

Lecture 1: free electron gas (Ch6)

- **Electrons as waves: motivation**
- **1D infinite potential well**
- **Putting N electrons into infinite potential well**
- **Fermi-dirac distribution**

Up to this point your solid state physics course has not discussed electrons explicitly, but electrons are the first offenders for most useful phenomena in materials. When we distinguish between metals, insulators, semiconductors, and semimetals, we are discussing conduction properties of the electrons. When we notice that different materials have different colors, this is also usually due to electrons. Interactions with the lattice will influence electron behavior (which is why we studied them independently first), but electrons are the primary factor in materials' response to electromagnetic stimulus (light or a potential difference), and this is how we usually use materials in modern technology.

Electrons as particles

A classical treatment of electrons in a solid, called the Drude model, considers valence electrons to act like billiard balls that scatter off each other and off lattice imperfections (including thermal vibrations). This model introduces important terminology and formalism that is still used to this day to describe materials' response to electromagnetic radiation, but it is not a good physical model for electrons in most materials, so we will not discuss it in detail.

Electrons as waves

In chapter 3, which discussed metallic bonding, the primary attribute was that **electrons are delocalized**.

In quantum mechanical language, when something is delocalized, it means that its position is ill defined which means that its **momentum** is more well defined. An object with a well defined momentum but an ill-defined position is a plane-wave, and in this chapter we will treat electrons like plane waves, defined by their momentum.

Another important constraint at this point is that electrons **do not** interact with each other, except for pauli exclusion (that is, two electrons cannot be in the same state, where a state is defined by a momentum and a spin).

To find the momentum and energy of the available quantum states, we solve a particle-in-a-box problem, where the box is defined by the boundaries of the solid.

Particle in a Box in one dimension

An electron of mass m is confined to a one-dimensional box of length L with infinitely high walls.

We need to solve Schrodinger's equation with the boundary conditions determined by the box

$$\mathcal{H}\psi_n = \frac{-\hbar^2}{2m} \frac{d^2\psi_n}{dx^2} = \epsilon_n\psi_n$$

Here, ψ_n is the wavefunction of the n -th solution, and ϵ_n is the energy associated with that eigenstate.

The boundary conditions (infinitely high walls) dictate that:

$$\psi_n(x = 0) = 0$$

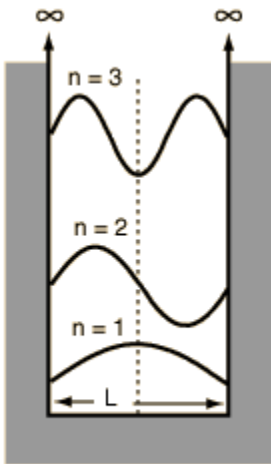
$$\psi_n(x = L) = 0$$

For all n.

A solution for the wavefunction which satisfies Schrodinger's equation and the boundary conditions is:

$$\psi_n = A \sin\left(\frac{n\pi x}{L}\right)$$

Where A is a constant. Here, each solution wavefunction corresponds to an integer number of half-wavelengths fitting inside the box: $\frac{1}{2}n\lambda_n = L$



x = 0 at left wall of box.

Now, plug our 'guess' back into Schrodinger's equation to get the eigenenergies:

$$\frac{\hbar^2}{2m} \frac{d^2\psi_n}{dx^2} = A \frac{\hbar^2}{2m} \left(\frac{n\pi}{L}\right)^2 \sin\left(\frac{n\pi x}{L}\right) = A \epsilon_n \sin\left(\frac{n\pi x}{L}\right)$$

$$\epsilon_n = \frac{\hbar^2}{2m} \left(\frac{n\pi}{L}\right)^2$$

Each energy level, n, defines a 'state' in which we can put two electrons into, one spin up and one spin down. Here is where the approximation/assumption comes in. We are assuming that our one-particle wavefunction is applicable to a many-electron system—that we do not change the wavefunction of one electron when we add others to the box. It turns out that this approximation works reasonably well for some

simple metals like sodium or copper, and the formalism developed here is an excellent framework for describing real many-electron systems where our hopeful assumption doesn't necessarily hold. For now, we are also assuming that the lattice is not there.

Lets say we have N electrons and we want to place them into available eigenstates, defined by N. There are two rules we need to follow.

- Only two electrons per n, one spin up and one spin down (pauli exclusion) (note: if we were not using electrons but some other fermion with a different spin, the number of electrons in each energy eigenstate would change accordingly)
- Lower energy levels get filled up first, sort of like pouring water into a container. We are looking to describe the *ground state* configuration, and you won't get to the ground state if you fill up higher energy levels first.

The **Fermi level** (ϵ_F) is defined as the highest energy level you fill up to, once you have used up all your electrons.

$$\epsilon_F = \frac{\hbar^2}{2m} \left(\frac{N\pi}{2L}\right)^2$$

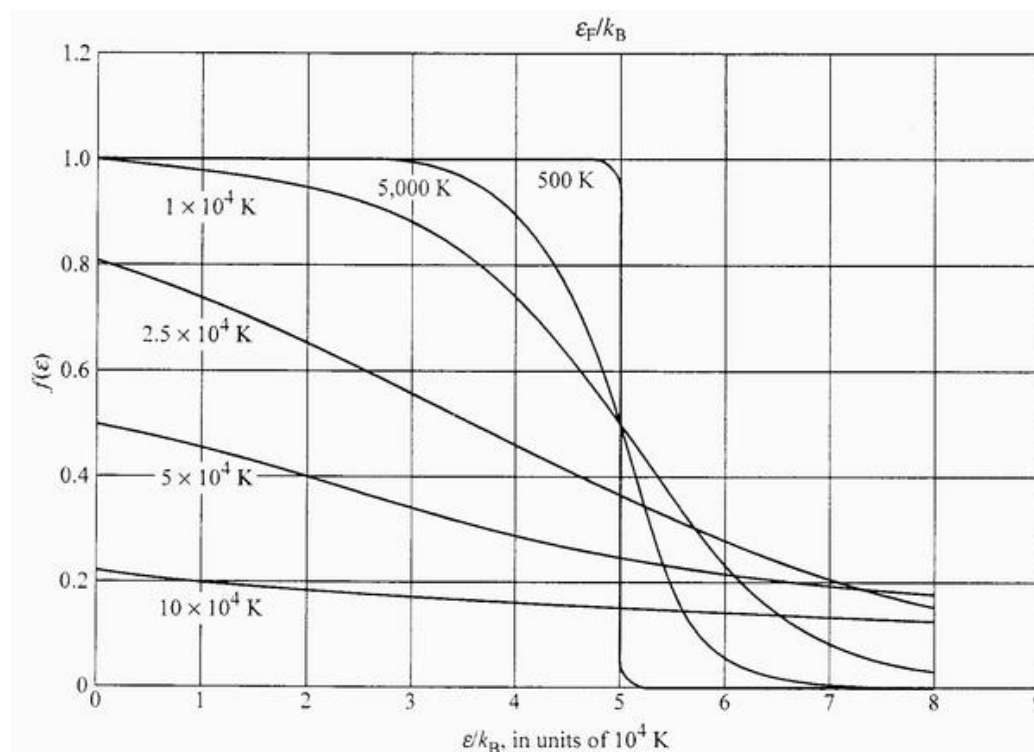
(we simply plugged in $N/2$ for n , since we use up two electrons for each state)

Effect of temperature

What has been described thus far is the zero-temperature ground state of a collection of electrons confined to a box in 1D. What finite temperature does is it slightly modifies the occupation probability for energies close to the Fermi level, and this is encompassed in the Fermi-Dirac distribution (also called the Fermi function). The probability that a given energy level, ϵ , is occupied by electrons at a given temperature is given by:

$$f(\epsilon) = \frac{1}{e^{(\epsilon-\mu)/k_B T} + 1}$$

The quantity μ is the **chemical potential** and it ensures that the number of particles come out correctly. At $T=0$, $\mu = \epsilon_F$, and at temperatures we typically encounter in solid state physics, it does not differ too much from that value.

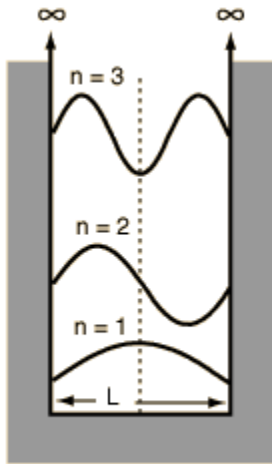


At zero temperature, the Fermi-Dirac distribution represents a sharp cutoff between states that are occupied by electrons and states that are unoccupied. At higher temperature, the Fermi-Dirac function introduces a small probability that states with energy higher than the chemical potential contain an electron and a symmetric small probability that states below the chemical potential lack an electron.

Lecture 2: free electron gas in 3D

- Review free electron gas in 1D
- Free electron gas in 3D
- Concepts: Density of states, Fermi momentum, Fermi velocity

Review: free electron gas in 1D



$x = 0$ at left wall of box.

- Why gas? Electrons in a solid are modeled as an ensemble of weakly interacting, identical, indistinguishable particles (also called a Fermi gas)
- Why this model for a metal? Electrons are stuck inside the chunk of the metal (box) and cannot escape (infinitely high walls) and are delocalized (wavelike)
- Infinite potential well \rightarrow boundary conditions select specific wavelengths of plane waves, those which fit a half-integer number of wavelengths into the length (L) of the box

$$\frac{1}{2}n\lambda_n = L$$

- These allowed wavelengths are associated with allowed momenta and energies

$$k_n = 2\pi/\lambda_n$$

$$\epsilon_n = \frac{\hbar^2}{2m}(k_n)^2 = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2$$

- Infinite potential well is exactly solvable for **one** particle. For a many-electron system (a crystalline solid), we make the ansatz that the series of one-particle solutions stated above still hold
 - Fill the lowest energy states first, until all N electrons are used up
 - Each state (defined by n in 1D) can hold two electrons, one spin up and one spin down
- Fermi energy (or Fermi level)—The highest occupied energy at $T=0$. In 1D it is given by:

$$\epsilon_F = \frac{\hbar^2}{2m}\left(\frac{N\pi}{2L}\right)^2$$

Free electron gas in three dimensions

This toy problem turns out to be applicable to many simple metals such as sodium or copper, and it is a generalization of the infinite potential well to three dimensions.

In three dimensions, the free particle Schrodinger equation is:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi_k(r) = \epsilon_k\psi_k(r)$$

The wavefunctions are marked by k instead of by n , and we will see why in a moment.

If we use boundary conditions that are a 3D generalization of the boundary conditions in 1D, we get standing wave solutions of the form:

$$\psi_{n_x, n_y, n_z}(\mathbf{r}) = A \sin\left(\frac{\pi n_x x}{L}\right) \sin\left(\frac{\pi n_y y}{L}\right) \sin\left(\frac{\pi n_z z}{L}\right)$$

$$\epsilon_{n_x, n_y, n_z} = \frac{\hbar^2}{2m} \left(\frac{\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2)$$

Where n_x, n_y, n_z are positive integers, and every eigenstate is defined by a unique number of half-periods of a sine wave in each of the x, y, and z direction (but not necessarily by a unique energy, because for example $(n_x, n_y, n_z) = (1, 2, 1)$ will have the same energy as $(n_x, n_y, n_z) = (1, 1, 2)$).

At this point, it is helpful to start over with a different formalism.

We consider plane wave wavefunctions of the form

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

And periodic boundary conditions of the form

$$\psi(x + L, y, z) = \psi(x, y, z)$$

$$\psi(x, y + L, z) = \psi(x, y, z)$$

$$\psi(x, y, z + L) = \psi(x, y, z)$$

Plugging the first one into the wavefunction we get:

$$e^{i(k_x(x+L)+k_y y+k_z z)} = e^{i(k_x x+k_y y+k_z z)}$$

$$e^{ik_x L} = 1$$

$$k_x = 0, \pm \frac{2\pi}{L}, \pm \frac{4\pi}{L}, \dots$$

And similar for k_y and k_z .

Plugging the plane wave wavefunction into schrodinger's equation we get:

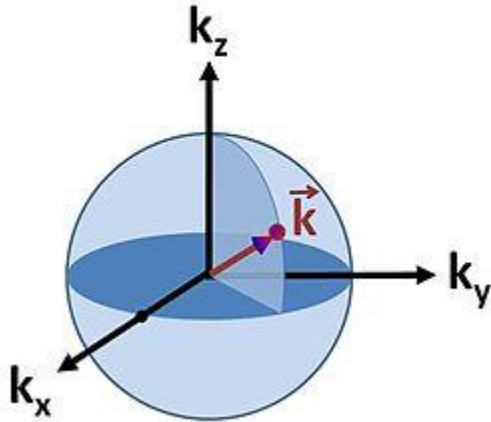
$$\frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2) = \epsilon_k = \frac{\hbar^2 k^2}{2m}$$

This is almost equivalent to the version of the eigenenergies that we had earlier, as that the values that k_x can take on can be expressed as $2n_x\pi/L$ (and similar for k_y and k_z). The factor of 2 comes from the fact that only the even sine wave solutions satisfy periodic boundary conditions. On the surface it seems like these two solutions give contradictory results, but what really matters for a materials electronic properties is what happens close to the Fermi energy, and you can work out that you make up the factor of 4 (in energy) with a factor of 2 shorter k_F (which will be defined shortly...)

As before, we take our N electrons and put them into the available states, filling lowest energy first. In 3D this is trickier because multiple states may have the same energy, even though they are marked by different k_x, k_y, k_z . In 3D, our rules for filling up electrons are:

- Every state is defined by a unique quantized value of (k_x, k_y, k_z)
- Every state can hold one spin up and one spin down electrons

- Fill low energy states first. In 3D, this corresponds to filling up a sphere in k space, one 'shell' at a time. Each shell is defined by a radius k , where $k^2 = k_x^2 + k_y^2 + k_z^2$, and every state in the shell has the same energy, although different combinations of k_x, k_y, k_z



When we have used up all our electrons, we are left with a filled sphere in k space with radius k_F (called the Fermi momentum) such that

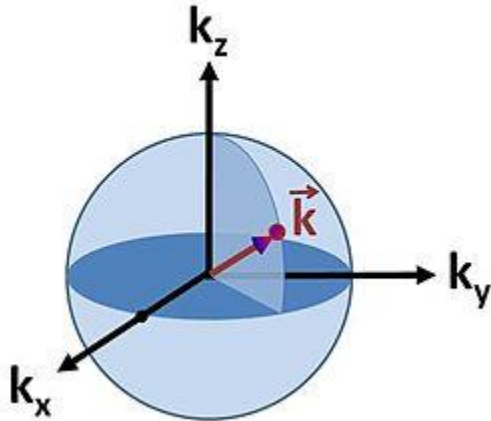
$$\epsilon_F = \frac{\hbar^2}{2m} k_F^2$$

This sphere in k-space has a volume $\frac{4}{3}\pi k_F^3$ and it is divided into voxels of volume $\left(\frac{2\pi}{L}\right)^3$

If we divide the total volume of the sphere by the volume of each 'box' and account for the fact that each box holds 2 electrons, we get back how many electrons we put in:

$$2 * \frac{\frac{4}{3}\pi k_F^3}{\left(\frac{2\pi}{L}\right)^3} = N = V k_F^3 / 3\pi^2$$

Lecture 3: free electron gas in 3D



Model: electron wavefunction is plane wave which obeys periodic boundary conditions (repeats every L in all dimensions, where L is the nominal size of the chunk of metal)

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} = e^{i(k_x x + k_y y + k_z z)}$$

With periodic boundary conditions, \mathbf{k} can only take on certain allowed values:

$$k_x, k_y, k_z = 0, \pm \frac{2\pi}{L}, \pm \frac{4\pi}{L}, \dots$$

These correspond to allowed values of energy:

$$\epsilon_{\mathbf{k}} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2) = \frac{\hbar^2 k^2}{2m}$$

We have N electrons to put into the allowed states, lowest energy first. If we move to coordinate space (k_x, k_y, k_z) , states with the same energy are located on a sphere with radius $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$

When we have used up all our electrons, we are left with a filled sphere in \mathbf{k} space with radius k_F (called the Fermi momentum) such that

$$\epsilon_F = \frac{\hbar^2}{2m} k_F^2$$

This sphere in \mathbf{k} -space has a volume $\frac{4}{3}\pi k_F^3$ and it is divided into voxels of volume $\left(\frac{2\pi}{L}\right)^3$

If we divide the total volume of the sphere by the volume of each 'box' and account for the fact that each box holds 2 electrons, we get back how many electrons we put in:

$$2 * \frac{\frac{4}{3}\pi k_F^3}{\left(\frac{2\pi}{L}\right)^3} = N = V k_F^3 / 3\pi^2$$

Here, $V = L^3$ is the volume of the solid. We can use this relationship to solve for k_F and show that it depends on electron density (N/V)

$$k_F = \left(\frac{3\pi^2 N}{V}\right)^{1/3}$$

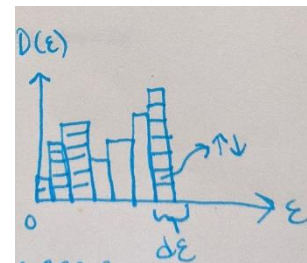
Plugging this back into the expression for ϵ_F we get:

$$\epsilon_F = \frac{\hbar^2}{2m} \left(\frac{3\pi^2 N}{V}\right)^{2/3}$$

At absolute zero, the Fermi sphere has a hard boundary between occupied and unoccupied states. At higher temperature, this boundary becomes fuzzier with increasing occupation permitted outside the initial boundary (think of a rocky planet like earth vs a gaseous planet like Jupiter). The width of this fuzziness is determined by the width of the Fermi-Dirac distribution at that temperature, and it is roughly proportional to $k_B T$. Notably, the vast majority of electrons in the Fermi gas are completely inert because they are buried deep inside the sphere. Only electrons close to the Fermi level are affected by temperature and participate in conduction. This is quite contrary to the conclusions of particle-like treatments of electrons in a metal which assume that all valence electrons participate in electronic properties.

As with phonons, the density of states is a useful quantity for electrons.

I like to think of Density of States as a series of “boxes” where electrons can live. Each box is defined by the coordinates which distinguish one electron from another. In the case of a 3D free electron gas, each box is defined by unique k_x, k_y, k_z and spin. Where the density comes in is at each energy interval $d\epsilon$ we consider ‘how many ‘boxes’ are there?’



It is defined as:

$$D(\epsilon) \equiv \frac{dN}{d\epsilon}$$

We can find it by expressing N in terms of ϵ and taking a derivative. We begin by considering a sphere in k -space with an arbitrary radius k and asking how many electrons that will hold

$$N(k) = V k^3 / 3\pi^2$$

The relationship between energy and momentum in a free electron gas is pretty straightforward too (unlike with phonons):

$$\epsilon = \frac{\hbar^2 k^2}{2m}$$

Solving for k , and plugging in above we get

$$N(\epsilon) = \frac{V}{3\pi^2} \left(\frac{2m\epsilon}{\hbar^2} \right)^{3/2}$$

Now we can just take the derivative with respect to energy and get:

$$D(\epsilon) \equiv \frac{dN}{d\epsilon} = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon^{1/2}$$

Thus, the density of electron states in 3D is a function of energy. If you have more electrons, you will end up with a higher density of states **at the Fermi energy**. It should be noted that as with phonons, the functional form of the density of states will depend on if you are thinking of a 1D, 2D, or 3D system.

We use density of states to calculate aggregate properties of a free electron gas, and it comes into play most situations we have an integral over energy. Some examples we will use in this class/HW are:

Total number of electrons: $\int_0^\infty d\epsilon D(\epsilon)f(\epsilon, T) = N$ (total # of electrons is determined by number of states at each energy multiplied by probability that each state is occupied at temperature T, integrated over all energy)

Total energy of electrons: $\int_0^\infty d\epsilon \epsilon D(\epsilon)f(\epsilon, T) = U_{tot}$ (total energy of electrons is determined by energy multiplied by number of states at each energy multiplied by probability that each state is occupied at temperature T, integrated over all energy)

Electron velocity

There are two ways of extracting an electrons' velocity in a Fermi gas.

- From the derivative of the energy vs k (equivalent to what we did for phonons): $v_g = \frac{1}{\hbar} \frac{\partial \epsilon_k}{\partial k} = \hbar k/m$
- By representing the linear momentum operator as $\mathbf{p} = -i\hbar\nabla$ and applying this to the plane-wave wavefunction to get $\mathbf{p} = \hbar\mathbf{k}$ and equating to $m\mathbf{v}$ to get $\mathbf{v} = \hbar\mathbf{k}/m$

The velocity of electrons at the fermi energy is called the Fermi velocity (v_F) and it is given by:

$$v_F = \frac{\hbar k_F}{m} = \frac{\hbar}{m} \left(\frac{3\pi^2 N}{V} \right)^{1/3}$$

Example: Sodium metal

Sodium metal is one electron beyond a full shell, so it has one valence electron per atom that becomes delocalized and contributes to the sea of conduction electrons that we have been representing as plane waves. Lets calculate some of the parameters we have discussed here

$$\text{Fermi momentum: } k_F = \left(\frac{3\pi^2 N}{V} \right)^{1/3}$$

This depends on the electron concentration. Sodium takes on a BCC structure with a conventional (cubic) unit cell dimension of 4.29Å.

There are 2 valence electrons in this conventional cell, so $N/V = 2.53 \times 10^{28}/m^3$

This gives $k_F = 9.1 \times 10^9 m^{-1}$

From this, we can get the Fermi energy: $\epsilon_F = \frac{\hbar^2}{2m} k_F^2$

$$\epsilon_F = 5.03 \times 10^{-19} \text{ joules}$$

It is convenient to divide by a factor of the electron charge to put energy in units of **electron volts (eV)**.

$$\epsilon_F = 3.1 \text{ eV}$$

For comparison, $k_B T$ at room temperature (300K) is $4.14 \times 10^{-21} \text{ Joules}$ or 0.026 eV. Thus, the energy scale of the temperature fuzzing is <1% of the highest energy level at room temperature.

Another way to think about this is to convert the Fermi energy to a Fermi temperature (T_F) by dividing by the Boltzmann constant.

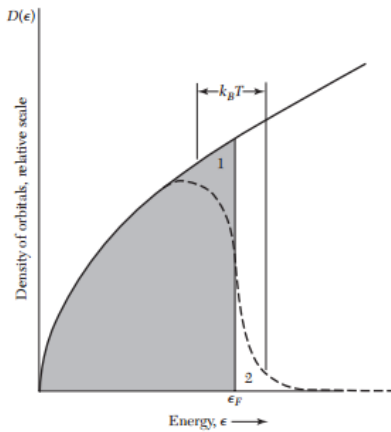
$$T_F = \frac{\epsilon_F}{k_B} = 36,342 \text{ K}$$

Physically, the Fermi temperature for a fermi gas is the temperature when the fermions begin to act like classical particles because they do not have to worry about available states already being occupied by electrons. For sodium, the Fermi temperature is waaaay above the melting temperature.

Finally, let's calculate the Fermi velocity for sodium:

$$v_F = \frac{\hbar}{m} \left(\frac{3\pi^2 N}{V} \right)^{1/3} = 1.05 \times 10^6 \text{ m/s}$$

This is $\sim 1/300$ the speed of light, so electrons would get places quite quickly if they didn't scatter.



Effect of temperature

Temperature introduces a 'cutoff' by the Fermi-dirac function

$$f(\epsilon) = \frac{1}{e^{(\epsilon-\mu)/k_B T} + 1}$$

Such that some states with $\epsilon > \epsilon_F \sim \mu$ can be occupied and some states with $\epsilon < \epsilon_F \sim \mu$. Temperature only affects states roughly within $k_B T$ of the Fermi energy. Another way to think of the effect of temperature is the fuzzing out of the boundary of the Fermi surface.

Lecture 4: Heat capacity of free electron gas

- Qualitative result
- Quantitative derivation
- Electron and lattice heat capacity

Heat Capacity of free electron gas

In chapter 5, you learned about lattice heat capacity—how inputting energy into a solid raises the temperature by exciting more vibrational modes. However, in metals, particularly at low temperature, this is not the whole story, because electrons can absorb heat as well.

Heat capacity: amount of energy that you must add to raise temperature by one unit (e.g. 1 K)

Qualitative derivation

In a free electron gas, only electrons with energy within $\sim k_B T$ of the Fermi level do anything. This represents a small fraction of the total electrons N , given by NT/T_F where T_F is the Fermi temperature which is usually $\sim 10^4 K$, well above the melting point of metals.

Thus, the total electronic thermal kinetic energy when electrons are heated from 0 to temperature T is

$$U_{el} \approx \left(\frac{NT}{T_F}\right) k_B T$$

The heat capacity is found from the temperature derivative:

$$C_{el} = \frac{\partial U}{\partial T} \approx N k_B T / T_F$$

This sketch of a derivation is intended only to achieve the proper temperature dependence: $C_{el} \propto T$, which we will show more rigorously in the next section

Quantitative derivation

This derivation of electron heat capacity is applicable to the regime when a Fermi gas **does not** behave like a classical gas—when $k_B T \ll \epsilon_F$

The change in internal energy when electrons are heated up to temperature T from 0K is given by:

$$\Delta U = U(T) - U(0) = \int_0^\infty d\epsilon \epsilon D(\epsilon) f(\epsilon) - \int_0^{\epsilon_F} d\epsilon \epsilon D(\epsilon)$$

Where $f(\epsilon)$ is the Fermi function, $f(\epsilon) = \frac{1}{e^{(\epsilon-\mu)/k_B T} + 1}$, which describes the occupation probability of a given energy level. It is equal to 1 for $\epsilon \ll \mu$ and 0 for $\epsilon \gg \mu$ and something in between 0 and 1 for $|\epsilon - \mu| \sim k_B T$. The parameter μ is called the “chemical potential”, and its value is temperature dependent and close to ϵ_F for most temperatures one might realistically encounter.

And $D(\epsilon)$ is the density of states, where for a 3D free electron gas, $D(\epsilon) = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \epsilon^{1/2}$

The integral terms above take each energy slice, multiply it by how many electrons have that energy (via the density of states multiplied by the Fermi function), and sum up over all the energies available to an electron. The second integral truncates at $\epsilon = \epsilon_F$ at its upper bound because at $T=0$, $\mu = \epsilon_F$, and the Fermi function is a step function which is equal to zero for $\epsilon > \epsilon_F$.

The Fermi energy is determined by the number of electrons and there are two ways to express this similarly to the integrals above:

$$N = \int_0^{\infty} d\epsilon D(\epsilon)f(\epsilon) = \int_0^{\epsilon_F} d\epsilon D(\epsilon)$$

The right-most integral is the total number of electrons at zero temperature, and the other integral is the total number of electrons at finite temperature. They must be equal since electrons (unlike phonons) cannot be spontaneously created.

We now multiply both integrals by ϵ_F , which is a constant. This is just a mathematical trick.

$$\int_0^{\infty} d\epsilon \epsilon_F D(\epsilon)f(\epsilon) = \int_0^{\epsilon_F} d\epsilon \epsilon_F D(\epsilon)$$

And split up the first integral:

$$\int_0^{\epsilon_F} d\epsilon \epsilon_F D(\epsilon)f(\epsilon) + \int_{\epsilon_F}^{\infty} d\epsilon \epsilon_F D(\epsilon)f(\epsilon) = \int_0^{\epsilon_F} d\epsilon \epsilon_F D(\epsilon)$$

Move all terms to RHS:

$$\int_0^{\epsilon_F} d\epsilon \epsilon_F D(\epsilon) (1 - f(\epsilon)) - \int_{\epsilon_F}^{\infty} d\epsilon \epsilon_F D(\epsilon)f(\epsilon) = 0$$

Use this to rewrite the expression for ΔU

$$\begin{aligned} \Delta U &= \int_0^{\infty} d\epsilon \epsilon D(\epsilon)f(\epsilon) - \int_0^{\epsilon_F} d\epsilon \epsilon D(\epsilon) \\ \Delta U &= \int_{\epsilon_F}^{\infty} d\epsilon \epsilon D(\epsilon)f(\epsilon) + \int_0^{\epsilon_F} d\epsilon \epsilon D(\epsilon)f(\epsilon) - \int_0^{\epsilon_F} d\epsilon \epsilon D(\epsilon) + \int_0^{\epsilon_F} d\epsilon \epsilon_F D(\epsilon) (1 - f(\epsilon)) \\ &\quad - \int_{\epsilon_F}^{\infty} d\epsilon \epsilon_F D(\epsilon)f(\epsilon) \\ \Delta U &= \int_{\epsilon_F}^{\infty} d\epsilon D(\epsilon)[\epsilon f(\epsilon) - \epsilon_F f(\epsilon)] + \int_0^{\epsilon_F} d\epsilon D(\epsilon)[\epsilon f(\epsilon) - \epsilon + \epsilon_F - \epsilon_F f(\epsilon)] \end{aligned}$$

$$\Delta U = \int_{\epsilon_F}^{\infty} d\epsilon D(\epsilon)(\epsilon - \epsilon_F)f(\epsilon) + \int_0^{\epsilon_F} d\epsilon D(\epsilon)(\epsilon_F - \epsilon)(1 - f(\epsilon))$$

The first integral describes the energy needed to take electrons from the Fermi level to higher energy levels, and the second integral describes the energy needed to excite electrons from lower energy levels up to the Fermi level.

The heat capacity is found by differentiating ΔU with respect to temperature, and the only terms in the integrals which have temperature dependence are $f(\epsilon)$

$$C_{el} = \frac{\partial \Delta U}{\partial T} = \int_0^\infty d\epsilon (\epsilon - \epsilon_F) D(\epsilon) \frac{\partial f(\epsilon, T)}{\partial T}$$

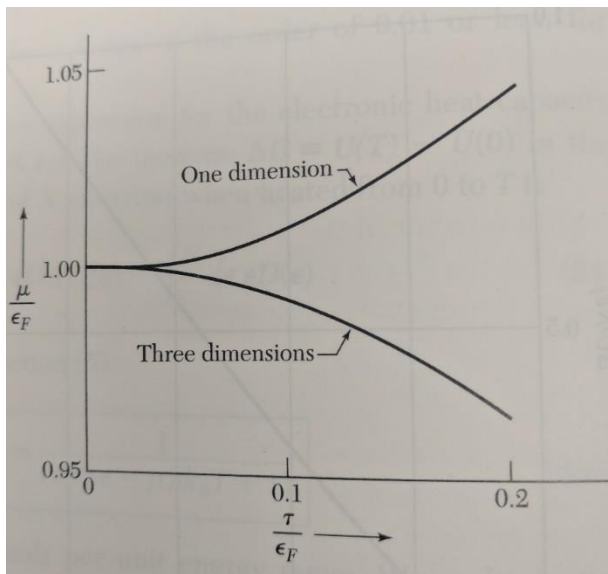
Where we have spliced the integrals back together after the temperature derivative produced the same integrand

At low temperature, $\mu \sim \epsilon_F$

And the temperature derivative of $f(\epsilon, T)$ is peaked close to ϵ_F , so the density of states can come out of the integral (this is another way of saying that only electrons very close to the Fermi level matter)

$$C_{el} \approx D(\epsilon_F) \int_0^\infty d\epsilon (\epsilon - \epsilon_F) \frac{\partial f(\epsilon, T)}{\partial T}$$

To solve this integral, first set $\mu = \epsilon_F$. This is a decent approximation for most ordinary metals at temperatures we might realistically encounter (remember that $\frac{k_B T}{\epsilon_F} = \frac{\tau}{\epsilon_F} \sim 0.01$ at room temperature)



$$\frac{\partial f}{\partial T} = \frac{\left(\frac{\epsilon - \epsilon_F}{k_B T}\right) e^{\frac{\epsilon - \epsilon_F}{k_B T}}}{\left(e^{\frac{\epsilon - \epsilon_F}{k_B T}} + 1\right)^2}$$

Define a new variable x and plug back into integral

$$x \equiv \frac{\epsilon - \epsilon_F}{k_B T}$$

$$C_{el} \approx D(\epsilon_F) k_B^2 T \int_{-\epsilon_F/k_B T}^\infty dx \frac{x^2 e^x}{(e^x + 1)^2}$$

Since we are working at low temperature, we can replace the lower bound of the integral by $-\infty$ because $k_B T \ll \epsilon_F$ (our starting assumption)

$$C_{el} \approx D(\epsilon_F) k_B^2 T \int_{-\infty}^\infty dx \frac{x^2 e^x}{(e^x + 1)^2} = D(\epsilon_F) k_B^2 T \frac{\pi^2}{3}$$

We can further express the Density of states at the Fermi energy in another way:

$$D(\epsilon_F) = \frac{3N}{2\epsilon_F} = 3N/2k_B T_F$$

This gives $C_{el} = \frac{1}{2} \pi^2 N k_B T / T_F$

This is very similar to our 'qualitative derivation' from earlier, except the prefactors are exact. Again, the key thing to remember is that for a 3D free electron gas, the heat capacity of electrons increases linearly with temperature.

Lecture 5:

- Heat capacity of electrons and phonons together
- Electrical conductivity

Review:

In the previous lecture we calculated heat capacity of a free electron gas:

Heat capacity: how much energy you must add to raise temperature by one unit

Assumptions in derivation:

- $k_B T \ll \epsilon_F$
- $\mu = \epsilon_F$

Result: $C_{el} = \frac{1}{2} \pi^2 N k_B T / T_F$

Putting it together: heat capacity from electrons and phonons

In a metal, both electrons and phonons contribute to the heat capacity, and their respective contributions can simply be added together to get the total. At low temperature ($T \ll \theta, T \ll T_F$) we can write an exact expression for the total heat capacity

$$C = C_{phonon} + C_{el} = \frac{12\pi^4}{5} N_{primitive\ cells} k_B \left(\frac{T}{\theta}\right)^3 + \frac{1}{2} \pi^2 N_{electrons} k_B T / T_F$$

This can be rewritten in terms of new constants, A and γ

$$C = AT^3 + \gamma T$$

At very low temperature, the electronic contribution (T-linear) will dominate and when the temperature increases a little, the phonon contribution to specific heat (T^3) will dominate. At room temperature, the phonon specific heat typically dominates over the electron contribution, even if we are outside the regime where the approximation $T \ll \theta$ holds.

$$\frac{C}{T} = \gamma + AT^2$$

If C/T is plotted as a function of T^2 , A will give the slope, and γ will give the y-intercept. This is actually observed in many/most metals.

Q: what is γ in an insulator?

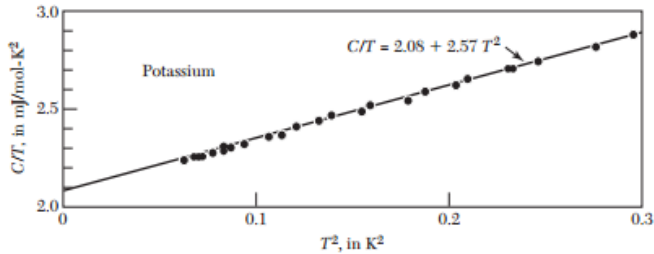


Figure 9 Experimental heat capacity values for potassium, plotted as C/T versus T^2 . (After W. H. Lien and N. E. Phillips.)

The experimental value of γ is very important because it is a customary way of extracting the effective electron mass. In real metals, the electrons do not always behave as if they have $m = m_e$. Sometimes they behave as if they have a heavier mass, and this is called the “effective

mass” m^* . In some compounds called ‘heavy fermion’ compounds, electrons can behave as if they have effective masses up to 1000x the free electron mass! An enhanced effective mass can be caused by interactions between electrons and other electrons or electrons and the periodic lattice potential. When heat capacity is used to extract an effective mass, this is called the “thermal mass”, m_{th} .

$$\frac{m^* = m_{th}}{m_e} = \frac{\gamma_{observed}}{\gamma_{free}}$$

γ is related to an electron mass because it is inversely proportional to the Fermi temperature

Significance of effective mass:

- Will affect materials’ response to electromagnetic fields
- Gives information about interactions inside solid (e.g. enhanced effective mass is caused by interactions with lattice, other electrons, etc)

Electrical conductivity (semi-classical treatment)

This is a slightly more sophisticated version of $V=IR$.

So far we have discussed electrons in a free electron gas in terms of their ground state, and in terms of thermal excitations. Now we will use these electrons for transporting heat and energy. The treatment in your textbook is a hybrid quantum/classical picture of metallic conduction.

- Quantum: fermi sphere in momentum-space,
- Classical: electron collisions

The Fermi sphere structure of electrons in a metal, with a hierarchy of energy states, provides a much more organized way of understanding electrical conduction than the real-space picture of electrons haphazardly zipping around and bumping into things.

The momentum of a free electron is related to its wavevector by

$$m\mathbf{v} = \hbar\mathbf{k}$$

In an electric field \mathbf{E} and magnetic field \mathbf{B} , the force on an electron (charge e) is given by:

$$\mathbf{F} = m \frac{d\mathbf{v}}{dt} = \hbar \frac{d\mathbf{k}}{dt} = -e(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B})$$

We set $B=0$ for now.

In the **absence of collisions**, the entire Fermi sphere will be accelerated by an electric field as a unit.

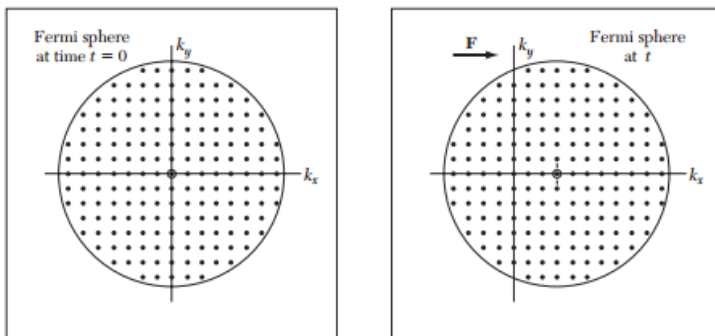
If a force $\mathbf{F}=-e\mathbf{E}$ is applied at $t=0$ to an electron gas, an electron with initial wavevector $\mathbf{k}(0)$ will end up at a final wavevector $\mathbf{k}(t)$

$$\mathbf{k}(t) - \mathbf{k}(0) = -e\mathbf{E}t/\hbar$$

This statement applies to every electron in the Fermi sea without regards to the specific momentum or energy that electron has, so a Fermi sphere centered at $\mathbf{k}=0$ at $t=0$ will have its center displaced by

$$\delta\mathbf{k} = -e\mathbf{E}t/\hbar$$

This also corresponds to a velocity kick $\delta\mathbf{v} = \hbar\delta\mathbf{k}/m$ (found by replacing derivatives in equation of motion by infinitesimal changes δk and δv)



The Fermi sphere does not accelerate indefinitely, because electrons eventually do scatter with lattice imperfections, impurities, or phonons. **This characteristic scattering time is called τ** , which gives a

'steady state' value of $\delta\mathbf{k} = -\frac{eE\tau}{\hbar} = m\delta\mathbf{v}/\hbar$

Lecture 6

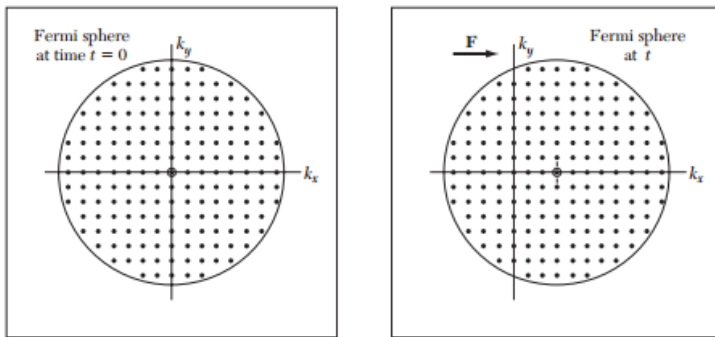
- Finish electrical conductivity

Electric conductivity

$m\mathbf{v} = \hbar\mathbf{k}$ for massive particles modeled as plane wave

Incremental change in \mathbf{k} corresponds to incremental velocity \mathbf{v} : $m\mathbf{v} = \hbar\delta\mathbf{k}$

- Electric field accelerates entire Fermi sphere as a unit, until scattering events reset velocities
- Average result of acceleration+scattering: small velocity kick to each electron, corresponding to steady state shift of fermi sphere, $\delta\mathbf{k}$



The Fermi sphere does not accelerate indefinitely, because electrons eventually do scatter with lattice imperfections, impurities, or phonons. This characteristic scattering time is called τ , which gives a 'steady state' value of $\delta\mathbf{k} = -\frac{e\mathbf{E}\tau}{\hbar} = m\mathbf{v}/\hbar$

Thus, the incremental velocity

imparted to electrons by the applied electric field is $\mathbf{v} = \frac{\hbar\delta\mathbf{k}}{m} = -e\mathbf{E}\tau/m$

If in a steady electric field, there are n electrons per unit volume, the current density (\mathbf{j}) is given by

$$\mathbf{j} = nq\mathbf{v} = ne^2\tau\mathbf{E}/m$$

This is a generalized version of Ohm's law, because \mathbf{j} is related to the current I , and electric field is related to a voltage or potential difference.

The electrical conductivity is defined by $\sigma = \frac{ne^2\tau}{m}$

And the resistivity (ρ) is defined as the inverse of conductivity

$$\rho = \frac{m}{ne^2\tau}$$

Resistivity is related to resistance (R) via a materials geometry, so resistivity is considered to be a more fundamental quantity because it does not depend on geometry

$$R = \frac{\rho\ell}{A}$$

Where ℓ is the length of the specimen, and A is the cross sectional area.

What is the physical origin of a finite τ ?

The derivation above stipulates that electrons **scatter**—bump into something and lose their momentum information—every interval τ , which in real materials tends to be on the order of 10^{-14} s, depending on temperature.

- At room temperature, phonons provide the primary scattering mechanism for electrons.
 - a perfect lattice will not scatter electrons and will not contribute to resistivity, but at higher temperature, a crystal lattice becomes increasingly ‘imperfect’ (because of increased atomic vibrations) which allows increased scattering off the lattice.
 - Or, if one views phonons as emergent particles with a certain energy and momentum, electrons scatter off these ‘particles’ such that the total energy and momentum is conserved.
 - This type of scattering happens every time interval τ_L , which depends on temperature
- At cryogenic temperature, electrons primarily scatter off impurities and other permanent defects in the crystalline lattice. This type of scattering happens every time interval τ_i , and is independent of temperature.

The scattering frequency (inverse of scattering time) is given by adding up scattering frequencies from each contribution:

$$\frac{1}{\tau} = \frac{1}{\tau_L} + \frac{1}{\tau_i}$$

This also implies that the contribution to resistivity from each type of scattering adds up linearly

$$\rho = \rho_L + \rho_i$$

Example: A copper sample has a residual resistivity (resistivity in the limit of $T=0$) of $1.7 \times 10^{-2} \mu\Omega \text{ cm}$. Find the impurity concentration.

Solution:

At zero temperature, only impurities contribute to resistivity

$$\rho = m/ne^2\tau$$

Solve for τ .

n is the electron concentration, and copper has 1 valence electron per atom. Copper forms an FCC structure (4 atoms per cubic cell) with a unit cell dimension of 3.61×10^{-10} m. Thus, $n = 8.5 \times 10^{28} \text{ m}^{-3}$

to solve for τ , first change the units of ρ . $\rho = 1.7 \times 10^{-10} \Omega \text{ m}$

$$\tau = \frac{m}{ne^2\rho} = 2.46 \times 10^{-12} \text{ s}$$

This can be used to solve for an average distance (ℓ) between collisions using

$$\ell = v_F\tau$$

where v_F is the Fermi velocity

$$v_F = \left(\frac{\hbar}{m}\right) \left(\frac{3\pi^2 N}{V}\right)^{1/3} = 1.6 \times 10^6 \text{ m/s}$$

$$\ell = 3.9 \mu\text{m}$$

Thus, given the T=0 resistivity of this specimen, the average spacing between impurities is $3.9 \mu\text{m}$ which means an electron would travel on average $\frac{(3.9 \times 10^{-6})}{(3.61 \times 10^{-10})} = 12,341$ unit cells before encountering an impurity

Lecture 7

- Electrons in a magnetic field
- Thermal conductivity

Electrons' motion in a magnetic field

In an electric field \mathbf{E} and magnetic field \mathbf{B} , the force on an electron (charge e) is given by:

$$\mathbf{F} = m \frac{d\mathbf{v}}{dt} = \hbar \frac{d\mathbf{k}}{dt} = -e(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B})$$

Again, we consider displacing the Fermi sphere by a momentum $\delta\mathbf{k}$ such that

$$m\mathbf{v} = \hbar\delta\mathbf{k}$$

Where \mathbf{v} is the incremental velocity kick that all electrons get.

We express acceleration in a slightly different way than we did previously to write expressions for motion in electric and magnetic field applied simultaneously (previously, we dropped the first term on the left because in the steady state, time derivatives are zero, but this notation is being introduced because it is needed to study time-varying fields, like in your homework):

$$m \left(\frac{d}{dt} + \frac{1}{\tau} \right) \mathbf{v} = -e(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B})$$

A special case of this problem arises when the magnetic field is applied along the z axis ($\mathbf{B} = B\hat{z}$):

$$m \left(\frac{d}{dt} + \frac{1}{\tau} \right) v_x = -e \left(E_x + \frac{Bv_y}{c} \right)$$

$$m \left(\frac{d}{dt} + \frac{1}{\tau} \right) v_y = -e \left(E_y - \frac{Bv_x}{c} \right)$$

$$m \left(\frac{d}{dt} + \frac{1}{\tau} \right) v_z = -e(E_z + 0)$$

In steady state, the time derivatives are zero, so the first terms on the left side disappear. These equations then become:

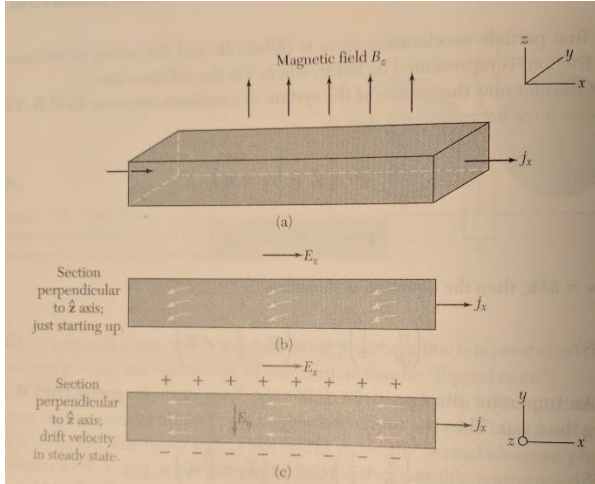
$$v_x = -\frac{e\tau}{m} E_x - \omega_c \tau v_y$$

$$v_y = -\frac{e\tau}{m} E_y + \omega_c \tau v_x$$

$$v_z = -\frac{e\tau}{m} E_z$$

Where $\omega_c = \frac{eB}{mc}$ is the cyclotron frequency. The cyclotron frequency describes the frequency of electrons' circular motion in a perpendicular magnetic field. It is notable independent of the electron's velocity or the spatial size of the circular orbit, and it only depends on a particle's charge-to-mass ratio.

Hall effect



The hall effect refers to a transverse voltage that develops when a current flows across a sample at the same time that a magnetic field is applied in the perpendicular direction. It is a very important characterization tool for assessing the number of charge carriers and their charge.

In general, the transverse electric field will be in the direction $\mathbf{j} \times \mathbf{B}$, and customarily, the current (\mathbf{j}) direction is set perpendicular to the magnetic field direction.

We consider a specific case where $\mathbf{B} = B\hat{z}$ and $\mathbf{j} = j\hat{x}$

When electrons flow with a velocity v_x perpendicular to the direction of the magnetic field, they will feel a force in the $\mathbf{v} \times \mathbf{B}$ direction, which is in the $-y$ direction. Thus, there will be an accumulation of negative charges on the $-y$ side of the sample, leading to an electric field in the $-y$ direction. Note that this electric field will tend to deflect directions in the opposite direction from the magnetic field, so a steady state situation is reached where the Lorentz force from the magnetic field perfectly balances the force from the electric field.

To write this more quantitatively:

Use: $v_y = -\frac{e\tau}{m}E_y + \omega_c\tau v_x$ and $v_x = -\frac{e\tau}{m}E_x - \omega_c\tau v_y$ and set $v_y = 0$ to reflect the steady state situation when there is no more y-deflection

$$E_y = -\omega_c\tau E_x = -\frac{eB\tau}{mc}E_x$$

A note about signs:

Electric current is defined as the flow of **positive** charges, so the electron velocity is in the opposite direction to the current (in this case, $-x$) [thus, $\times \mathbf{B} = vB\hat{y}$]

The negative sign is explicitly included in the definition of force (that an electron feels in a perpendicular magnetic field), so electrons will accelerate to the $-y$ side of the sample

The direction of the electric field is defined as the direction of the force that a **positive** test charge will feel, so electric field direction always points from positive to negative charges (towards $-y$ in this case)

A **hall coefficient** is defined as

$$R_H = \frac{E_y}{j_x B}$$

Use $j_x = \frac{ne^2\tau E_x}{m}$ and $E_y = -\frac{eB\tau}{m}E_x$ to evaluate

$$R_H = \frac{-eB\tau/mE_x}{ne^2\tau E_x B/mc} = -\frac{1}{nec}$$

The two free parameters here are n (the electron concentration) and the sign of e .

In some metals, the dominant charge carriers are electrons, and in other metals, it is the voids left behind by electrons, which are called holes (and are mathematically equivalent to positrons, physical particles which are positively charged electrons). The sign of the hall coefficient distinguishes between those two cases.

Additionally, the number of mobile charge carriers in a metal might be different from the number of valence electrons you think you have, and hall coefficient measurements can detect that too.

In some metals, both electrons and holes can be charge carriers, each with different densities, and in those cases, interpretation of the hall coefficient can be tricky.

Thermal conductivity of metals

Ch5 considers thermal conductivity if heat could **only** be carried by phonons. In metals, heat can be carried by electrons too.

Definition of thermal conductivity (in 1

dimension): $j_u = -K \frac{dT}{dx}$



Where j_u is the flux of thermal energy (or the energy transmitted across unit area per unit time), K is the thermal conductivity coefficient, and dT/dx is the temperature gradient across the specimen.

This equation is analogous to the form of ohms law stated as $\mathbf{j} = \sigma \mathbf{E}$, noting that in 1 dimension, $E = -\frac{dV}{dx}$, where V is the electric potential (or voltage)

For phonons, the thermal conductivity coefficient was given by $K = \frac{1}{3} C v \ell$ and we can identify analogous quantities for metals (C = heat capacity, v is a characteristic velocity e.g. the speed of sound, and ℓ is the mean free path between collision)

To get the thermal conductivity coefficient for a metal, we simply replace every quantity in the equation above by the equivalent concept for electrons.

For phonons, v is the sound velocity—the group velocity that acoustic phonons follow. For electrons, the equivalent quantity is the **Fermi velocity** (v_F)—the group velocity of electrons at the Fermi energy (most electrons in a metal are inert, except for those that happen to have energy within $\sim k_B T$ of the Fermi energy).

C is the heat capacity per unit volume, and earlier in this chapter we calculated heat capacity for electrons

It turns out that **in pure/clean metals, electrons are more effective at transporting heat than phonons**, but in metals with many impurities, the two types of thermal conductivity are comparable. In general, the two contributions to thermal conductivity are independent and can be added.

$$K_{total} = K_{electron} + K_{phonon}$$

Lecture 8

- Thermal conductivity of metals (+ lattice)
- Wiedemann Franz law
- Example: Thermoelectrics
- Begin Ch7: energy bands

Thermal conductivity of metals

Definition of thermal conductivity (in 1 dimension): $j_u = -K \frac{dT}{dx}$

Where j_u is the flux of thermal energy (or the energy transmitted across unit area per unit time), K is the thermal conductivity coefficient, and dT/dx is the temperature gradient across the specimen.

This equation is analogous to the form of ohms law stated as $\mathbf{j} = \sigma \mathbf{E}$, noting that in 1 dimension, $E = -\frac{dV}{dx}$, where V is the electric potential (or voltage)

For phonons, the thermal conductivity coefficient was given by $K = \frac{1}{3} C v \ell$ (C = heat capacity per unit volume, v is a characteristic velocity e.g. the speed of sound, and ℓ is the mean free path between collision)

To get the thermal conductivity coefficient for a metal, we simply replace every quantity in the equation above by the equivalent concept for electrons.

$$C_{lattice} \rightarrow C_{electrons}$$

$$v_s \rightarrow v_F$$

$$\ell \rightarrow \ell$$

Earlier in this chapter we calculated heat capacity for electrons

$$C_{el} = \frac{1}{2} \pi^2 N k_B T / T_F$$

$$T_F = \frac{\epsilon_F}{k_B} = \frac{1}{2} \frac{m v_F^2}{k_B}$$

$$C_{el} = \frac{\pi^2 N k_B^2 T}{m v_F^2}$$

This is the total heat capacity, and we need to divide by a factor of V to get the **heat capacity per volume** (reminder: $n=N/V$)

$$C = \frac{\pi^2 n k_B^2 T}{m v_F^2}$$

Plugging this in to the expression for the thermal conductivity coefficient:

$$K_{el} = \frac{\pi^2 n_{electrons} k_B^2 T}{3m v_F^2} v_F \ell$$

Compare this to the phonon thermal conductivity at **low temperature** (when ℓ does not depend on temperature)

$$K_{ph} = \frac{4\pi^4}{5} n_{primitive\ cells} k_B \left(\frac{T}{\theta}\right)^3 v_s \ell$$

And the phonon thermal conductivity at **high temperature** when C does not depend on T, but $\ell \propto \sim 1/T$

$$K_{ph} \propto n_{primitive\ cells} k_B v_s / T$$

Or the high-temperature phonon thermal conductivity in the “**dirty limit**” when impurities set ℓ (average impurity distance is given the symbol D), rather than phonon-phonon scattering setting ℓ

$$K_{ph} = 3n_{primitive\ cells} k_B D$$

We can further express the electronic thermal conductivity in terms of the mean scattering time $\tau = \ell/v_F$

$$K_{el} = \frac{\pi^2 n_{electrons} k_B^2 T \tau}{3m}$$

It turns out that **in pure/clean metals, electrons are more effective at transporting heat than phonons**, but in metals with many impurities, the two types of thermal conductivity are comparable. In general, the two contributions to thermal conductivity are independent and can be added.

$$K_{total} = K_{electron} + K_{phonon}$$

Wiedemann-Franz law

Since the same electrons carry both electric current and heat, there is an expected ratio between thermal conductivity and electrical conductivity:

$$\frac{K}{\sigma} = \frac{\pi^2 k_B^2 T n \tau / 3m}{n e^2 \tau / m} = \frac{\pi^2}{3} \left(\frac{k_B}{e}\right)^2 T$$

Interestingly, materials' dependent parameters such as τ , n , and m drop out of this ratio.

The Lorentz number L is defined as

$$L = \frac{K}{\sigma T} = \frac{\pi^2}{3} \left(\frac{k_B}{e}\right)^2 = 2.45 \times 10^{-8} \text{Watt } \Omega/\text{K}^2$$

Most simple metals have values of L roughly in this range (see table 6.5 in textbook)

The Wiedemann-Franz law is a very useful metric in contemporary research for assessing how much an exotic material behaves like a ‘simple’ or ‘textbook’ metal which is expected to follow W-F law.

Deviation from W-F law in temperature regimes where it should apply are used as evidence that a given material has ‘abnormal’ behavior.

Example: thermoelectrics

Thermoelectrics are materials that can convert waste heat into electricity (and vis versa: use a voltage to affect a temperature change), and they are defined by a figure of merit ZT . The larger ZT , the better, but most of the best thermoelectrics have $ZT \sim 1-2$.

$$ZT = \frac{\sigma S^2 T}{K}$$

Where S is the Seebeck coefficient. This is a materials property which describes the degree to which a temperature gradient produces an electric potential: $S = -\Delta V / \Delta T$

σ is the electrical conductivity and K is the thermal conductivity. According to the equation above, one can increase ZT for a given material (fixed S) by increasing σ or decreasing K .

But as we learned in the previous section, **electrons** carry both charge and heat, and there is a specific ratio between the two, so there is no way to simultaneously raise one and lower the other.

A trick that people often employ is manipulating the **phonon** thermal conductivity.

$$K_{total} = K_{electron} + K_{phonon}$$

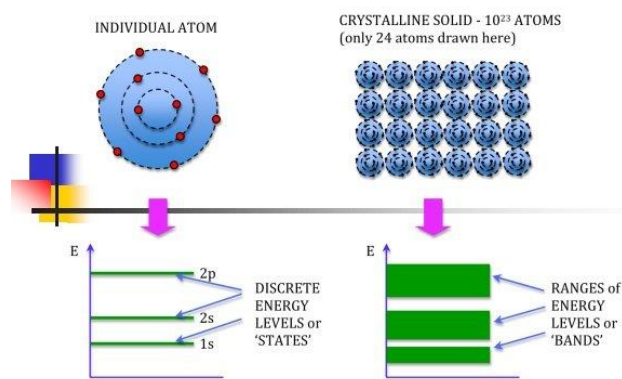
For example, by creating nanostructured materials (small D ; phonons scatter off the boundaries and have difficulty conducting heat), people can suppress the phonon contribution to thermal conductivity

$$K_{ph} = 3n_{primitive\ cells}k_B D$$

Without hurting electrical conductivity too much.

The other option to make a small D determine phonon thermal conductivity is to put in a bunch of impurities, but that can negatively affect electrical conductivity (by producing a small τ) and degrade the figure of merit.

Ch7—energy bands



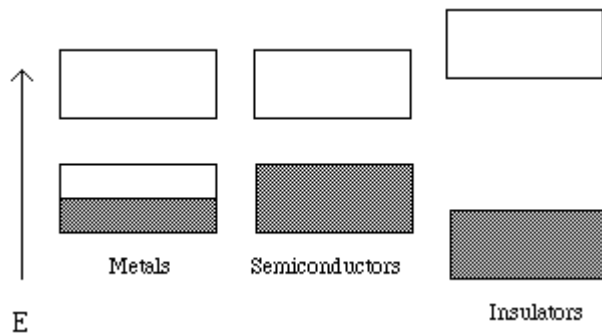
In this chapter, we will introduce the lattice back into the discussion, in several different ways. When we do this, the concept of 'band gaps' emerge, which are energies where electrons are forbidden. These band gaps occur between 'bands' where electrons are allowed. This is analogous to the energy states of individual atoms where there are 'forbidden' energies between discrete allowed energies. In solids, instead of discrete energy levels, there are 'bands' of allowed energies where allowed

energy levels are so close together that they can be treated as being continuous.

Lecture 9

- Nearly free electron model
- Bloch functions
- Begin Kronig-Penney model

In this chapter, we will introduce the lattice back into the discussion, in several different ways. When we do this, the concept of '**band gaps**' emerge, which are energies where electrons are forbidden. These band gaps occur between bands where electrons do live. All types of materials have band gaps, and it is the location of the Fermi energy (highest occupied electron level) relative to a band gap which determines if a material is a metal or an insulator. The size of the band gap determines if the material is



an insulator or a semiconductor.

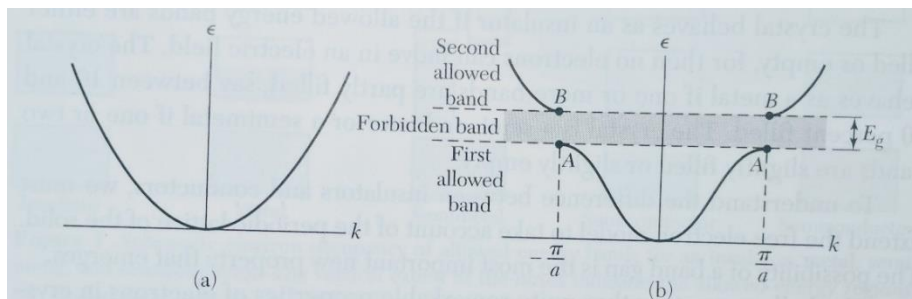
The logic of this chapter is as follows. The *existence* of band gaps is taken to be an experimental fact (which it is), and the mechanism behind their appearance is demonstrated in several different ways.

Nearly free electron model

The nearly free electron model starts with a free electron description from the previous chapter and perturbs it slightly by introducing a lattice potential. The lattice potential can be considered to be a lattice of positively charged ion cores (because some of the electrons are off being free)

The electron wavefunctions for a 3D particle in a box (omitting boundary conditions for now) are given by: $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}$ and they represent waves propagating with momentum $\mathbf{p} = \hbar\mathbf{k}$.

The relationship between energy and momentum is $\epsilon_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m} (k_x^2 + k_y^2 + k_z^2)$. This is applicable for a 'free' electron. When we add the lattice potential, the relationship between energy and momentum is different at some specific wavevectors.



As we learned in Ch2, a **wave** that encounters a periodic potential will get **bragg reflected** if the change in its wavevector is equal to a reciprocal lattice vector. It doesn't matter if the wave is a

light wave (e.g. x-ray) or a quantum particle with wave duality (e.g. neutrons or in this case electrons)

The Bragg condition can be written as:

$$(\mathbf{k} + \mathbf{G})^2 = k^2$$

In one dimension, this simplifies to: $k = \pm \frac{1}{2}G = \pm n\pi/a$, where $G = \frac{2\pi n}{a}$ is a reciprocal lattice vector and n is an integer.

To get this result, consider the following two scalar sums $(k \pm G)^2 = k^2$ to reflect the two ways to add \mathbf{k} and \mathbf{G} in 1D

The first band gap opens at $k = \pm \frac{\pi}{a}$ and $-\frac{\pi}{a} < k < \frac{\pi}{a}$ defines the first Brillouin zone.

The wavefunction at $k = \pm \frac{\pi}{a}$ is not a traveling wave, but rather, is a standing wave. This can be seen in the image above by noting that the group velocity is given by $v_g = \frac{1}{\hbar} \frac{\partial \epsilon(k)}{\partial k}$. This equation holds for any propagating particle that one can define an energy vs momentum relation for.

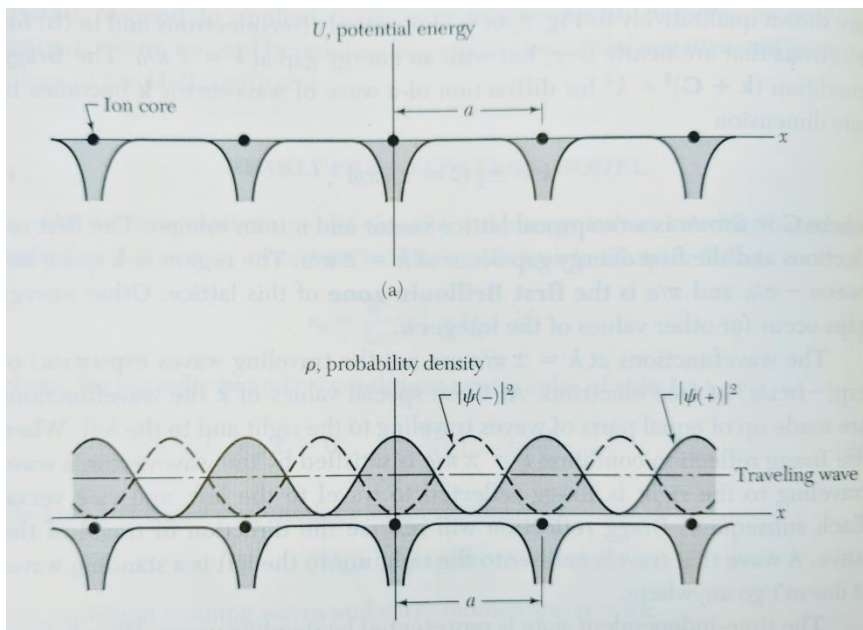
A standing wave can be expressed as the sum of two traveling waves (the +/- case in the eqn below):

Reminder (Euler's formula): $e^{\pm \frac{i\pi x}{a}} = \cos\left(\frac{\pi x}{a}\right) \pm i \sin\left(\frac{\pi x}{a}\right)$

$$\psi(+)=e^{\frac{i\pi x}{a}}+e^{-\frac{i\pi x}{a}}=2\cos(\pi x/a)$$

$$\psi(-)=e^{\frac{i\pi x}{a}}-e^{-\frac{i\pi x}{a}}=2i\sin(\pi x/a)$$

An electron wave with $k = \pi/a$ traveling to the right will be Bragg reflected to travel to the left, and together, they will add to produce a standing wave. This explains the zero electron group velocity at the Brillouin zone boundary, but it



does not explain directly why a gap forms. To do that we consider charge density. As a reminder, $\psi^*\psi = |\psi|^2$ is a probability density (a physically measurable quantity) related to the probability of finding electrons at a specific coordinate. It is related to charge density ρ

For a traveling wave, $\rho = e^{-ikx}e^{ikx} = 1$, meaning it gives uniform charge density

For the standing waves, we

get the following:

$$\rho(+)=|\psi(+)|^2 \propto \cos^2 \pi x/a$$

This function is peaked at $x = 0, a, 2a, \dots$ where the ion cores are located. This makes sense; you expect negative charges to pile up on top of the positive charges. The other solution concentrates electrons away from the ion cores:

$$\rho(-) = |\psi(-)|^2 \propto \sin^2 \pi x/a$$

In both cases, when we normalize the wavefunctions we get the correct prefactors (normalization is done by integrating over unit length ($x=0$ to $x=a$)):

$$\psi(+) = \sqrt{\frac{2}{a}} \cos \pi x/a$$

$$\psi(-) = \sqrt{\frac{2}{a}} \sin \pi x/a$$

The energy gap comes from calculating the expectation value of potential energy over these charge distributions

Suppose the potential energy of an electron in the crystal at point x is given by $U(x) = U \cos 2\pi x/a$

The first order energy difference between the two standing wave states is:

$$\begin{aligned} E_g &\equiv \langle U+ \rangle - \langle U- \rangle = \int_0^a dx U(x) [|\psi(+)|^2 - |\psi(-)|^2] \\ &= \frac{2}{a} \int_0^a dx U \cos\left(\frac{2\pi x}{a}\right) \left[\cos^2 \frac{\pi x}{a} - \sin^2 \frac{\pi x}{a}\right] = U \end{aligned}$$

This integral can be solved to give the answer given, and this is left as a homework problem. The energy difference (energy gap) between the two wavefunctions for standing waves is the energy gap, and its magnitude depends on the strength of the ion potential. This is one way to explain band gaps, and several more will be discussed in this chapter.

Lecture 10

- **Review: nearly free electron model**
- **Bloch waves**
- **Kronig Penney model**

Reminder: Ch7 uses concepts relating to reciprocal lattice a lot, so please review Ch2 if you are a bit rusty on that

Nearly free electron model

Summary:

- Electrons whose wavelength is a half integer multiple of the crystal lattice (reminder: $\lambda = 2\pi/k$) get **Bragg reflected**.
More generally, the Bragg reflection condition corresponds to $(\mathbf{k} + \mathbf{G})^2 = k^2$
In one dimension, this simplifies to: $k = \pm \frac{1}{2}G = \pm n\pi/a$, where $G = \frac{2\pi n}{a}$ is a reciprocal lattice vector and n is an integer.
- The sum of a wave moving in one direction and a wave of the same wavelength moving in the other direction is a **standing wave**.
- The two standing wave solutions (physically corresponding to electron density bunching up at the atomic positions and between atoms) yield different expectation values of potential energy, and the band gap is the energy difference between these solutions
- Size of band gap related to strength of atomic potential

Bloch functions

Bloch's theorem is one of the most important principles in solid state physics. It states that the solution to the Schrödinger equation with a periodic potential (i.e. a crystal) must have a specific form:

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

Where $u_{\mathbf{k}}(\mathbf{r})$ has the period of the lattice such that it is invariant under translation by a lattice vector (\mathbf{T}): $u_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r} + \mathbf{T})$

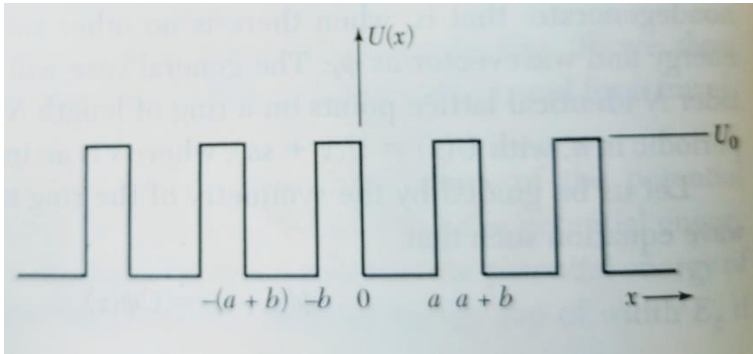
Eigenfunctions of this form are called Bloch functions. They consist of a product of a plane wave and a function which shares the periodicity of the lattice. Proofs of Bloch's theorem are provided in auxiliary reading materials; for the purposes of this course we will use the result axiomatically.

Kronig-Penney model

The Kronig Penney model is one of very few exactly solvable problems in quantum mechanics, and it demonstrates another explanation of why band gaps exist. The Kronig Penney model is characterized by a periodic potential in a 1D schrodinger equation:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + U(x)\psi(x) = \epsilon\psi(x)$$

The potential is described with the drawing below.



Potential barriers of height U_0 and width b are arranged in an infinite line with periodicity $(a+b)$.

The solution is found by writing down the solution to Schrodinger equation in the barrier region and in the free region, applying continuity boundary conditions, and enforcing adherence to Bloch's theorem.

In the region $0 < x < a$, where the potential is zero, free-electron-like solutions arise:

$$\psi_1(x) = Ae^{iKx} + Be^{-iKx}$$

K is related to energy via $\epsilon = \hbar^2 K^2 / 2m$

In the barrier region ($-b < x < 0$), the solution is given by:

$$\psi_2(x) = Ce^{Qx} + De^{-Qx}$$

Where $U_0 - \epsilon = \frac{\hbar^2 Q^2}{2m}$

First, we apply continuity conditions. At $x=0$, both ψ and its first derivative must be continuous. This gives us two equations:

$$A + B = C + D$$

$$iK(A - B) = Q(C - D)$$

Second, we need to ensure compliance with Bloch's theorem:

$$\psi(a < x < a + b) = \psi(-b < x < 0)e^{ik(a+b)}$$

This provides a wavevector k used to label solutions

Out of the x -range, select $x=a$ and $x=-b$ to enforce boundary conditions

$$\psi_1(x = a) = \psi_2(x = -b)e^{ik(a+b)}$$

$$Ae^{iKa} + Be^{-iKa} = (Ce^{-Qb} + De^{Qb})e^{ik(a+b)}$$

Do the same for the derivatives to get the last equations:

$$iK(Ae^{iKa} - Be^{-iKa}) = Q(Ce^{-Qb} - De^{Qb})e^{ik(a+b)}$$

We now have 4 variables and 4 equations, which can be solved. The way to solve them is to arrange them in a 4x4 matrix and equate its determinant to zero. The textbook just presents the solution.

$$\left[\frac{Q^2 - K^2}{2QK} \right] \sinh Qb \sin Ka + \cosh Qb \cos Ka = \cos k(a + b)$$

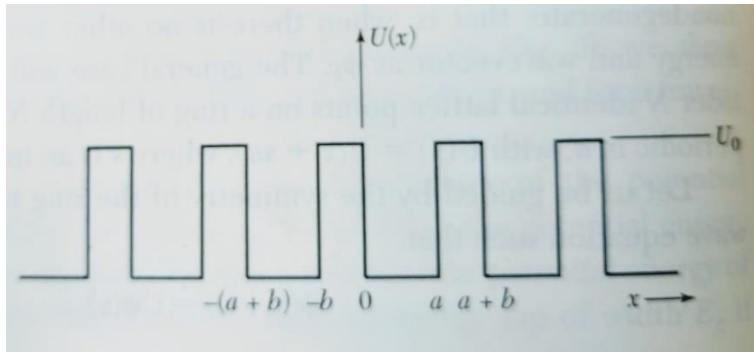
Lecture 11

- **Finish Kronig Penney model**

The Kronig Penney model is characterized by a periodic potential in a 1D schrodinger equation:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + U(x)\psi(x) = \epsilon\psi(x)$$

The potential is described with the drawing below.



To solve, we write down 'guesses' for wavefunctions in the barrier regions and between barriers, and apply continuity of the wavefunction, continuity of the derivative of the wavefunction, and Bloch's theorem.

In the region $0 < x < a$, where the potential is zero, free-electron-like solutions arise:

$$\psi_1(x) = Ae^{iKx} + Be^{-iKx}$$

K is related to energy via $\epsilon = \hbar^2 K^2 / 2m$

In the barrier region ($-b < x < 0$), the solution is given by:

$$\psi_2(x) = Ce^{Qx} + De^{-Qx}$$

Where $U_0 - \epsilon = \frac{\hbar^2 Q^2}{2m}$

At $x=0$, both ψ and its first derivative must be continuous (true for $x=a$, $a+b$, etc, but we choose 0 for easy math).

Continuity of wavefunction: $A + B = C + D$

Continuity of first derivative of wavefunction: $iK(A - B) = Q(C - D)$

Then, need to ensure compliance with Bloch's theorem. We do this by ensuring that the value of the wavefunction (and its first derivative) is equal (modulo a phase factor) at $x=-b$ and $x=a$ (representing two left-hand corners of barriers, separated by one period, $a+b$)

$$\psi_1(x = a) = \psi_2(x = -b)e^{ik(a+b)}$$

lowercase k in the plane wave term is used to label the solutions

Wavefunction continuity: $Ae^{iKa} + Be^{-iKa} = (Ce^{-Qb} + De^{Qb})e^{ik(a+b)}$

Derivative continuity: $iK(Ae^{iKa} - Be^{-iKa}) = Q(Ce^{-Qb} - De^{Qb})e^{ik(a+b)}$

We now have 4 variables and 4 equations, which can be solved. The way to solve them is to arrange them in a 4x4 matrix and equate its determinant to zero. The textbook just presents the solution. This is an expression for the energy eigenvalues (reminder: K is related to energy via $\epsilon = \hbar^2 K^2 / 2m$)

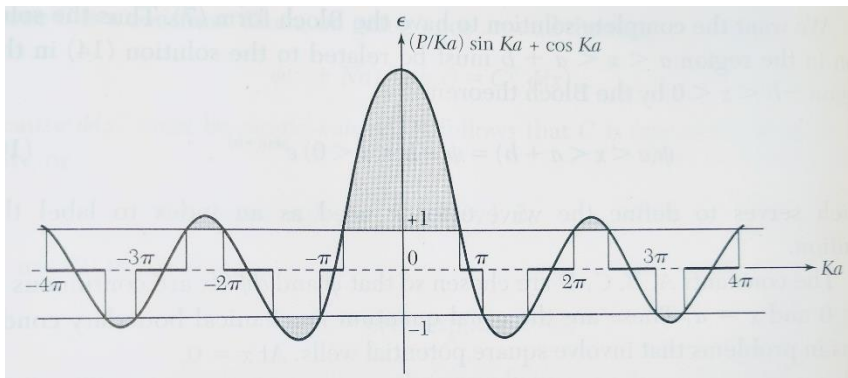
$$\left[\frac{Q^2 - K^2}{2QK} \right] \sinh Qb \sin Ka + \cosh Qb \cos Ka = \cos k(a + b)$$

To simplify, we consider the limit where $b \rightarrow 0$ and $U_0 \rightarrow \infty$ in such a way that the product $\frac{Q^2 b a}{2} \equiv P$ stays constant.

Essentially, we are just taking the limit that the potential barriers become delta functions. In this limit, $Q \gg K$ (another way of saying that electron energies, which are related to K) are much smaller than the potential barriers; in the limit that the electrons' energies are much higher than the potential barriers, they don't much notice them) and $Qb \ll 1$

In this limit we get a simpler solution:

$$\left(\frac{P}{Ka} \right) \sin Ka + \cos Ka = \cos ka$$



Note that the limits on the right hand side are ± 1 , which means that a solution **does not** exist for values of K where the left hand side is larger than 1 or smaller than -1.

Ka is related to energy via $\epsilon = \hbar^2 K^2 / 2m$. Thus, values of Ka where a solution does not exist represent energies that

electrons cannot occupy—a band gap

We can also visualize the solutions in the usual way, in terms of E vs k, and that is shown in the diagram below. Note that energy (ϵ) of electrons is given in units of $\hbar^2 \pi^2 / 2ma^2$ which is equivalent to K being in units of π/a

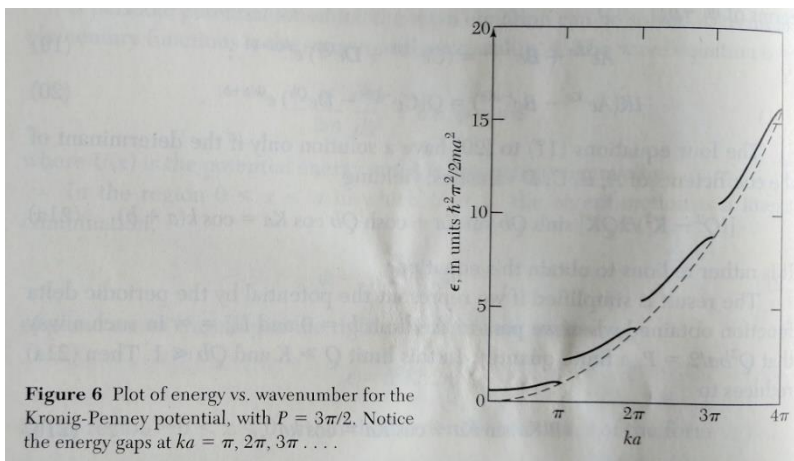


Figure 6 Plot of energy vs. wavenumber for the Kronig-Penney potential, with $P = 3\pi/2$. Notice the energy gaps at $ka = \pi, 2\pi, 3\pi, \dots$

Physics 140B Lecture 12

- Central equation
- Crystal momentum
- Solutions to central equation

Wave equation of electrons in a periodic potential

Earlier we made an approximation of the lattice potential, and now we will express it in a more correct way, in terms of Fourier components. We know that the potential energy of the lattice must be invariant under translation by a lattice vector. In this derivation, we consider a 1 dimensional lattice with unit cell a . The potential energy of one electron in a lattice of positive charges is given by:

$$U(x) = \sum_G U_G e^{iGx}$$

We want the potential to be a real function:

$$U(x) = \sum_{G>0} U_G (e^{iGx} + e^{-iGx}) = 2 \sum_{G>0} U_G \cos Gx$$

The Shrodinger equation for electrons in this 1D crystal is given by:

$$\left(\frac{p^2}{2m} + U(x) \right) \psi(x) = \left(\frac{1}{2m} \left(-i\hbar \frac{d}{dx} \right)^2 + \sum_G U_G e^{iGx} \right) \psi(x) = \epsilon \psi(x)$$

Because $U(x)$ is a periodic potential, the solutions to the eigenfunctions or wavefunctions ψ must in the form of Bloch functions. The wavefunctions can also be expressed as a Fourier series:

$$\psi = \sum_k C(k) e^{ikx}$$

As before, k can take on the values $2\pi n/L$, where n is an integer, to satisfy periodic boundary conditions.

Substitute this into the Schrödinger equation:

$$\left(\frac{-\hbar^2}{2m} \frac{d^2}{dx^2} + \sum_G U_G e^{iGx} \right) \sum_k C(k) e^{ikx} = \epsilon \sum_k C(k) e^{ikx}$$

$$\sum_k \frac{\hbar^2}{2m} k^2 C(k) e^{ikx} + \sum_G \sum_k U_G C(k) e^{i(k+G)x} = \epsilon \sum_k C(k) e^{ikx}$$

Equate same k -values on both sides of the equation

$$(\lambda_k - \epsilon) C(k) + \sum_G U_G C(k - G) = 0$$

$$\text{Here, } \lambda_k = \frac{\hbar^2 k^2}{2m}$$

And the translation by $-G$ in the 2nd term is legit because in a crystal lattice, the wavevectors k are defined modulo a reciprocal lattice vector G (i.e. $k=k+G$).

The equation above is called the **central equation** and it is a useful form of Schrodinger's equation in a periodic lattice.

A return to Bloch's theorem

In the central equation, the wavefunctions are of the form:

$$\psi_k(x) = \sum_G C(k-G)e^{i(k-G)x} = (\sum_G C(k-G)e^{-iGx})e^{ikx} \equiv e^{ikx}u_k(x)$$

The last step rearranges the equation to follow the notation written earlier for a Bloch wave

Because $u_k(x)$, as it is written above, originated from a fourier series over reciprocal lattice vectors, it is also invariant under translation by a crystal lattice vector \mathbf{T} . $u_k(x) = u_k(x+T)$

We can verify this by plugging in $x+T$:

$$u_k(x+T) = \sum_G C(k-G)e^{-iG(x+T)}$$

By definition of the reciprocal lattice, $GT = 2\pi n \rightarrow e^{-iGT} = 1$

$$u_k(x+T) = \sum_G C(k-G)e^{-iGx} = u_k(x)$$

Thus, solutions to the central equation satisfy Bloch's theorem.

Crystal momentum of an electron

The Bloch wavefunctions are labeled by an index k (as the free electron wavefunctions were earlier), and this quantity, is called crystal momentum. A few comments about crystal momentum

- $e^{ik \cdot T}$ is the phase factor which multiplies a Bloch function when we make a translation by a lattice vector \mathbf{T}
- If the lattice potential vanishes in the central equation, we are left with $\psi_k(\mathbf{r}) = e^{ik \cdot \mathbf{r}}$ just like in the free electron case
- Crystal momentum ($\hbar k$) is like regular momentum in that it enters into conservation laws that govern collisions (e.g. electrons with momentum $\hbar k$ colliding with a phonon with momentum $\hbar q$)
- Crystal momentum is different from regular momentum in that it is defined only modulo a reciprocal lattice vector \mathbf{G} . Thus, if an electron collides with a phonon and is kicked into momentum k' , this is expressed in the following way, $\mathbf{k} + \mathbf{q} = \mathbf{k}' + \mathbf{G}$

Solutions of the Central Equation

The central equation represents a set of linear equations that connect the coefficients $C(k-G)$ for all reciprocal vectors G .

$$(\lambda_k - \epsilon)C(k) + \sum_G U_G C(k-G) = 0$$

Consider a specific case where g denotes the shortest G , and $U_g = U_{-g} = U$. That is, only two fourier components survive, $G = \pm g$. The central equation will be approximated by 5 equations. It helps to use dummy variable k' in central equation and cycle through values of k' for which the fourier components of interest are present:

$$(\lambda_{k'} - \epsilon)C(k') + \Sigma_G U_G C(k' - G) = 0$$

$$k' = k - 2g: (\lambda_{k-2g} - \epsilon)C(k - 2g) + \Sigma_G U_G C(k - 2g - G) = 0$$

$$\text{In the second term, only the } G = -g \text{ coefficient survives: } (\lambda_{k-2g} - \epsilon)C(k - 2g) + U_{-g}C(k - g) = 0$$

$$\text{Substitute } U_{-g} = U: (\lambda_{k-2g} - \epsilon)C(k - 2g) + UC(k - g) = 0$$

Lets do one more.

$$k' = k - g: (\lambda_{k-g} - \epsilon)C(k - g) + \Sigma_G U_G C(k - g - G) = 0$$

Both the $G=-g$ and $G=g$ fourier components survive

$$(\lambda_{k-g} - \epsilon)C(k - g) + U_g C(k - 2g) + U_{-g}C(k) = 0 \rightarrow (\lambda_{k-g} - \epsilon)C(k - g) + UC(k - 2g) + UC(k) = 0$$

There will be three more equations derived by a similar logic, and the end result is a matrix expression of the following form:

$$\begin{pmatrix} \lambda_{k-2g} - \epsilon & U & 0 & 0 & 0 \\ U & \lambda_{k-g} - \epsilon & U & 0 & 0 \\ 0 & U & \lambda_k - \epsilon & U & 0 \\ 0 & 0 & U & \lambda_{k+g} - \epsilon & U \\ 0 & 0 & 0 & U & \lambda_{k+2g} - \epsilon \end{pmatrix} \begin{pmatrix} C(k - 2g) \\ C(k - g) \\ C(k) \\ C(k + g) \\ C(k + 2g) \end{pmatrix} = 0$$

The solution is found by setting the determinant of the matrix to zero and solving for ϵ_k . Each root lies on a different energy band. The solutions give a set of energy eigenvalues ϵ_{nk} where k is the wavevector and n is the index for ordering the bands (lowest energy, next lowest, etc).

Lecture 13

- Solutions to central equation
- Kronig Penney model solved using central equation
- Approximate solutions to central equation near zone boundary

Solutions of the Central Equation

The central equation represents a set of linear equations that connect the coefficients $C(k-G)$ for all reciprocal vectors G .

$$(\lambda_k - \epsilon)C(k) + \sum_G U_G C(k - G) = 0$$

- Consider a specific case where g denotes the shortest G , and $U_g = U_{-g} = U$. Only two fourier components survive, $G = \pm g$.
- The central equation will be approximated by 5 equations, which contain $U_{\pm g}$. Note: 5 equations implies 5 terms in the fourier expansion of the electron wavefunction ($C(k)$ terms), but there are only two terms in the fourier expansion of the potential energy. This is fine; they don't need to have the same number of fourier components.
- It helps to use dummy variable k' in central equation and cycle through values of k' for which the fourier components of interest are present:

$$(\lambda_{k'} - \epsilon)C(k') + \sum_G U_G C(k' - G) = 0$$

$$k' = k - 2g: (\lambda_{k-2g} - \epsilon)C(k - 2g) + \sum_G U_G C(k - 2g - G) = 0$$

In the second term, only the $G = -g$ coefficient involves $k \pm g$ so it is the only one we keep:

$$(\lambda_{k-2g} - \epsilon)C(k - 2g) + U_{-g}C(k - g) = 0$$

$$\text{Substitute } U_{-g} = U: (\lambda_{k-2g} - \epsilon)C(k - 2g) + UC(k - g) = 0$$

Lets do one more.

$$k' = k - g: (\lambda_{k-g} - \epsilon)C(k - g) + \sum_G U_G C(k - g - G) = 0$$

Both the $G=-g$ and $G=g$ fourier components survive

$$(\lambda_{k-g} - \epsilon)C(k - g) + U_g C(k - 2g) + U_{-g}C(k) = 0 \rightarrow (\lambda_{k-g} - \epsilon)C(k - g) + UC(k - 2g) + UC(k) = 0$$

There will be three more equations derived by a similar logic, and the end result is a matrix expression of the following form:

$$\begin{pmatrix} \lambda_{k-2g} - \epsilon & U & 0 & 0 & 0 \\ U & \lambda_{k-g} - \epsilon & U & 0 & 0 \\ 0 & U & \lambda_k - \epsilon & U & 0 \\ 0 & 0 & U & \lambda_{k+g} - \epsilon & U \\ 0 & 0 & 0 & U & \lambda_{k+2g} - \epsilon \end{pmatrix} \begin{pmatrix} C(k-2g) \\ C(k-g) \\ C(k) \\ C(k+g) \\ C(k+2g) \end{pmatrix} = 0$$

The solution is found by setting the determinant of the matrix to zero and solving for ϵ_k . Each root lies on a different energy band. The solutions give a set of energy eigenvalues ϵ_{nk} where k is the wavevector and n is the index for ordering the bands (lowest energy, next lowest, etc).

Kronig-Penney model in reciprocal space

We will now apply the central equation to the Kronig-Penney model, which we already found solutions for. We will consider the version that we found the limit of in the end, where the potential consists of a series of delta functions:

$$U(x) = 2\sum_{G>0} U_G \cos Gx = Aa\sum_s \delta(x - sa)$$

The second sum is over all integers s between 0 and $1/a$. The boundary conditions are set to be periodic over a ring of unit length, and we use this to derive the fourier coefficients.

$$U_G = \int_0^1 dx U(x) \cos Gx = Aa\sum_s \int_0^1 dx \delta(x - sa) \cos Gx = Aa\sum_s \cos Gsa = A$$

All this was to show that all U_G are equal for delta function potential. Note: the factor of a drops out because s is defined between 0 and $1/a$ and $\cos Gsa=1$, so the term in the sum becomes $1 \cdot 1/a = 1/a$.

Write the central equation:

$$(\lambda_k - \epsilon)C(k) + A\sum_n C\left(k - \frac{2\pi n}{a}\right) = 0$$

In the equation above, $\lambda_k = \hbar^2 k^2 / 2m$, and the sum is over all integers n . We want to solve for $\epsilon(k)$. The derivation will be a bit vague as the purpose of this exercise was to set up the central equation for a situation where we *do* keep a large number of fourier components.

Define $f(k) = \sum_n C\left(k - \frac{2\pi n}{a}\right)$

This is just another expression for the potential energy term of this problem, since all fourier coefficients are equal. Plug into central equation and solve for $C(k)$ and also plug in $\lambda_k = \hbar^2 k^2 / 2m$

$$C(k) = \frac{-\left(\frac{2mA}{\hbar^2}\right)f(k)}{k^2 - \left(\frac{2m\epsilon}{\hbar^2}\right)}$$

Because the expression for $f(k)$ is a sum over all coefficients C

$$f(k) = f\left(k - \frac{2\pi n}{a}\right)$$

For any n . This is just another way of saying that the potential energy term is invariant under translation by a reciprocal lattice vector.

This allows us to re-express the C coefficients in the following way for any n :

$$C\left(k - \frac{2\pi n}{a}\right) = -\left(\frac{2mA}{\hbar^2}\right)f(k) \left[\left(k - \frac{2\pi n}{a}\right)^2 - \frac{2m\epsilon}{\hbar^2}\right]^{-1}$$

Sum both sides over n:

$$\sum_n C\left(k - \frac{2\pi n}{a}\right) = -\left(\frac{2mA}{\hbar^2}\right)f(k) \sum_n \left[\left(k - \frac{2\pi n}{a}\right)^2 - \frac{2m\epsilon}{\hbar^2}\right]^{-1}$$

$$f(k) = -\left(\frac{2mA}{\hbar^2}\right)f(k) \sum_n \left[\left(k - \frac{2\pi n}{a}\right)^2 - \frac{2m\epsilon}{\hbar^2}\right]^{-1}$$

When we sum both sides over all n, it cancels f(k) giving:

$$\frac{\hbar^2}{2mA} = -\sum_n \left[\left(k - \frac{2\pi n}{a}\right)^2 - \frac{2m\epsilon}{\hbar^2}\right]^{-1}$$

The sum can be carried out by noting that $\pi \cot \pi z = \frac{1}{z} + 2z \sum_{n=1}^{\infty} \frac{1}{z^2 - n^2}$

The term in the summation becomes $\frac{a^2 \sin Ka}{4Ka (\cos ka - \cos Ka)}$

Where $K^2 = 2m\epsilon/\hbar^2$

The final result is

$$\left(\frac{mAa^2}{2\hbar^2}\right) \frac{1}{Ka} \sin Ka + \cos Ka = \cos ka$$

This agrees with the earlier result with $P \equiv mAa^2/2\hbar^2$

Approximate solutions to central equation near a zone boundary

This section of the chapter uses the central equation to extract the following quantities, assuming only two fourier coefficients of the lattice potential in 1D

- Band gap
- Wavefunction at Brillouin zone boundary
- Band dispersion (ϵ vs k) near zone boundary

Physics 140B Lecture 14

- **Approximate solutions to central equation near zone boundary**
- **Empty Lattice Approximation**

Approximate solutions to central equation near a zone boundary

This section of the chapter uses the central equation to extract the following quantities, assuming only two Fourier coefficients of the lattice potential in 1D

- Band gap
- Wavefunction at Brillouin zone boundary
- Band dispersion (ϵ vs k) near zone boundary

Suppose that the Fourier components, $U_G = U$, of the potential energy are small relative to the kinetic energy of free electrons at the Brillouin zone boundary. This is intended to be applicable to any form of the potential, not necessarily that in the Kronig-Penney model. We consider an electron with wavevector at the zone boundary, $k = \pm \frac{\pi}{a}$. Note that this is also equal to $\frac{1}{2}G$

$$\text{Here: } k^2 = \left(\frac{1}{2}G\right)^2 \text{ and } (k - G)^2 = \left(\frac{1}{2}G - G\right)^2 = \left(\frac{1}{2}G\right)^2$$

Thus, the kinetic energy of the original wave (wavevector k) is equal to the kinetic energy of the wavevector translated by a reciprocal lattice vector (wavevector $k-G$) since for free electrons, $\epsilon \propto k^2$. Putting all this together, $k = \pm \frac{1}{2}G$ yields the same kinetic energy

If $C(\frac{1}{2}G)$ is an important coefficient, so is $C(-\frac{1}{2}G)$. Let's consider only those equations in the central equation that contain both coefficients $C(\frac{1}{2}G)$ and $C(-\frac{1}{2}G)$ and neglect others. The two surviving equations of the central equation are:

$$(\lambda - \epsilon)C\left(\frac{1}{2}G\right) + UC\left(-\frac{1}{2}G\right) = 0$$

$$(\lambda - \epsilon)C\left(-\frac{1}{2}G\right) + UC\left(\frac{1}{2}G\right) = 0$$

These have nontrivial solutions if the determinant of the following matrix is zero:

$$\begin{vmatrix} \lambda - \epsilon & U \\ U & \lambda - \epsilon \end{vmatrix} = 0$$

This gives:

$$\begin{aligned} (\lambda - \epsilon)^2 &= U^2 \\ \epsilon &= \lambda \pm U = \frac{\hbar^2}{2m} \left(\frac{1}{2}G\right)^2 \pm U \end{aligned}$$

The modified energy eigenvalues have two roots (at specific wavevector $k = \pm G$) separated from one another by an energy difference $2U$ —the band gap.

We can also extract the C coefficients (reminder: these are fourier series coefficients), which will allow us to write the electron wavefunction at the the Brillouin zone boundary. As a reminder, electron wavefunctions can be written in terms of C coefficients in the following way: $\psi_k = \sum_G C(k - G)e^{i(k-G)x}$

Using the first central equation above and the energy eigenvalues:

$$\frac{C(-\frac{1}{2}G)}{C(\frac{1}{2}G)} = \frac{\epsilon - \lambda}{U} = \pm 1$$

This gives the the wavefunction near the zone boundary as:

$$\psi(x) = e^{iGx/2} \pm e^{-iGx/2}$$

This is the same as the guess at the start of the chapter (nearly free electron model, standing wave).

Rewrite the equations of the central equation as a function of k

$$(\lambda_k - \epsilon)C(k) + UC(k - G) = 0$$

$$(\lambda_{k-G} - \epsilon)C(k - G) + UC(k) = 0$$

Put in matrix form and set determinant equal to zero:

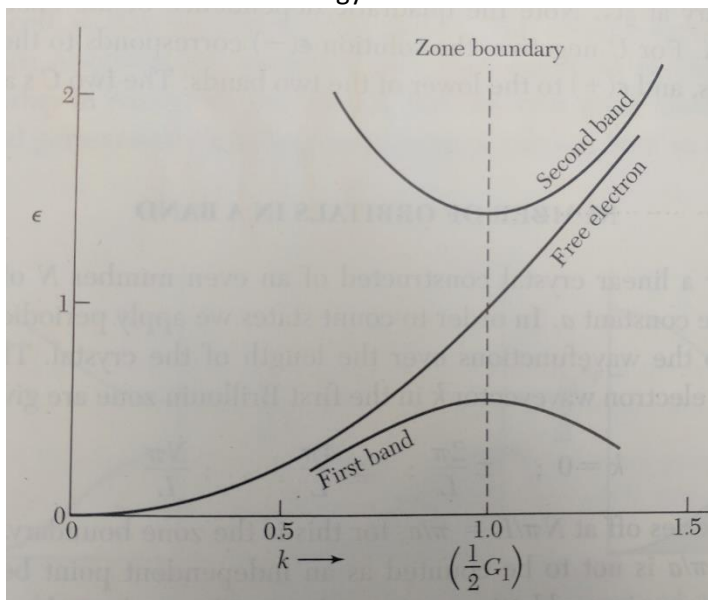
$$\begin{vmatrix} \lambda_k - \epsilon & U \\ U & \lambda_{k-G} - \epsilon \end{vmatrix} = 0$$

$$\epsilon^2 - \epsilon(\lambda_{k-G} + \lambda_k) + \lambda_k \lambda_{k-G} - U^2 = 0$$

There are two solutions:

$$\epsilon = \frac{1}{2}(\lambda_{k-G} + \lambda_k) \pm \left[\frac{1}{4}(\lambda_{k-G} - \lambda_k)^2 + U^2 \right]^{1/2}$$

Each root describes an energy band.



Expand in terms of $\tilde{K} \equiv k - \frac{1}{2}G$

In this notation, $\lambda_k \rightarrow \lambda_{\tilde{K} + \frac{1}{2}G}$

And $\lambda_{k-G} \rightarrow \lambda_{\tilde{K} - \frac{1}{2}G}$

$$\epsilon_{\tilde{K}} = \frac{\hbar^2}{2m} \left(\frac{1}{4}G^2 + \tilde{K}^2 \right) \pm \left[4\lambda \left(\frac{\hbar^2 \tilde{K}^2}{2m} \right) + U^2 \right]^{1/2}$$

Where $\lambda \equiv \left(\frac{\hbar^2}{2m} \right) \left(\frac{1}{2}G \right)^2$

In the limit $\frac{\hbar^2 G \tilde{K}}{2m} \ll U$ the solutions can be approximated as:

$$\begin{aligned}\epsilon_{\tilde{K}} &\approx \frac{\hbar^2}{2m} \left(\frac{1}{4} G^2 + \tilde{K}^2 \right) \pm U \left[1 + 2 \left(\frac{\lambda}{U^2} \right) \left(\frac{\hbar^2 \tilde{K}^2}{2m} \right) \right] \\ &= \epsilon(\pm) + \frac{\hbar^2 \tilde{K}^2}{2m} \left(1 \pm \frac{2\lambda}{U} \right)\end{aligned}$$

Where $\epsilon(\pm) = \frac{\hbar^2}{2m} \left(\frac{1}{2} G \right)^2 \pm U$

These are the roots for energy as a function of crystal momentum close to the Brillouin zone boundary.

Empty lattice approximation

The empty lattice approximation introduces the *periodicity* of the lattice but with zero potential. In the textbook, this concept is used to introduce the consequences of the fact that dispersion relations are *fully defined* in the first Brillouin zone. When all energy vs momentum information is presented in the first Brillouin zone only, it is called the ‘reduced zone scheme.’ Let’s begin with a free-electron-like dispersion:

$$\epsilon_k = \frac{\hbar^2 k^2}{2m}$$

If k happens to be outside of the first Brillouin zone, one can translate it back into the first Brillouin zone via a reciprocal lattice vector. The procedure is to look for a \mathbf{G} such that \mathbf{k}' in the first Brillouin zone satisfies:

$$\mathbf{k}' + \mathbf{G} = \mathbf{k}$$

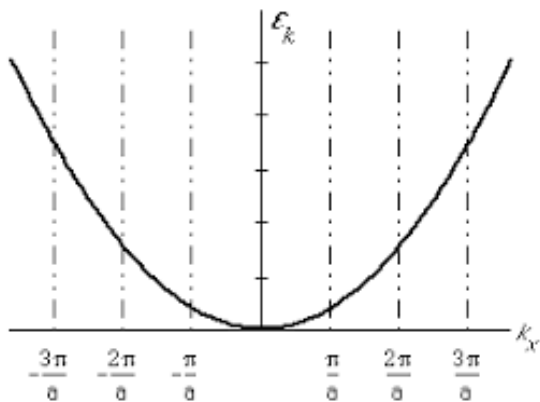


Figure 2. Free electron like dispersion (1D) drawn over several Brillouin zones

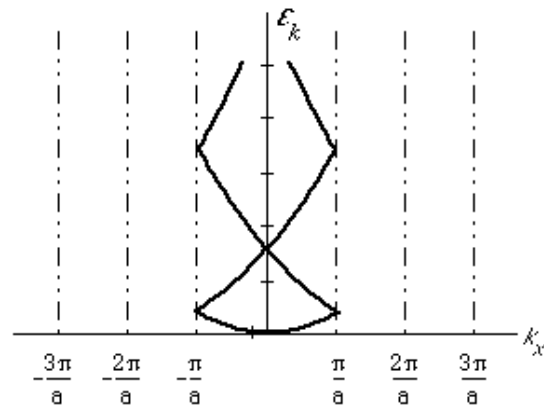


Figure 1. Free electron dispersion (1D) folded into first Brillouin zone

In three dimensions, the free electron energy can be written as:

$$\epsilon(k_x, k_y, k_z) = \frac{\hbar^2}{2m} (\mathbf{k} + \mathbf{G})^2 = \frac{\hbar^2}{2m} \left[(k_x + G_x)^2 + (k_y + G_y)^2 + (k_z + G_z)^2 \right]$$

An example of this so-called ‘reduced zone scheme’ in three dimensions is discussed on p 176-177 of the textbook.

Lecture 15

- Empty lattice approximation
- Number of orbitals in a band
- Direct vs indirect band gap

Empty lattice approximation

The empty lattice approximation introduces the *periodicity* of the lattice but with zero potential. In the textbook, this concept is used to introduce the consequences of the fact that dispersion relations are *fully defined* in the first Brillouin zone. When all energy vs momentum information is presented in the first Brillouin zone only, it is called the 'reduced zone scheme'.

If k happens to be outside of the first Brillouin zone, one can translate it back into the first Brillouin zone via a reciprocal lattice vector. The procedure is to look for a \mathbf{G} such that \mathbf{k}' in the first Brillouin zone satisfies:

$$\mathbf{k}' + \mathbf{G} = \mathbf{k}$$

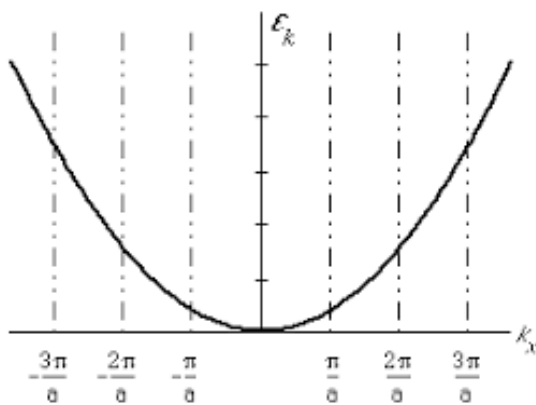


Figure 2. Free electron like dispersion (1D) drawn over several Brillouin zones (extended zone scheme)

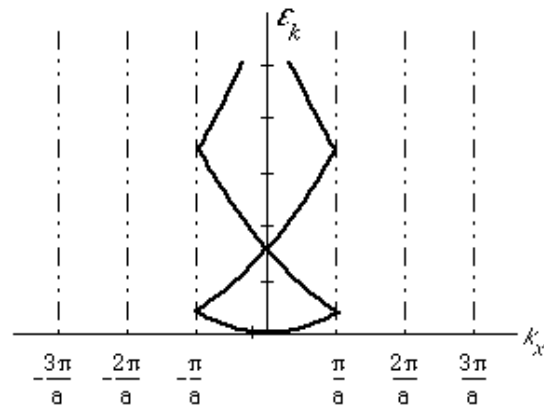


Figure 1. Free electron dispersion (1D) folded into first Brillouin zone (reduced zone scheme)

In three dimensions, the free electron energy can be written as:

$$\epsilon(k_x, k_y, k_z) = \frac{\hbar^2}{2m} (\mathbf{k} + \mathbf{G})^2 = \frac{\hbar^2}{2m} [(k_x + G_x)^2 + (k_y + G_y)^2 + (k_z + G_z)^2]$$

An example of this so-called 'reduced zone scheme' in three dimensions is discussed on p 176-177 of the textbook.

Number of orbitals in a band

Consider a 1D crystal that consists of N primitive cells

The allowed values of k in the **first Brillouin zone** are $k = 0, \pm \frac{2\pi}{L}, \pm \frac{4\pi}{L}, \dots, \pm \frac{N\pi}{L}$

We know that $\frac{N\pi}{L}$ is the proper cutoff because $k = \pm \pi/a$ at the zone boundary, and $a = L/N$

The number of permissible k -values in the first Brillouin zone is: $\# = \frac{\text{span of allowed } k}{\text{spacing between each } k} = \frac{\frac{2N\pi}{L}}{\frac{2\pi}{L}} = N$

This means that each primitive cell contributes one independent value of k to each energy band. This result applies to 2 and 3 dimensions as well.

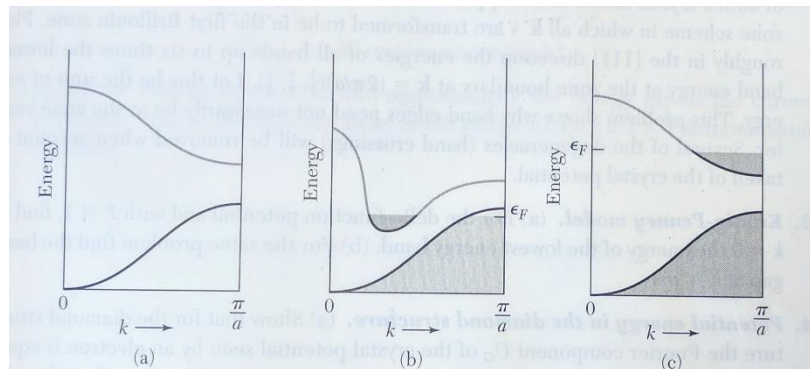
When we account for two independent orientations of electron spin, there are **2N independent orbitals in each energy band**. The implications of these are as follows:

- If there is one valence electron per primitive cell, there will be N valence electrons in total, and the band will be half filled with electrons
- If there are 2 valence electrons per primitive cell, there will be $2N$ valence electrons in total. The band can potentially be fully filled up to the band gap
- If there are 3 valence electrons per primitive cell, there will be $3N$ valence electrons total. The first band will be fully filled, and the remaining N electrons will go into the next band above the band gap

Metals vs insulators

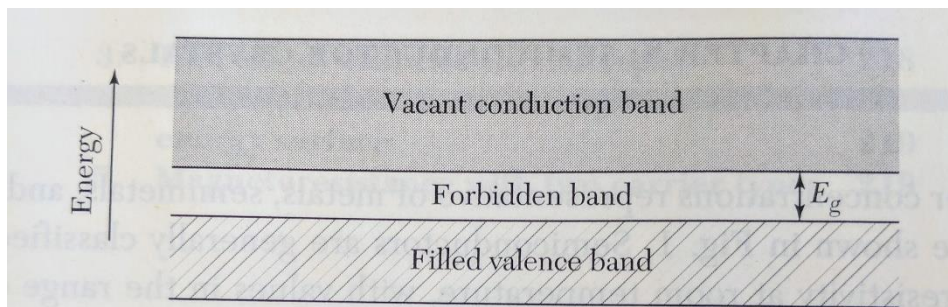
An insulator (or semiconductor) has electron states filled up to the band gap, such that small excitations (e.g. temperature) are typically insufficient to promote an electron into a permissible state. A crystal can be an insulator if there are an even number of valence electrons per primitive cell. However, not all crystals that have this property are insulators (e.g. the entire 2nd column of the periodic table has 2 valence electrons, but they are metals). In 1D, even number of electrons always corresponds to an insulator.

A crystal will be a metal if it has an odd number of valence electrons.



In the image to the right, the pictures from left to right correspond to an insulator, a metal, and a metal.

Ch8: Direct vs indirect band gap



In the previous chapter we learned about how band gaps can arise from the periodic ionic potential. In a semiconductor (or an insulator), electrons

are filled up to the top of a band, called the valence band, and a range of forbidden energies (the band gap) separates the valence band from the conduction band. The only difference between a

semiconductor and an insulator is in the *size* of the band gap, with insulators having a larger one. The distinction is usually defined in that the band gap of semiconductors is *smaller* than optical frequencies.

The easiest way to measure the magnitude of a band gap is optically, by shining light of different frequencies onto a semiconductor and seeing the threshold frequency for absorption

- Photon is absorbed → its frequency is *larger* than the band gap
- Photon is not absorbed (transmitted) → its frequency is *smaller* than the band gap

There are two types of absorption processes, direct and indirect:

- Direct absorption: The threshold for absorption measures the band gap, $E_g = \hbar\omega_g$. This occurs when there is a **direct band gap**, meaning the top of the valence band has the same k as the bottom of the conduction band. An example of a direct band gap semiconductor is GaAs, and direct band gap semiconductors are required for light-emitting applications (e.g. LEDs)

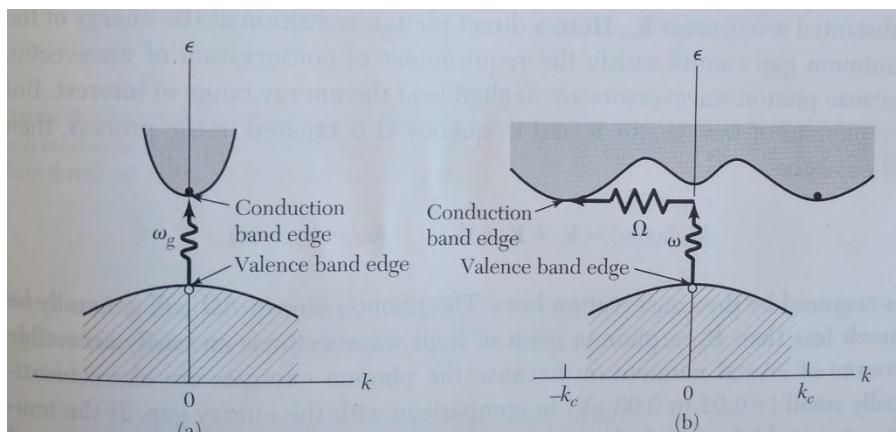
In a direct band gap semiconductor, the bottom of conduction band is located at the same k as the top of the valence band

- Indirect absorption: this occurs when the maximum of the valence band is separated from the minimum of the conduction band by a substantial momentum k_c (**indirect band gap**). Examples of indirect band gap semiconductors are silicon and germanium. A photon with frequency of the minimum band gap cannot traverse the band gap without a momentum kick. A phonon of wavevector \mathbf{K} and frequency Ω can provide this momentum kick. Phonon energies are typically much smaller than band gap energies, and this process is more probable at higher temperature when the needed phonon mode is populated.

$$\mathbf{k}(\text{photon}) = \mathbf{k}_c + \mathbf{K} \approx 0$$

$$\hbar\omega = E_g + \hbar\Omega$$

In an indirect band gap semiconductor, the bottom of conduction band is located at a different k as the top of the valence band



Equations of motion

In this section, we consider the effect of applied electric and magnetic fields on the motion of electrons in a *semiconductor*. This is an extension of the content at the end of Ch 6, but with two major differences

- We are talking

about semiconductors, which are electrically insulating at $T=0$

- The lattice exists

The starting point is the relationship between an electron's group velocity and its ϵ vs k

$$v_g = \hbar^{-1} d\epsilon/dk$$

$$\mathbf{v} = \hbar^{-1} \nabla_{\mathbf{k}} \epsilon(\mathbf{k})$$

The work $\delta\epsilon$ done on the electron by the electric field \mathbf{E} in time interval δt is given by: $\delta\epsilon = -eE v_g \delta t$

Lecture 16

Equations of motion

In this section, we consider the effect of applied electric and magnetic fields on the motion of electrons in a *semiconductor*. This is an extension of the content at the end of Ch 6, but with two major differences

- We are talking about semiconductors, which are electrically insulating at $T=0$
- The lattice exists, and interactions between electrons and lattice will be considered

The starting point is the relationship between an electron's group velocity and its ϵ vs k

$$v_g = \hbar^{-1} d\epsilon/dk$$
$$\mathbf{v} = \hbar^{-1} \nabla_{\mathbf{k}} \epsilon(\mathbf{k})$$

generally: $\hbar \frac{d\mathbf{k}}{dt} = \mathbf{F}$

This is the same equation as we used in Ch6.

The derivation of this is shown in a few lines, but will not be presented in class.

For an electric field, the equation becomes:

$$\hbar \frac{d\mathbf{k}}{dt} = -e\mathbf{E}$$

For a magnetic field:

$$\hbar \frac{d\mathbf{k}}{dt} = -e\mathbf{v} \times \mathbf{B}$$
$$\frac{d\mathbf{k}}{dt} = -\frac{e}{\hbar^2} \nabla_{\mathbf{k}} \epsilon \times \mathbf{B}$$

According to the second equation, an electron in a magnetic field moves in \mathbf{k} -space in a direction **normal** to the gradient of ϵ vs \mathbf{k} . Thus, electrons in a magnetic field move on a surface of constant energy. The motion in \mathbf{k} space is also on a plane normal to the direction of \mathbf{B} .

Derivation of $\hbar \dot{\mathbf{k}} = \mathbf{F}$ (Note: this is shown here for completeness but not covered in class)

Consider Bloch functions $\psi_{\mathbf{k}}$ which are defined by wavevector \mathbf{k} and have energy $\epsilon_{\mathbf{k}}$

$$\psi_{\mathbf{k}} = \sum_{\mathbf{G}} C(\mathbf{k} + \mathbf{G}) e^{i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}}$$

The expectation value of the momentum of an electron in Bloch state \mathbf{k} is given by:

$$\mathbf{p}_{el} = \langle \psi_{\mathbf{k}} | -i\hbar \nabla | \psi_{\mathbf{k}} \rangle = \sum_{\mathbf{G}} \hbar(\mathbf{k} + \mathbf{G}) |C(\mathbf{k} + \mathbf{G})|^2 = \hbar(\mathbf{k} + \sum_{\mathbf{G}} \mathbf{G} |C(\mathbf{k} + \mathbf{G})|^2)$$

The last step uses the following relation:

$$\sum_{\mathbf{G}} |C(\mathbf{k} + \mathbf{G})|^2 = 1$$

We now examine the following scenario:

- The application of an external force changes the electron momentum from \mathbf{k} to $\mathbf{k} + \Delta\mathbf{k}$, and this extra momentum is transferred to the lattice
- Insulating crystal, electrostatically neutral
- A weak external force is applied over a time interval such that the total impulse given is $\mathbf{J} = \int \mathbf{F} dt$

If the conduction electrons were free-electron-like ($m_{eff} = m_{electron}$), the total momentum imparted to the crystal system would appear in the change in momentum of conduction electrons

$$\mathbf{J} = \Delta \mathbf{p}_{tot} = \Delta \mathbf{p}_{el} = \hbar \Delta \mathbf{k}$$

If the conduction electrons interact with the periodic potential of the crystal lattice, we have the following relation:

$$\mathbf{J} = \Delta \mathbf{p}_{tot} = \Delta \mathbf{p}_{el} + \Delta \mathbf{p}_{lat}$$

We already have an expression for $\Delta \mathbf{p}_{el}$ from earlier:

$$\Delta \mathbf{p}_{el} = \hbar \Delta \mathbf{k} + \sum_{\mathbf{G}} \hbar \mathbf{G} [(\nabla_{\mathbf{k}} |C(\mathbf{k} + \mathbf{G})|^2) \cdot \Delta \mathbf{k}]$$

$\Delta \mathbf{p}_{lat}$ is derived in the following way:

- An electron which collides with the lattice transfers momentum to the lattice
- If an incident electron with momentum $\hbar \mathbf{k}$ and its final momentum is $\hbar(\mathbf{k} + \mathbf{G})$, the lattice will acquire momentum $-\hbar \mathbf{G}$

The momentum transferred to the lattice when the electron state goes from $\psi_{\mathbf{k}}$ to $\psi_{\mathbf{k}+\Delta \mathbf{k}}$ is:

$$\Delta \mathbf{p}_{lat} = -\hbar \sum_{\mathbf{G}} \mathbf{G} [(\nabla_{\mathbf{k}} |C(\mathbf{k} + \mathbf{G})|^2) \cdot \Delta \mathbf{k}]$$

The total momentum change is then:

$$\mathbf{J} = \Delta \mathbf{p}_{el} + \Delta \mathbf{p}_{lat} = \hbar \Delta \mathbf{k} + \sum_{\mathbf{G}} \hbar \mathbf{G} [(\nabla_{\mathbf{k}} |C(\mathbf{k} + \mathbf{G})|^2) \cdot \Delta \mathbf{k}] - \hbar \sum_{\mathbf{G}} \mathbf{G} [(\nabla_{\mathbf{k}} |C(\mathbf{k} + \mathbf{G})|^2) \cdot \Delta \mathbf{k}] = \hbar \Delta \mathbf{k}$$

Same as free electrons

From the original definition of \mathbf{J} ($\mathbf{J} = \int \mathbf{F} dt$), we get back $\hbar \frac{d\mathbf{k}}{dt} = \mathbf{F}$

Holes

In solid state physics there are many examples of emergent particles or quasiparticles that do not exist outside the solid. Sometimes, those quasiparticles are analogues of particles found in free space, and sometimes they are unique particles that have never been observed outside of crystalline solids. Last quarter, you learned about one type of emergent particle, the phonon, which is analogous to a photon, other than their dispersion relations. In semiconductors, another important particle appears—the hole. The hole is analogous to a positron, a positively charged electron. A hole is technically the *absence* of an electron, but it is much more straightforward to think of it as an electron-like particle with positive charge.

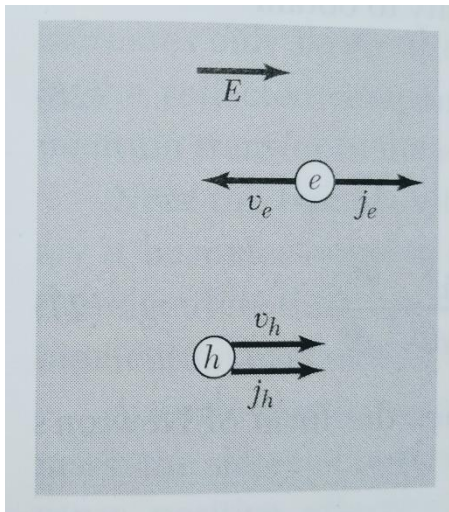


Figure 1. Drift velocity and electric current for electrons and holes.

In an electric or magnetic field, a hole acts exactly as if it has a charge $+e$, which is why it makes sense to think of it like a positron. Additionally, holes have *negative* effective masses, which happens because their response to EM fields is opposite to electrons, as shown in the image to the left.

We will see another rationale for this negative effective mass in the section on effective mass below.

Effective mass

The dispersion relation for free electrons, $\epsilon = \hbar^2 k^2 / 2m$ shows that the 2nd derivative of ϵ with respect to k is related inversely to mass. This relationship holds even if the dispersion is not free-electron like to give an ‘effective mass’—the mass that accurately describes how charged particles respond to

temperature and external fields. Electronic heat capacity in the limit of $T \rightarrow 0$ gave one way to quantify the ‘effective mass’ of electrons inside a crystal that is different from the free electron mass (applicable only to a metal). The method presented in this chapter (2nd derivative of dispersion relation) is broadly applicable to both metals and insulators.

Lets consider effective mass where electron bands no longer behave free-electron-like, near the Brillouin zone boundary. In the previous chapter, the dispersion relation near the zone boundary can be written in the following way for the conduction band:

$$\epsilon(K) = \epsilon_c + \left(\frac{\hbar^2}{2m_e} \right) K^2$$

$$\frac{m_e}{m} = \frac{1}{\frac{2\lambda}{U} - 1}$$

K is the wavevector measured from the zone boundary, and m_e is the effective mass

For the valence band, the dispersion relations can be written as:

$$\epsilon(K) = \epsilon_v - \left(\frac{\hbar^2}{2m_h} \right) K^2$$

$$\frac{m_h}{m} = \frac{1}{\frac{2\lambda}{U} + 1}$$

More generally, we arrive at this definition of effective mass starting with the expression for group velocity:

$$\frac{dv_g}{dt} = \hbar^{-1} \frac{d^2\epsilon}{dk dt} = \hbar^{-1} \frac{d^2\epsilon}{dk^2} \frac{dk}{dt}$$

$$\frac{dv_g}{dt} = \left(\frac{1}{\hbar^2} \frac{d^2\epsilon}{dk^2} \right) F$$

$$F = \frac{\hbar^2}{\frac{d^2\epsilon}{dk^2}} \frac{dv_g}{dt} = ma$$

The effective mass is thus defined:

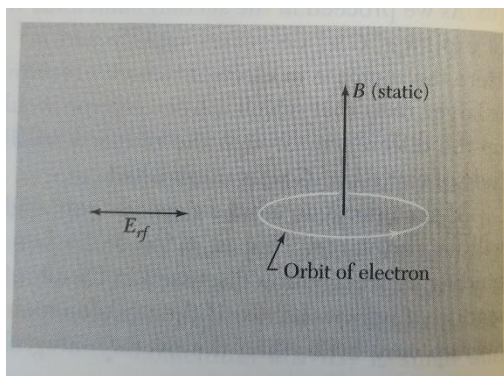
$$\frac{1}{m^*} = \frac{1}{\hbar^2} \frac{d^2\epsilon}{dk^2}$$

In general, ϵ may be a complicated function of k , such that we need to introduce a mass tensor to account for the different electron response for fields applied along different directions. In the equations below, μ and ν are cartesian coordinates.

$$\left(\frac{1}{m^*} \right)_{\mu\nu} = \frac{1}{\hbar^2} \frac{d^2\epsilon_k}{dk_\mu dk_\nu}$$

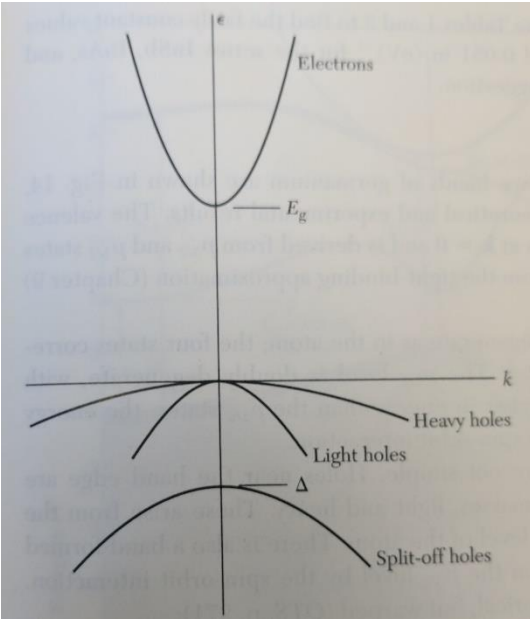
$$\frac{dv_\mu}{dt} = \left(\frac{1}{m^*} \right)_{\mu\nu} F_\nu$$

Measuring effective mass in semiconductors



One way of measuring effective masses in semiconductors is via cyclotron resonance. When a magnetic field is applied, charge carriers are accelerated in helical orbits about the magnetic field axis (opposite direction for electrons and holes). The angular rotation frequency (ω_c) is called the cyclotron frequency.

$$\omega_c = \frac{eB}{m^*c}$$



This is measured by applying a radio frequency (RF) electric field along the axis orthogonal to the magnetic field. There is resonant (sharply increased) absorption when the RF frequency is equal to the cyclotron frequency.

The band structure is often more complex than we have discussed thus far, as shown in the image to the left for a direct band gap semiconductor. The conduction band dispersion relation is given by:

$$\epsilon_c = E_g + \hbar^2 k^2 / 2m_e$$

Where “zero” energy is taken to be the top of the valence band.

There are **three** valence bands: heavy hole, light holes, and a hole-like band that is split off from the others by spin orbit coupling (relativistic interaction between electron’s spin and its motion). The dispersions of these three bands are approximated as follows:

$$\epsilon_v(hh) \approx -\frac{\hbar^2 k^2}{2m_{hh}}$$

$$\epsilon_v(lh) \approx -\frac{\hbar^2 k^2}{2m_{lh}}$$

$$\epsilon_v(soh) \approx -\Delta - \frac{\hbar^2 k^2}{2m_{soh}}$$

Silicon and Germanium

Silicon and Germanium are both indirect band gap semiconductors for which the top of the valence band is located at $\mathbf{k}=0$. They both have the diamond crystal structure which is based on a fcc lattice. Constant-energy surfaces are not spherical or centered around $\mathbf{k}=0$.

The band dispersion can be written in the following way:

$$\epsilon(\mathbf{k}) = Ak^2 \pm [B^2k^4 + C^2(k_x^2k_y^2 + k_y^2k_z^2 + k_z^2k_x^2)]^{1/2}$$

The split-off band is given by $-\Delta + Ak^2$

Experimentally, constants A-C, in units of $\hbar^2/2m$ have been found to be the following for Ge:

$$A=-13.38, |B|=8.48, |C|=13.5, \Delta = 0.29 \text{ eV}$$

Each band has a spheroidal (stretched sphere) constant-energy surface oriented along the $\langle 111 \rangle$ and equivalent crystal axes. Electrons

in each spheroid constant energy surface have two effective masses, a transverse (m_t) and a longitudinal (m_l). For a static magnetic field at an angle θ with the longitudinal axis of the spheroid, the cyclotron mass (m_c) is given by:

$$\frac{1}{m_c^2} = \frac{\cos^2 \theta}{m_t^2} + \frac{\sin^2 \theta}{m_t m_l}$$

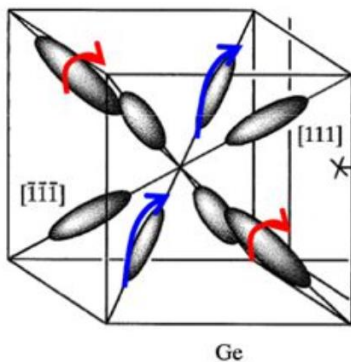
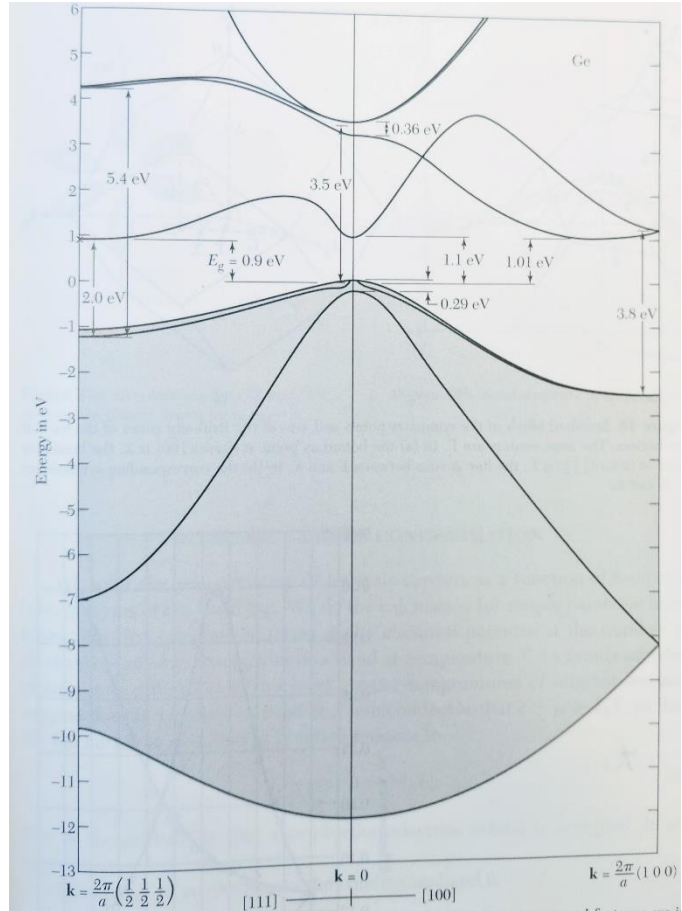


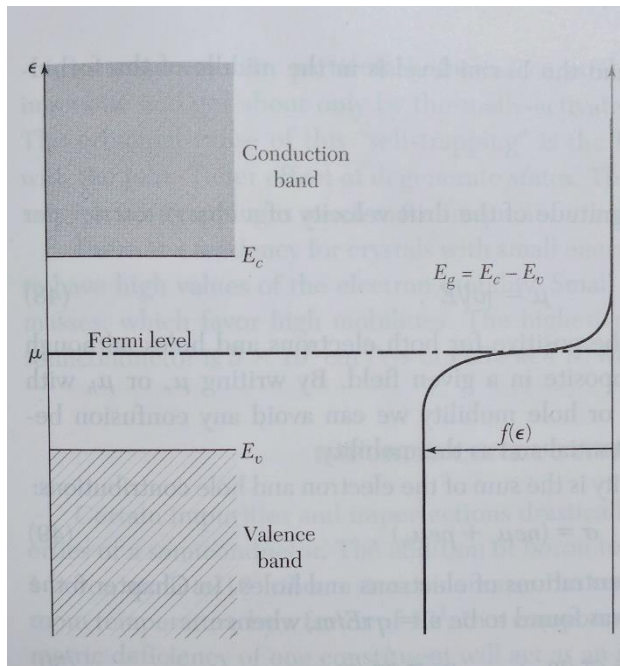
Figure 2. Transverse in red, longitudinal in blue.

Physics 140B—Lecture 17

- Intrinsic carrier concentration
- Mobility

Intrinsic Carrier Concentration

At zero temperature, a semiconductor is insulating, but at elevated temperatures, there is some small possibility of exciting electrons across the band gap. These excited electrons and the holes they leave behind constitute the ‘intrinsic carriers’ in a semiconductor. Assume we are in a temperature regime where $\epsilon - \mu \gg k_B T$ so that the Fermi-dirac function ($f_e = \frac{1}{e^{(\epsilon-\mu)/k_B T} + 1}$) can reduce to simply an exponential:



$$f_e \sim e^{(\mu-\epsilon)/k_B T}$$

This allows us to approximate the probability of occupation in the conduction band at temperature T. The energy of an electron in the conduction band is given by:

$$\epsilon_k = E_c + \hbar^2 k^2 / 2m_e$$

The density of electron states in 3D is given by:

$$D_e(\epsilon) = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} (\epsilon - E_c)^{1/2}$$

This allows us to calculate the concentration of electrons in the conduction band

$$\begin{aligned} n &= \int_{E_c}^{\infty} D_e(\epsilon) f_e(\epsilon) d\epsilon \\ &= \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} e^{\mu/k_B T} \int_{E_c}^{\infty} (\epsilon - E_c)^{1/2} e^{-\epsilon/k_B T} d\epsilon \\ &= 2 \left(\frac{m_e k_B T}{2\pi \hbar^2} \right)^{3/2} e^{(\mu-E_c)/k_B T} \end{aligned}$$

We can use the same logic to calculate the concentration of holes in the valence band.

First, the occupation number of holes is given by $f_h = 1 - f_e$ (at every energy if a state is *unoccupied* by an electron, this is the same as being occupied by a hole). Second, the density of states has the same energy dependence, and the only difference is that it uses a hole effective mass and references the energy of the top of the valence band, E_v

$$f_h = 1 - \frac{1}{e^{\frac{\epsilon-\mu}{k_B T} + 1}} = \frac{1}{e^{\frac{\mu-\epsilon}{k_B T} + 1}} \approx e^{(\epsilon-\mu)/k_B T}$$

$$p = \int_{-\infty}^{E_v} D_h(\epsilon) f_h(\epsilon) d\epsilon = 2 \left(\frac{m_h k_B T}{2\pi \hbar^2} \right)^{3/2} e^{(E_v - \mu)/k_B T}$$

The product of these is given by:

$$np = 4 \left(\frac{k_B T}{2\pi \hbar^2} \right)^3 (m_e m_h)^{3/2} e^{-E_g/k_B T}$$

Which uses the relation $E_g = E_c - E_v$. This result does not rely on knowing the chemical potential, μ , and it will hold when we consider impurities in semiconductors as well. At a given temperature, the product of the number of electrons and holes is constant.

In an intrinsic semiconductor (i.e. a pure semiconductor with no carriers intentionally added), the electron and hole concentrations are equal:

$$n_i = p_i = 2 \left(\frac{k_B T}{2\pi \hbar^2} \right)^{3/2} (m_e m_h)^{3/4} e^{-E_g/2k_B T}$$

We can also use this to solve for the chemical potential, by setting the original expressions for n and p, which contained μ , equal to each other:

$$e^{2\mu/k_B T} = \left(\frac{m_h}{m_e} \right)^{3/2} e^{E_g/k_B T}$$

$$\mu = \frac{1}{2} E_g + \frac{3}{4} k_B T \ln m_h/m_e$$

If $m_h = m_e$, then $\mu = \frac{1}{2} E_g$ and the Fermi level is exactly in the middle of the band gap. At T=0, it will be in the middle of the band gap regardless of electron and hole masses.

Intrinsic mobility

The mobility is defined as the ratio between a charge carriers drift velocity and the applied electric field. The drift velocity is the velocity that charged particles will travel under applied electric field, related to current, but excluding some factors (charge, density). Mobility is always defined to be **positive** for electrons and holes, even though drift velocities are in different directions.

$$\mu \equiv |v|/E$$

The electrical conductivity for a system with different charge carrier species can be defined in terms of mobility:

$$\sigma = ne\mu_e + pe\mu_h$$

Where μ_e and μ_h are mobilities for electrons and holes respectively. As a reminder, $\mathbf{j} = \sigma \mathbf{E}$, where \mathbf{j} is a current density. As before, n and p are electron and hole concentrations.

The drift velocity of charge q is related to electric field and mean collision time (τ) via (see Ch6):

$$v = q\tau E/m$$

Which allows for the following definitions of mobility:

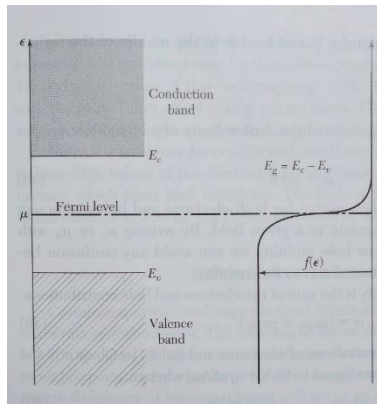
$$\mu_e = \frac{e\tau_e}{m_e}, \mu_h = e\tau_h/m_h$$

Mobilities have a slight temperature dependence, going as a power law, so the temperature dependence of σ is dominated by the temperature dependence of charge concentrations which vary exponentially with temperature as found in the previous section.

Physics 140B lecture 18

- Review intrinsic semiconductors
- Dopants in semiconductors
- p-n junction

Intrinsic semiconductors



When $k_B T \ll E_g$

$$f_e(\epsilon) \sim e^{(\mu - \epsilon)/k_B T}$$

Can calculate the number of thermally excited electrons (n) in the conduction band:

$$n = \int_{E_c}^{\infty} D_e(\epsilon) f_e(\epsilon) d\epsilon = 2 \left(\frac{m_e k_B T}{2\pi \hbar^2} \right)^{3/2} e^{(\mu - E_c)/k_B T}$$

Can similarly calculate the number of holes (p) left behind in the valence band

$$p = 2 \left(\frac{m_h k_B T}{2\pi \hbar^2} \right)^{3/2} e^{(E_v - \mu)/k_B T}$$

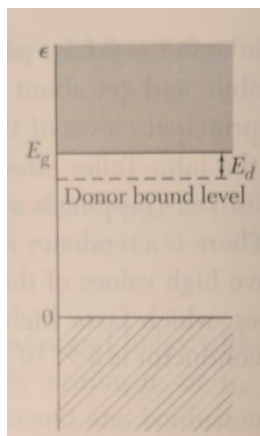
The product of these is given by:

$$np = 4 \left(\frac{k_B T}{2\pi \hbar^2} \right)^3 (m_e m_h)^{3/2} e^{-E_g/k_B T}$$

Which uses the relation $E_g = E_c - E_v$

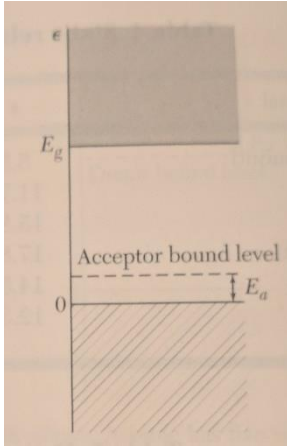
Impurity conductivity: donors and acceptors

A chunk of pure semiconductor does not a transistor make, and to engineer useful devices out of semiconductors, it is necessary to introduce small numbers of substitutional impurities. This is called 'doping'. Doping is important because it allows to precisely control the **sign** and **abundance** of charge carriers, crucial for engineering devices. Doping allow another method, besides raising the temperature, for introducing electrons into the conduction band and holes into the valence band



Silicon and Germanium both have 4 valence electrons, and in the crystal structure (diamond structure) each atom bonds with 4 other atoms. If one silicon atom is replaced by another atom that has 5 valence electrons (e.g. P, As, Sb), that impurity atom will attempt to bond with its four neighbors, but it will have one electron left over. This type of impurity is called a **donor** because it can donate its extra electron to be mobile in the crystal.

One way to model this donor atom and its extra valence electron is as a hydrogen-like atom, with the extra electron feeling a coulomb potential $\sim e/\epsilon r$, where ϵ is the **static dielectric constant** of the medium (silicon) which is used to replace the permittivity



of free space in the normal expression for coulomb potential, and better describe the coulomb potential that this electron actually feels (the other electrons in the Si specimen essentially 'screen out' some of the positive charge of the donor nucleus).

An expression for the binding energy of the extra donor electron can be expressed in terms of the ground state of the Bohr model:

$$E_d = \frac{e^4 m_e}{2(4\pi\epsilon\epsilon_0\hbar)^2} = \frac{13.6 m_e}{\epsilon^2 m} eV$$

Where m_e is the effective mass of the donor electron and m is the free electron mass.

Typical values for donor binding energies are several meV to several 10s of meV.

The typical distance of the donor atom from the ionic core on the ground state can also be expressed in terms of hydrogen atom concepts. The ground state typical distance is modified from the Bohr radius.

$$a_d = \frac{4\pi\epsilon\epsilon_0\hbar^2}{m_e e^2} = \left(\frac{0.53 \epsilon}{m_e/m}\right) \text{\AA}$$

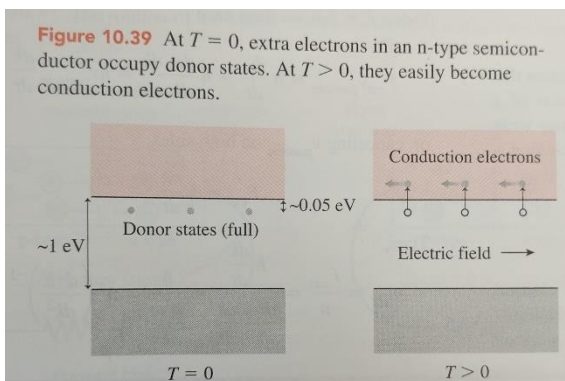
Silicon atoms can also be replaced by trivalent atoms (e.g. B, Al, Ga), such that the impurity is short one electron. This is called **hole doping** and the doped impurity is called an **acceptor**. All of the equations above apply, except replacing effective masses with the hole version and the model is 'antihydrogen' with a negatively charged nucleus and a positively charged bound particle.

At zero temperature, impurities do not contribute to conductivity, because there is no thermal energy to excite into the conduction band (doped electrons) or valence band (doped holes). If there are no acceptors present, the concentration of mobile electrons in the conduction band is given by:

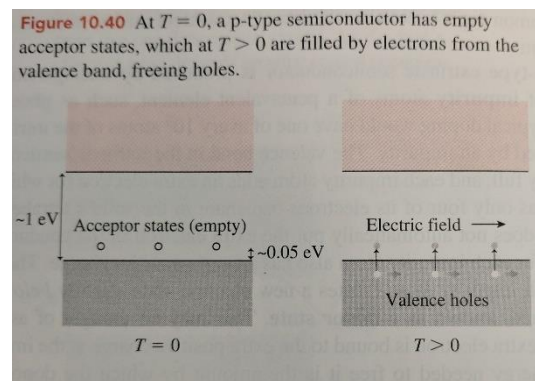
$$n_d \approx (n_0 N_d)^{\frac{1}{2}} e^{-E_d/2k_B T}$$

Where $n_0 \equiv 2 \left(\frac{m_e k_B T}{2\pi\hbar^2}\right)^{3/2}$, N_d is the concentration of donor atoms, and it is assumed that $k_B T \ll E_d$.

A similar expression can be written for acceptors:



2. From Harris Modern Physics



1. From Harris Modern Physics

$$p_a \approx (p_0 N_a)^{\frac{1}{2}} e^{-E_a/2k_B T}$$

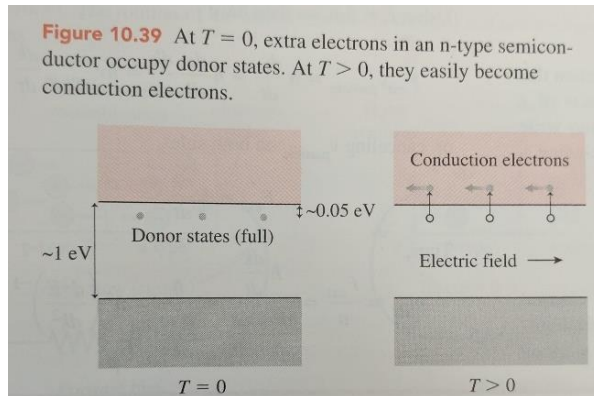
Once doped electrons are excited into the conduction band (or doped holes into the valence band) they can contribute to conductivity.

Physics 140B lecture 19

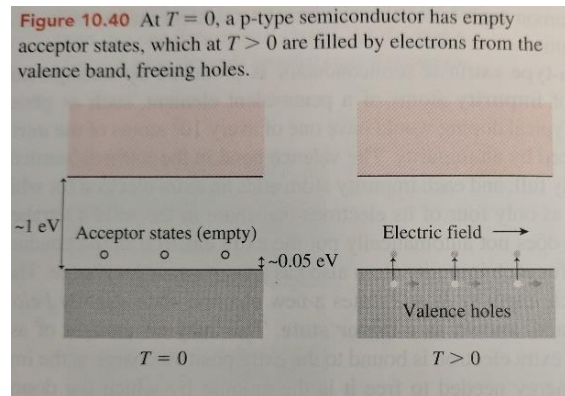
- Review doped semiconductors
- The pn junction + applications

Doped semiconductors

Donor levels, acceptor levels, and their response to electric field at finite temperature is shown in the drawings below.



2. From Harris Modern Physics



1. From Harris Modern Physics

At zero temperature, impurities do not contribute to conductivity, because there is no thermal energy to excite into the conduction band (doped electrons) or valence band (doped holes). If there are no acceptors present, the concentration of mobile electrons in the conduction band is given by:

$$n \approx (n_0 N_d)^{\frac{1}{2}} e^{-E_d/2k_B T}$$

Where $n_0 \equiv 2 \left(\frac{m_e k_B T}{2\pi\hbar^2} \right)^{3/2}$, N_d is the concentration of donor atoms, and it is assumed that $k_B T \ll E_d$.

Note that this limits the regime of applicability to the equation above essentially to cryogenic temperatures only. At room temperature, n_d is comparable to N_d

A similar expression can be written for acceptors:

$$p \approx (p_0 N_a)^{\frac{1}{2}} e^{-E_a/2k_B T}$$

Once doped electrons are excited into the conduction band (or doped holes into the valence band) they can contribute to conductivity.

p-n junctions and diodes

Doping is the first step for making a useful device out of a semiconductor, but one also need to combine semiconductors that are doped in different ways. The simplest semiconductor device is a p-n junction, simply slapping a p-type and an n-type semiconductor together. Although the idea is simple, this technology is very powerful, forming the basis of semiconductor diodes, LEDs, solar cells, photodetectors, CCDs, and diode lasers (e.g. laser pointers).

Terminology:

p-type: hole-doped semiconductor; acceptor states sit close to valence band inside band gap

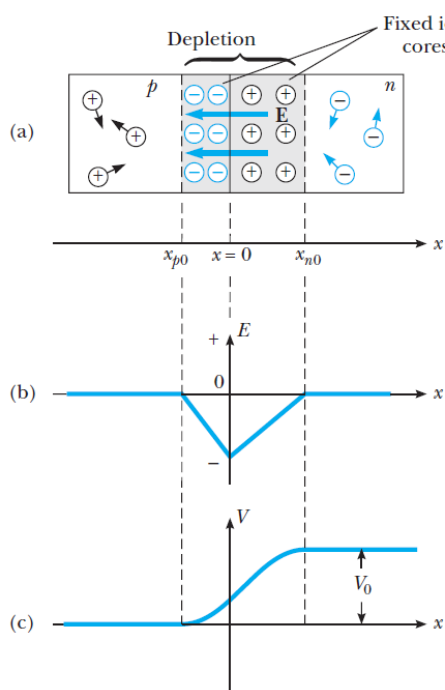
n-type: electron-doped semiconductor; donor states sit close to conduction band inside band gap

intrinsic semiconductor: semiconductor with no intentional doping

minority carriers: in a p-type semiconductor, electrons are minority carriers, and in an n-type semiconductor, holes are minority carriers. Minority carriers respond to an electric field opposite to majority carriers

The p-n junction

The most basic semiconductor device—the diode—arises from joining a p-type semiconductor to an n-type semiconductor.



When these two types of semiconductors are brought together, mobile electrons diffuse to the p-type side and mobile holes diffuse to the n-type side. This leaves behind charged ion cores because the n-type side is now short some electrons and the p-type side is now short some holes. These charged ion cores set up an electric field, or equivalently, a potential barrier which opposes further charge diffusion. This region of charged ion cores which develops at a p-n junction is called a **depletion region** because it is depleted of mobile charge carriers.

One feature of the p-n junction is that it is able to function as a **diode**—a device which passes current only in one direction. Diodes are able to introduce a one way current into a circuit, and semiconductor diodes are the basis of many familiar devices (light-emitting diodes (LEDs), solar cells, diode lasers, photodiodes and CCDs).

The current-voltage (I-V) relationship for an ideal diode is given by:

$$I = I_0(e^{\frac{eV}{k_B T}} - 1)$$

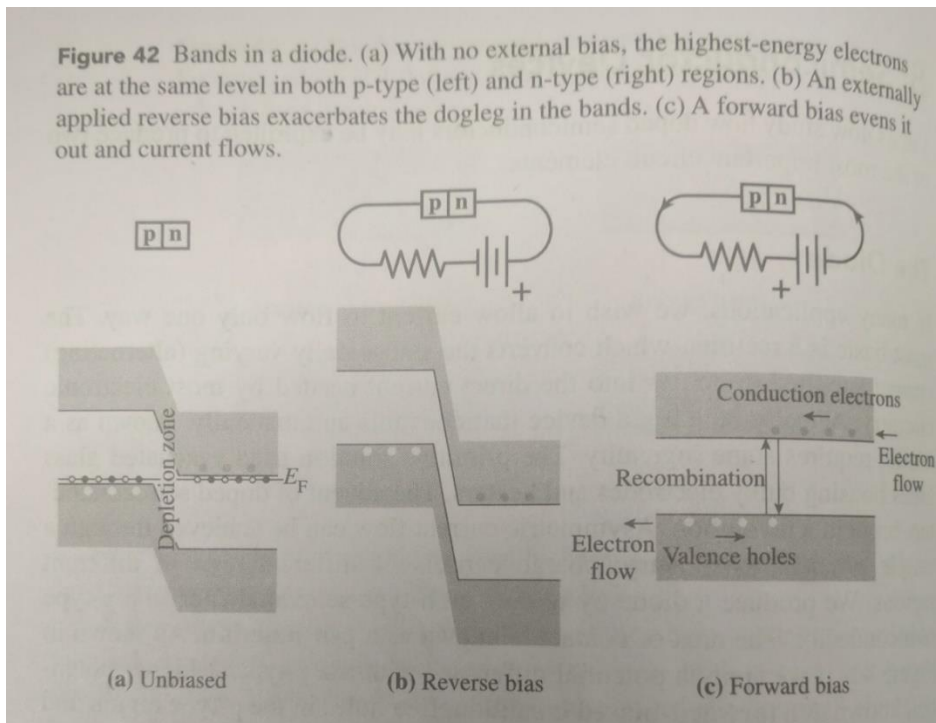
3. p-n junction with depletion region, resulting electric field, and resulting potential. From *Modern Physics* by Serway/Moses/Moyer

This I-V curve has the following attributes:

- For negative bias: current reaches saturation current (also called leakage current) $-I_0$ as negative bias increases; I_0 is typically very small, $\approx 20\mu A$
- For positive bias: current increases exponentially; this is the part where a diode acts like a one-way valve
- At zero bias: there is zero current

Another way to visualize the behavior of a diode is with energy level diagrams. A few preliminary facts:

- In an intrinsic semiconductor, the fermi energy (E_F) is located inside the energy gap (pretty much in the center, see lecture 17).
- With hole doping, E_F moves towards the valence band (you can think of this in terms of hole doping=*decreasing* the number of electrons, and E_F depends on number of electrons)
- With electron doping, E_F moves towards the conduction band (electron doping=*more* electrons=*higher* E_F)
- When a p-type and n-type semiconductor are brought into contact, their Fermi levels align. This puts the conduction/valence bands at different energies on the p and n-doped side. Electrons on the n-doped side now have to overcome an energy barrier to get to the p-doped side. This is precisely from the voltage set up in the depletion region



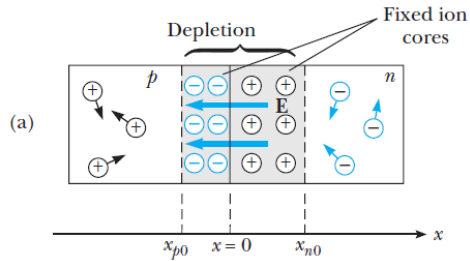
- Reverse bias increases the energy barrier
- Forward bias decreases the barrier. Electrons and holes can now diffuse towards each other and flow in a circuit or recombine

Physics 140B Lecture 20

- Pn junctions
- Begin Ch 9

The p-n junction

The most basic semiconductor device—the diode—arises from joining a p-type semiconductor to an n-type semiconductor.



1. p-n junction with depletion region, resulting electric field, and resulting potential. From *Modern Physics* by Serway/Moses/Moyer

When these two types of semiconductors are brought together, mobile electrons diffuse to the p-type side and mobile holes diffuse to the n-type side. This leaves behind charged ion cores because the n-type side is now short some electrons and the p-type side is now short some holes. These charged ion cores set up an electric field opposes further charge diffusion. This region of charged ion cores which develops at a p-n junction is called a **depletion region** because it is depleted of mobile charge carriers.

One feature of the p-n junction is that it is able to function as a **diode**—a device which passes current only in one direction.

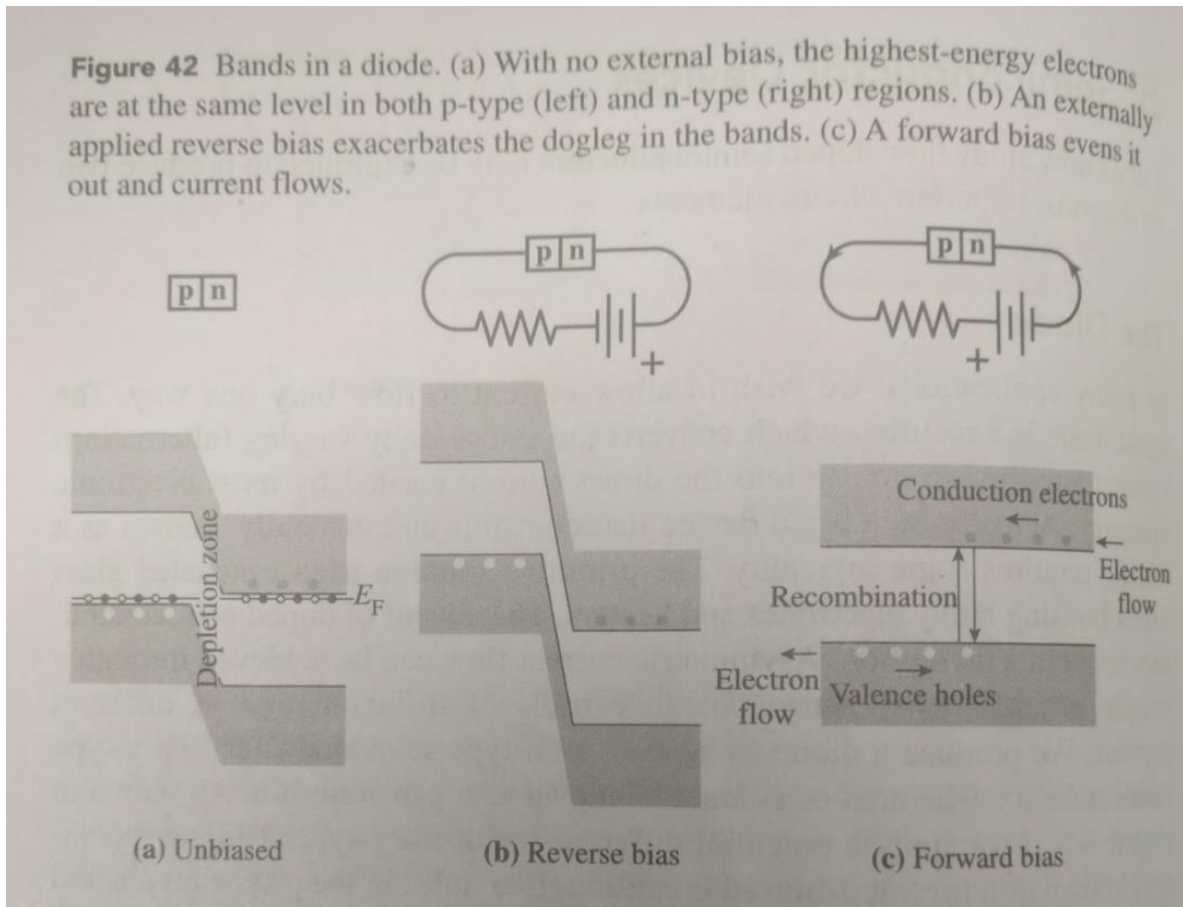
The current-voltage (I-V) relationship for an ideal diode is given by:

$$I = I_0 \left(e^{\frac{eV}{k_B T}} - 1 \right)$$

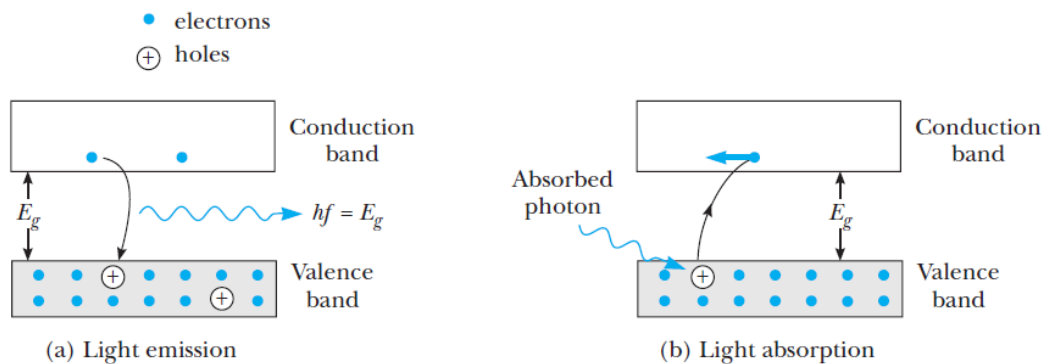
Another way to visualize the behavior of a diode is with energy level diagrams. A few preliminary facts:

- In an intrinsic semiconductor, the fermi energy (E_F) is located inside the energy gap (pretty much in the center, see lecture 17).
- With hole doping, E_F moves towards the valence band (you can think of this in terms of hole doping=*decreasing* the number of electrons, and E_F depends on number of electrons)
- With electron doping, E_F moves towards the conduction band (electron doping=*more* electrons=*higher* E_F)
- When a p-type and n-type semiconductor are brought into contact, their Fermi levels align. This puts the conduction/valence bands at different energies on the p and n-doped side. Electrons on the n-doped side now have to overcome an energy barrier to get to the p-doped side. This is precisely from the voltage set up in the depletion region
- Reverse bias increases the energy barrier
- Forward bias *decreases* the barrier. Electrons and holes can now diffuse towards each other and flow in a circuit or recombine

- When electrons and holes recombine they can emit light or heat. Many of the applications of p-n junctions take advantage of the fact that semiconductors can absorb or emit light (note: although silicon is the favored semiconductor material for many types of devices, it is not ideal for applications involving light because it has an *indirect bandgap*. Nevertheless, most solar cells use silicon, although no LEDs do)



2. From Harris Modern Physics

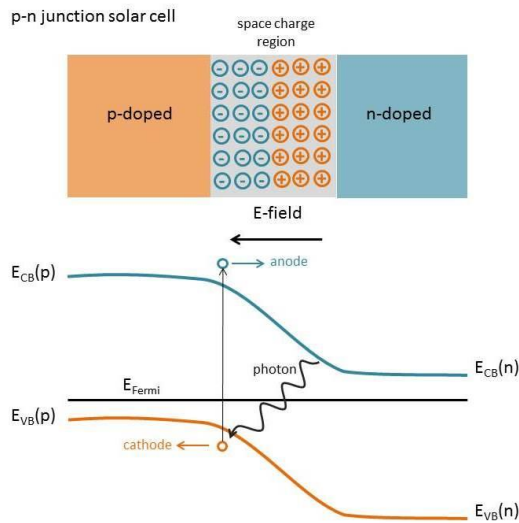


3. Light emission and absorption in a semiconductor across a band gap. From Modern Physics by Serway/Moses/Moyer

Application of p-n junctions: solar cells

The most plentiful source of alternative energy on earth is the sun, and solar cells or photovoltaic cells are essentially semiconductor p-n junctions which can turn this energy source into electricity. Three things happen in a solar cell in this order:

1. Photon hits p-n junction in depletion region creating an electron-hole pair, if its energy is large enough to exceed the band gap
2. Electron and hole are swept in *opposite* directions by electric field in depletion region. This is an important step, because it is important to separate the charges so they don't recombine
3. Electron and hole are swept out of the semiconductor into an external circuit so that electricity can be used for something

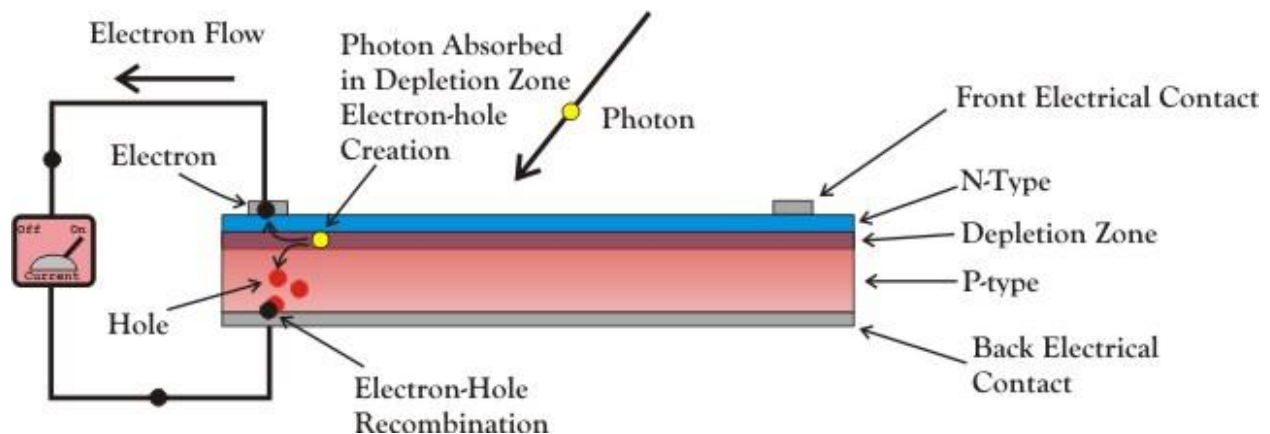


Photodiodes

A photodiode is very similar to a photovoltaic cell in that it converts light to electricity. However, the purpose of a photodiode is to *measure* light, not use it to create useable power.

Thus, photodiodes are operated with **reverse bias** across the p-n junction. This has the effect of making the depletion region wider (higher probability of detecting a photon) and making the junction capacitance smaller (faster response time).

4. Image source: <http://www.science-kick.com/solar-cell-type/working-principle-of-junction-solar-cells>



5 Image source: <http://www.imagesco.com/articles/photovoltaic/photovoltaic-pg4.html>

Physics 140B lecture 21

- Reduced, extended, and repeating zone scheme
- Construction of Fermi surfaces

Ch9-Fermi surfaces and metals

A Fermi surface is a surface of constant energy ϵ_F in \mathbf{k} -space. It separates occupied states from unoccupied states. In Ch6, the free-electron model yielded a spherical Fermi surface centered around $\mathbf{k}=(0,0,0)$ with radius k_F . In this chapter, we learn that the Fermi surfaces in real metals are sometimes more complicated.

Reduced, extended, and repeating zone scheme

If an electron has momentum \mathbf{k}' outside the first Brillouin zone, there exists a reciprocal lattice vector \mathbf{G} such that $\mathbf{k}=\mathbf{k}'+\mathbf{G}$ is in the first Brillouin zone. The Bloch function is written as:

$$\psi_{\mathbf{k}'}(\mathbf{r}) = e^{i\mathbf{k}'\cdot\mathbf{r}}u_{\mathbf{k}'}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}(e^{-i\mathbf{G}\cdot\mathbf{r}}u_{\mathbf{k}'}(\mathbf{r})) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r}) = \psi_{\mathbf{k}}(\mathbf{r})$$

Where $u_{\mathbf{k}}(\mathbf{r}) = e^{-i\mathbf{G}\cdot\mathbf{r}}u_{\mathbf{k}'}(\mathbf{r})$. Both $e^{-i\mathbf{G}\cdot\mathbf{r}}$ and $u_{\mathbf{k}'}(\mathbf{r})$ both have the periodicity of the lattice, so $u_{\mathbf{k}}(\mathbf{r})$ does as well.

Similarly, any energy $\epsilon_{\mathbf{k}'}$ with \mathbf{k}' outside the first Brillouin zone, is equal to $\epsilon_{\mathbf{k}}$ inside the first Brillouin zone, where $\mathbf{k}=\mathbf{k}'+\mathbf{G}$

Two wavefunctions at the same \mathbf{k} but different energy are independent of each other, and are labeled with a symbol n , to denote that they refer to different bands

$$\psi_{n,\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_n(\mathbf{k} + \mathbf{G})e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$$

Similar to how we can translate bands outside the first Brillouin zone back into the first Brillouin zone, we can also translate bands from the first Brillouin zone to another Brillouin zone. Any energy $\epsilon_{\mathbf{k}}$ inside the first Brillouin zone is equivalent to $\epsilon_{\mathbf{k}+\mathbf{G}}$ outside the first Brillouin zone.

For example, if the energy band in a simple cubic lattice has the form below (calculated from tight binding model, which will be presented in this chapter):

$$\epsilon_{\mathbf{k}} = -\alpha - 2\gamma(\cos k_x a + \cos k_y a + \cos k_z a)$$

Where α and γ are constants, we can translate \mathbf{k} by a reciprocal lattice vector of

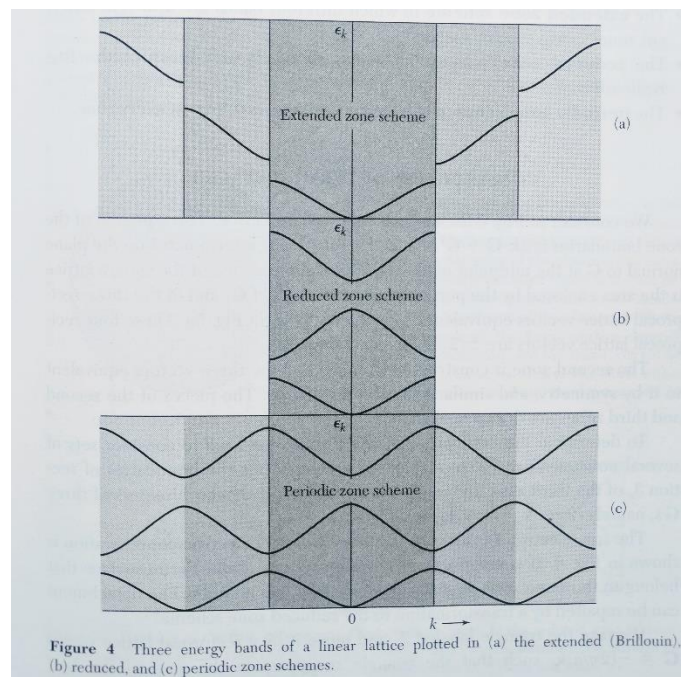


Figure 4 Three energy bands of a linear lattice plotted in (a) the extended (Brillouin), (b) reduced, and (c) periodic zone schemes.

our choice, e.g. $\mathbf{G} = \frac{2\pi}{a} \hat{x}$, and get the same result

$$\cos k_x a \rightarrow \cos \left(k_x + \frac{2\pi}{a} \right) a = \cos(k_x a + 2\pi) = \cos(k_x a)$$

There are three ways of drawing equivalent energy-vs-momentum dispersion relations for electrons:

- Extended zone scheme: different bands in different Brillouin zones
- Reduced zone scheme: all bands drawn in first Brillouin zone
- Periodic zone scheme: energy band is drawn in every zone

Construction of Fermi surfaces

The first step to constructing a Fermi surface is constructing all of the Brillouin zones. As a reminder, a Brillouin zone is a Wigner-Seitz cell in reciprocal space, constructed by picking one reciprocal lattice point to be the origin, drawing lines connecting it to all nearby points, and constructing a cell from the perpendicular bisectors of these lines.

The equation of the zone boundaries is given by: $2\mathbf{k} \cdot \mathbf{G} + G^2 = 0$. The \mathbf{k} that solves this equation for a specific \mathbf{G} corresponds to a plane normal to \mathbf{G} on the midpoint of \mathbf{G}

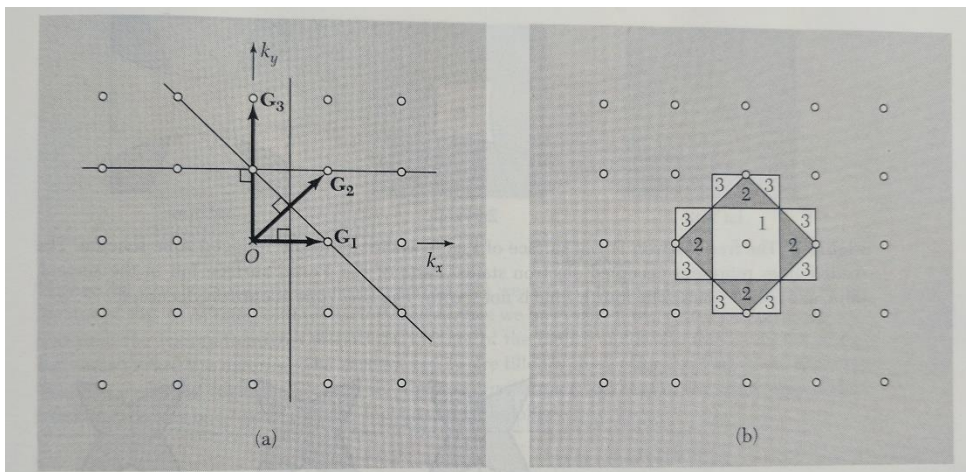
For a 2D square lattice, the first Brillouin zone is constructed from the shortest reciprocal lattice vectors, $\mathbf{G}_1 = \pm \left(\frac{2\pi}{a} \right) \hat{k}_x, \pm \left(\frac{2\pi}{a} \right) \hat{k}_y$

The second Brillouin zone is constructed from the 2nd longest reciprocal lattice vectors, equivalent by symmetry:

$$\mathbf{G}_2 = \left(\frac{2\pi}{a} \right) \hat{k}_x + \left(\frac{2\pi}{a} \right) \hat{k}_y \text{ (and equivalent)}$$

For the third Brillouin zone, non-equivalent reciprocal lattice vectors have to be considered: =

$$\left(\frac{2\pi}{a} \right) \hat{k}_x, \left(\frac{4\pi}{a} \right) \hat{k}_y, \left(\frac{2\pi}{a} \right) (\hat{k}_x + \hat{k}_y)$$



Normally one looks up what Brillouin zones (especially higher ones) look like, but if you are constructing them yourself, here are some double-checks:

- All BZs must have the same total area (check that disjointed regions of

BZ 2 & 3 have the same area as one)

- Disjointed regions of higher BZs must translate into the first BZ via reciprocal lattice vectors without overlap (e.g. translation only, no rotation)

Physics 140B Lecture 22

- **Constructing a Fermi surface in multiple BZs**
- **Quantization of electron orbits in magnetic field**

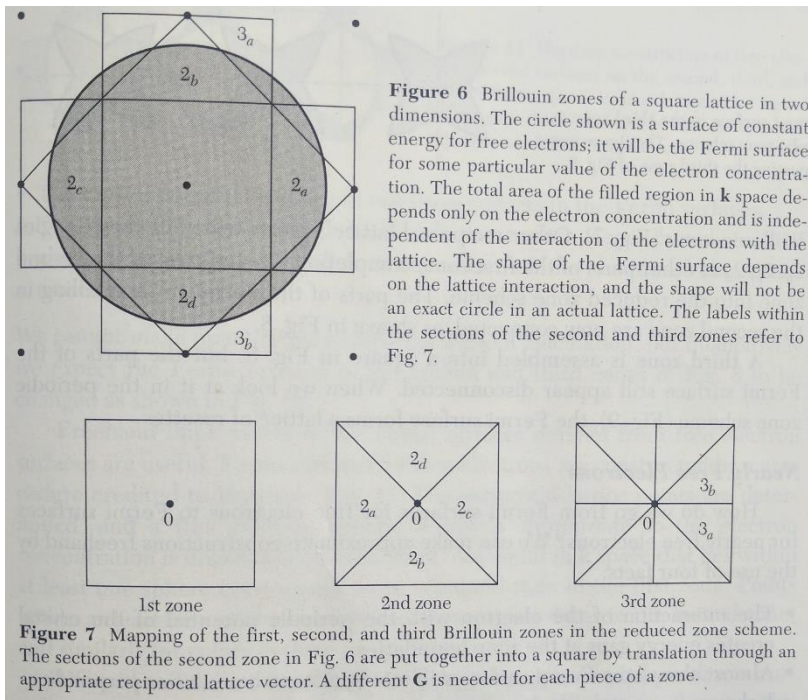
In the previous lecture we constructed the first three Brillouin zones in a square lattice (drawn below). Now add electrons to make a Fermi surface!

Starting off with a free-electron model, the Fermi surface in 2D is a circle centered around $(k_x, k_y) = (0,0)$.

If this circle is small enough to fit entirely within the first Brillouin zone we are done.

If the circle overfills the first Brillouin zone, we do the following:

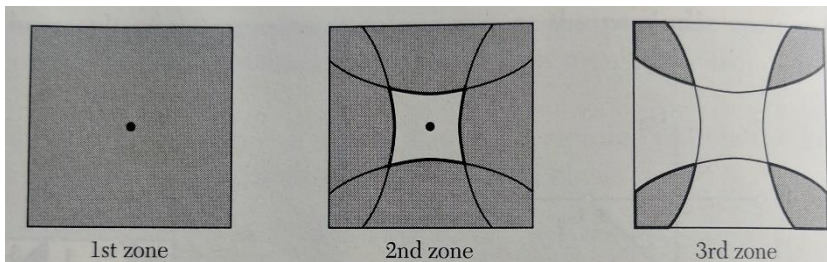
- Superimpose the circle on all the Brillouin zones (extended zone scheme)



- Consider the portion of the Fermi surface that is inside each Brillouin zone, and translate this back into the first Brillouin zone.

For the example in Kittel, the free electron Fermi surface entirely fills the first Brillouin zone, partially fills the 2nd, and has a small incursion into the 3rd and 4th. Only the first 3 Brillouin zones are shown.

Although the filled areas of the 3rd Brillouin zone look disconnected, they can be shown to form connected propeller shapes in the repeating zone scheme.

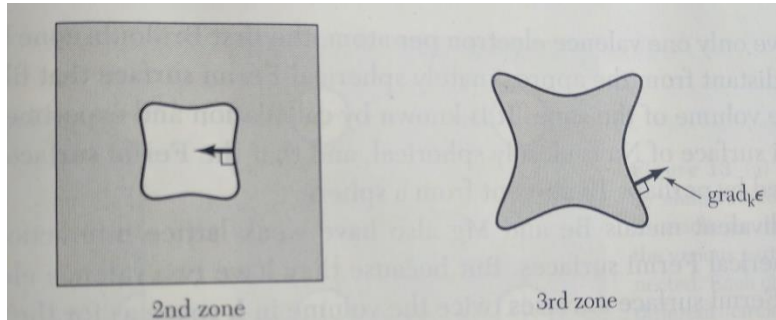


When this procedure is extended to the *nearly free electron model*, where the lattice potential is not ignored, the following considerations are used

- Interactions of electrons with periodic ionic potential opens gaps at Brillouin zone boundaries
- Fermi surface will almost always intersect Brillouin zone boundary perpendicular
- Crystal potential will round sharp corners of fermi surface

- The total volume enclosed by Fermi surface depends only on electron concentrations, and will be the same for the free electron case and when the ionic potential is turned on (nearly free electron)

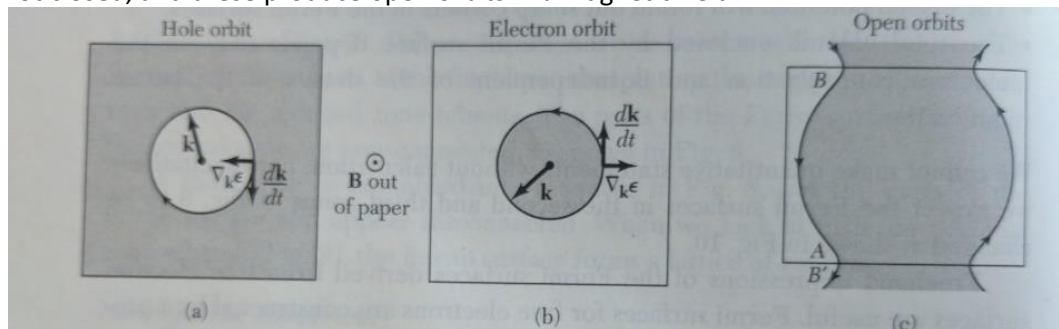
Qualitatively, the Fermi surfaces in the 2nd and 3rd Brillouin zone, for the same example as above, will change slightly to the following shapes:



Notice that electrons almost fill the second zone, except for a small empty region in the center. This is considered to be a 'hole-like' Fermi surface, because the enclosed surface constitutes the absence of electrons. In the third zone, electrons fill a minority of the area. This is considered to be an electron-like

Fermi-surface

This will have implications for how electrons move in a magnetic field, which electrons circulating in opposite directions for electron-like and hole-like band. It is also possible to have a Fermi surface that is not closed, and these produce open orbits in a magnetic field.



Quantization of electron orbits in magnetic field

A metal's Fermi surface is a fundamental property that determines much of its behavior, so measuring Fermi surfaces is very important. There are various ways to measure either the entire size and shape of the Fermi surface(s), or some aspect thereof. This lecture, and your textbook focuses on methods that take advantage of **quantization of electron orbits** in a magnetic field (we will discuss other methods of measuring Fermi surfaces in later lectures).

The total momentum of a particle in a magnetic field is given by:

$$\mathbf{p} = \mathbf{p}_{kin} + \mathbf{p}_{field} = \hbar\mathbf{k} + q\mathbf{A}/c$$

Where q is the charge and \mathbf{A} is the vector potential defined as $\mathbf{B} = \nabla \times \mathbf{A}$

We assume that orbits in a magnetic field are quantized by the Bohr-Sommerfeld relation (the assumption that the only quantum orbits that survive are those for which a half integer or integer number of wavelengths fit around a circular orbit; this is the basis, for example, of the Bohr model of the hydrogen atom):

$$\oint \mathbf{p} \cdot d\mathbf{r} = (n + \gamma)2\pi\hbar$$

Where n is an integer and γ is a phase factor that has a value of $\frac{1}{2}$ for free electrons. Plugging in the earlier expression for the momentum of a quantum particle:

$$\oint \mathbf{p} \cdot d\mathbf{r} = \oint \hbar\mathbf{k} \cdot d\mathbf{r} + \frac{q}{c} \oint \mathbf{A} \cdot d\mathbf{r}$$

From earlier chapters, the equation of motion for a particle of charge q in a magnetic field is given by:

$$\hbar \frac{d\mathbf{k}}{dt} = \frac{q}{c} \frac{d\mathbf{r}}{dt} \times \mathbf{B}$$

Integrating this expression with respect to time gives:

$$\hbar\mathbf{k} = \frac{q}{c} \mathbf{r} \times \mathbf{B}$$

Physics 140B lecture 23

- Quantization of magnetic flux
- De Haas van Alphen effect

In the previous lecture we left off saying that electron orbits in a magnetic field are quantized and the momentum of an electron in a magnetic field has a kinetic component and a field component

$$\oint \mathbf{p} \cdot d\mathbf{r} = \oint \hbar \mathbf{k} \cdot d\mathbf{r} + \frac{q}{c} \oint \mathbf{A} \cdot d\mathbf{r} = (n + \gamma) 2\pi \hbar$$

- $n = \text{integer}$
- $\gamma = \text{phase factor} = 1/2$ for electrons
- $\mathbf{B} = \nabla \times \mathbf{A}$

From earlier chapters, the equation of motion for a particle of charge q in a magnetic field is given by:

$$\hbar \frac{d\mathbf{k}}{dt} = \frac{q}{c} \frac{d\mathbf{r}}{dt} \times \mathbf{B}$$

Integrating this expression with respect to time gives:

$$\hbar \mathbf{k} = \frac{q}{c} \mathbf{r} \times \mathbf{B}$$

We can plug this expression into one of the earlier path integrals:

$$\oint \hbar \mathbf{k} \cdot d\mathbf{r} = \frac{q}{c} \oint \mathbf{r} \times \mathbf{B} \cdot d\mathbf{r} = -\frac{q}{c} \mathbf{B} \cdot \oint \mathbf{r} \times d\mathbf{r} = -\frac{2q}{c} \Phi$$

Where Φ is the magnetic flux (magnetic field \times area) contained in the orbit in real space. The transition from the 2nd step to the 3rd takes advantage of the following vector algebra relation: $\mathbf{r} \times \mathbf{B} \cdot d\mathbf{r} = \mathbf{B} \cdot (d\mathbf{r} \times \mathbf{r}) = -\mathbf{B} \cdot (\mathbf{r} \times d\mathbf{r})$. Going from the 3rd step to the 4th takes advantage of the fact that $\oint \mathbf{r} \times d\mathbf{r} = 2 \times$ area enclosed by loop, which can be seen graphically by considering a circular contour in which \mathbf{r} is pointing radially and $d\mathbf{r}$ is pointing tangentially.

The other path integral can be written as:

$$\frac{q}{c} \oint \mathbf{A} \cdot d\mathbf{r} = \frac{q}{c} \int \nabla \times \mathbf{A} \cdot d\boldsymbol{\sigma} = \frac{q}{c} \int \mathbf{B} \cdot d\boldsymbol{\sigma} = \frac{q}{c} \Phi$$

Where $d\boldsymbol{\sigma}$ is the area element in real space. Going from the first to the second step uses the divergence theorem.

$$\text{Thus } \oint \mathbf{p} \cdot d\mathbf{r} = -\frac{2q}{c} \Phi + \frac{q}{c} \Phi = -\frac{q}{c} \Phi = (n + \gamma) 2\pi \hbar$$

The quantization condition for electrons in a magnetic field is thus given by:

$$\Phi_n = (n + \gamma) \left(\frac{2\pi \hbar c}{e} \right)$$

The quantity in parenthesis is called the flux quantum, and has a value of $\frac{2\pi \hbar c}{e} = 4.14 \times 10^{-7} \text{ gauss cm}^2$. It is conventionally also called Φ_0 . What this formula tells us is that electron orbits in

a magnetic field sweep out quantized areas, or equivalently, circular orbits have only certain allowed radii for a given B.

The concept of flux relies on a specified area in real space. We can connect this to an area in momentum space in the following way:

A line element Δr in the plane normal to \mathbf{B} is related to a length element in k-space, Δk via $\Delta r = \left(\frac{\hbar c}{eB}\right) \Delta k$ (this comes from the relation $\hbar \mathbf{k} = \frac{q}{c} \mathbf{r} \times \mathbf{B}$).

The area S_n in k-space is thus related to the area A_n in real space via:

$$A_n = \left(\frac{\hbar c}{eB}\right)^2 S_n$$

The magnetic flux can be written as:

$$\Phi = BA_n = \left(\frac{\hbar c}{e}\right)^2 \frac{1}{B} S_n = (n + \gamma) \frac{2\pi\hbar c}{e}$$

Solve for S_n

$$S_n = (n + \gamma) \frac{2\pi e}{\hbar c} B$$

There can be orbits for different B with the same area in k-space, but different values of n:

$$S \left(\frac{1}{B_{n+1}} - \frac{1}{B_n} \right) = \frac{2\pi e}{\hbar c}$$

The consequence of this result is that **the population of orbits near the Fermi surface oscillates as 1/B, and the period of this oscillation in measured quantities can be used to construct the Fermi surface.**

De Haas van Alphen effect

The de Haas van Alphen (dHvA) effect is the **oscillation** of the **magnetic moment** of a metal as a function of magnetic field intensity. This is easiest to observe

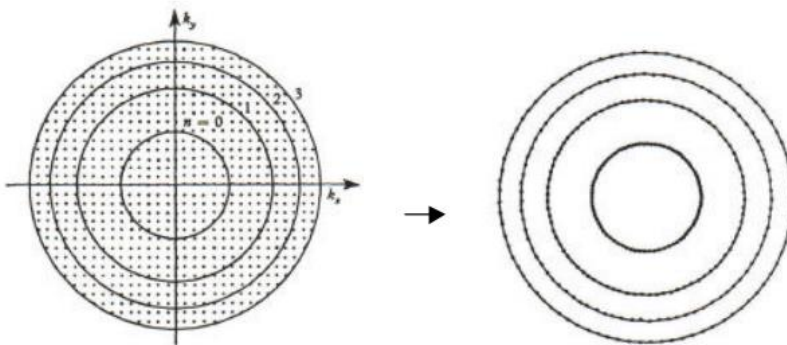
- In clean specimens, so that electrons do not scatter off impurities before completing an orbit
- At low temperature, so that population oscillations in a magnetic field are not averaged out by thermal population of adjacent orbitals

dHvA measures the magnetic moment (more specifically, the susceptibility of the electron gas to be magnetized/polarized by the applied field), which oscillates in inverse magnetic field, but oscillatory effects are also seen in resistivity (Shubnikov de Haas effect) and the Hall voltage.

Consider a 2D electron gas in the x-y plane where the magnetic field is applied along z. The following things happen:

- When magnetic field is turned on, electrons can only 'live' at quantized energy levels set by the magnetic field. Instead of density of states being a constant function of energy, as is normally the case for a 2D electron gas, the density of states is a series of delta functions

- The allowed energy levels are separated by energy $\hbar\omega_c = \frac{e\hbar}{m^*c} B$; the energies where electrons can 'live' changes as a function of magnetic field



For the 2D electron gas for which the area between successive orbits in k-space is given by:

$$\Delta S = S_n - S_{n-1} = 2\pi eB/\hbar c$$

The area in k-space (neglecting spin) occupied by a single orbital is $\left(\frac{2\pi}{L}\right)^2$, from chapter 6, where

L is the length of the side of the specimen. The number of orbitals (i.e. states) that coalesce in a single magnetic level is thus

$$D = \left(\frac{2\pi eB}{\hbar c}\right) \left(\frac{L}{2\pi}\right)^2 = \rho B$$

Where $\rho \equiv eL^2/2\pi\hbar c$. These magnetic levels are called **Landau Levels**.

For a system of N electrons at absolute zero, Landau levels are filled up to a magnetic quantum number s (a positive integer). If any electrons are left over, they will go into level $s+1$, which will be partially filled. **The fermi level will lie in the $s+1$ Landau level if there are electrons at this level.** As magnetic field increases, electrons move to lower levels. When $s+1$ is vacated, the Fermi level moves down abruptly to lower level s . This transfer of electrons to a lower Landau level can occur because the degeneracy D increases as B increases. As B increases, there are critical values of B where the quantum number of the highest level decreases

It turns out that the number of electrons in the highest occupied level (where E_F is located) is an oscillatory function of $1/B$. This can be seen intuitively because as B increases the highest landau level empties out slower because there are fewer lower landau levels.

To show the periodicity of the *energy* as a function of B , we use quantization of energy levels in terms of cyclotron orbits:

$$E_n = \left(n - \frac{1}{2}\right) \hbar\omega_c$$

The total energy of electrons in fully occupied levels is given by:

$$\sum_{n=1}^s D\hbar\omega_c \left(n - \frac{1}{2}\right) = \frac{1}{2} D\hbar\omega_c s^2$$

(sum of finite arithmetic series)

The total energy of electrons in the partially occupied $s+1$ level is:

$$\hbar\omega_c(s + \frac{1}{2})(N - sD)$$

where sD is the number of electrons in levels 1 through s .

The total energy of all electrons is the sum of the two, and it is plotted in the figure above

$$E_N = \frac{1}{2}D\hbar\omega_c s^2 + \hbar\omega_c(s + \frac{1}{2})(N - sD)$$

The magnetic moment μ of a system at absolute zero is given by:

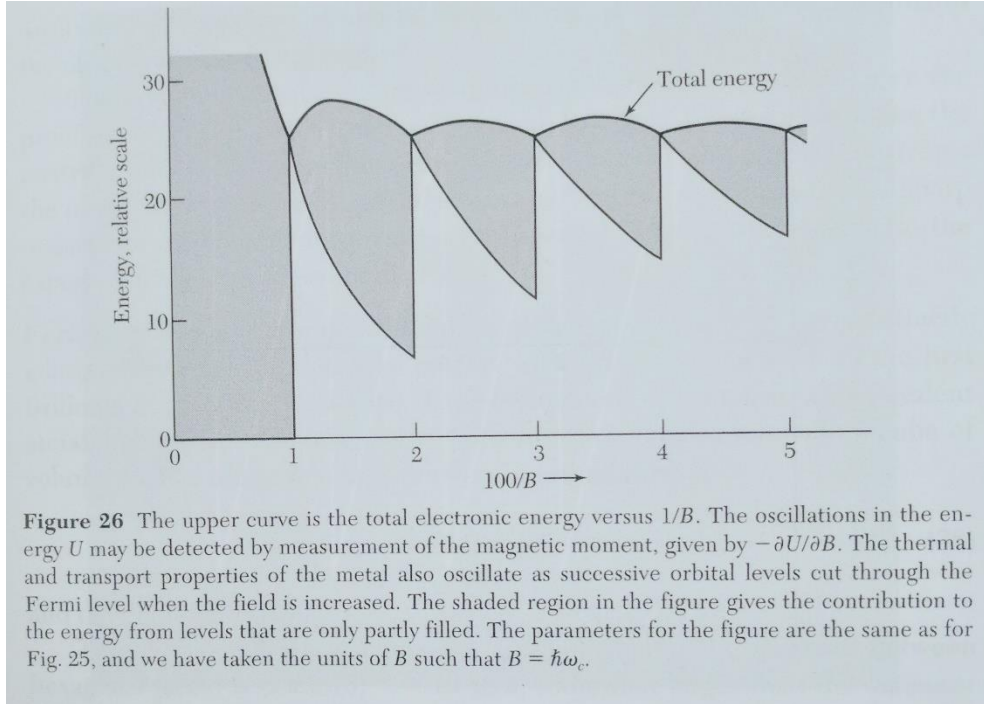


Figure 26 The upper curve is the total electronic energy versus $1/B$. The oscillations in the energy U may be detected by measurement of the magnetic moment, given by $-\partial U/\partial B$. The thermal and transport properties of the metal also oscillate as successive orbital levels cut through the Fermi level when the field is increased. The shaded region in the figure gives the contribution to the energy from levels that are only partly filled. The parameters for the figure are the same as for Fig. 25, and we have taken the units of B such that $B = \hbar\omega_c$.

$$\mu = -\frac{\partial U}{\partial B}$$

Because the total energy is an oscillatory function of $1/B$, its derivative is as well. This oscillatory magnetic moment of a metal at low temperature is the dHvA effect. The oscillations are periodic in $1/B$ such that:

$$\Delta\left(\frac{1}{B}\right) = \frac{2\pi e}{\hbar c S}$$

Experimentally, one would measure magnetic moment as a function of external magnetic field at very low temperature, plot the data as magnetic moment vs $1/B$, take a Fourier transform of the graph to extract the frequency or frequencies in $1/B$, and use the relation above to extract S , the orbit in k -space. Then, we use $\pi k_F^2 = S$ to extract k_F and the size of the Fermi surface (assuming it is circular).

Physics 140B Lecture 24

- Finish de Haas Van Alphen effect
- 3D metals
- Other method of measuring Fermi surfaces and band dispersions: angle-resolved photoemission spectroscopy

Electron orbits are quantized in a magnetic field, such that most measured quantities that relate to density of states at E_F (e.g. resistivity, hall number, magnetic moment) have an oscillatory dependence on $1/B$:

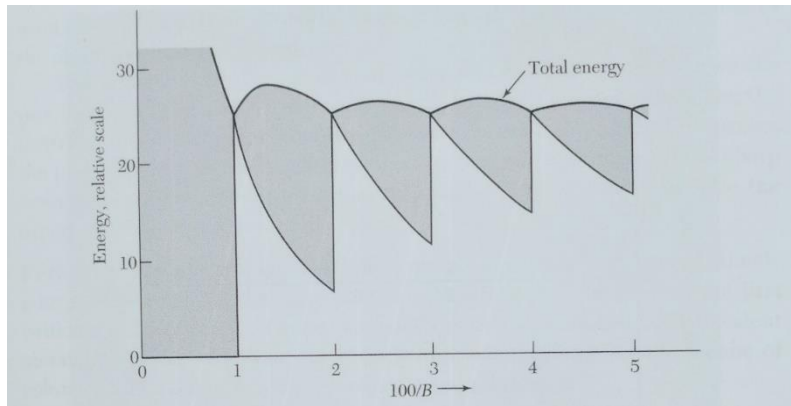
$$\Delta\left(\frac{1}{B}\right) = \frac{2\pi e}{\hbar c S}$$

Where $\Delta\left(\frac{1}{B}\right)$ is the period when measured quantity is plotted as a function of $1/B$ and S is the area of the (2D) Fermi surface.

De Haas Van Alphen effect measures magnetic moment (μ) as a function of magnetic field, which is derived via the following formula:

$$\mu = -\frac{\partial U}{\partial B}$$

In the previous lecture, we saw that for a magnetic field B , the allowed energies and allowed orbits in k -space are given by:



- $S_n = \left(n - \frac{1}{2}\right) \frac{2\pi e}{\hbar c} B$
- $E_n = \left(n - \frac{1}{2}\right) \hbar \omega_c = \left(n - \frac{1}{2}\right) \frac{e\hbar}{m^*c} B$

Where n is an integer starting at 1. These are called **Landau levels**.

- Landau levels each have allowed occupancy

$$D = \left(\frac{2\pi e B}{\hbar c}\right) \left(\frac{L}{2\pi}\right)^2 = \rho B$$

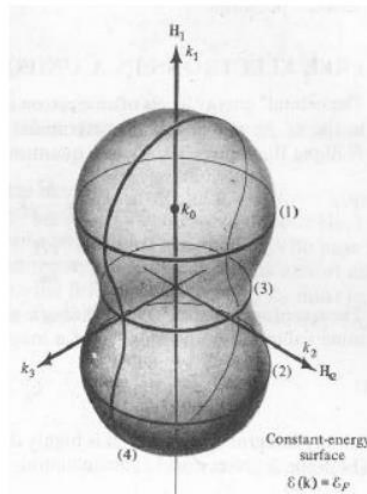
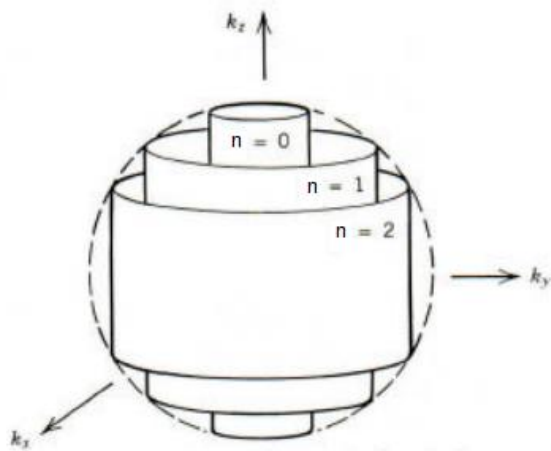
- The Landau levels up to $n=s$ are completely filled and the $s+1$ level is partially filled.
- As magnetic field increases, Landau levels $1\dots s$ can hold more electrons, so the $s+1$ level empties out. Eventually it empties out and s is the highest filled energy level.
- E_F is always at the highest energy level that contains electrons ($s+1$ until it empties out)

The total energy of the electrons in a magnetic field is given by the equation below where the first term comes from the totally filled levels (up to s) and the 2nd term comes from the partially filled ($s+1$) level:

$$E_N = \frac{1}{2} D \hbar \omega_c s^2 + \hbar \omega_c \left(s + \frac{1}{2}\right) (N - sD) = U$$

This is an oscillatory function of inverse magnetic field, because the 2nd term is, and it is plotted below. The magnetic moment in the dHvA effect is a derivative of this with respect to magnetic field, so it is also oscillatory in 1/B.

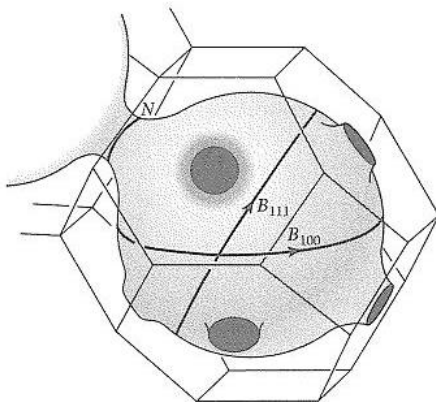
Generalizing to three dimensions



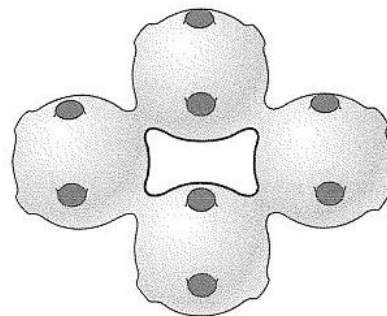
In 3D, everything we discussed still applies, except electron orbits are quantized to cylinders ('Landau tubes') instead of circles. The S that is detected corresponds to **extremum** perpendicular left, if magnetic

orbits (largest and smallest) in the plane to magnetic field. For the Fermi surface on the field is along k₁, the observed orbits are (1), (2), and (3), and if it is along k₂, the orbit observed is (4). The magnetic field can be placed in arbitrary directions to map out the full 3D shape of the Fermi surface.

In noble metals (e.g. copper, gold), the Fermi surface has an almost spherical shape, but with sections 'sucked in' towards Brillouin zone boundaries. Orbits around the fattest part of the Fermi surface are called belly orbits, and orbits around the smallest parts are called neck orbits. In the repeated zone scheme, there is also a dog's bone hole-like orbit.



Cu / Au



Example: Fermi surface of gold

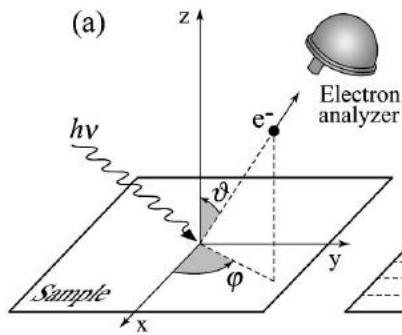
One orbit of gold corresponds to a period of $2 \times 10^{-9} \text{ gauss}^{-1}$. We can use this to calculate the area of the extremal orbit:

$$S = \frac{2\pi e/\hbar c}{\Delta(\frac{1}{B})} = \frac{9.55 \times 10^7}{2 \times 10^{-9}} = 4.8 \times 10^{16} \text{ cm}^{-2}$$

This can be converted to a value of k_F (assuming orbit is circular), giving $k_F = 1.24 \times 10^8 \text{ cm}^{-1}$

This is very close to the 'free electron' value in Ch6 (1.2×10^8).

Another way of measuring Fermi surfaces: Angle-resolved photoemission spectroscopy (ARPES)



ARPES is an experiment based on the photoelectric effect (UV light in, electrons out), which can measure Fermi surface and electron band dispersions directly.

Electrons that are photoemitted from the crystalline solid carry information about how they were moving inside the solid, which is expressed in the following conservation laws:

$$E_{kin} = h\nu - \phi - |E_B|$$

$$p_{\parallel} = \hbar k_{\parallel} = \sqrt{2mE_{kin}} \cdot \sin \vartheta$$

With the following definitions:

E_{kin} =kinetic energy of photoemitted electrons (measured in ARPES experiment)

$h\nu$ =photon energy of light used to photoemit electrons (chosen/controlled in experiment)

ϕ =work function (minimum energy to produce photoemission; known quantity)

E_B =binding energy of electrons when they were inside the solid (defined such that $E_B = 0$ at E_F ; quantity that one wants to know from experiment)

$p_{\parallel} = \hbar k_{\parallel}$ =electrons' momentum inside the sample, parallel to plane (quantity one wants to know from experiment)

m =free electron mass (known quantity; free electron mass is used instead of effective mass because this part of the equation considers the electron after it has left the solid and is now a free electron)

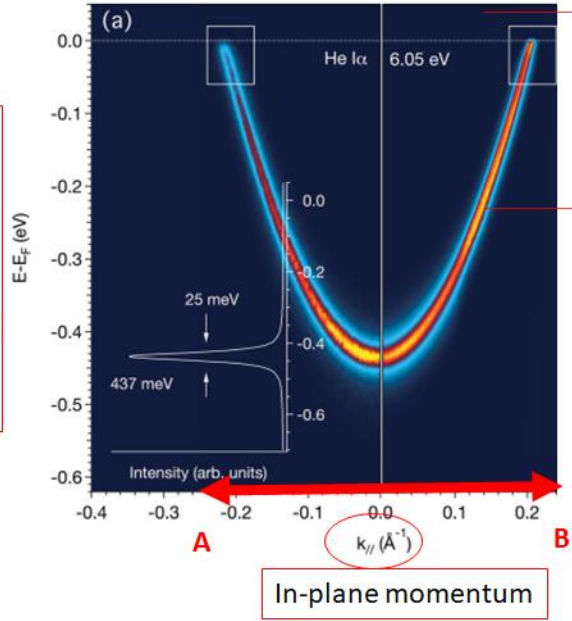
ϑ =photoemission angle relative to sample normal (measured in ARPES experiment)

Summarizing the information above, the measured quantities are electrons' kinetic energy and emission angle (E_{kin} vs ϑ), which can be connected via known/controlled quantities to the binding energy and momentum they had back when they were inside the crystal (E_B vs k_{\parallel}). The latter is precisely the same as ε vs k that we calculated in previous chapters. Considering only $E_B = 0$ (at the Fermi level) allows ARPES to measure Fermi surfaces, but measuring which values of \mathbf{k} yield photoelectrons.

Band structure: simple metal (Cu 111 surface)

orange/red=high intensity
(many

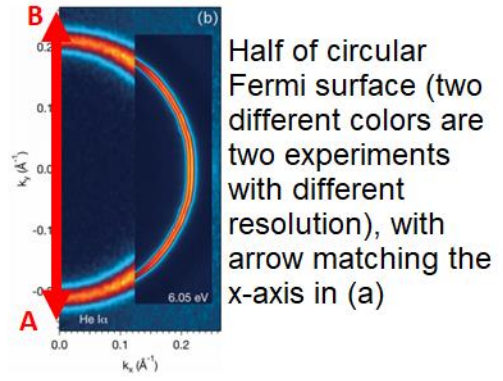
Electron binding energy
 $|E_B| = E - E_F$



Fermi-Dirac cutoff

$$F(E) = \frac{1}{e^{(E-E_F)/k_B T} + 1}$$

$$\epsilon_k = E(k) = \frac{\hbar^2 k^2}{2m^*}$$



PRB **87**, 075113 (2013)

Physics 140B Lecture 25

- One way of calculating energy bands: tight binding model
- Begin superconductivity

Tight Binding model

The tight binding model is based on combining wavefunctions of individual atomic orbitals.

Suppose an isolated atom has potential $U(\mathbf{r})$ and is in an s-state (spherically symmetric), represented by wavefunction $\phi(\mathbf{r})$. Now suppose that there is a crystal of these atoms, and the presence of other atoms doesn't much affect the single-atom wavefunction. The wavefunction of an electron in this whole crystal can be expressed as:

$$\psi_{\mathbf{k}}(\mathbf{r}) = \sum_j C_{\mathbf{k},j} \phi(\mathbf{r} - \mathbf{r}_j)$$

Where $C_{\mathbf{k},j} = N^{-1/2} e^{i\mathbf{k}\cdot\mathbf{r}_j}$, ensuring that the wavefunctions comply with Bloch's theorem. The sum is over all lattice points and N is the total number of atoms. This gives

$$\psi_{\mathbf{k}}(\mathbf{r}) = N^{-1/2} \sum_j e^{i\mathbf{k}\cdot\mathbf{r}_j} \phi(\mathbf{r} - \mathbf{r}_j)$$

The first order energy is found by calculating the diagonal matrix elements of the Hamiltonian

$$\langle \mathbf{k} | H | \mathbf{k} \rangle = N^{-1} \sum_j \sum_m e^{i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_m)} \langle \phi_m | H | \phi_j \rangle$$

Where $\phi_m \equiv \phi(\mathbf{r} - \mathbf{r}_m)$

Define a new variable $\boldsymbol{\rho}_m = \mathbf{r}_m - \mathbf{r}_j$

$$\langle \mathbf{k} | H | \mathbf{k} \rangle = \sum_m e^{i\mathbf{k}\cdot\boldsymbol{\rho}_m} \int dV \phi^*(\mathbf{r} - \boldsymbol{\rho}_m) H \phi(\mathbf{r})$$

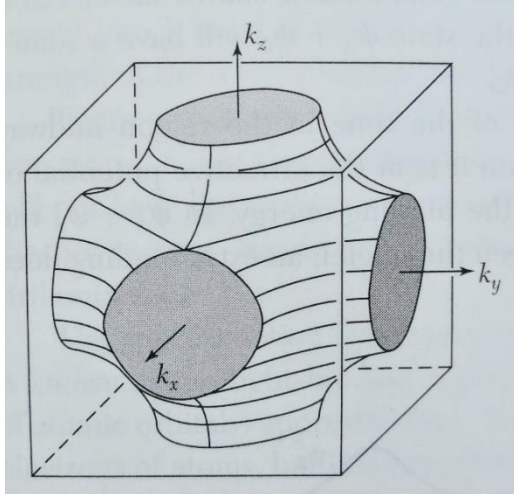
Now, we neglect all of the integrals above **except** for those on the same atom and those between nearest neighbors (separated by $\boldsymbol{\rho}$). This is the tight binding part of the tight binding model: only considering orbital overlap with adjacent atoms assumes that electrons do not make excursion far from their original atom and are hence, tightly bound. Note that it is perfectly acceptable, and sometimes necessary, to consider second nearest neighbors (2nd most closest atom) or even third and fourth, in the tight binding model. However, the solved examples in the book only involve the nearest neighbors.

$$\int dV \phi^*(\mathbf{r}) H \phi(\mathbf{r}) = -\alpha$$

$$\int dV \phi^*(\mathbf{r} - \boldsymbol{\rho}) H \phi(\mathbf{r}) = -\gamma$$

γ can be determined by assuming some specific form of ϕ . For example, for two hydrogen atoms in 1s states, $\gamma = 2 \left(1 + \frac{\rho}{a_0}\right) e^{-\rho/a_0}$ where a_0 is the Bohr radius

Thus:



$$\langle \mathbf{k} | H | \mathbf{k} \rangle = -\alpha - \gamma \sum_m e^{-i\mathbf{k} \cdot \boldsymbol{\rho}_m} = \epsilon_{\mathbf{k}}$$

To proceed further, we need information about the crystal structure. For a simple cubic structure, $\boldsymbol{\rho}_m = (\pm a, 0, 0); (0, \pm a, 0); (0, 0, \pm a)$

$$\text{Thus, } \epsilon_{\mathbf{k}} = -\alpha - 2\gamma(\cos k_x a + \cos k_y a + \cos k_z a)$$

A constant energy surface is shown on the left.

For a BCC crystal structure with 8 nearest neighbors, the dispersion is given by:

$$\epsilon_{\mathbf{k}} = -\alpha - 8\gamma \cos \frac{1}{2} k_x a \cos \frac{1}{2} k_y a \cos \frac{1}{2} k_z a$$

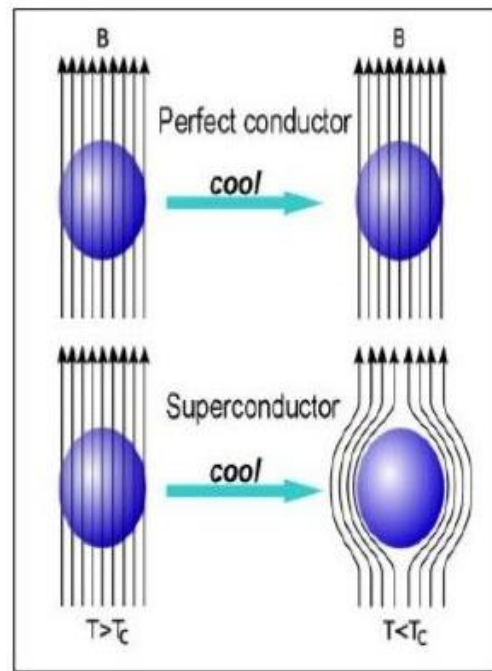
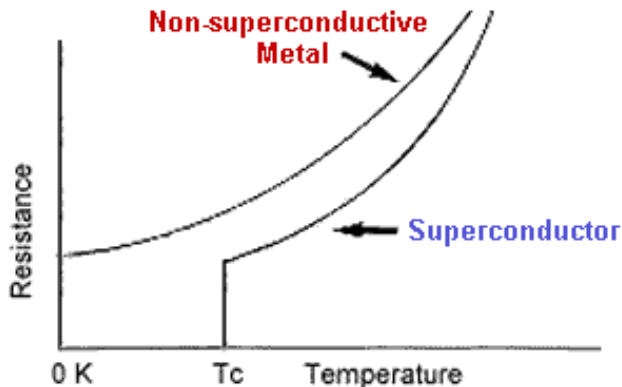
For a FCC structure with 12 nearest neighbors, the dispersion is given by:

$$\epsilon_{\mathbf{k}} = -\alpha - 4\gamma(\cos \frac{1}{2} k_y a \cos \frac{1}{2} k_z a + \cos \frac{1}{2} k_z a \cos \frac{1}{2} k_x a + \cos \frac{1}{2} k_x a \cos \frac{1}{2} k_y a)$$

Superconductivity: experimental survey

Superconductors are characterized by two experimental observables: zero resistivity that onsets suddenly at a transition temperature (T_c) and an expulsion of magnetic field (Meissner effect). T_c 's as high as 260K have been reported and confirmed (under very high pressure), but superconductivity is generally considered a low temperature phenomenon, with 'typical' T_c 's on the order of 1-10K.

Superconductivity is a macroscopic quantum state that forms the ground state of some but not all materials. Many elements on the periodic table become superconductors at low temperature, some only under high pressure.



Notably a superconductor is different from a hypothetical perfect conductor.

- A metal completely free of defects at 0K (something that doesn't exist) will be a perfect conductor because there is nothing to scatter electrons, but it will reach the zero resistance state **gradually**, not suddenly as in a superconductor
- A perfect conductor will want to maintain its present magnetization in the presence of a magnetic field, whereas a superconductor will expel magnetic field always
 - If a magnetic field is turned on, both a superconductor and a perfect conductor will set up currents to expel it
 - If a magnetic field is already turned on and a superconductor is cooled below its transition temperature, it will suddenly expel its magnetic field, while a perfect conductor

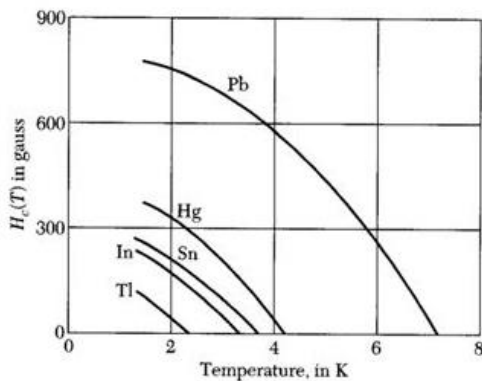
There are some superconductors that scientists 'understand' (microscopic mechanism of superconductivity is explained) and others that are still being researched for the purpose of uncovering the mechanism of their behavior. Many of the topics in this chapter apply to some superconductors, but a few apply only to superconductors for which there is a microscopic mechanism.

Current applications of superconductors include

- MRI machines (to produce a large magnetic field in a solenoid with no heating)
- Detectors for astrophysics (superconductor kept exactly at transition temperature, and any impinging particle will heat it up slightly and produce finite resistivity)
- Definition of the volt (using Josephson junctions, a superconducting device)
- Sensitive detectors of small magnetic fields (also using Josephson junctions)
- The most mature quantum computing technology so far (e.g. d-wave)

Properties shared by all superconductors

Destruction of superconductivity by magnetic field



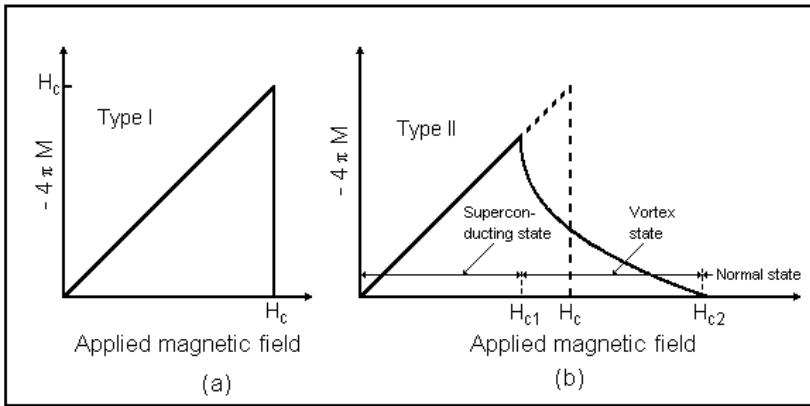
A sufficiently large magnetic field will destroy superconductivity, and for now this critical magnetic field is called H_c . H_c is a function of temperature, and is highest at zero temperature (when superconductivity is strongest) and zero at T_c . Generally, higher T_c corresponds to higher $H_c(T = 0)$, but not always.

Meissner effect

The Meissner effect is the property that a superconductor expels magnetic field in the superconducting state

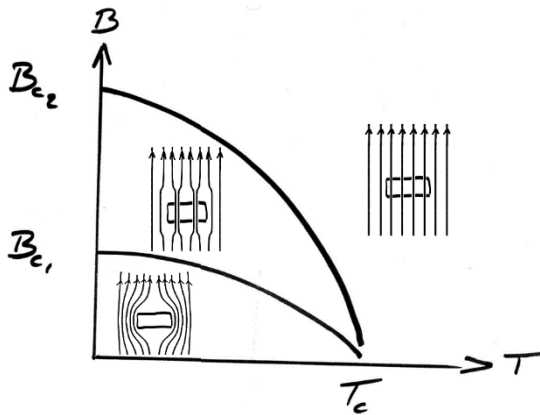
$$B = B_a + 4\pi M = 0$$

Where B is the magnetic field inside the superconductor, B_a is the applied magnetic field (sometimes called H), and M is the magnetization of the specimen. Note that this relation is only exact in specimens that have a specific shape so that the demagnetizing field is irrelevant.



For some superconductors, the

relationship above is obeyed for all magnetic fields, up until the magnetic field is so large that it kills superconductivity (H_c , more on that later). These are called **type-I superconductors**. For other types of superconductors (**type-II superconductors**), the relationship above is obeyed up until an intermediate magnetic field, H_{c1} , at which point the superconductor expels some but not all of its magnetic field.



Superconductivity is not completely destroyed until a higher magnetic field, H_{c2} . For a type-II superconductor in the vortex state ($B_a > H_{c1}$), magnetic field is allowed to enter via small filaments called 'vortices'. Each filament

contains one flux quantum of magnetic flux ($\frac{h}{2e}$ in SI units or $\frac{hc}{2e}$ in CGS). It is called a vortex because the magnetic core is surrounded by a circulating supercurrent (a superconducting current).

Physics 140B lecture 26

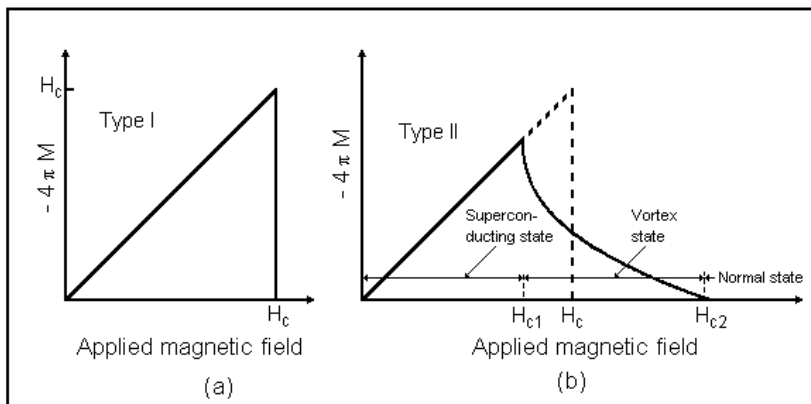
- Experimental survey of superconductor properties
- Magnetic fields and currents inside superconductor

Note: there are superconductors which have a microscopic explanation (e.g. elements on the periodic table that become superconducting, NbN used in many superconducting electromagnets) and superconductors that do not have a microscopic explanation (usually called 'unconventional superconductors'). Most things in today's lecture apply to both types of superconductors. The microscopic mechanism for 'conventional' superconductors is called Bardeen-Cooper-Schrieffer (BCS) theory.

Meissner effect

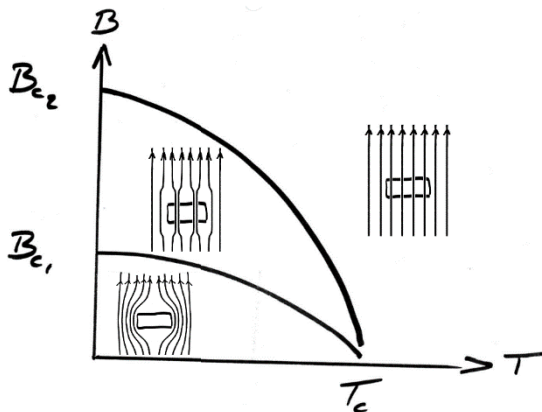
The Meissner effect is the property that a superconductor expels magnetic field in the superconducting state

$$B = B_a + 4\pi M = 0$$



Where B is the magnetic field inside the superconductor, B_a is the applied magnetic field (sometimes called H), and M is the magnetization of the specimen. Note that this relation is only exact in specimens that have a specific shape so that the demagnetizing field is irrelevant.

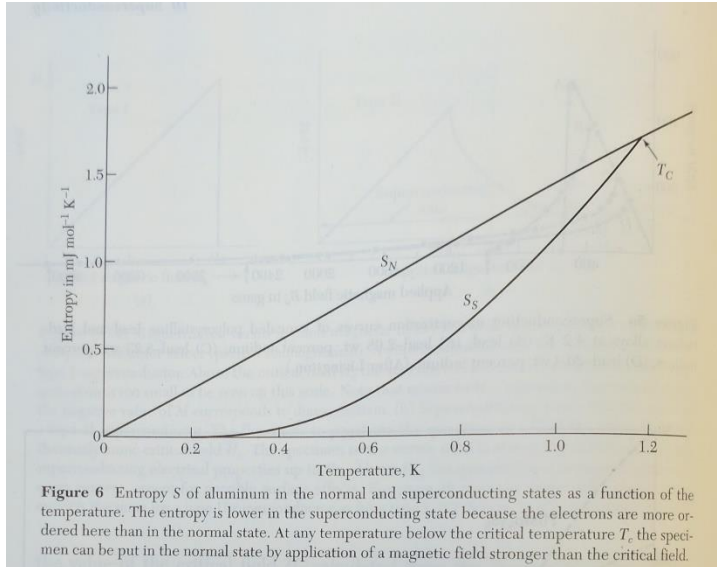
For some superconductors, the



relationship above is obeyed for all magnetic fields, up until the magnetic field is so large that it kills superconductivity (H_c , more on that later). These are called **type-I superconductors**. For other types of superconductors (**type-II superconductors**), the relationship above is obeyed up until an intermediate magnetic field, H_{c1} , at which point the superconductor expels some but not all of its magnetic field. Superconductivity is not completely destroyed until a higher magnetic field, H_{c2} . For a type-II superconductor in the vortex state ($B_a > H_{c1}$), magnetic field is allowed to enter via small filaments called 'vortices'. Each filament contains one flux quantum of magnetic flux ($\frac{h}{2e}$ in SI units or $\frac{hc}{2e}$ in CGS). It is called a vortex because the magnetic core is surrounded by a circulating supercurrent (a superconducting current).

Heat Capacity

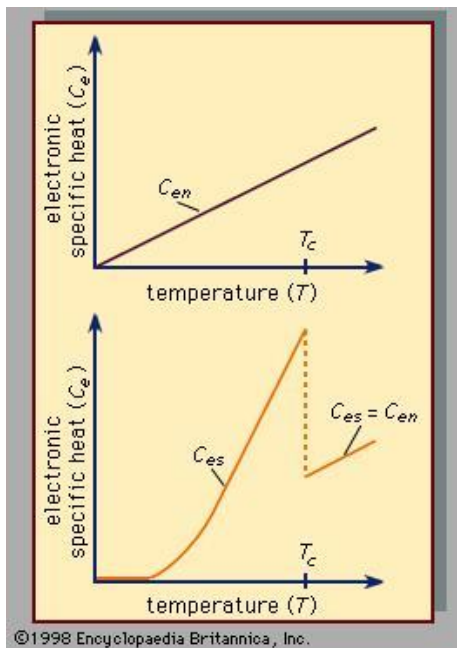
Superconductivity constitutes a thermodynamic phase transition, which means that there is a discontinuity in free energy (U) or some derivative thereof. Evidence of these discontinuities is measured via heat capacity, and heat capacity (at constant volume) can be connected to entropy via the following relations:



$$C_v(T) \equiv \left(\frac{\partial U}{\partial T} \right)_{V,N}$$

$$\begin{aligned} \left(\frac{\partial U}{\partial T} \right)_{V,N} &= \left(\frac{\partial U}{\partial S} \right)_{V,N} \left(\frac{\partial S}{\partial T} \right)_{V,N} \\ &= T \left(\frac{\partial S}{\partial T} \right)_{V,N} \end{aligned}$$

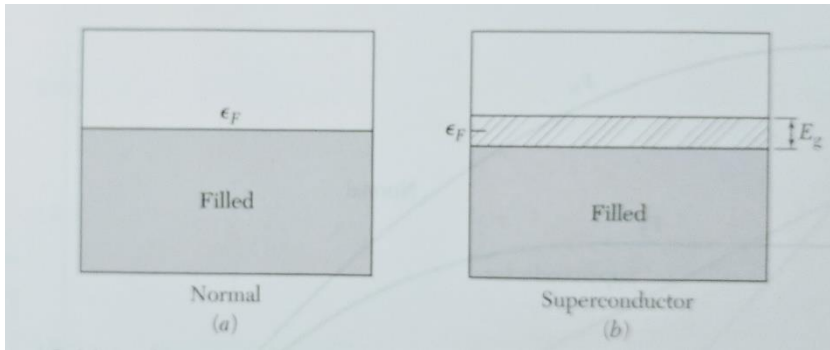
The entropy as a function of temperature is plotted to the left. There is a discontinuity at T_c , which means that there is a step function in its derivative, which is measured by heat capacity. The temperature dependence of heat capacity *below* T_c can reveal information about another property of superconductors, the superconducting energy gap.



Energy gap

Superconductors are characterized by an energy gap which is tied to the Fermi surface. This gap is different than in semiconductors because it arises from superconductivity, not from electron-lattice interactions, and it appears all around the Fermi surface, not just at the Brillouin zone boundaries.

The energy gap in a superconductor is centered around E_F by definition. In some superconductors, the gap is isotropic around the entire Fermi surface or Fermi surfaces, but in others, it may be different in different directions (even zero at some points or lines on the Fermi surface called nodes) or on different Fermi surfaces.



In a superconductor, the charge carrier unit is not a single electron, but a pair of electrons called a Cooper pair, and the energy gap can physically be thought of as the energy required to break a Cooper pair.

The presence of this energy gap has consequences for other physical observables. For

example, if the superconducting gap is non-zero everywhere on the Fermi surface, the superconductor is transparent to photons with energy smaller than the gap. Moreover, the energy gap (both magnitude and presence/absence of nodes) manifests in details of the temperature dependence of heat capacity below T_c .

Isotope effect

In simple metallic superconductors, for which there is a microscopic theory, they are characterized by the fact that the superconducting transition temperature varies depending on the atomic mass (isotope) of the specimen. The variation is as follows:

$$M^\alpha T_c = \text{const}$$

Where M is the average atomic mass, and α is the isotope effect coefficient (a materials dependent quantity) typically ~ 0.5 but a bit smaller.

London equation

For a superconductor with full expulsion of the magnetic field, the magnetic susceptibility is $\chi = \frac{M}{B_a} = -\frac{1}{4\pi}$ (in CGS units) or $\chi = -1$ (in SI). How do we set up electromagnetic fields inside the superconductor to make this happen? The London equation describes this, and it is applicable to all superconductors.

One way to arrive at the London equation is to postulate that in the superconducting state, the current density, \mathbf{j} , is proportional to the vector potential of the magnetic field, \mathbf{A} ($\mathbf{B} = \nabla \times \mathbf{A}$):

In SI units, this proportionality is written as:

$$\mathbf{j} = -\frac{1}{\mu_0 \lambda_L^2} \mathbf{A}$$

Where μ_0 is the permeability of free space, and λ_L is a constant whose physical significance will become clear in a little bit. The equation above is the London equation, and we can express it in another way by taking the curl of both sides

$$\nabla \times \mathbf{j} = -\frac{1}{\mu_0 \lambda_L^2} \mathbf{B}$$

The London equation is assumed to be written in the London gauge where $\nabla \cdot \mathbf{A} = 0$ and $\mathbf{A}_n = 0$ on any external surface through which no external current is fed. This also implies that $\nabla \cdot \mathbf{j} = 0$ and $\mathbf{j}_n = 0$

Using the Maxwell equation $\nabla \times \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} = \mu_0 \mathbf{j}$ and assuming there is no time varying electric field, we get

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$$

We can take the curl of both sides to obtain

$$\nabla \times \nabla \times \mathbf{B} = -\nabla^2 \mathbf{B} = \mu_0 \nabla \times \mathbf{j}$$

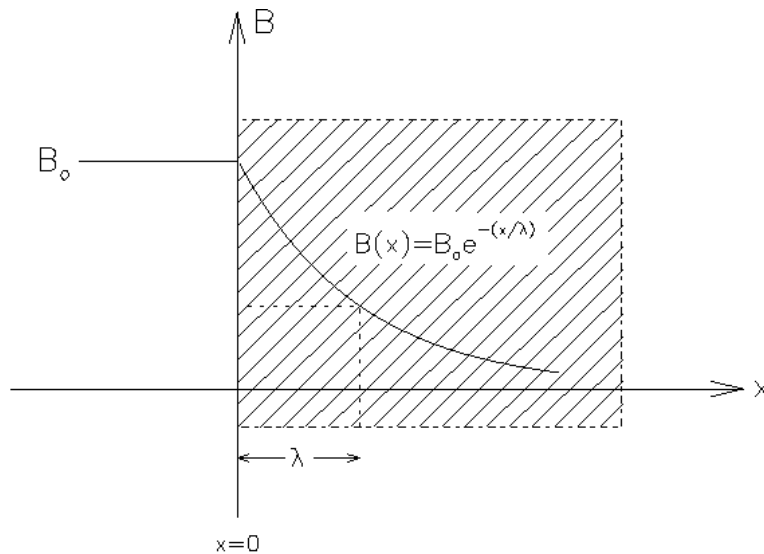
(the second step uses the vector calculus identity $\nabla \times \nabla \times \mathbf{B} = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}$ where the first term on the right is zero because there are no magnetic monopoles)

This can be combined with the London equation to give:

$$\nabla^2 \mathbf{B} = \frac{\mathbf{B}}{\lambda_L^2}$$

Notably, a uniform solution ($B(r) = B_0 = \text{const}$) is not allowed unless $B_0 = 0$, and this perfectly captures the fact that a superconductor will not support an internal magnetic field. The equation from

a few steps ago ensures that $\mathbf{j} = 0$ in a region where $B=0$.



The solution to the equation above is a decaying exponential. Assuming a semi-infinite superconductor with magnetic field $B(0)$ on the surface, the field inside the superconductor is given by:

$$B(x) = B(0)e^{-x/\lambda_L}$$

Thus, λ_L measures the depth of penetration of the magnetic field and it is known as the London penetration depth. Quantitatively, this materials-dependent length

scale varies from 15-300 nm at zero temperature

We can also find a similar equation for current density:

$$\nabla \times \nabla^2 \mathbf{B} = \frac{\nabla \times \mathbf{B}}{\lambda_L^2}$$

$$\nabla^2 (\nabla \times \mathbf{B}) = \frac{\nabla \times \mathbf{B}}{\lambda_L^2}$$

$$\nabla^2 (\mu_0 \mathbf{j}) = \frac{\mu_0 \mathbf{j}}{\lambda_L^2}$$

$$\nabla^2 \mathbf{j} = \frac{\mathbf{j}}{\lambda_L^2}$$

Thus, current also decays exponentially from the surface of a superconductor, with a length set by the London penetration depth. This means that:

- If a superconductor is placed in a magnetic field, 'screening currents' will spontaneously occur in the superconductor to oppose this magnetic field, and these screening currents are confined only to the surface of a superconductor over a depth set by λ_L
- If a current deliberately flows through a superconductor, it will also be confined only to the surface regions

The London penetration depth is related to material parameters in the following way:

$$\lambda_L = \left(\frac{\epsilon_0 m c^2}{n_s e^2} \right)^{1/2} = \left(\frac{m}{\mu_0 n_s e^2} \right)^{1/2}$$

Where m is the effective electron mass, μ_0 and ϵ_0 are the permeability and permittivity of free space ($c^2 = 1/\mu_0 \epsilon_0$), and n_s is the superfluid density—the density of electrons (N/V) which participate in superconductivity. Note that a tiny fraction of valence electrons in a metal are involved in superconductivity.

Physics 140B Lecture 27

- Superconducting coherence length
- Bardeen-Cooper-Schrieffer theory

Coherence length

The London penetration depth (λ_L) is one fundamental length in a superconductor, and the coherence length (ξ) is another. The coherence length can have several (not unrelated) physical interpretations

- In a superconductor, the charge carrier unit is two electrons called a 'Cooper pair', and these two electrons are not necessarily adjacent to each other in real space. The coherence length can be thought of as describing the physical 'size' of a Cooper pair
- If superconductivity is disturbed or destroyed locally, the coherence length represent a 'healing length' over which it will recover.
- If a superconductor is interfaced with a non-superconductor, superconductivity will be suppressed slightly over a coherence length of the interface (and superconductivity will also penetrate into the non-superconductor—the proximity effect)

A spatial variation in the state of an electronic system requires kinetic energy, as we will see in a moment. We compare a plane wave with a strongly modulated wavefunction. This derivation relates most closely to the second interpretation of the coherence length.

Plane wave: $\psi = e^{ikx}$

Modulated function: $\phi(x) = 2^{-\frac{1}{2}}(e^{i(k+q)x} + e^{ikx})$

Probability density of plane wave is uniform in space: $\psi^* \psi = e^{-ikx} e^{ikx} = 1$

Probability density of other wavefunction is modulated with wavevector q: $\phi^* \phi = \frac{1}{2}(e^{-i(k+q)x} + e^{-ikx})(e^{i(k+q)x} + e^{ikx}) = \frac{1}{2}(2 + e^{iqx} + e^{-iqx}) = 1 + \cos qx$

The kinetic energy of the plane wave is $\epsilon = \frac{\hbar^2 k^2}{2m}$

The kinetic energy of the modulated distribution is given by:

$$\int dx \phi^* \left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \right) \phi = \frac{1}{2} \frac{\hbar^2}{2m} [(k+q)^2 + k^2] \approx \frac{\hbar^2 k^2}{2m} + \frac{\hbar^2}{2m} kq$$

Assuming $q \ll k$

The increase in energy in the modulated wavefunction is $\frac{\hbar^2}{2m} kq$, and if this exceeds the superconducting gap, E_g , superconductivity is destroyed. Note that most of the superconductivity literature uses the symbol Δ to denote the superconducting gap, but your textbook uses E_g . We can solve for a critical value of q for this to happen, assuming the relevant k is k_F , because superconductivity is an instability of the Fermi surface.

$$\frac{\hbar^2}{2m} k_F q_0 = E_g$$

The intrinsic coherence length, $\xi_0 = 1/q_0$, and we can solve for it:

$$\xi_0 = \frac{\hbar^2 k_F}{2m E_g} = \frac{\hbar v_f}{2E_g}$$

Coherence length varies from 1-1600 nm, with lower T_c superconductors often having longer coherence length. As a reference E_g is typically $\sim 0.1-10$ meV

Superconductivity (at least the kind that we understand) is a macroscopic quantum phenomenon that is quite robust against impurities (at least non-magnetic ones), but impurities do decrease the effective coherence length. If the mean free path (measured in the normal state) is given by ℓ and is smaller than ξ_0 , the two length scales in superconductors are written in the following way:

Coherence length: $\xi = (\xi_0 \ell)^{1/2}$

Magnetic penetration depth: $\lambda = \lambda_L \left(\frac{\xi_0}{\ell} \right)^{1/2}$

(the blue color above indicates that we didn't cover this in class)

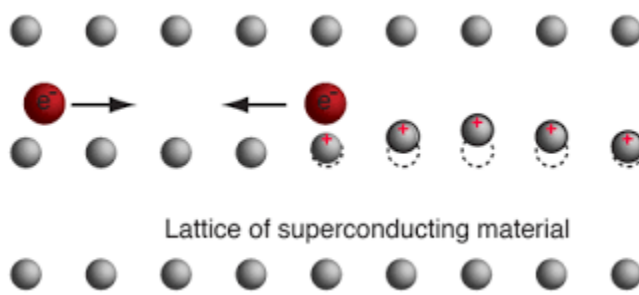
For all superconductors, the ratio $\frac{\lambda}{\xi} = \kappa$ or $\frac{\lambda_L}{\xi_0} = \kappa_0$ determines whether a superconductor is type I (full expulsion of magnetic field) or type II (partial expulsion of magnetic field with magnetic field entering superconductor via vortices):

Type I: $\kappa < 1/\sqrt{2}$ (ξ is larger)

Type II: $\kappa > 1/\sqrt{2}$ (λ is larger)

BCS theory overview

- Attractive interaction between electron leads to the formation of Cooper pairs—a pair of electrons with opposite momentum and opposite spin ($\mathbf{k} \uparrow, -\mathbf{k} \downarrow$) (for superconductors explained by BCS theory. ‘unconventional’ superconductors can have same-spin electrons or a more complicated momentum relationship between the two electrons). Cooper pairs are the charge carrier unit in a superconductor
- This attractive interaction is mediated by **electron-phonon** interaction. A cartoon is shown below; it has many deficiencies, but is a useful starting point. The general idea is the following



- Electrons are much lighter than atoms and hence move much faster
- First electron attracts positively charged atoms in the lattice toward it as it zips by
- Second electron comes before the atoms have relaxed back to their original position and is attracted by

the extra positive charge in the region where the first electron was; hence the motion of the lattice (phonons) mediating attraction between electrons

- The role of the lattice in mediating superconductivity manifests in the following experimental observables
 - Isotope effect
 - Higher Debye temperature = higher T_c

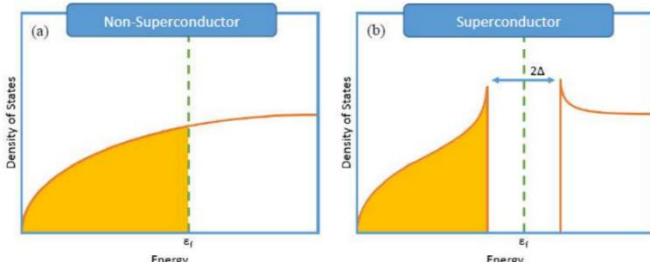


Figure 1. image source:
<https://www.osti.gov/servlets/purl/1239346>

- Only a small number of electrons are involved in superconductivity—roughly those within E_g of the Fermi energy.
 - T_c varies exponentially with density of states at E_F ($D(E_F)$)
 - Importance of Fermi energy and Fermi surfaces in determining superconducting properties
- The following equations describe the superconducting state in the ‘weak

coupling’ limit, meaning all interactions are as weak as possible

- $T_c = 1.14 \Theta e^{-1/UD(\epsilon_F)}$ where Θ is the Debye temperature (involvement of phonons), $D(\epsilon_F)$ is the density of states at the Fermi energy (only electrons close to the Fermi energy matter), and U is an attractive electron-phonon interaction.
- $E_g(T = 0) = 3.52 k_B T_c$
- The many-electron wavefunction in the superconducting state has a magnitude (Δ) and a phase (ϕ): $\psi \sim \Delta e^{i\phi}$. The magnitude depends on the material and the temperature. The phase is spontaneously chosen below T_c and reflects the fact that superconductivity relies both on the formation of Cooper pairs (which relates to the magnitude) and the long-range phase coherence of these Cooper pairs, to form a macroscopic quantum state.

Below is a basic, but more complete overview of BCS theory. We will not get through all of it in the last lecture, but it is here for your information. **The information below will not be on the final exam or HW8.**

The first key idea in BCS theory is that there is an effective attraction for electrons near the Fermi energy. Normally, electrons repel each other:

$$V(\mathbf{r} - \mathbf{r}') = \frac{e^2}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|}$$

While this is true for free electrons, in a metal, this interaction is screened, and the Thomas-Fermi model is the simplest model to describe this, giving an effective interaction of the form:

$$V_{TF}(\mathbf{r} - \mathbf{r}') = \frac{e^2}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} e^{-|\mathbf{r} - \mathbf{r}'|/r_{TF}}$$

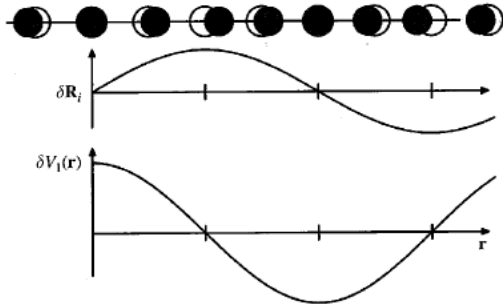
Where r_{TF} is the Thomas-Fermi screening length. This gives a repulsive force that is much shorter range.

The second ingredient is electrons interacting with the lattice.

As we saw in chapter 5, phonons in a crystal lattice can be treated as a set of quantum harmonic oscillators. The Hamiltonian for this can be written as:

$$H = \sum_{\mathbf{q}, \lambda} \hbar \omega_{\mathbf{q}\lambda} \left(a_{\mathbf{q}\lambda}^{\dagger} a_{\mathbf{q}\lambda} + \frac{1}{2} \right) = \sum_{\mathbf{q}, \lambda} \hbar \omega_{\mathbf{q}\lambda} \left(n_{\mathbf{q}\lambda} + \frac{1}{2} \right)$$

Where \mathbf{q} is the wavevector of the phonon and λ labels the phonon mode (branch), in the case that there is more than one phonon with the same \mathbf{q} . In a 3D solid with N atoms per unit cell, there are 3N phonon branches in total. The harmonic oscillator ladder operators can be used to calculate atomic displacement for an atom located at position \mathbf{R}_i due to phonons:



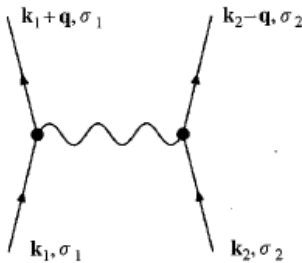
$$\delta \mathbf{R}_i = \sum_{\mathbf{q}, \lambda} \mathbf{e}_{\mathbf{q}\lambda} \left(\frac{\hbar}{2M\omega_{\mathbf{q}\lambda}} \right)^{\frac{1}{2}} (a_{\mathbf{q}\lambda}^{\dagger} + a_{\mathbf{q}\lambda}) e^{i\mathbf{q} \cdot \mathbf{R}_i}$$

Where M is the atomic mass (assuming only one atom per unit cell), and $\mathbf{e}_{\mathbf{q}\lambda}$ is the unit vector of the **atomic displacement** of phonon mode $\mathbf{q}\lambda$.

This modulation in the atomic position will also modulate the atomic potential of the lattice:

$$\delta V_1(\mathbf{r}) = \sum_i \frac{\partial V_1(\mathbf{r})}{\partial \mathbf{R}_i} \delta \mathbf{R}_i$$

This is a periodic modulation of the lattice with wavelength $2\pi/q$, and an electron moving through this potential will experience diffraction. If it is initially in Bloch state $\psi_{\mathbf{k}}(\mathbf{r})$ it will be diffracted to Bloch state $\psi_{\mathbf{k}-\mathbf{q}}(\mathbf{r})$. The extra momentum has been provided by the phonon. One can either think of this process as creating a phonon with momentum \mathbf{q} or annihilating one with momentum $-\mathbf{q}$. A second electron can also interact with this phonon, and by reciprocity, these electrons then interact with each other through exchange of a phonon.



The effective interaction between electrons due to exchange of a phonon

is given by:

$$V_{eff}(\mathbf{q}, \omega) = |g_{\mathbf{q}\lambda}|^2 \frac{1}{\omega^2 - \omega_{\mathbf{q}\lambda}^2}$$

Where $g_{\mathbf{q}\lambda}$ is the matrix element for scattering an electron from state \mathbf{k} to state $\mathbf{k}+\mathbf{q}$. It quantifies the strength of electron-phonon coupling for phonon $\mathbf{q}\lambda$, or the probability for an electron to be scattered by phonon $\mathbf{q}\lambda$. It turns out that $g_{\mathbf{q}\lambda}$ is of order $\sqrt{m/M}$ where m is the electron mass and M is the atomic mass. Because electrons are much lighter than atoms, we can think of electrons and phonons being weakly coupled, which allows for this simplified derivation to be valid.

Further simplifications can be made to the equation above to achieve an approximate solution that captures the key physics:

- Replace $g_{q\lambda}$ by g_{eff} , a q-independent average value for electron-phonon coupling
- Replace $\omega_{q\lambda}$ by a 'typical' phonon frequency, which is usually taken to be the debye frequency ω_D

This gives an effective electron-phonon interaction as:

$$V_{eff} = |g_{eff}|^2 \frac{1}{\omega^2 - \omega_D^2}$$

Notice that this interacting is attractive (negative) for frequencies less than ω_D . Because superconductivity is a low temperature phenomenon, typically only low frequencies ($\ll \omega_D$) will be relevant. Thus, the interaction is always attractive, and can be rewritten in its final form, $V_{eff} = -|g_{eff}|^2$. The key insight of BCS theory is that electrons near the Fermi surface are susceptible to attractive interactions with one another.

The corresponding term in the Hamiltonian, which we will use in the next section, is given by:

$$H_1 = -|g_{eff}|^2 \sum c_{k_1+q\sigma_1}^+ c_{k_2-q\sigma_2}^+ c_{k_1\sigma_1} c_{k_2\sigma_2}$$

Which corresponds to scattering an electron from momentum \mathbf{k}_1 and spin σ_1 to momentum $\mathbf{k}_1+\mathbf{q}$ and an electron from momentum \mathbf{k}_2 and spin σ_2 to momentum $\mathbf{k}_2-\mathbf{q}$

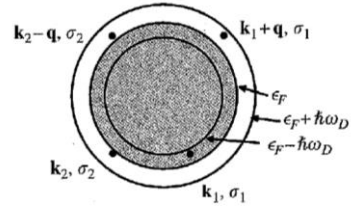


Fig. 6.6 The effective electron-electron interaction near the Fermi surface. The electrons at \mathbf{k}_1, σ_1 and \mathbf{k}_2, σ_2 are scattered to $\mathbf{k}_1 + \mathbf{q}, \sigma_1$ and $\mathbf{k}_2 - \mathbf{q}, \sigma_2$. The interaction is attractive provided that all of the wave vectors lie in the range where ϵ_k is within energy $\pm \hbar\omega_D$ of the Fermi energy.

Cooper pairs

The next step after the presence of an attractive interaction is the formation of pairs. The charge carrier unit in a superconductor is not a single electron, but a pair of electrons called a Cooper pair.

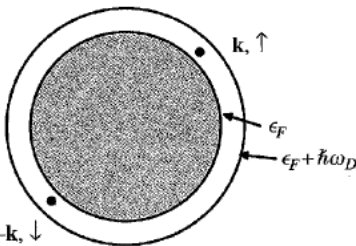


Fig. 6.7 The Cooper problem: two electrons outside a fully occupied Fermi sea. The interaction is attractive provided that the electron energies are in the range $\epsilon_F < \epsilon_k < \epsilon_F + \hbar\omega_D$.

Consider a spherical Fermi surface where all states $k < k_F$ are occupied. Now consider placing two extra electrons outside the Fermi surface. The two-particle wavefunction of these extra electrons is:

$$\Psi(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2) = e^{ik_{cm} \cdot R_{cm}} \varphi(\mathbf{r}_1 - \mathbf{r}_2) \phi_{\sigma_1, \sigma_2}^{spin}$$

Where R_{cm} is the center of mass position of the pair, $\hbar k_{cm}$ is the total momentum of the pair. φ is the spatial part of the two-electron wavefunction and ϕ is the spin part. The spin part is usually a singlet:

$$\phi_{\sigma_1, \sigma_2}^{spin} = \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

Fermion antisymmetry implies that $\Psi(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2) = -\Psi(\mathbf{r}_2, \sigma_2, \mathbf{r}_1, \sigma_1)$. Since the spin function is odd with respect to exchange of particles, the spatial part must be even.

Expanding $\varphi(\mathbf{r}_1 - \mathbf{r}_2)$ in terms of Bloch waves gives:

$$\varphi(\mathbf{r}_1 - \mathbf{r}_2) = \sum_{\mathbf{k}} \varphi_{\mathbf{k}} e^{i\mathbf{k} \cdot (\mathbf{r}_1 - \mathbf{r}_2)}$$

Where $\varphi_{\mathbf{k}}$ are some expansion coefficients to be found, given the constraint $C = \sum_{\mathbf{k}} \varphi_{\mathbf{k}}$

The full pair wave function can be written as a sum of slater determinants

$$\Psi(\mathbf{r}_1, \sigma_1, \mathbf{r}_2, \sigma_2) = \sum_{\mathbf{k}} \varphi_{\mathbf{k}} \begin{vmatrix} \psi_{\mathbf{k}\uparrow}(\mathbf{r}_1) & \psi_{\mathbf{k}\downarrow}(\mathbf{r}_2) \\ \psi_{-\mathbf{k}\uparrow}(\mathbf{r}_1) & \psi_{-\mathbf{k}\downarrow}(\mathbf{r}_2) \end{vmatrix}$$

Where terms like $\psi_{\mathbf{k}\downarrow}(\mathbf{r}_2)$ represent single particle bloch states. The form of the slater determinant implies that Cooper pairs involve electrons at \mathbf{k} and $-\mathbf{k}$ which have opposite spins.

Plugging in this trial wavefunction into the schrodinger equation, including both a kinetic energy term and the interaction term from earlier yields:

$$E\varphi_{\mathbf{k}} = 2\epsilon_{\mathbf{k}}\varphi_{\mathbf{k}} - |g_{eff}|^2 \sum_{\mathbf{k}'} \varphi_{\mathbf{k}'}$$

Where $\mathbf{q} = \mathbf{k}' - \mathbf{k}$

Using $C = \sum_{\mathbf{k}} \varphi_{\mathbf{k}}$ we can solve the equation above to yield

$$\varphi_{\mathbf{k}} = -C |g_{eff}|^2 \frac{1}{E - 2\epsilon_{\mathbf{k}}}$$

Self consistency requires that

$$C = \sum_{\mathbf{k}} \varphi_{\mathbf{k}} = -C |g_{eff}|^2 \sum_{\mathbf{k}} \frac{1}{E - 2\epsilon_{\mathbf{k}}}$$

Or equivalently:

$$1 = -|g_{eff}|^2 \sum_{\mathbf{k}} \frac{1}{E - 2\epsilon_{\mathbf{k}}}$$

We can convert the sum into an integral over energy via the density of states, and we take this integral only up to $\hbar\omega_D$, because interactions cease to be attractive beyond this point. The density of states is taken outside the integral because it is assumed that it does not change much over the narrow energy range being considered.

$$1 = -|g_{eff}|^2 D(\epsilon_F) \int_0^{\hbar\omega_D} d\epsilon \frac{1}{E - 2\epsilon}$$

The integral can be solved and the results can be arranged to give

$$-E = 2\hbar\omega_D e^{-1/|g_{eff}|^2 D(\epsilon_F)}$$

This equation implies that a **bound state** exists, meaning that it is more energetically favorable for electrons to be bound in a pair than to be separated. The energy scale is set by both the debye

frequency, and the density of states at ϵ_F multiplied by the electron-phonon coupling parameter. Note that your textbook uses U to refer to $|g_{eff}|^2$.

Superconducting gap

The derivation of the BCS gap parameter, Δ , is very much abridged.

The BCS gap parameter can be expressed as the expectation value of the Cooper pair operator, which corresponds to removing electrons of opposite spin from opposite sides of the fermi surface to join the superconducting condensate:

$$\Delta = |g_{eff}|^2 \sum_{\mathbf{k}} \langle c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \rangle = |g_{eff}| \sum_{\mathbf{k}} \frac{\Delta}{2E_{\mathbf{k}}}$$

Where $E_{\mathbf{k}} = \sqrt{(\epsilon_{\mathbf{k}} - \mu)^2 + |\Delta|^2}$. Physically, $\pm E_{\mathbf{k}}$ represent band dispersions in the superconducting state. The expression above is given without derivation, but the derivation can be found in the resource mentioned earlier. Δ can be cancelled from the first and third term above to yield:

$$1 = \frac{|g_{eff}|^2}{2} \sum_{\mathbf{k}} \frac{1}{\sqrt{(\epsilon_{\mathbf{k}} - \mu)^2 + |\Delta|^2}}$$

This expression can be used to find the magnitude of the BCS gap parameter, Δ . As before, we transform the sum over \mathbf{k} into an integral over energy using the density of states, and we assume that our energy range is small enough such that the density of states can be approximated as a constant, $D(\epsilon_F)$. We have also set the chemical potential μ to be zero.

$$1 = |g_{eff}|^2 D(\epsilon_F) \int_0^{\hbar\omega_D} \frac{1}{\sqrt{\epsilon^2 + |\Delta|^2}} d\epsilon$$

This integral can be solved approximately to yield:

$$|\Delta| = 2\hbar\omega_D e^{-1/|g_{eff}|^2 D(\epsilon_F)}$$

This is similar to an expression in your textbook for T_c , and you will connect the two in your homework. Physically, the BCS gap parameter $|\Delta|$ is related to the order parameter of the superconducting state. It is zero above T_c , and increases below T_c until it saturates at low temperature. Its temperature dependence is captured by the following equation, also given without derivation:

$$1 = |g_{eff}|^2 D(\epsilon_F) \int_0^{\hbar\omega_D} \frac{1}{\sqrt{\epsilon^2 + |\Delta|^2}} \tanh \frac{\sqrt{\epsilon^2 + |\Delta|^2}}{2k_B T} d\epsilon$$

By taking the limit where $\Delta \rightarrow 0$, one can obtain the expression for T_c in your textbook:

$$k_B T_c = 1.13 \hbar\omega_D e^{-1/|g_{eff}|^2 D(\epsilon_F)}$$

This has very similar form to the equation for the gap because both describe the robustness of the superconducting state

At $T=0$, one can also derive the relationship between $\Delta(T = 0)$ and T_c :

$$2\Delta(T = 0) = 3.52k_B T_c$$

Note that all the expressions with very specific numerical prefactors are only applicable in a 'weak coupling' regime, but they are nevertheless a useful benchmark. Note that in relation to the gap parameter that Kittel uses, $E_g = 2\Delta$