can clearly see the energy gap of around $1.1 eV$ between the bands.

How Many States in the Brillouin Zone?

The Brillouin zone consists of all wavevectors \mathbf{k} that lie within the Wigner-Seitz cell of the reciprocal lattice Λ^* . How many quantum states does it hold? Well, if the spatial lattice Λ is infinite in extent then \mathbf{k} can take any continuous value and there are an infinite number of states in the Brillouin zone. But what if the spatial lattice is finite in size?

In this section we will count the number of quantum states in the Brillouin zone of a finite spatial lattice Λ . We will find a lovely answer: the number of states is equal to N , the number of lattice sites.

Recall that the lattice Λ consists of all vectors $\mathbf{r} = \sum_i n_i \mathbf{a}_i$ where \mathbf{a}_i are the primitive lattice vectors and $n_i \in \mathbf{Z}$. For a finite lattice, we simply restrict the value of these integers to be

$$0 \leq n_i \leq N_i$$

for some N_i . The total number of lattice sites is then $N = N_1 N_2 N_3$ (assuming a three-dimensional lattice). The total volume of the lattice is VN where $V = |\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)|$ is the volume of the unit cell.

The basic physics is something that we've met before: if we put a particle in a box, then the momentum $\hbar\mathbf{k}$ becomes quantised. This arises because of the boundary conditions that we place on the wavefunction. It's simplest to think about a finite, periodic lattice where we require that the wavefunction inherits this periodicity, so that

$$\psi(\mathbf{x} + N_i \mathbf{a}_i) = \psi(\mathbf{x}) \quad \text{for each } i = 1, 2, 3 \quad (2.5)$$

But we know from Bloch's theorem that energy eigenstates take the form $\psi_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k}\cdot\mathbf{x}} u_{\mathbf{k}}(\mathbf{x})$ where $u_{\mathbf{k}}(\mathbf{x} + \mathbf{a}_i) = u_{\mathbf{k}}(\mathbf{x})$. This means that the periodicity condition (2.5) becomes

$$e^{iN_i \mathbf{k} \cdot \mathbf{a}_i} = 1 \quad \Rightarrow \quad \mathbf{k} = \sum_i \frac{m_i}{N_i} \mathbf{b}_i$$

where $m_i \in \mathbf{Z}$ and \mathbf{b}_i are the primitive vectors of the reciprocal lattice defined in (2.2). This is sometimes called the *Born-von Karmen* boundary condition.

This is the quantisation of momentum that we would expect in a finite system. The states are now labelled by integers $m_i \mathbf{Z}$. Each state can be thought of as occupying a volume in \mathbf{k} -space, given by

$$\frac{|\mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3)|}{N_1 N_2 N_3} = \frac{V^*}{N}$$

where V^* is the volume of the Brillouin zone. We see that the number of states that live inside the Brillouin zone is precisely N , the number of sites in the spatial lattice.

3. Fermi Surfaces

In the previous chapter we have seen how the single-electron energy states form a band structure in the presence of a lattice. Our goal now is to understand the consequences of this, so that we can start to get a feel for some of the basic properties of materials.

There is one feature in particular that will be important: materials don't just have one electron sitting in them. They have lots. A large part of condensed matter physics is concerned with understanding the collective behaviour of this swarm of electrons. This can often involve the interactions between electrons giving rise to subtle and surprising effects. However, for our initial foray into this problem, we will make a fairly brutal simplification: we will ignore the interactions between electrons. Ultimately, much of the basic physics that we describe below is unchanged if we turn on interactions, although the reason for this turns out to be rather deep.

3.1 Fermi Surfaces

Even in the absence of any interactions, electrons still are still affected by the presence of others. This is because electrons are fermions, and so subject to the *Pauli exclusion principle*. This is the statement that only one electron can sit in any given state. As we will see below, the Pauli exclusion principle, coupled with the general features of band structure, goes some way towards explaining the main properties of materials.

Free Electrons

As a simple example, suppose that we have no lattice. We take a cubic box, with sides of length L , and throw in some large number of electrons. What is the lowest energy state of this system? Free electrons sit in eigenstates with momentum $\hbar\mathbf{k}$ and energy $E = \hbar^2 k^2 / 2m$. Because we have a system of finite size, momenta are quantised as $k_i = 2\pi n_i / L$. Further, they also carry one of two spin states, $|\uparrow\rangle$ or $|\downarrow\rangle$.

The first electron can sit in the state $\mathbf{k} = 0$ with, say, spin $|\uparrow\rangle$. The second electron can also have $\mathbf{k} = 0$, but must have spin $|\downarrow\rangle$, opposite to the first. Neither of these electrons costs any energy. However, the next electron is not so lucky. The minimum energy state it can sit in has $n_i = (1, 0, 0)$. Including spin and momentum there are a total of six electrons which can carry momentum $|\mathbf{k}| = 2\pi/L$. As we go on, we fill out a ball in momentum space. This ball is called the *Fermi sea* and the boundary

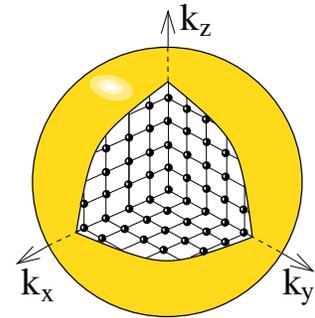


Figure 18: The Fermi surface

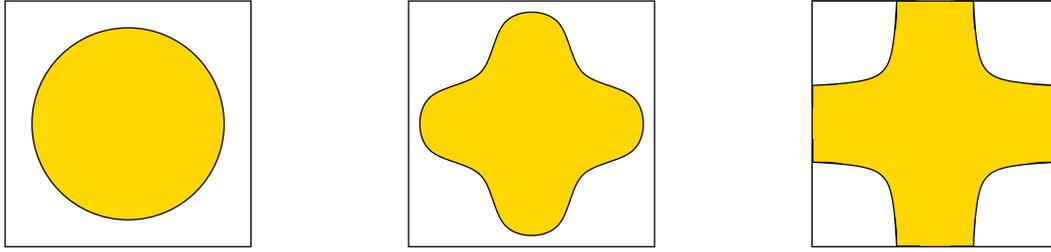


Figure 19: Fermi surfaces for valence $Z = 1$ with increasing lattice strength.

of the ball is called the *Fermi surface*. The states on the Fermi surface are said to have *Fermi momentum* $\hbar k_F$ and *Fermi energy* $E_F = \hbar^2 k_F^2 / 2m$. Various properties of the free Fermi sea are explored in the lectures on [Statistical Physics](#).

3.2 Metals vs Insulators

Here we would like to understand what becomes of the Fermi sea and, more importantly, the Fermi surface in the presence of a lattice. Let's recapitulate some important facts that we'll need to proceed:

- A lattice causes the energy spectrum to split into bands. We saw in Section 2.6 that a Bravais lattice with N sites results in each band having N momentum states. These are either labelled by momenta in the first Brillouin zone (in the reduced zone scheme) or by momentum in successive Brillouin zones (in the extended zone scheme).
- Because each electron carries one of two spin states, each band can accommodate $2N$ electrons.
- Each atom of the lattice provides an integer number of electrons, Z , which are free to roam the material. These are called *valence electrons* and the atom is said to have *valence* Z .

From this, we can piece the rest of the story together. We'll discuss the situation for two-dimensional square lattices because it's simple to draw the Brillouin zones. But everything we say carries over for more complicated lattices in three-dimensions.

Suppose that our atoms have valence $Z = 1$. There are then N electrons, which can be comfortably housed inside the first Brillouin zone. In the left-hand of Figure 19 we have drawn the Fermi surface for free electrons inside the first Brillouin zone. However, we know that the effect of the lattice is to reduce the energy at the edges

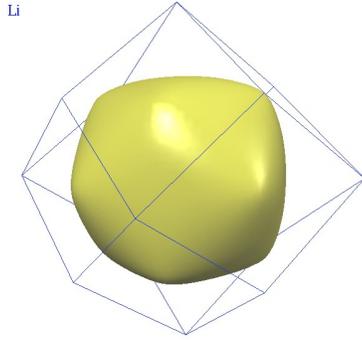


Figure 20: Lithium.

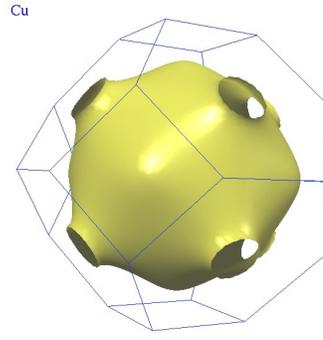


Figure 21: Copper.

of the Brillouin zone. We expect, therefore, that the Fermi surface — which is the equipotential E_F — will be distorted as shown in the middle figure, with states closer to the edge of the Brillouin zone filled preferentially. Note that the area inside the Fermi surface remains the same.

If the effects of the lattice get very strong, it may be that the Fermi surface touches the edge of the Brillouin zone as shown in the right-hand drawing in Figure 19. Because the Brillouin zone is a torus, if the Fermi surface is to be smooth then it must hit the edge of the Brillouin zone at right-angles.

This same physics can be seen in real Fermi surfaces. Lithium has valence $Z = 1$. It forms a BCC lattice, and so the Brillouin zone is FCC. Its Fermi surface is shown above, plotted within its Brillouin zone². Copper also has valency $Z = 1$, with a FCC lattice and hence BCC Brillouin zone. Here the effects of the lattice are somewhat stronger, and the Fermi surface touches the Brillouin zone.

In all of these cases, there are unoccupied states with arbitrarily small energy above E_F . (Strictly speaking, this statement holds only in the limit $L \rightarrow \infty$ of an infinitely large lattice.) This means that if we perturb the system in any way, the electrons will easily be able to respond. Note, however, that only those electrons close to the Fermi surface can respond; those that lie deep within the Fermi sea are locked there by the Pauli exclusion principle and require much larger amounts of energy if they wish to escape.

This is an important point, so I'll say it again. In most situations, only those electrons which lie on the Fermi surface can actually do anything. This is why Fermi surfaces play such a crucial role in our understanding of materials.

²This, and other pictures of Fermi surfaces, are taken from <http://www.phys.ufl.edu/fermisurface/>.

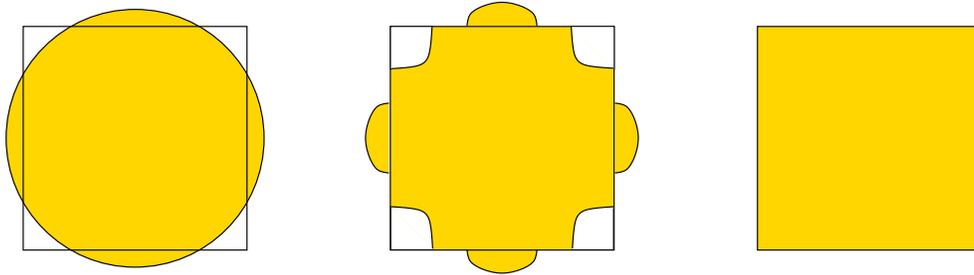


Figure 22: Fermi surfaces for valence $Z = 2$ with increasing lattice strength, moving from a metal to an insulator.

Materials with a Fermi surface are called *metals*. Suppose, for example, that we apply a small electric field to the sample. The electrons that lie at the Fermi surface can move to different available states in order to minimize their energy in the presence of the electric field. This results in a current that flows, the key characteristic of a metal. We'll discuss more about how electrons in lattices respond to outside influences in Section 5

Before we move on, a couple of comments:

- The Fermi energy of metals is huge, corresponding to a temperature of $E_F/k_B \sim 10^4 K$, much higher than the melting temperature. For this reason, the zero temperature analysis is a good starting point for thinking about real materials.
- Metals have a very large number of low-energy excitations, proportional to the area of the Fermi surface. This makes metals a particularly interesting theoretical challenge.

Let's now consider atoms with valency $Z = 2$. These have $2N$ mobile electrons, exactly the right number to fill the first band. However, in the free electron picture, this is not what happens. Instead, they partially fill the first Brillouin zone and then spill over into the second Brillouin zone. The resulting Fermi surface, drawn in the extended zone scheme, is shown in left-hand picture of Figure 22

If the effects of the lattice are weak, this will not be greatly changed. Both the first and second Brillouin zones will have available states close to the Fermi surface as shown in the middle picture. These materials remain metals. We sometimes talk

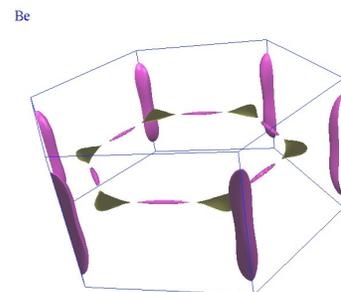


Figure 23: Beryllium

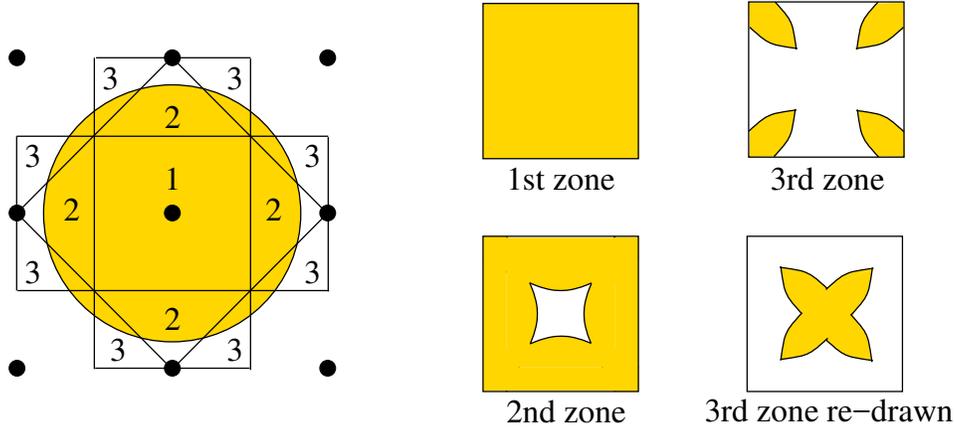


Figure 24: Fermi surfaces for valence $Z = 3$.

of electrons in the second band, and holes (i.e. absence of electrons) in the first band. We will discuss this further in Section 5. Beryllium provides an example of a metal with $Z = 2$; its Fermi surface is shown in the figure, now plotted in the reduced zone scheme. It includes both an electron Fermi surface (the cigar-like shapes around the edge) and a hole Fermi surface (the crown in the middle).

Finally, if the effects of the lattice become very strong, the gap between the two bands is large enough to overcome the original difference in kinetic energies. This occurs when the lowest lying state in the second band is higher than the highest state in the first. Now the electrons fill the first band. The second band is empty. The Fermi sea looks like the right-hand picture in Figure 22. This is qualitatively different from previous situations. There is no Fermi surface and, correspondingly, no low-energy excitations. Any electron that wishes to change its state can only do so by jumping to the next band. But that costs a finite amount of energy, equal to the gap between bands. This means that all the electrons are now locked in place and cannot respond to arbitrarily small outside influences. We call such materials *insulators*. (Sometimes they are referred to as *band insulators* to highlight the fact that it is the band structure which prevents the electrons from moving.)

This basic characterisation remains for higher valency Z . Systems with partially filled bands are metals; systems with only fully-filled bands are insulators. Note that a metal may well have several fully-filled bands, before we get to a partially filled band. In such circumstances, we usually differentiate between the fully-filled lower bands — which are called *valence bands* — and the partially filled *conduction band*.

The Fermi surfaces may exist in several different bands. An example of a Fermi surface for $Z = 3$ is shown in Figure 24, the first three Brillouin zones are shown separately in the reduced zone scheme. At first glance, it appears that the Fermi surface in the 3rd Brillouin zone is disconnected. However, we have to remember that the edges of the Brillouin zone are identified. Re-drawn, with the origin taken to be $\mathbf{k} = (\pi/a, \pi/a)$, we see the Fermi surface is connected, taking the rosette shape shown.

Looking Forwards

We have seen how band structure allows us to classify all materials as metals or insulators. This, however, is just the beginning, the first chapter in a long and detailed story which extends from physics into materials science. To whet the appetite, here are three twists that we can add to this basic classification.

- For insulators, the energy required to reach the first excited state is set by the band gap Δ which, in turn, is determined by microscopic considerations. Materials whose band gap is smaller than $\Delta \lesssim 2 \text{ eV}$ or so behave as insulators at small temperature, but starts to conduct at higher temperatures as electrons are thermally excited from the valence band to the conduction band. Such materials are called *semiconductors*. They have the property that their conductivity increases as the temperature increases. (This is in contrast to metals whose conductivity decreases as temperature increases.) John Bardeen, Walter Brattain and William Shockley won the 1956 Nobel prize for developing their understanding of semiconductors into a working transistor. This, then, changed the world.
- There are some materials which have $Z = 1$ but are, nonetheless, insulators. An example is nickel oxide NiO . This contradicts our predictions using elementary band structure. The reason is that, for these materials, we cannot ignore the interactions between electrons. Roughly speaking, the repulsive force dominates the physics and effectively prohibits two electrons from sitting on the same site, even if they have different spins. But with only one spin state allowed per site, each band houses only N electrons. Materials with this property are referred to as *Mott insulators*. Nevill Mott, Cavendish professor and master of Caius, won the 1977 Nobel prize, in part for this discovery.
- For a long time band insulators were considered boring. The gap to the first excited state means that they can't do anything when prodded gently. This attitude changed relatively recently when it was realised that you can be boring in different ways. There is a topological classification of how the phase of the quantum states winds as you move around the Brillouin zone. Materials in which

this winding is non-trivial are called *topological insulators*. They have wonderful and surprising properties, most notably on their edges where they come alive with interesting and novel physics. David Thouless and Duncan Haldane won the 2016 Nobel prize for their early, pioneering work on this topic.

More generally, there is a lesson above that holds in a much wider context. Our classification of materials into metals and insulators hinges on whether or not we can excite a multi-electron system with an arbitrarily small cost in energy. For insulators, this is not possible: we require a finite injection of energy to reach the excited states. Such systems are referred to as *gapped*, meaning that there is finite energy gap between the ground state and first excited state. Meanwhile, systems like metals are called *gapless*. Deciding whether any given quantum system is gapped or gapless is one of the most basic questions we can ask. It can also be one of the hardest. For example, the question of whether a quantum system known as *Yang-Mills theory* has a gap is one of the six unsolved millenium maths problems.

4. Graphene

Graphene is a two-dimensional lattice of carbon atoms, arranged in a honeycomb structure as shown in the figure. Although it is straightforward to build many of these layers of these lattices — a substance known as graphite — it was long thought that a purely two-dimensional lattice would be unstable to thermal fluctuations and impossible to create. This changed in 2004 when Andre Geim and Konstantin Novoselov at the University of Manchester succeeded in isolating two-dimensional graphene. For this, they won the 2010 Nobel prize. As we now show, the band structure of graphene is particularly interesting.

First, some basic lattice facts. We described the honeycomb lattice in Section 2.1. It is not Bravais. Instead, it is best thought of as two triangular sublattices. We define the primitive lattice vectors

$$\mathbf{a}_1 = \frac{\sqrt{3}a}{2}(\sqrt{3}, 1) \quad \text{and} \quad \mathbf{a}_2 = \frac{\sqrt{3}a}{2}(\sqrt{3}, -1)$$

where a the distance between neighbouring atoms, which in graphene is about $a \approx 1.4 \times 10^{10}$ m. These lattice vectors are shown in the figure.

Sublattice A is defined as all the points $\mathbf{r} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2$ with $n_i \in \mathbf{Z}$. These are the red dots in the figure. Sublattice B is defined as all points $\mathbf{r} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + \mathbf{d}$ with $\mathbf{d} = (-a, 0)$. These are the white dots.

The reciprocal lattice is generated by vectors \mathbf{b}_j satisfying $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$. These are

$$\mathbf{b}_1 = \frac{2\pi}{3a}(1, \sqrt{3}) \quad \text{and} \quad \mathbf{b}_2 = \frac{2\pi}{3a}(1, -\sqrt{3})$$

This reciprocal lattice is also triangular, rotated 90° from the original. The Brillouin zone is constructed in the usual manner by drawing perpendicular boundaries between the origin and each other point in the reciprocal lattice. This is shown in the figure. We shortly see that the corners of the Brillouin zone carry particular interest. It naively appears that there are 6 corners, but this should really be viewed as two sets of three. This follows because any points in the Brillouin zone which are connected by a reciprocal lattice vector are identified. Representatives of the two, inequivalent corners of the Brillouin zone are given by

$$\mathbf{K} = \frac{1}{3}(2\mathbf{b}_1 + \mathbf{b}_2) = \frac{2\pi}{3a}\left(1, \frac{1}{\sqrt{3}}\right) \quad \text{and} \quad \mathbf{K}' = \frac{1}{3}(\mathbf{b}_1 + 2\mathbf{b}_2) = \frac{2\pi}{3a}\left(1, -\frac{1}{\sqrt{3}}\right) \quad (4.1)$$

These are shown in the figure above.

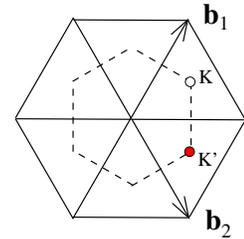


Figure 25:

Tight Binding for Graphene

The carbon atoms in graphene have valency $Z = 1$, with the p_z -atomic orbital abandoned by their parent ions and free to roam the lattice. In this context, it is usually called the π -orbital. We therefore write down a tight-binding model in which this electron can hop from one atomic site to another. We will work only with nearest neighbour interactions which, for the honeycomb lattice, means that the Hamiltonian admits hopping from a site of the A-lattice to the three nearest neighbours on the B-lattice, and vice versa. The Hamiltonian is given by

$$H = -t \sum_{\mathbf{r} \in \Lambda} \left[|\mathbf{r}; A\rangle \langle \mathbf{r}; B| + |\mathbf{r}; A\rangle \langle \mathbf{r} + \mathbf{a}_1; B| + |\mathbf{r}; A\rangle \langle \mathbf{r} + \mathbf{a}_2; B| + \text{h.c.} \right] \quad (4.2)$$

where we're using the notation

$$|\mathbf{r}; A\rangle = |\mathbf{r}\rangle \quad \text{and} \quad |\mathbf{r}; B\rangle = |\mathbf{r} + \mathbf{d}\rangle \quad \text{with} \quad \mathbf{d} = (-a, 0)$$

Comparing to (??), we have set $E_0 = 0$, on the grounds that it doesn't change any of the physics. For what it's worth, $t \approx 2.8 \text{ eV}$ in graphene, although we won't need the precise value to get at the key physics.

The energy eigenstates are again plane waves, but now with a suitable mixture of A and B sublattices. We make the ansatz

$$|\psi(\mathbf{k})\rangle = \frac{1}{\sqrt{2N}} \sum_{\mathbf{r} \in \Lambda} e^{i\mathbf{k}\cdot\mathbf{r}} \left(c_A |\mathbf{r}; A\rangle + c_B |\mathbf{r}; B\rangle \right)$$

Plugging this into the Schrödinger equation, we find that c_A and c_B must satisfy the eigenvalue equation

$$\begin{pmatrix} 0 & \gamma(\mathbf{k}) \\ \gamma^*(\mathbf{k}) & 0 \end{pmatrix} \begin{pmatrix} c_A \\ c_B \end{pmatrix} = E(\mathbf{k}) \begin{pmatrix} c_A \\ c_B \end{pmatrix} \quad (4.3)$$

where

$$\gamma(\mathbf{k}) = -t \left(1 + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2} \right)$$

The energy eigenvalues of (4.3) are simply

$$E(\mathbf{k}) = \pm |\gamma(\mathbf{k})|$$

We can write this as

$$E(\mathbf{k})^2 = t^2 \left| 1 + e^{i\mathbf{k}\cdot\mathbf{a}_1} + e^{i\mathbf{k}\cdot\mathbf{a}_2} \right|^2 = t^2 \left| 1 + 2e^{3ik_x a/2} \cos\left(\frac{\sqrt{3}k_y a}{2}\right) \right|^2$$

Expanding this out, we get the energy eigenvalues

$$E(\mathbf{k}) = \pm t \sqrt{1 + 4 \cos\left(\frac{3k_x a}{2}\right) \cos\left(\frac{\sqrt{3}k_y a}{2}\right) + 4 \cos^2\left(\frac{\sqrt{3}k_y a}{2}\right)}$$

Note that the energy spectrum is a double cover of the first Brillouin zone, symmetric about $E = 0$. This doubling can be traced to the fact the the honeycomb lattice consists of two intertwined Bravais lattices. Because the carbon atoms have valency $Z = 1$, only the lower band with $E(\mathbf{k}) < 0$ will be filled.

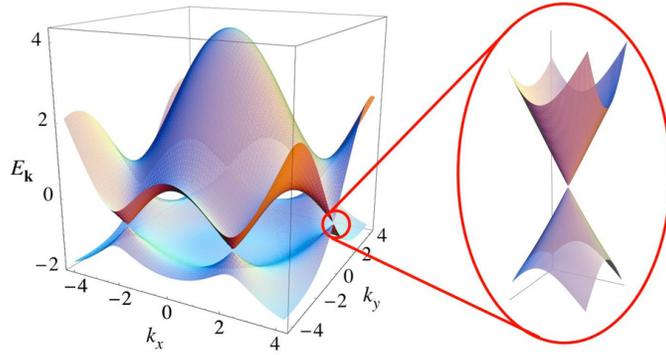


Figure 26: The band structure of graphene.

The surprise of graphene is that these two bands meet at special points. These occur on the corners $\mathbf{k} = \mathbf{K}$ and $\mathbf{k} = \mathbf{K}'$ (4.1), where $\cos(3k_x a/2) = -1$ and $\cos(\sqrt{3}k_y a/2) = 1/2$. The resulting band structure is shown in Figure 26³. Because the lower band is filled, the Fermi surface in graphene consists of just two points, \mathbf{K} and \mathbf{K}' where the bands meet. It is an example of a *semi-metal*.

Emergent Relativistic Physics

The points $\mathbf{k} = \mathbf{K}$ and \mathbf{K}' where the bands meet are known as *Dirac points*. To see why, we linearise about these points. Write

$$\mathbf{k} = \mathbf{K} + \mathbf{q}$$

A little Taylor expansion shows that in the vicinity of the Dirac points, the dispersion relation is linear

$$E(\mathbf{k}) \approx \pm \frac{3ta}{2} |\mathbf{q}|$$

³The image is taken from the [exciting-code website](#).

But this is the same kind of energy-momentum relation that we meet in relativistic physics for massless particles! In that case, we have $E = |\mathbf{p}|c$ where p is the momentum and c is the speed of light. For graphene, we have

$$E(\mathbf{k}) \approx \hbar v_F |\mathbf{q}|$$

where $\hbar\mathbf{q}$ is the momentum measured with respect to the Dirac point and $v_F = 3ta/2\hbar$ is the speed at which the excitations propagate. In graphene, v_F is about 300 times smaller than the speed of light. Nonetheless, it remains true that the low-energy excitations of graphene are governed by the same equations that we meet in relativistic quantum field theory. This was part of the reason for the excitement about graphene: we get to test ideas from quantum field theory in a simple desktop experiment.

We can tease out more of the relativistic structure by returning to the Hamiltonian (4.2). Close to the Dirac point $\mathbf{k} = \mathbf{K}$ we have

$$\begin{aligned} \gamma(\mathbf{k}) &= -t \left(1 - 2e^{3iq_x a/2} \cos \left(\frac{\pi}{3} + \frac{\sqrt{3}q_y a}{2} \right) \right) \\ &= -t \left(1 - 2e^{3iq_x a/2} \left[\frac{1}{2} \cos \left(\frac{\sqrt{3}q_y a}{2} \right) - \frac{\sqrt{3}}{2} \sin \left(\frac{\sqrt{3}q_y a}{2} \right) \right] \right) \\ &\approx -t \left(1 - 2 \left(1 + \frac{3iq_x a}{2} + \dots \right) \left(\frac{1}{2} - \frac{3q_y a}{4} + \dots \right) \right) \\ &\approx v_F \hbar (iq_x - q_y) \end{aligned}$$

This means that the Hamiltonian in the vicinity of the Dirac point $\mathbf{k} = \mathbf{K}$ takes the form

$$H = v_F \hbar \begin{pmatrix} 0 & iq_x - q_y \\ -iq_x - q_y & 0 \end{pmatrix} = -v_F \hbar (q_x \sigma^y + q_y \sigma^x) \quad (4.4)$$

where σ^x and σ^y are the Pauli matrices. But this is the *Dirac equation* for a massless particle moving in two-dimensions, sometimes referred to as the *Pauli equation*. (Note: our original choice of orientation of the honeycomb lattice has resulted in a slightly annoying expression for the Hamiltonian. Had we rotated by 90° to begin with, we would be left with the nicer $H = \hbar v_F \mathbf{q} \cdot \boldsymbol{\sigma}$ where $\boldsymbol{\sigma} = (\sigma^x, \sigma^y)$.)

There's something of an irony here. In the original Dirac equation, the 2×2 matrix structure comes about because the electron carries spin. But that's not the origin of the matrix structure in (4.4). Indeed, we've not mentioned spin anywhere in our discussion. Instead, in graphene the emergent "spin" degree of freedom arises from the existence of the two A and B sublattices.

We get a very similar equation in the vicinity of the other Dirac point. Expanding $\mathbf{k} = \mathbf{K}' + \mathbf{q}'$, we get the resulting Hamiltonian

$$H = -v_F \hbar (q_x \sigma^y - q_y \sigma^x)$$

The difference in minus sign is sometimes said to be a different *handedness* or *helicity*. You will learn more about this in the context of high energy physics in the lectures on [Quantum Field Theory](#).

As we mentioned above, we have not yet included the spin of the electron. This is trivial: the discussion above is simply repeated twice, once for spin $|\uparrow\rangle$ and once for spin $|\downarrow\rangle$. The upshot is that the low-energy excitations of graphene are described by four massless Dirac fermions. One pair comes from the spin degeneracy of the electrons; the other from the existence of two Dirac points \mathbf{K} and \mathbf{K}' , sometimes referred to as the *valley degeneracy*.

5. Dynamics of Bloch Electrons

In this section, we look more closely at how electrons moving in a lattice environment react to external forces. We call these electrons *Bloch electrons*. We start by describing how some familiar quantities are redefined for Bloch electrons.

For simplicity, consider an insulator and throw in one further electron. This solitary electron sits all alone in an otherwise unoccupied band. The possible states available to it have energy $E(\mathbf{k})$ where \mathbf{k} lies in the first Brillouin zone. (The energy should also have a further discrete index which labels the particular band the electron is sitting in, but we'll suppress this in what follows). Despite its environment, we can still assign some standard properties to this electron.

5.1 Velocity

The average velocity \mathbf{v} of the electron is

$$\mathbf{v} = \frac{1}{\hbar} \frac{\partial E}{\partial \mathbf{k}} \quad (5.1)$$

First note that this is simply the *group velocity* of a wavepacket (a concept that we've met previously in the lectures on [Electromagnetism](#)). However, the “average velocity” means something specific in quantum mechanics, and to prove (5.1) we should directly compute $\mathbf{v} = \frac{1}{m} \langle \psi | -i\hbar \nabla | \psi \rangle$.

Bloch's theorem ensures that the electron eigenstates take the form

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

with \mathbf{k} in the Brillouin zone. As with the energy, we've suppressed the discrete band index on the wavefunction. The full wavefunction satisfies $H\psi_{\mathbf{k}}(\mathbf{x}) = E(\mathbf{k})\psi_{\mathbf{k}}(\mathbf{x})$, so that $u_{\mathbf{k}}(\mathbf{x})$ obeys

$$H_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{x}) = E(\mathbf{k}) u_{\mathbf{k}}(\mathbf{x}) \quad \text{with} \quad H_{\mathbf{k}} = \frac{\hbar^2}{2m} (-i\nabla + \mathbf{k})^2 + V(\mathbf{x}) \quad (5.2)$$

We'll use a slick trick. Consider the Hamiltonian $H_{\mathbf{k}+\mathbf{q}}$ which we expand as

$$H_{\mathbf{k}+\mathbf{q}} = H_{\mathbf{k}} + \frac{\partial H_{\mathbf{k}}}{\partial \mathbf{k}} \cdot \mathbf{q} + \frac{1}{2} \frac{\partial^2 H_{\mathbf{k}}}{\partial k^i \partial k^j} q^i q^j \quad (5.3)$$

For small \mathbf{q} , we view this as a perturbation of $H_{\mathbf{k}}$. From our results of first order perturbation theory, we know that the shift of the energy eigenvalues is

$$\Delta E = \langle u_{\mathbf{k}} | \frac{\partial H_{\mathbf{k}}}{\partial \mathbf{k}} \cdot \mathbf{q} | u_{\mathbf{k}} \rangle$$

But we also know the exact result: it is simply $E(\mathbf{k} + \mathbf{q})$. Expanding this to first order in \mathbf{q} , we have the result

$$\langle u_{\mathbf{k}} | \frac{\partial H_{\mathbf{k}}}{\partial \mathbf{k}} | u_{\mathbf{k}} \rangle = \frac{\partial E}{\partial \mathbf{k}}$$

But this is exactly what we need. Using the expression (5.2) for $H_{\mathbf{k}}$, the left-hand side is

$$\frac{\hbar^2}{m} \langle u_{\mathbf{k}} | (-i\nabla + \mathbf{k}) | u_{\mathbf{k}} \rangle = \frac{\hbar}{m} \langle \psi_{\mathbf{k}} | -i\hbar\nabla | \psi_{\mathbf{k}} \rangle = \hbar \mathbf{v}$$

This gives our desired result (5.1).

It is perhaps surprising that eigenstates in a crystal have a fixed, average velocity. One might naively expect that the particle would collide with the crystal, bouncing all over the place with a corresponding vanishing average velocity. Yet the beauty of Bloch's theorem is that this is not what happens. The electrons can quite happily glide through the crystal structure.

A Filled Band Carries Neither Current nor Heat

Before we go on, we can use the above result to prove a simple result: a completely filled band does not contribute to the current. This is true whether the filled band is part of an insulator, or part of a metal. (In the latter case, there will also be a partially filled band which will contribute to the current.)

The current carried by each electron is $\mathbf{j} = -e\mathbf{v}$ where $-e$ is the electron charge. From (5.1), the total current of a filled band is then

$$\mathbf{j} = -\frac{2e}{\hbar} \int_{\text{BZ}} \frac{d^3k}{(2\pi)^3} \frac{\partial E}{\partial \mathbf{k}} \quad (5.4)$$

where the overall factor of 2 counts the spin degeneracy. This integral vanishes. This follows because $E(\mathbf{k})$ is a periodic function over the Brillouin zone and the total derivative of any periodic function always integrates to zero.

Alternatively, if the crystal has an *inversion symmetry* then there is a more direct proof. The energy satisfies $E(\mathbf{k}) = E(-\mathbf{k})$, which means that $\partial E(\mathbf{k})/\partial \mathbf{k} = -\partial E(-\mathbf{k})/\partial \mathbf{k}$ and the contributions to the integral cancel between the two halves of the Brillouin zone.

The same argument shows that a filled band cannot transport energy in the form of heat. The heat current is defined as

$$\mathbf{j}_E = 2 \int_{\text{BZ}} \frac{d^3k}{(2\pi)^3} E \mathbf{v} = \frac{1}{\hbar} \int_{\text{BZ}} \frac{d^3k}{(2\pi)^3} \frac{\partial(E^2)}{\partial \mathbf{k}}$$

which again vanishes when integrated over a filled band. This means that the electrons trapped in insulators can conduct neither electricity nor heat. Note, however, that while there is nothing else charged that can conduct electricity, there are other degrees of freedom – in particular, phonons – which can conduct heat.

5.2 The Effective Mass

We define the effective mass tensor to be

$$m_{ij}^* = \hbar^2 \left(\frac{\partial^2 E}{\partial k^i \partial k^j} \right)^{-1}$$

where we should view the right-hand side as the inverse of a matrix.

For simplicity, we will mostly consider this isotropic systems, for which $m_{ij}^* = m^* \delta_{ij}$ and the effective mass of the electron is given by

$$m^* = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2} \right)^{-1} \quad (5.5)$$

where the derivative is now taken in any direction. This definition reduces to something very familiar when the electron sits at the bottom of the band, where we can Taylor expand to find

$$E = E_{\min} + \frac{\hbar^2}{2m^*} |\mathbf{k} - \mathbf{k}_{\min}|^2 + \dots$$

This is the usual dispersion relation or a non-relativistic particle.

The effective mass m^* has more unusual properties higher up the band. For a typical band structure, m^* becomes infinite at some point in the middle, and is negative close to the top of the band. We'll see how to interpret this negative effective mass in Section 5.4.

In most materials, the effective mass m^* near the bottom of the band is somewhere between 0.01 and 10 times the actual mass of the electron. But there are exceptions. Near the Dirac point, graphene has an infinite effective mass by the definition (5.5), although this is more because we've used a non-relativistic definition of mass which is rather daft when applied to graphene. More pertinently, there are substances known, appropriately, as *heavy fermion materials* where the effective electron mass is around a 1000 times heavier than the actual mass.

A Microscopic View on the Effective Mass

We can get an explicit expression for the effective mass tensor m_{ij} in terms of the microscopic electron states. This follows by continuing the slick trick we used above, now thinking about the Hamiltonian (5.3) at second order in perturbation theory. This time, we find the inverse mass matrix is given by

$$(m^*)_{ij}^{-1} = \frac{\delta_{ij}}{m} + \frac{1}{m^2} \sum_{n \neq n'} \frac{\langle \psi_{n,\mathbf{k}} | p_i | \psi_{n',\mathbf{k}} \rangle \langle \psi_{n,\mathbf{k}} | p_j | \psi_{n',\mathbf{k}} \rangle - \text{h.c.}}{E_n(\mathbf{k}) - E_{n'}(\mathbf{k})}$$

where n labels the band of each state. Note that the second term takes the familiar form that arises in second order perturbation theory. We see that, microscopically, the additional contributions to the effective mass come from matrix elements between different bands. Nearby bands of a higher energy give a negative contribution to the effective mass; nearby bands of a lower energy give a positive contribution.

5.3 Semi-Classical Equation of Motion

Suppose now that we subject the electron to an external potential force of the form $\mathbf{F} = -\nabla U(\mathbf{x})$. The correct way to proceed is to add $U(\mathbf{x})$ to the Hamiltonian and solve again for the eigenstates. However, in many circumstances, we can work semi-classically. For this, we need that $U(\mathbf{x})$ is small enough that it does not distort the band structure and, moreover, does not vary greatly over distances comparable to the lattice spacing.

We continue to restrict attention to the electron lying in a single band. To proceed, we should think in terms of wavepackets, rather than plane waves. This means that the electron has some localised momentum \mathbf{k} and some localised position \mathbf{x} , within the bounds allowed by the Heisenberg uncertainty relation. We then treat this wavepacket as if it was a classical particle, where the position \mathbf{x} and momentum $\hbar\mathbf{k}$ depend on time. This is sometimes referred to as a *semi-classical* approach.

The total energy of this semi-classical particle is $E(\mathbf{k}) + U(\mathbf{x})$ where $E(\mathbf{k})$ is the band energy. The position and momentum evolve such that the total energy is conserved. This gives

$$\frac{d}{dt} \left(E(\mathbf{k}(t)) + U(\mathbf{x}(t)) \right) = \frac{\partial E}{\partial \mathbf{k}} \cdot \frac{d\mathbf{k}}{dt} + \nabla U \cdot \frac{d\mathbf{x}}{dt} = \mathbf{v} \cdot \left(\hbar \frac{d\mathbf{k}}{dt} + \nabla U \right) = 0$$

which is satisfied when

$$\hbar \frac{d\mathbf{k}}{dt} = -\nabla U = \mathbf{F} \tag{5.6}$$

This should be viewed as a variant of Newton’s equation, now adapted to the lattice environment. In fact, we can make it look even more similar to Newton’s equation. For an isotropic system, the effective “mass times acceleration” is

$$m^* \frac{d\mathbf{v}}{dt} = \frac{m^*}{\hbar} \frac{d}{dt} \left(\frac{\partial E}{\partial \mathbf{k}} \right) = \frac{m^*}{\hbar} \left(\frac{d\mathbf{k}}{dt} \cdot \frac{\partial}{\partial \mathbf{k}} \right) \frac{\partial E}{\partial \mathbf{k}} = \hbar \frac{d\mathbf{k}}{dt} = \mathbf{F} \quad (5.7)$$

where you might want to use index notation to convince yourself of the step in the middle where we lost the effective mass m^* . It’s rather nice that, despite the complications of the lattice, we still get to use some old equations that we know and love. Of course, the key to this was really the definition (5.5) of what we mean by effective mass m^* .

An Example: Bloch Oscillations

Consider a Bloch electron, exposed to a constant electric field \mathcal{E} . The semi-classical equation of motion is

$$\hbar \dot{\mathbf{k}} = -e\mathcal{E} \quad \Rightarrow \quad k(t) = k(0) - \frac{e\mathcal{E}}{\hbar} t$$

So the crystal momentum \mathbf{k} increases linearly. At first glance, this is unsurprising. But it leads to a rather surprising effect. This is because \mathbf{k} is really periodic, valued in the Brillouin zone. Like a character in a 1980s video game, when the electron leaves one edge of the Brillouin zone, it reappears on the other side. 

 We can see what this means in terms of velocity. For a typical one-dimensional band structure shown on the right, the velocity $\mathbf{v} \sim \mathbf{k}$ in the middle of the band, but $\mathbf{v} \sim -\mathbf{k}$ as the particle approaches the edge of the Brillouin zone. In other words, a constant electric field gives rise to an oscillating velocity, and hence an oscillating current! This surprising effect is called *Bloch oscillations*.

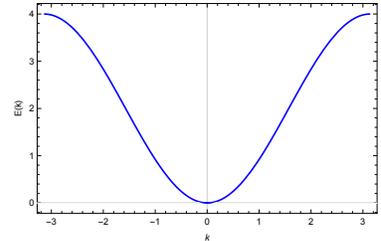


Figure 27:

As an example, consider a one-dimensional system with a tight-binding form of band structure

$$E = -C \cos(ka)$$

Then the velocity in a constant electric field oscillates as

$$v(k) = \frac{Ca}{\hbar} \sin(ka) = -\frac{Ca}{\hbar} \sin \left(\frac{e\mathcal{E}a}{\hbar} t \right)$$

The Bloch frequency is $\omega = e\mathcal{E}a/\hbar$. If we construct a wavepacket from several different energy eigenstates, then the position of the particle will similarly oscillate back and forth. This effect was first predicted by Leo Esaki in 1970.

Bloch oscillations are somewhat counterintuitive. They mean that a DC electric field applied to a pure crystal does *not* lead to a DC current! Yet we've all done experiments in school where we measure the DC current in a metal! This only arises because a metal is not a perfect crystal and the electrons are scattered by impurities or thermal lattice vibrations (phonons) which destroy the coherency of Bloch oscillations and lead to a current.

Bloch oscillations are delicate. The system must be extremely clean so that the particle does not collide with anything else over the time necessary to see the oscillations. This is too much to ask in solid state crystals. However, Bloch oscillations have been observed in other contexts, such as cold atoms in an artificial lattice. The time variation of the velocity of Caesium atoms in an optical lattice is shown in the figure⁴.

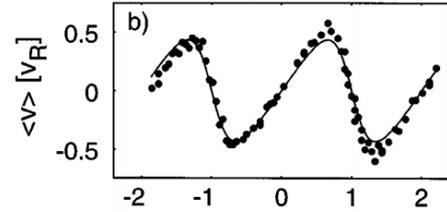


Figure 28:

5.4 Holes

Consider a totally filled band, and remove one electron. We're left with a vacancy in the otherwise filled band. In a zen-like manoeuvre, we ascribe properties to the absence of the particle. Indeed, as we will now see, this vacancy moves as if it were itself an independent particle. We call this particle a *hole*.

Recall that our definition (5.5) means that the effective mass of electrons is negative near the top of the band. Indeed, expanding around the maximum, the dispersion relation for electrons reads

$$E(\mathbf{k}) = E_{\max} + \frac{\hbar^2}{2m^*} |\mathbf{k} - \mathbf{k}_{\max}|^2 + \dots$$

and the negative effective mass $m^* < 0$ ensures that electrons have less energy as they move away from the maximum.

Now consider filling all states except one. As the hole moves away from the maximum, it costs more energy (because we're subtracting less energy!). This suggests that we should write the energy of the hole as

$$E_{\text{hole}}(\mathbf{k}) = -E(\mathbf{k}) = -E_{\max} + \frac{\hbar^2}{2m_{\text{hole}}^*} |\mathbf{k} - \mathbf{k}_{\max}|^2 + \dots$$

where

$$m_{\text{hole}}^* = -m^*$$

⁴This data is taken from “*Bloch Oscillations of Atoms in an Optical Potential*” by Dahan et. al., Phys. Rev. Lett. vol 76 (1996), which reported the first observation of this effect.

so that the effective mass of the hole is positive near the top of the band, but becomes negative if the hole makes it all the way down to the bottom.

The hole has other properties. Suppose that we take away an electron with momentum \mathbf{k} . Then the resulting hole can be thought of as having momentum $-\mathbf{k}$. This suggests that we define

$$\mathbf{k}_{\text{hole}} = -\mathbf{k} \quad (5.8)$$

However, the velocity of the hole is the same as that of the missing electron

$$\mathbf{v}_{\text{hole}} = \frac{1}{\hbar} \frac{\partial E_{\text{hole}}}{\partial \mathbf{k}_{\text{hole}}} = \frac{1}{\hbar} \frac{\partial E}{\partial \mathbf{k}} = \mathbf{v}$$

This too is intuitive, since the hole is moving in the same direction as the electron that we took away.

The definitions above mean that the hole obeys the Newtonian force law with

$$m_{\text{hole}}^* \frac{d\mathbf{v}_{\text{hole}}}{dt} = -\mathbf{F} = \mathbf{F}_{\text{hole}} \quad (5.9)$$

At first sight, this is surprising: the hole experiences an opposite force to the electron. But there's a very simple interpretation. The force that we typically wish to apply to our system is an electric field \mathcal{E} which, for an electron, gives rise to

$$\mathbf{F} = -e\mathcal{E}$$

The minus sign in (5.9) is simply telling us that the hole should be thought of as carrying charge $+e$, the opposite of the electron,

$$\mathbf{F}_{\text{hole}} = +e\mathcal{E}$$

We can also reach this same conclusion by computing the current. We saw in (5.4) that a fully filled band carries no current. This means that the current carried by a partially filled band is

$$\mathbf{j} = -2e \int_{\text{filled}} \frac{d^3k}{(2\pi)^3} \mathbf{v}(\mathbf{k}) = +2e \int_{\text{unfilled}} \frac{d^3k}{(2\pi)^3} \mathbf{v}(\mathbf{k})$$

The filled states are electrons carrying charge $-e$; the unfilled states are holes, carrying charge $+e$.

Finally, it's worth mentioning that the idea of holes in band structure provides a fairly decent analogy for anti-matter in high-energy physics. There too the electron has a positively charged cousin, now called the positron. In both cases, the two particles can come together and annihilate. In solids, this releases a few eV of energy, given by the gap between bands. In high-energy physics, this releases a million times more energy, given by the rest mass of the electron.

5.5 Drude Model Again

The essence of Bloch's theorem is that electrons can travel through perfect crystals unimpeded. And yet, in the real world, this does not happen. Even the best metals have a resistance, in which any current degrades and ultimately relaxes to zero. This happens because metals are not perfect crystals, and the electrons collide with impurities and vacancies, as well as thermally vibrations called phonons.

We can model these effects in our semi-classical description by working with the electron equation of motion called the *Drude model*

$$m^* \dot{\mathbf{v}} = -e\mathcal{E} - \frac{m^*}{\tau} \mathbf{v} \quad (5.10)$$

Here \mathcal{E} is the applied electric field and τ is the *scattering time*, which should be thought of as the average time between collisions.

We have already met the Drude model in the lectures on [Electromagnetism](#) when we tried to describe the conductivity in metals classically. We have now included the quantum effects of lattices and the Fermi surface yet, rather remarkably, the equation remains essentially unchanged. The only difference is that the effective mass m^* will depend on \mathbf{k} , and hence on \mathbf{v} , if the electron is not close to the minimum of the band.

In equilibrium, the velocity of the electron is

$$\mathbf{v} = -\frac{e\tau}{m^*} \mathcal{E} \quad (5.11)$$

The proportionality constant is called the *mobility*, $\mu = |e\tau/m^*|$. The total current density $\mathbf{j} = -en\mathbf{v}$ where n is the density of charge carriers. The equation (5.11) then becomes $\mathbf{j} = \sigma\mathcal{E}$ where σ is the conductivity,

$$\sigma = \frac{e^2\tau n}{m^*} \quad (5.12)$$

We also define the resistivity $\rho = 1/\sigma$. This is the same result that we found in our earlier classical analysis, except the mass m is replaced by the effective mass m^* .

There is, however, one crucial difference that the existence of the Fermi surface has introduced. When bands are mostly unfilled, it is best to think of the charge carriers in terms of negatively charged electrons, with positive effective mass m^* . But when bands are mostly filled, it is best to think of the charge carriers in terms of positively charged holes, also with positive mass m_{hole}^* . In this case, we should replace the Drude model (5.10) with the equivalent version for holes,

$$m_{\text{hole}}^* \dot{\mathbf{v}} = +e\mathcal{E} - \frac{m_{\text{hole}}^*}{\tau} \mathbf{v} \quad (5.13)$$

This means that certain materials can appear to have positive charge carriers, even though the only things actually moving are electrons. The different sign in the charge carrier doesn't show up in the conductivity (5.12), which depends on e^2 . To see it, we need to throw in an extra ingredient.

Hall Resistivity

The standard technique to measure the charge of a material is to apply a magnetic field \mathbf{B} . Classically, particles of opposite charges will bend in opposite directions, perpendicular to \mathbf{B} . In a material, this results in the classical *Hall effect*.

We will discuss the motion of Bloch electrons in a magnetic field in much more detail in Section ?? (And we will discuss the Hall effect in much more detail in [other lectures](#).) Here, we simply want to show how this effect reveals the difference between electrons and holes. For electrons, we adapt the Drude model (5.10) by adding a Lorentz force,

$$m^* \dot{\mathbf{v}} = -e(\mathcal{E} + \mathbf{v} \times \mathbf{B}) - \frac{m^*}{\tau} \mathbf{v}$$

We once again look for equilibrium solutions with $\dot{\mathbf{v}} = 0$. Writing $\mathbf{j} = -nev$, we now must solve the vector equation

$$\frac{1}{ne} \mathbf{j} \times \mathbf{B} + \frac{m^*}{ne^2\tau} \mathbf{j} = \mathbf{E}$$

The solution to this is

$$\mathbf{E} = \rho \mathbf{j}$$

where the resistivity ρ is now a 3×3 matrix. If we take $\mathbf{B} = (0, 0, B)$, then we have

$$\rho = \begin{pmatrix} \rho_{xx} & \rho_{xy} & 0 \\ -\rho_{xy} & \rho_{xx} & 0 \\ 0 & 0 & \rho_{xx} \end{pmatrix}$$

where the diagonal, *longitudinal resistivity* is $\rho_{xx} = 1/\sigma$ where σ is given in (5.12). The novelty is the off-diagonal, Hall resistivity

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

We often define the *Hall coefficient* R_H as

$$R_H = \frac{\rho_{xy}}{B} = \frac{1}{ne}$$

This, as promised, depends on the charge e . This means that if we were to repeat the above analysis for holes (5.13) rather than electrons, we would find a Hall coefficient which differs by a minus sign.

There are metals – such as beryllium and magnesium – whose Hall coefficient has the “wrong sign”. We drew the Fermi surface for beryllium in Section 3.2; it contains both electrons and holes. In this case, we should add to two contributions with opposite signs. It turns out that the holes are the dominant charge carrier.