

PHY3102 - Solid State Physics I

Dr. A. P. Hibbins

If you take a bale of hay and tie it to the tail of a mule and then strike a match and set the bale of hay on fire, and if you then compare the energy expended shortly thereafter by the mule with the energy expended by yourself in the striking of the match, you will understand the concept of amplification.

William Shockley

1 Crystal Lattices

Copied from Physics Of Crystals, may be irrelevant The crystal structure is described by the lattice and the basis. To find the lattice structure first identify the atoms in identical environments and then define co-ordinate axes between these atoms. Note the axes need not be perpendicular and are not unique.

Note that every atom on the lattice points must be identical or else the crystal will not remain invariant as the lattice is translated.

The **primitive unit cell** is the smallest possible area which can be repeated to fill the whole space. The primitive unit cell only contains one lattice point within it (though it may share many with neighbouring cells). The area is given by $\vec{a} \times \vec{b}$ for 2d co-ordinates and $\vec{a} \cdot (\vec{b} \times \vec{c})$ (note that this is the scalar triple product).

And so the position vector of any lattice point is given by: $\vec{r} = u\vec{a} + v\vec{b}$ where u and v are integers. That is to say, any lattice point can be described by an integer translation along the lattice array.

Areas which are larger than the primitive unit cell, but can still be translated to fill all the space, are known as non-primitive lattice vectors.

Unit cell: When repeated through space without rotation and without gaps or overlaps, reconstructs the pattern to infinity.

Primitive unit cell: Is the smallest possible unit cell.

Non-Primitive unit cell: Contain extra lattice points not at the corners.

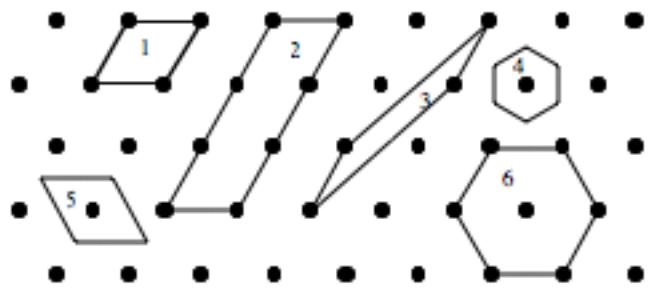
1, 3, 4 and 5 are primitive unit cells, 2 is a non-primitive unit cell (as it does not enclose the smallest area) and 6 is not a unit cell at all (as it cannot be repeated to fill all space without overlapping). **4 is the Wigner-Seitz primitive unit cell.**

1.1 The Wigner-Seitz unit cell

Copied from Physics Of Crystals, may be irrelevant

The Wigner-Seitz cell is found by choosing a lattice point and drawing lines from it to its nearest neighbours. The drawing perpendicular bisects to these vectors, then repeat this to the next nearest neighbours until an area is fully enclosed. This area is the Wigner-Seitz primitive unit cell.

- 1) The diagram shows a portion of a 2D lattice. Which of the outlines represent (i) primitive unit cells, (ii) non-primitive unit cells; (iii) not unit cells at all.



It can also be found by drawing points half-way between the lattice point and each of its neighbours, and then joining these points up. That is to say it is the region of space that is closer to the chosen lattice point than to any other.

1.2 The basis

Copied from Physics Of Crystals, may be irrelevant

The lattice alone does not fully describe the crystal structure, the basis is necessary too, to describe what atoms are attached to the lattice and how they are arranged. The lattice only describes the periodicity of the pattern.



Figure 1: The black points could be part of the the lattice whilst the grey points would be described by the basis. Note that the black points' position on the lattice is also specified by the basis (even if they are only at the origin). The roles could also be reversed with no consequences.

In this case the basis would be two carbon atoms (in the case of graphite) at $C(0,0)$ and $C(\frac{2}{3}, \frac{1}{3})$.

A basis of two atoms means that each lattice point has two atoms associated with it.

1.3 Lattice symmetries

Copied from Physics Of Crystals, may be irrelevant A Bravais Lattice is an infinite array of points that looks the same from any point. There are a finite number of distinct Bravais lattices and so they serve as template lattices for all crystals.

All lattices possess translational symmetry - i.e. the lattice remains the same when moved through integer numbers of lattice vectors. Note this means that all the atoms on lattice points must be of the same type, or else the crystal would not be invariant during translation. There are other types of symmetry:

Reflectional: The lattice is the same when reflected through a line through the lattice.

Rotation: The lattice is the same when rotated a certain number of degrees about a lattice point. Six-fold would be 60° (e.g. in the Wigner-Seitz cell for the graphite crystal), Four-fold would be 90° (e.g. in the square lattice) etc.

Inversion: The lattice is the same when lattice points are moved from their relative position to one lattice point, to the opposite side of the point.

1.4 Types of lattices:

Copied from Physics Of Crystals, may be irrelevant There are five distinct 2D lattices and 14 distinct 3D lattices.

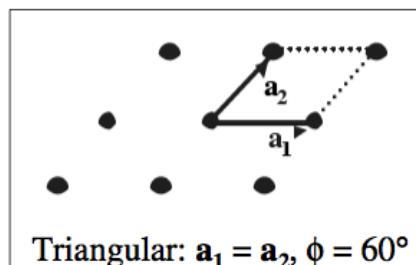
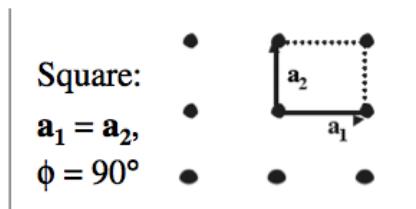
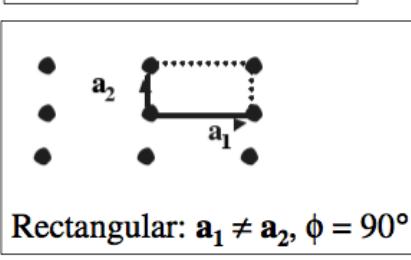
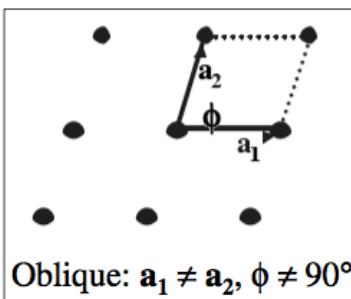
1.4.1 2D lattices:

Copied from Physics Of Crystals, may be irrelevant



2. 2D and 3D lattices

Five distinct 2D lattices:



12

The oblique lattice has the least symmetry as there is no link between a_1 , a_2 and ϕ . Whilst the square and triangular lattices are the most symmetrical.

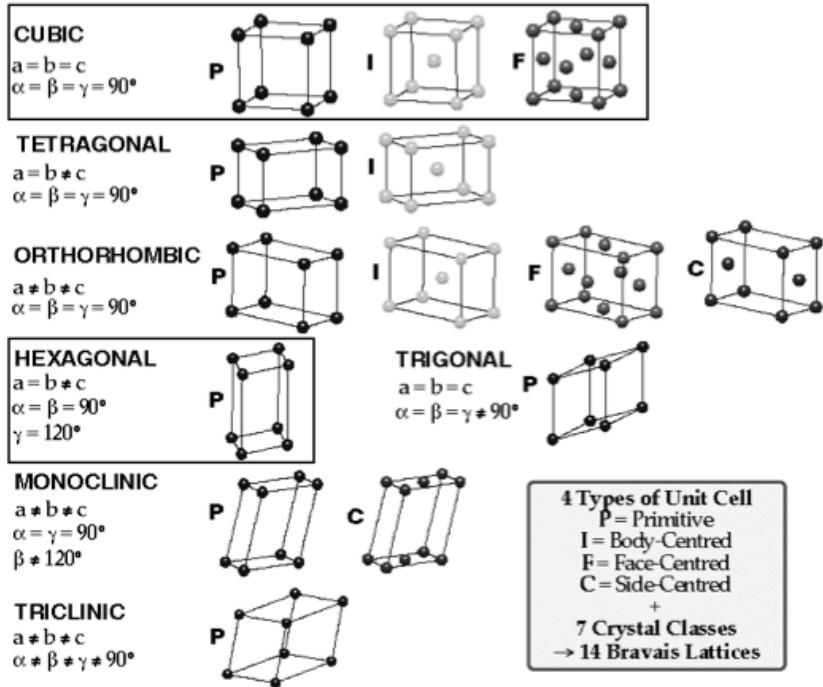
1.4.2 3D lattices:

Copied from Physics Of Crystals, may be irrelevant Where the top lattices have the most symmetry (cubic) and the bottom lattices have the least.

We will only study the cubic and hexagonal lattices.

2. 2D and 3D lattices

...and 14 distinct 3D lattices:



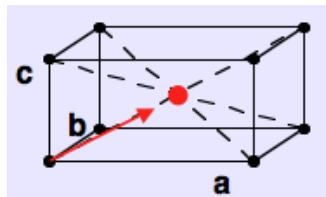
2 Miller Indices

Copied from Physics Of Crystals, may be irrelevant

2.1 Defining position and direction

Copied from Physics Of Crystals, may be irrelevant The position of any lattice point is given by $\mathbf{r} = u\mathbf{a} + v\mathbf{b} + w\mathbf{c}$, where u, v and w are integers (a general point in space is written the same way but when u, v and w need not be integers).

Directions are defined in terms of the vector \mathbf{r} from the origin, in the direction required.



E.g. for a point at the centre of an orthorhombic lattice. It has fractional co-ordinates $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, and the associated direction passing through the point (from the origin) is given by $[\frac{1}{2} \frac{1}{2} \frac{1}{2}]$.

2.2 Defining sets of planes

Copied from Physics Of Crystals, may be irrelevant Miller indices are used to define sets of planes. To find the Miller indices associated with a set of planes:

1. Identify the plane in the set which passes closest to the origin without actually passing through it.
2. Determine the positions at which the chosen plane crosses each of the co-ordinate axes defined by the primitive lattice vectors, writing them in fractional co-ordinates.
3. The Miller indices are the reciprocals of these fractional co-ordinates.

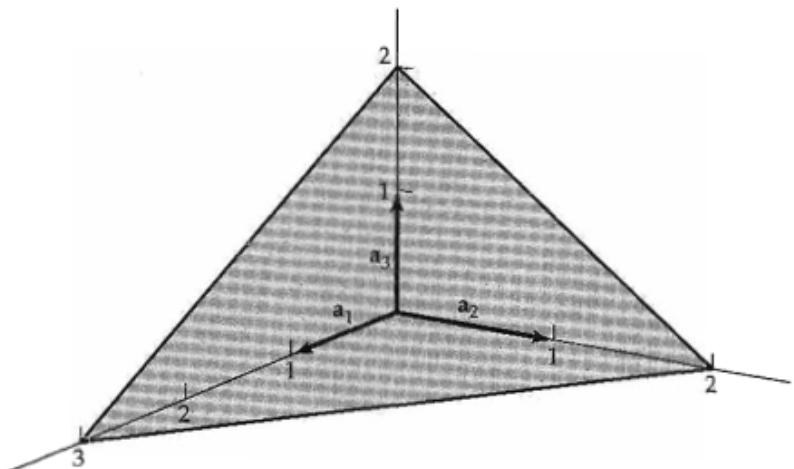


Figure 13 This plane intercepts the a_1 , a_2 , a_3 axes at $3a_1$, $2a_2$, $2a_3$. The reciprocals of these numbers are $\frac{1}{3}$, $\frac{1}{2}$, $\frac{1}{2}$. The smallest three integers having the same ratio are 2, 3, 3, and thus the indices of the plane are (233).

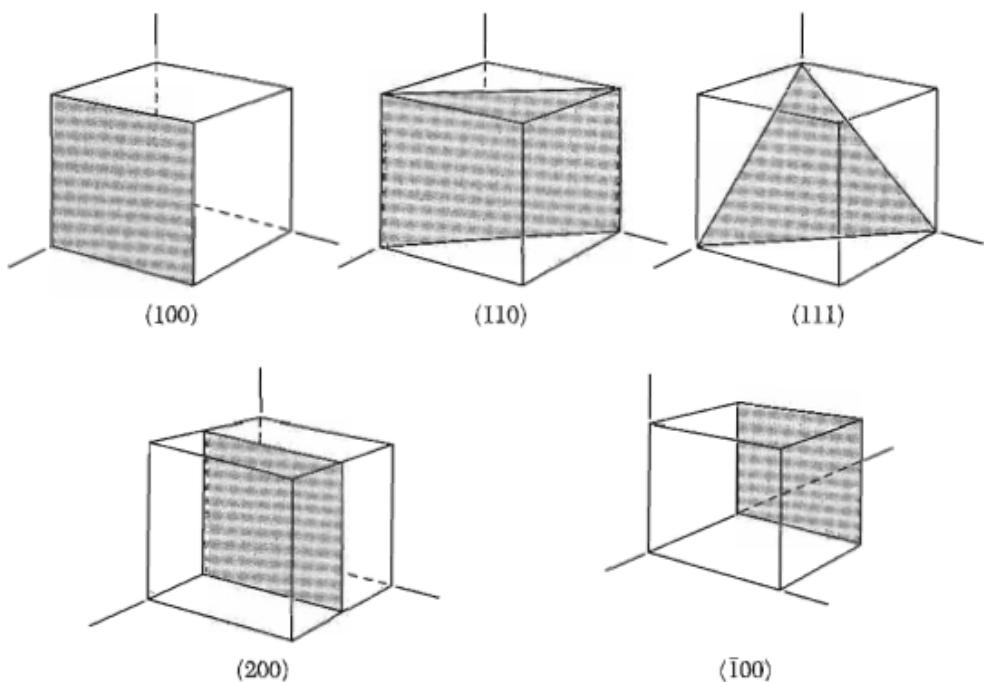


Figure 14 Indices of important planes in a cubic crystal. The plane (200) is parallel to (100) and to (100).

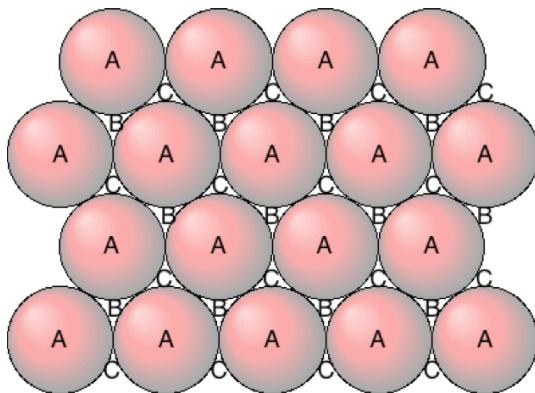
Note that:

- A bar above a number, in the Miller indices, designates it as negative.
- The indices represent a set of planes (with the same ratio of co-ordinate intercepts), not just one.
- Sets of 2D lines in a lattice can be represented similarly.
- In crystals of high symmetry, several sets of planes may be equivalent. E.g. in a cubic crystal (100), (010) and (001) are equivalent (all defining faces). These would be referred to as a group as {100}. The same thing can be done for a set of equivalent directions, which would be written as ⟨100⟩.

3 Typical crystal structures: Close-Packed Structures

Copied from Physics Of Crystals, may be irrelevant Structures with similar-sized atoms (such as the elemental solids) and non-directional/non-saturable bonds like to have the highest co-ordination number possible (as every bond lowers the energy).

For a set of rigid spheres in a plane, this is achieved by a triangular lattice (consider snooker balls). Note that A, B and C are all triangular lattices slightly displaced from one another. The length of the primitive lattice vectors is equal to the diameter of the spheres.



As layers are added they can be added on to either the B or C sites, for the first layer it matters not which sites are chosen. However, the next choice over whether to lay the third layer either back on the A sites or on to the C sites leads to two distinct lattice types. The ABCABC pattern creates a face-centered cubic (FCC or “cubic close packed”) lattice. This is more common.

The ABABAB pattern produces a hexagonal close-packed structure.

3.1 The Hexagonal Close-Packed structure

Copied from Physics Of Crystals, may be irrelevant The primitive unit vector is hexagonal with a two atom basis. One is at $(0, 0, 0)$, the other is at $(\frac{2}{3}, \frac{1}{3}, \frac{1}{2})$. Note that the unit vectors \mathbf{a} and \mathbf{b} are not perpendicular and that the second atom is **not** centered.

Finding the value of \mathbf{c} in terms of \mathbf{a} :

$$\left| \frac{2}{3}\vec{a} + \frac{1}{3}\vec{b} + \frac{1}{2}\vec{c} \right| = a \therefore \left(\frac{2}{3}\vec{a} + \frac{1}{3}\vec{b} + \frac{1}{2}\vec{c} \right) \cdot \left(\frac{2}{3}\vec{a} + \frac{1}{3}\vec{b} + \frac{1}{2}\vec{c} \right) = a^2$$

Since \vec{a} and \vec{b} are not perpendicular: $\therefore \frac{4}{9}a^2 + \frac{1}{9}a^2 + \frac{1}{4}c^2 + 2 \cdot \frac{2}{9}(\vec{a} \cdot \vec{b}) = a^2$

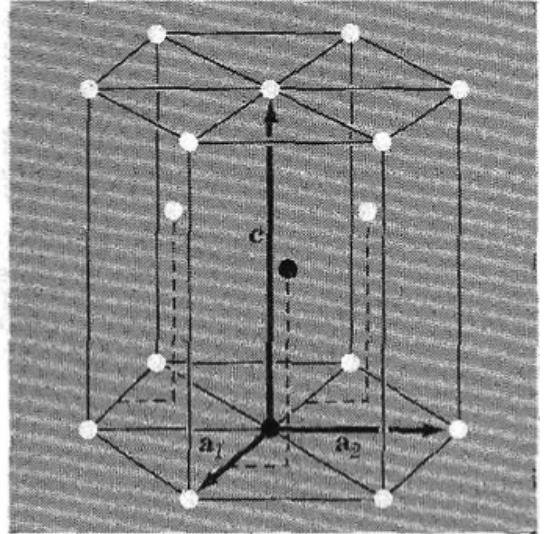


Figure 21 The primitive cell has $a_1 = a_2$, with an included angle of 120° . The c axis (or a_3) is normal to the plane of a_1 and a_2 . The ideal hcp structure has $c = 1.633 a$. The two atoms of one basis are shown as solid circles. One atom of the basis is at the origin; the other atom is at $\frac{2}{3}\vec{a}_1 + \frac{1}{3}\vec{a}_2 + \frac{1}{2}\vec{a}_3$.

$$\therefore \frac{4}{9}a^2 + \frac{1}{9}a^2 + \frac{1}{4}c^2 + \frac{4}{9}a^2\cos(120^\circ) = a^2 \quad \therefore \frac{4}{9}a^2 + \frac{1}{9}a^2 + \frac{1}{4}c^2 - \frac{2}{9}a^2 = a^2 \\ \therefore c = 2a\sqrt{\frac{2}{3}} = 1.63a$$

3.2 The Packing Fraction

Copied from Physics Of Crystals, may be irrelevant The packing fraction is the fraction of the volume actually taken up by the spheres (atoms). To calculate the packing fraction we use the primitive unit cell since its volume is easy to calculate and we know how many atoms are in the volume.

For the primitive unit cell: $V = \vec{a} \cdot (\vec{b} \times \vec{c})$

For the Hexagonal CP structure:

$$\vec{a} = a\vec{i}, \vec{b} = -a\vec{i}\sin(30^\circ) + a\vec{j}\cos(30^\circ), \vec{c} = 2\sqrt{\frac{2}{3}}a\vec{k}$$

So $\vec{a} \cdot (\vec{b} \times \vec{c}) = a\vec{i} \cdot \left((-a\vec{i}\sin(30^\circ) + a\vec{j}\cos(30^\circ)) \times \left(2\sqrt{\frac{2}{3}}a\vec{k}\right) \right)$, taking cross-product of unit-vectors:

$$= 2\sqrt{\frac{2}{3}}a^3\vec{i} \cdot ((-\vec{i}\sin(30^\circ) + \vec{j}\cos(30^\circ)) \times \vec{k}) = 2\sqrt{\frac{2}{3}}a^3\vec{i} \cdot (\vec{j}\sin(30^\circ) + \vec{i}\cos(30^\circ)) \\ = 2\sqrt{\frac{2}{3}}a^3\cos(30^\circ) = \sqrt{2}a^3$$

So to calculate the packing fraction: Packing Fraction = $\frac{\text{(Volume of two spheres, radius } \frac{a}{2}\text{)}}{\text{Volume of primitive unit cell}}$

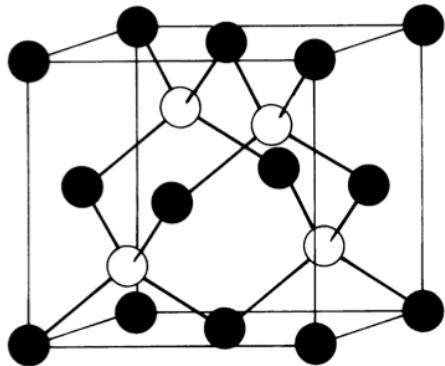
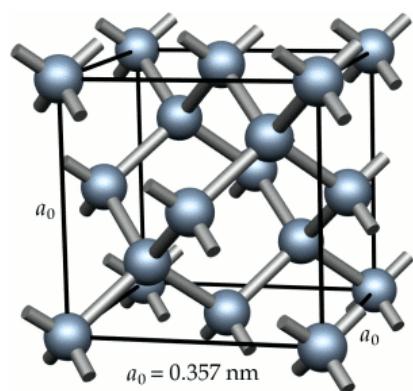
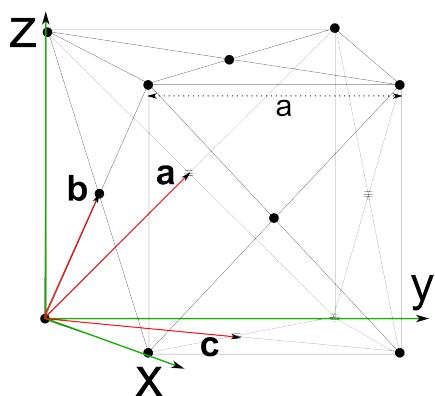
$$= \frac{2 \times \frac{4\pi(\frac{a}{2})^3}{3}}{\sqrt{2}a^3} = \frac{\pi}{3\sqrt{2}} = 0.740$$

This value is the largest packing fraction possible for rigid spheres, and is the same for the Face-Centred Cubic structure (ABCABC).

To solve for the packing fraction in general:

1. Calculate the smallest distance between atoms, i.e. the distance to the nearest neighbour specified in the basis.
2. Use this distance as the diameter of the spheres, and so calculate the radius (remember to take the modulus if in multiple directions).
3. Use the radius to calculate the volume per atom and multiply this by the number of atoms (lattice points multiplied by basis) to obtain the numerator of the packing fraction, the denominator is the total volume of the lattice cell.

Applying this to Diamond:



Diamond has an FCC structure with a diatomic basis of carbon atoms at $(0, 0, 0)$ and $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. The total number of lattice points is $\frac{1}{8} \cdot 8 + \frac{1}{2} \cdot 6 = 4$.

But there is a two-atom basis and so the number of atoms per lattice cell is $2 \cdot 4 = 8$.

Let a be the side length of the cube. We will let the diameter of the spheres be equal to the distance to the nearest neighbour, and so $2r = \sqrt{(\frac{1}{4}a)^2 + (\frac{1}{4}a)^2 + (\frac{1}{4}a)^2}$

$$\therefore 2r = \frac{a\sqrt{3}}{4} \quad \therefore r = \frac{a\sqrt{3}}{8}$$

$$V_{\text{cell}} = a^3 \text{ as it is a cubic lattice.}$$

$$\therefore \text{Packing Fraction} = \frac{8 \cdot \frac{4}{3}\pi \left(\frac{\sqrt{3}}{8}a\right)^3}{a^3} = 8 \cdot \frac{4}{3}\pi \cdot \frac{3^{\frac{3}{2}}}{8^3} = 0.34$$

3.3 Co-ordination number

Copied from Physics Of Crystals, may be irrelevant The co-ordination number is the number of nearest neighbours. For the HCP structure there are 6 in the same plane, plus three in each of the neighbouring planes, providing a total of 12 nearest neighbours, and so the co-ordination number of HCP structures is 12.

Co, He, Be, Mg, Zn and many other metals all crystallize in the HCP structure.

4 Laue Conditions for diffraction

Copied from Physics Of Crystals, may be irrelevant An alternative way to state the condition:

$$\Delta \vec{k} \cdot \vec{R} = \vec{G} \cdot \vec{R} = 2n\pi$$

given that $\vec{R} = u\vec{a} + v\vec{b} + w\vec{c}$ is to say:

$$\vec{G} \cdot \vec{a} = 2\pi h, \quad \vec{G} \cdot \vec{b} = 2\pi k, \quad \vec{G} \cdot \vec{c} = 2\pi l$$

where h, k and l are integers.

These three equations are the Laue conditions, and we shall see that the integers h, k and l are the Miller indices of the scattering planes.

This condition is satisfied if we define that:

$$\vec{G}_{hkl} = h\vec{a}^* + k\vec{b}^* + l\vec{c}^*$$

where:

$$\begin{aligned} \vec{a}^* \cdot \vec{a} &= 2\pi & \vec{a}^* \cdot \vec{b} &= 0 & \vec{a}^* \cdot \vec{c} &= 0 \\ \vec{b}^* \cdot \vec{b} &= 2\pi & \vec{b}^* \cdot \vec{c} &= 0 & \vec{b}^* \cdot \vec{a} &= 0 \\ \vec{c}^* \cdot \vec{c} &= 2\pi & \vec{c}^* \cdot \vec{a} &= 0 & \vec{c}^* \cdot \vec{b} &= 0 \end{aligned}$$

We need to find values of \vec{a}^*, \vec{b}^* and \vec{c}^* that satisfy these 9 conditions as this will satisfy the Laue conditions.

We can show this will work since:

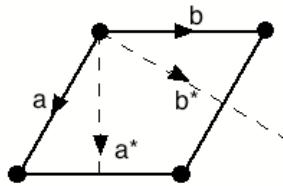
$$\vec{G} \cdot \vec{R} = (h\vec{a}^* + k\vec{b}^* + l\vec{c}^*) \cdot (u\vec{a} + v\vec{b} + w\vec{c}) = 2\pi(hu + kv + lw) = 2\pi \times \text{integer}$$

4.1 Geometric interpretation

Copied from Physics Of Crystals, may be irrelevant \vec{a}^* is perpendicular to the primitive lattice vectors \vec{b} and \vec{c} of the direct lattice, and the same for b and c with the other vectors.

So for a 2D triangular lattice:

From the dot product we see that the projection of \vec{a}^* in the direction of \vec{a} is $\frac{2\pi}{a}$.



4.2 Mathematical Derivation

Copied from Physics Of Crystals, may be irrelevant We need to construct a vector \vec{a}^* which is perpendicular to \vec{b} and \vec{c} and satisfying $\vec{a}^* \cdot \vec{a} = 2\pi$.

To obtain a perpendicular vector we take the cross product and then must divide by the triple scalar product to cancel out the other one that will form (to get 1), and so multiply this by 2π to satisfy the condition.

By considering this we obtain:

$$\vec{a}^* = 2\pi \frac{\vec{b} \times \vec{c}}{\vec{a} \cdot (\vec{b} \times \vec{c})}$$

$$\vec{b}^* = 2\pi \frac{\vec{c} \times \vec{a}}{\vec{a} \cdot (\vec{b} \times \vec{c})}$$

$$\vec{c}^* = 2\pi \frac{\vec{a} \times \vec{b}}{\vec{a} \cdot (\vec{b} \times \vec{c})}$$

Note the cyclic pattern in the changing cross product order for the different vectors, the triple scalar product remains unchanged since this is constant regardless of the order of the vectors. Note that the dimensions are $\frac{1}{\text{Length}}$ and the term in the denominator is the volume of the primitive unit cell of the real space lattice.

4.3 For 2D lattices:

Copied from Physics Of Crystals, may be irrelevant \vec{a}^* must be perpendicular to \vec{b} and \vec{b}^* must be perpendicular to \vec{a} .

We can either:

- Set \vec{c} to be a unit vector perpendicular to the plane and ignore the third reciprocal lattice vector \vec{c}^* .

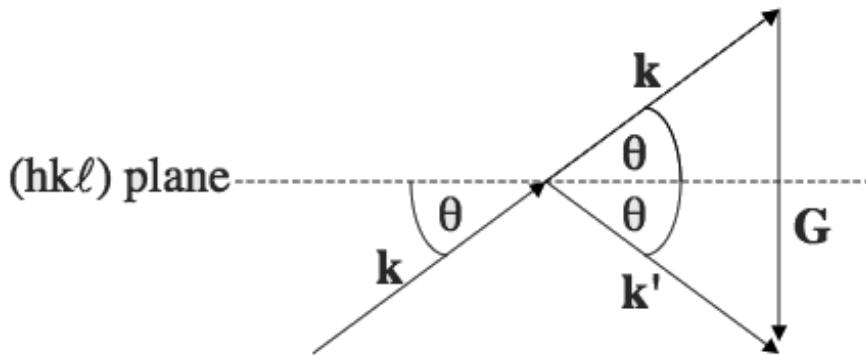
Or:

- Let \vec{c} be perpendicular to the plane and let it's length tend to infinity.

The latter method is preferable since it is more elegant. It works because if there is an infinite distance between planes in a certain dimension then the other planes do not exist. The equations for \vec{a}^* and \vec{b}^* still work because the magnitude of \vec{c} cancels out, meanwhile the magnitude of \vec{c} is present in the denominator of the equation for \vec{c}^* and so \vec{c}^* tends to zero.

4.4 Relationship with diffracting planes

Copied from Physics Of Crystals, may be irrelevant The expression $\vec{k}' - \vec{k} = \vec{G}$ can be represented in the form of a vector triangle.

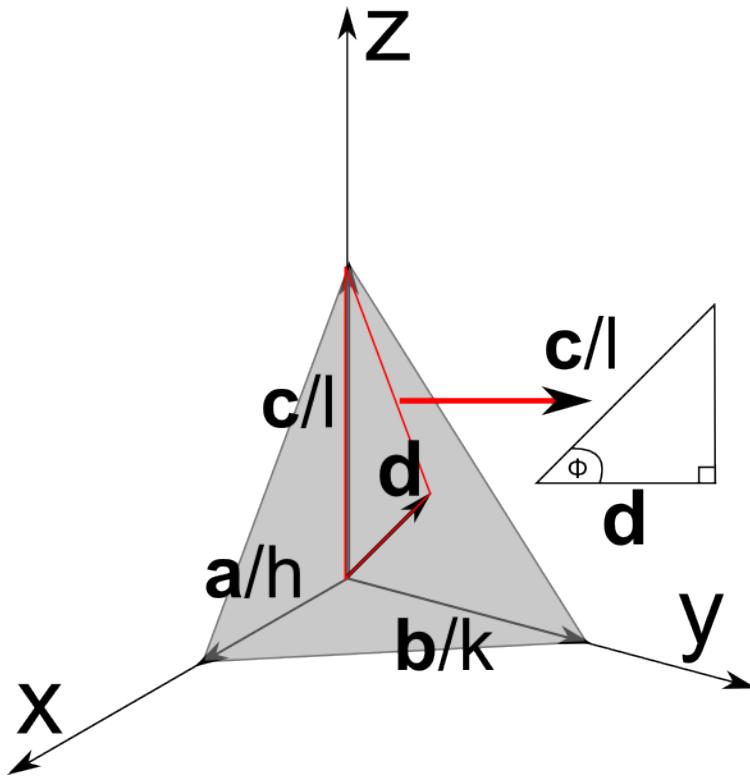


Note that \vec{G} is perpendicular to the scattering planes.

4.5 Connection with Miller indices

Copied from Physics Of Crystals, may be irrelevant The (hkl) plane closest to the origin without passing through it, another hkl plane passes through the origin and so \vec{d} is a vector perpendicular to the planes with a length equal to the plane separation.

The co-ordinates are found from the Miller indices since the co-ordinates are the reciprocals of the Miller indices, multiplied by the vector length.



Note that:

$$d \cdot \frac{c}{l} = d \left(\frac{c}{l} \cos(\phi) \right) = d^2$$

This is the same for other triangles and so:

$$d \cdot \frac{a}{h} = d \cdot \frac{b}{k} = d \cdot \frac{c}{l} = d^2$$

We note that this looks similar to the Laue condition which defined \vec{G} :

$$\begin{aligned}\vec{G} \cdot \vec{a} &= 2\pi h & \vec{d} \cdot \vec{a} &= d^2 h \\ \vec{G} \cdot \vec{b} &= 2\pi k & \vec{d} \cdot \vec{b} &= d^2 k \\ \vec{G} \cdot \vec{c} &= 2\pi l & \vec{d} \cdot \vec{c} &= d^2 l\end{aligned}$$

And so:

$$\vec{G}_{hkl} = \frac{2\pi}{d^2} \vec{d} = \frac{2\pi}{d} \hat{d}$$

This proves that \vec{G} is normal to the scattering planes and shows how the magnitude of \vec{G} is related to the plane spacing.

\vec{G} is the scattering wave vector $\vec{k}' - \vec{k}$ which gives rise to Bragg reflections involving planes having Miller indices (hkl) .

It's direction is **perpendicular** to these planes it's magnitude is $\frac{2\pi}{d}$ where d is the distance between adjacent planes.

4.6 For a cubic lattice

Copied from Physics Of Crystals, may be irrelevant Here:

$$\vec{a} = a\hat{i}, \quad \vec{b} = a\hat{j}, \quad \vec{c} = a\hat{k}$$

So:

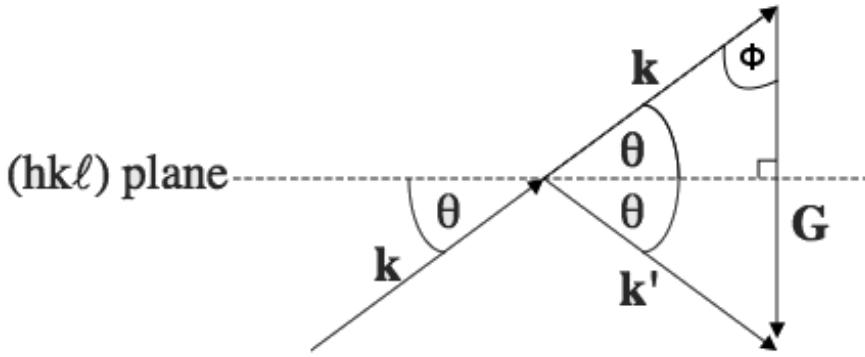
$$\vec{a}^* = \frac{2\pi}{a} \hat{i}, \quad \vec{b}^* = \frac{2\pi}{a} \hat{j}, \quad \vec{c}^* = \frac{2\pi}{a} \hat{k}$$

$$\vec{G}_{hkl} = \left(\frac{2\pi}{a} \right) (h\hat{i} + k\hat{j} + l\hat{k})$$

We can now find the distance between the neighbouring (hkl) planes:

$$d = \frac{2\pi}{|\vec{G}_{hkl}|} = \frac{2\pi}{\left(\frac{2\pi}{a} \right) \sqrt{h^2 + k^2 + l^2}}$$

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}}$$



4.7 Relationship with the Bragg Condition

Copied from Physics Of Crystals, may be irrelevant The diffracting planes bisect \vec{G} perpendicularly. This means that:

$$\vec{k} \cdot \vec{G} = \frac{-1}{2} G^2$$

But:

$$\vec{k} \cdot \vec{G} = \frac{\cos(\phi)}{|\vec{k}| |\vec{G}|}, \quad \text{but } \cos(\phi) = \sin(\theta)$$

So equivalently:

$$\begin{aligned} k \sin(\theta) &= \frac{G}{2} \\ \therefore \frac{2\pi}{\lambda} \sin(\theta) &= \frac{1}{2} \cdot \frac{2\pi}{d} \\ \therefore 2d \sin(\theta) &= \lambda \end{aligned}$$

Note the usually present n term is effectively included in the d term since higher order diffraction is taken care of in the Miller indices. I.e. 2nd order reflection from (100) planes would be termed (200) reflection, and so the distance is halved (which is equivalent to having an n of 2 in the traditional equation $n\lambda = 2dsin(\theta)$).

4.8 Brillouin Zones

Copied from Physics Of Crystals, may be irrelevant The Brillouin zones are related to the Wigner-Seitz cell of the reciprocal lattice.

The first Brillouin zone is simply the Wigner-Seitz cell of the reciprocal lattice, whilst the second Brillouin zone is the area covered when taken to the next nearest neighbours, minus the first Brillouin zone area.

In general the n th Brillouin zone is the region of reciprocal space reached from the chosen lattice point by crossing exactly $n - 1$ perpendicular bisectors.

This has the consequence that all Brillouin zones must have the same area (in 2D) or volume (in 3D).

The significance of the Brillouin zone is that any vector \vec{k} that ends at the edge of a Brillouin zone (i.e. the boundary) satisfies the Bragg condition and so gives rise to a diffracted beam.

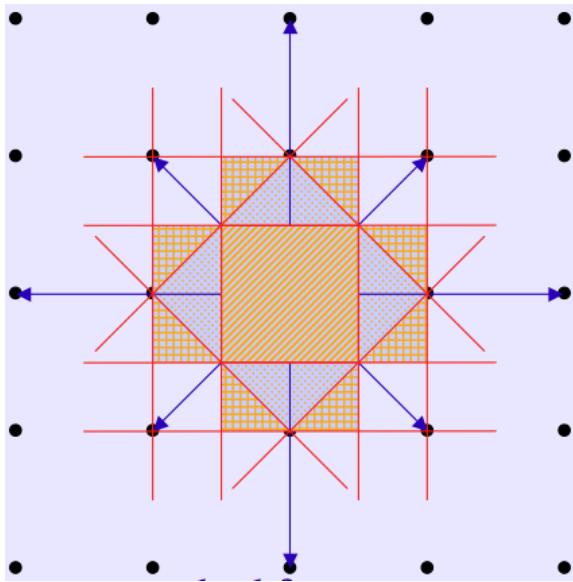


Figure 2: The first Brillouin zone is the central rectangle, the second Brillouin zone is the blue triangles, the third Brillouin zone is the checkered triangles.

4.9 Information in the reciprocal lattice

Copied from Physics Of Crystals, may be irrelevant The reciprocal lattice is the lattice of points in reciprocal space (k-space) which represents waves with the same periodicity as the direct lattice.

Note that:

- The reciprocal lattice vector \vec{G}_{hkl} is the scattering wave vector $\vec{k}' - \vec{k}$ which gives rise to Bragg reflections involving planes having Miller indices (hkl) .
- Its direction is perpendicular to these planes and its magnitude is $\frac{2\pi}{d}$ where d is the perpendicular distance between adjacent planes.

The reciprocal lattice tells us:

- All the scattering wavevectors $\Delta\vec{k}$ that give rise to Bragg reflections.
- The orientation (Miller indices) of the scattering planes.
- The separation between adjacent scattering planes.
- Through the construction of Brillouin zones, the incident wavevectors \vec{k} that will produce Bragg reflections.

4.10 X-ray Crystallography

Copied from Physics Of Crystals, may be irrelevant In X-ray crystallography we know the details of the incident waves but wish to analyse the structure of the sample crystal.

The diffraction gives a series of peaks at certain angles (2θ , where θ is the grazing angle) from the incident beam.

These allow us to construct the reciprocal lattice:

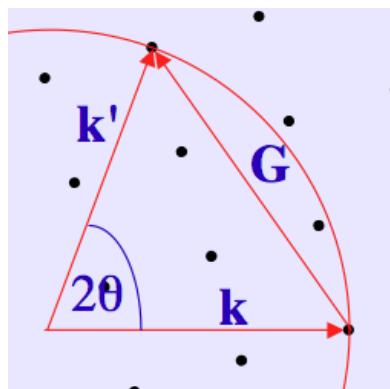
- Get the magnitude $\frac{2\pi}{d}$ from the Bragg condition: $2dsin(\theta) = n\lambda$.
- The orientation is in the direction of $\vec{k}' - \vec{k}$.

Once we know the reciprocal lattice we can derive the direct lattice since the reciprocal of the reciprocal lattice is the direct lattice.

4.11 Ewald Sphere construction

Copied from Physics Of Crystals, may be irrelevant The Ewald sphere is a graphical way of determining whether a wavevector \vec{k} satisfies the Bragg condition (and so produces Bragg reflection).

1. Draw reciprocal lattice.
2. Draw wavevector \vec{k} such that it ends on a reciprocal lattice point.
3. Draw a sphere (or circle in 2D) of radius k , centred at the start of the wave vector \vec{k} .
4. Diffraction will occur if the Ewald sphere passes through any other lattice points.



5 The Free-Electron Model

Copied from Physics Of Crystals, may be irrelevant

Here we neglect potential energy, and so the electrons are free except for the confinement within the crystal.

The dispersion relation is given by:

$$E = \frac{p^2}{2m} \therefore \omega = \frac{\hbar k^2}{2m}$$

The particle velocity is the group velocity:

$$v_g = \frac{d\omega}{dk} = \frac{1}{\hbar} \frac{dE}{dk} = \frac{\hbar k}{m} = \frac{p}{m} = v$$

The mass is inversely proportional to the curvature:

$$m = \frac{\hbar^2}{\frac{d^2E}{dk^2}}$$

The electrons are confined in a crystal with sides L_x , L_y and L_z . By considering with one dimension:

$$-\frac{\hbar^2}{2m} \frac{\partial^2\Psi}{\partial x^2} = E\Psi$$

with a boundary condition that Ψ is zero at the walls, the solution is:

$$\Psi = A\sin(k_x x) = A\sin\left(\frac{n\pi x}{L_x}\right)$$

where n is an integer.

In 2D:

$$\Psi = A\sin(k_x x)\sin(k_y y) = A\sin\left(\frac{n_x\pi x}{L_x}\right)\sin\left(\frac{n_y\pi y}{L_y}\right)$$

In 3D:

$$\Psi = A\sin(k_x x)\sin(k_y y)\sin(k_z z) = A\sin\left(\frac{n_x\pi x}{L_x}\right)\sin\left(\frac{n_y\pi y}{L_y}\right)\sin\left(\frac{n_z\pi z}{L_z}\right)$$

When we work out the quantized values of k we find there is $(\frac{\pi}{L})^3$ volume of k-space per allowed point in 3D, $(\frac{\pi}{L})^2$ area of k-space per allowed point in 2D and $\frac{\pi}{L}$ length of k-space per allowed point in 1D.

The **Pauli Exclusion Principle** states that only two electrons (one spin-up and one spin-down) can go into each of these states, and so the states get filled up to some energy, called the **Fermi Energy**, E_F .

The Fermi Energy, E_F , and the number of allowed states within kT of E_F both affect the properties of the solids.

5.1 The Density Of States

We have a cubic (or square or equally spaced) array of allowed states in 3D (or 2D or 1D) k-space.

We want to know the number of allowed states in a given energy range (between E and $E + dE$), and so in 3D:

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

So we use the allowed states between k and $k + dk$, the individual components don't matter so long as $|\vec{k}|$ is in this range.

This produces a sphere of radius k in 3D, a circle of radius k in 2D and a line up to k in 1D (so a length of $2k$).

However we can ignore the negative k-components, as these do not produce new quantum states (they are not measurable as distinct from positive k-components) and so we deal with

only the positive octant of the sphere (or the positive quadrant of the circle or positive half of the line), as using the whole sphere would overcount.

We know there are two allowed states for every $(\frac{\pi}{L})^3$ k-space in 3D, $((\frac{\pi}{L})^2$ k-space in 2D and $\frac{\pi}{L}$ k-space in 1D).

We also know that the volume of the positive octant of the sphere, area of the positive quadrant of the circle and length of the positive half space are given by:

$$V = \frac{\frac{4\pi k^3}{3}}{8} = \frac{\pi k^3}{6}$$

$$A = \frac{\pi k^2}{4}$$

$$L = k$$

So the number of states from 0 to k is given by:

$$\frac{\text{space from 0 to } k}{\text{space per state}}$$

Note that this includes two possible spin orientations.

So the number of states from 0 to k is:

3D	2D	1D
$N(0 \rightarrow k) = \frac{\frac{\pi k^3}{6}}{\frac{1}{2}(\frac{\pi}{L})^3} = \frac{V k^3}{3\pi^2}$	$N(0 \rightarrow k) = \frac{\frac{\pi k^2}{4}}{\frac{1}{2}(\frac{\pi}{L})^2} = \frac{A k^2}{2\pi}$	$N(0 \rightarrow k) = \frac{k}{\frac{\pi}{2L}} = \frac{2Lk}{\pi}$

Using that:

$$E = \frac{\hbar^2 k^2}{2m} \therefore k = \sqrt{\frac{2mE}{\hbar^2}}$$

3D	2D	1D
$N(0 \rightarrow E) = \frac{V}{3\pi^2} \left(\frac{2mE}{\hbar^2}\right)^{\frac{3}{2}}$	$N(0 \rightarrow E) = \frac{A}{2\pi} \left(\frac{2mE}{\hbar^2}\right)$	$N(0 \rightarrow E) = \frac{2L}{\pi} \left(\frac{2mE}{\hbar^2}\right)^{\frac{1}{2}}$

And so the density of states is the number of states per unit energy range:

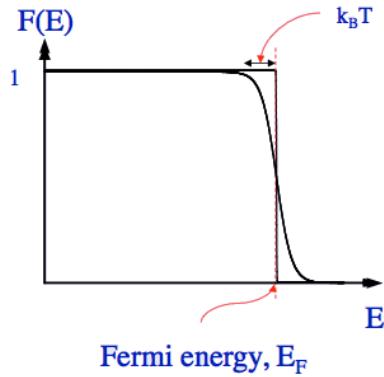
3D	2D	1D
$D(E) = \frac{dN}{dE} = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} E^{\frac{1}{2}}$	$D(E) = \frac{dN}{dE} = \frac{A}{2\pi} \left(\frac{2m}{\hbar^2}\right)$	$D(E) = \frac{dN}{dE} = \frac{L}{\pi} \left(\frac{2m}{\hbar^2}\right)^{\frac{1}{2}} E^{-\frac{1}{2}}$

5.2 The Fermi Distribution

So $D(E)$ tells us how many states we have available, however it does not specify which states are occupied by electrons, to do this we require a probability distribution function.

At $T = 0$ the electrons will fill the lowest energy states available. But there are only two electrons allowed per state (due to spin) and so the probability of the state being occupied is 1 up to a certain energy, E_F and 0 beyond this (as all the electrons are distributed).

However, at $T > 0$ electrons can be thermally excited to higher energy states if there are **empty** states within kT of the Fermi energy, E_F .



The function that describes this behaviour is:

$$F(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)}$$

Where $F(E)$ is the probability of the state being occupied. As $T \rightarrow 0$ then if $E > E_F$ then $F(E) \rightarrow 0$ whilst if $E_F > E$ then $F(E) \rightarrow 1$, which is the simple Fermi distribution described earlier.

5.3 Fermi Energy Estimation

The total number of states from $E = 0$ to $E = E_F$ must be equal to the total number of electrons in the system.

In 3D:

$$N_e = \frac{V}{3\pi^2} \left(\frac{2mE_F}{\hbar^2} \right)^{\frac{3}{2}} \quad \therefore E_F = \frac{\hbar^2}{2m} \left(3\pi^2 \frac{N_e}{V} \right)^{\frac{2}{3}}$$

In 2D:

$$N_e = \frac{A}{2\pi} \left(\frac{2mE_F}{\hbar^2} \right) \quad \therefore E_F = \frac{\pi\hbar^2 N_e}{Am}$$

For a metal there is usually one free electron for every atom and so usually one electron for every $(3 \times 10^{-10})^3 \text{ m}^3$ of volume (the typical volume of a primitive unit cell).

So as an estimate:

$$\frac{N_e}{V} = \frac{1}{(3 \times 10^{-10})^3} = 4 \times 10^{-28} \text{ m}^{-3} \quad \therefore E_f = 4 \text{ eV}$$

This is over 100 times kT at room temperature, and so the step like $t = 0$ behaviour of the Fermi function appears to be a good approximation. Electrons at E_F have velocities between 10^6 and 10^7 ms^{-1} (at which relativistic effects begin to become significant).

The Fermi temperature is given by:

$$T_F = \frac{E_F}{k} \approx 50,000 \text{ K}$$

5.4 When is a thin film 2D or 3D?

If L_z is very small in comparison to L_x and L_y (i.e. it is a thin film) then the energy levels associated with x and y are more closely spaced than those associated with z, since $k_x = \frac{n_x \pi}{L_x}$.

This means that the electrons will fill the increasing n_x and n_y values whilst at $n_z = 1$ until the number of electrons or temperature is increased significantly to provide enough energy for them to fill the $n_z = 2$ state. I.e. they will fill up in a staircase manner.

5.5 Heat Capacity

In the classical gas *every* particle acquires extra thermal energy as the temperature is increased, and so *every* particle contributes to the heat capacity:

$$C = \frac{3}{2} N_e k$$

Whilst in the Fermi gas, only electrons within kT of E_F acquire extra thermal energy. This is approximately a fraction $\frac{T}{T_F}$ of N .

So we expect something like:

$$C = \frac{3}{2} N_e k \frac{T}{T_F}$$

At room temperature $\frac{T}{T_F} \approx \frac{1}{100}$. This is what is observed and it explains the previous problem of the classical heat capacity being 100 times larger than the observations.

5.6 Electrical Conductivity

In equilibrium we now know that the electrons of a solid will fill up the allowed states in k-space up to an energy E_F . If we apply an electric field then the equation of motion is:

$$\frac{d\vec{p}}{dt} = -e\vec{E} \therefore \frac{d\vec{k}}{dt} = \frac{-e\vec{E}}{\hbar}$$

since $\vec{p} = \hbar\vec{k}$.

After a time t , an electron with wavevector \vec{k} will obtain an extra $\Delta\vec{k}$ from the electric field:

$$\Delta\vec{k} = \frac{-e\vec{E}t}{\hbar}$$

obtained by integrating the earlier expression with respect to time.

The electric field causes all the electrons to transfer from k-state to k-state in unison. I.e. they all move to an adjacent k-state in the direction of the electric field at once.

If the electrons were in a vacuum, or an infinite perfect crystal at absolute zero then the Fermi sphere would continue to shift this way forever as the electrons are accelerated indefinitely.

However, in a real solid it is assumed that the electrons scatter, on average, after a time τ called the **relaxation time**.

So each electron acquires *on average*:

$$\Delta \vec{k} = \frac{-e\vec{E}\tau}{\hbar}$$

The electrical conductivity is defined by the equation:

$$\vec{j} = \sigma \vec{E}$$

where $\vec{j} = n_e q \vec{v}_d$ and $n_e = \frac{N_e}{V}$ and σ is the inverse of the electrical resistivity.

Where \vec{v}_d is the drift velocity - the extra velocity due to the field which is related to $\Delta \vec{k}$ by:

$$\begin{aligned} \vec{v}_d &= \frac{\Delta \vec{p}}{m} = \frac{\hbar \Delta \vec{k}}{m} = \frac{-e \vec{E} \tau}{m} \\ \therefore \vec{j} &= \frac{n_e e^2 \vec{E} \tau}{m} \quad \therefore \sigma = \frac{n_e e^2 \tau}{m} \end{aligned}$$

This is the same formula as for the classical gas model, however the mean free path will be different.

In the classical model, the mean free path was calculated from:

$$M.F.P. = v_{\text{thermal}} \tau$$

where electrons were assumed to move at the thermal velocity and so:

$$\frac{1}{2} m v_{\text{thermal}}^2 = \frac{3}{2} kT \quad \therefore v_{\text{thermal}} = \sqrt{\frac{3kT}{m}} \approx 10^5 \text{ ms}^{-1}$$

Remember this gave a mean free path approximately equal to a lattice constant using a typical room temperature of value σ .

Note that the drift velocity caused by the electric field is very small compared with this, only a few mm s⁻¹.

In the free-electron fermi-gas model the velocity to use is the Fermi velocity:

$$M.F.P. = v_F \tau$$

where $\frac{1}{2} m v_F^2 = E_F$ and so $v_F = \sqrt{\frac{2E_F}{m}} = 10^6 \text{ ms}^{-1}$.

Consequently the value of τ obtained from the measurement of electrical conductivity gives a mean free path of approximately 10-100 lattice constants.

If we measure low temperature conductivity in very pure crystals we obtain a mean free path approximately equal to the size of the crystal, that is to say that the only scattering is from the edges of the crystal! (The scattering from the lattice disappears for some reason I didn't write down, I guess it's due to the long τ or something).

Pauli paramagnetism is another example of the success of the free electron model however I don't think it is examined.

So in summary the free-electron model succeeds in explaining:

- The small heat capacity with linear T dependence.
- Predicts contribution to the paramagnetic susceptibility (Pauli paramagnetism) in metals.
- Gives the same successful formula for the electrical conductivity as the classical theory

BUT it gives an unexpectedly long mean free path (which is correct however) and provides no explanation for insulators, which have no free electrons (which the free-electron model does not explain).

6 X-ray Diffraction Questions

Copied from Physics Of Crystals, may be irrelevant To solve the X-ray diffraction questions you must remember that:

$$\lambda = 2dsin(\theta)$$

as the n is taken in to the d as part of the Miller indices.

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}}$$

where a is the lattice constant and h, k, l are the Miller indices of the reflecting planes.

Bravais Lattice	Conditions on Miller Indices	
Simple Cubic	All h, k, l combinations allowed	From these equations we see
Body Centred Cubic	$h + k + l$ must be even	
Face Centred Cubic	h, k and l must be all even or all odd that:	

$$sin(\theta) \propto \sqrt{h^2 + k^2 + l^2}$$

So this means that the ratios of the $sin(\theta)$ values will be different depending on the allowed hkl values (as these will change the rate of change of the square rooted function).

Direct Lattice	Simple Cubic	Body-Centred Cubic	Face-Centred Cubic
Reciprocal Lattice (cube lattice parameter)	Simple Cubic ($2\pi/a$)	Face-Centred Cubic ($4\pi/a$)	Body-Centred Cubic ($4\pi/a$)
Planes, spacing	(100) a	No	No
	(110) $a/\sqrt{2}$	Yes	No
	(111) $a/\sqrt{3}$	No	Yes
	(200) $a/2$	Yes	Yes
	(210) $a/\sqrt{5}$	No	No
	(211) $a/\sqrt{6}$	Yes	No
	(220) $a/\sqrt{8}$	Yes	Yes
	(221) $a/3$	No	No
	(300) $a/3$	No	No
	(310) $a/\sqrt{10}$	Yes	No
	(311) $a/\sqrt{11}$	No	Yes
	(222) $a/\sqrt{12}$	Yes	Yes
	(320) $a/\sqrt{13}$	No	No
	(321) $a/\sqrt{14}$	Yes	No
	(400) $a/\sqrt{16}$	Yes	Yes
Rule:	All h, k, l	Sum $(h + k + l)$ even	h, k, l all odd or all even
Ratio of sines of Bragg angles	1: $\sqrt{2}$: $\sqrt{3}$:2: $\sqrt{5}$: $\sqrt{6}$: $\sqrt{8}$:...	1: $\sqrt{2}$: $\sqrt{3}$:2: $\sqrt{5}$: $\sqrt{6}$: $\sqrt{7}$: $\sqrt{8}$:...	1:2: $\sqrt{3}$: $\sqrt{8}$ / $\sqrt{3}$:...

So for the BCC lattice the ratio of the $\sin(\theta)$ of the second reflecting plane to the first is:

$$\frac{\sqrt{2^2 + 0^2 + 0^2}}{\sqrt{1^2 + 1^2 + 0^2}} = \frac{2}{\sqrt{2}} = \sqrt{2} = 1.414$$

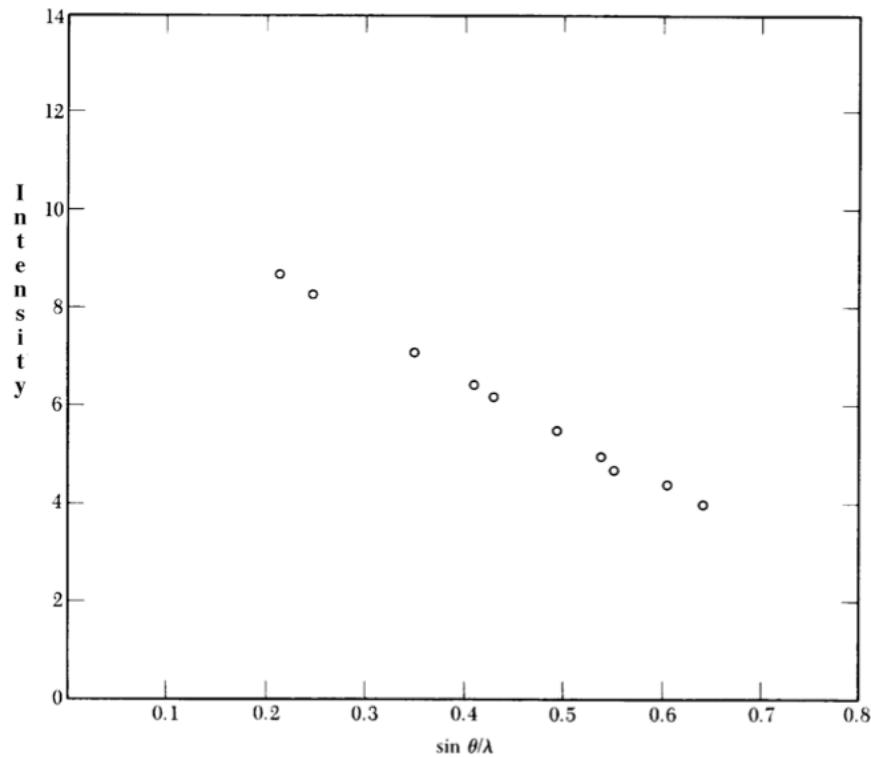
Whilst for the FCC lattice the ratio of the $\sin(\theta)$ of the second reflecting plane to the first is:

$$\frac{\sqrt{2^2 + 0^2 + 0^2}}{\sqrt{1^2 + 1^2 + 1^2}} = \frac{2}{\sqrt{3}} = 1.155$$

Note that the ratios of the second to first peaks for BCC and SC are the same, and so we would have to go to the third and second peak ratios to distinguish these.

So using the following example:

We observe that the first peak is at approximately 0.21 whilst the second is at 0.24. This a ratio of $\frac{0.24}{0.21} = 1.14$ which is closest to 1.155 and thus we deduce that the Bravais lattice is



Face-Centred Cubic.

It is then easy to label each of the peaks with the corresponding reflecting plane since we know that they will increase in the order of $\sqrt{h^2 + k^2 + l^2}$, so in this case it will go (111), (200), (220), etc.

Finally given that the wavelength used was 0.713 Å (though in this case this information is unnecessary as the horizontal scale is $\frac{\sin(\theta)}{\lambda}$) we can determine the lattice constant, a :

$$a = \frac{\lambda}{2\sin(\theta)} \sqrt{h^2 + k^2 + l^2} = \frac{1}{2} \sqrt{3} \frac{1}{0.21} = 4.12 \text{ Å}$$

End of Physics Of Crystals excerpts

7 Free electron model

7.1 Drude Model

The Drude model is the simplest of all the models of the electrical and thermal conductivity in metals. It neglects the periodic potential created by the regular positions of the ions in the lattice, and assumes that the potential energy is zero. It also ignores electron-electron interactions. The ions are assumed to be stationary, and the electrons interact only with the ions, like a pinball machine.

The electrons move with an average drift velocity v_d and collide with the ions after an average collision time τ .

This results in a constant average velocity:

$$\sum F_x = m \frac{d^2x}{dt^2} = -eE + F_{\text{coll}} = 0 \quad \text{but } F_{\text{coll}} = \frac{-mv}{\tau} \quad \therefore v = v_d = \frac{-eE\tau}{m}$$

if n electrons per unit time move with velocity \mathbf{v} then the current density they give rise to will be parallel to \mathbf{v} . In a time dt the electrons will advance by a distance vdt in the direction of \mathbf{v} . So $nAvdt$ electrons will pass an area A perpendicular to the direction of flow. Since each electron has a charge $-e$, the charge crossing the surface A in time dt is $-nevAdt$. $\mathbf{I} = \frac{d\mathbf{Q}}{dt} = -nevA$ and $\mathbf{j} = \frac{\mathbf{I}}{A}$ and so the current density is:

$$\mathbf{j} = -nev\mathbf{v}$$

Combining these equations yields Ohm's Law and the definition of electrical conductivity, σ :

$$J = \left(\frac{ne^2\tau}{m} \right) E = \sigma E$$

This is one correct result from the Drude model.

However, when the model is used to calculate the thermal conductivity this is not the case. Each monatomic gas molecule at a temperature T has an energy $\frac{3}{2}kT$ and so the total energy per mole is:

$$\epsilon'_{el} = \frac{3}{2}N_A kT = \frac{3}{2}RT \quad \therefore C_{el} = \frac{d(\epsilon'_{el})}{dT} = \frac{3}{2}R$$

But experimental observations show that the electronic specific heat is $\approx 1\%$ of this value expected from the classical theory, and does not explain its temperature dependence.

The Drude model fails at low temperature, as the actual effective mean free path is much higher than the interatomic spacing.

However, the Drude model does provide a reasonably accurate description of the experimentally discovered Wiedemann-Franz law, that the ratio of thermal and electrical conductivity is constant for a given temperature. Mathematically the law is stated as:

$$\frac{\kappa}{\sigma} = LT$$

The Drude theory gives:

$$L = \frac{3k^2}{2e^2}$$

Add derivation later

This is actually too small by a factor of 2.

What the Drude theory gets right

Ohm's Law

Wiedemann-Franz Law

What is wrong

Heat capacity*

Magnetic susceptibility*

Mean free paths of electrons in pure, cold specimens

*These are actually due to the failure of the classical Maxwellian distribution, not the free electron approximation.

8 A better model: The Quantum (Fermi) free electron model

Here we still assume electron-electron interactions are insignificant. Look for solutions using the 1D TISE:

$$\frac{-\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + U\psi(\mathbf{r}) = \varepsilon\psi(\mathbf{r})$$

With $U = 0$ this comes to $\nabla^2 \psi = -k^2 \psi$.

The solutions are $\psi = Ae^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}$ i.e. travelling plane waves propagating like a free particle, and $\varepsilon = \frac{\hbar^2 k^2}{2m}$ parabolic energy "bands".

Using periodic boundary conditions for a cubic solid with arbitrary side length $L = Qa$ we can define the set of allowed wave vectors as:

$$k_x = \frac{2\pi n_x}{L}, \quad k_y = \frac{2\pi n_y}{L}, \quad k_z = \frac{2\pi n_z}{L}, \quad n_x, n_y, n_z = 0, \pm 1, \pm 2 \dots$$

The volume of the first Brillouin Zone is $(\frac{2\pi}{L})^3$.

The Density Of States (DOS) is defined as $G(\varepsilon) = \frac{dN_s}{d\varepsilon}$. This is often needed to calculate the number of charge carriers that can contribute to the conduction process.

The number of states in the range $k \rightarrow k + dk$ is equal to the total volume of the k-space divided by the volume of one k-state (the volume of the first Brillouin zone).

So the number of allowed states in the shell (3D) $k \rightarrow k + dk$ is:

$$G(k)dk = 2 \times \left(\frac{4\pi k^2 dk}{(2\pi/L)^3} \right)$$

The factor of two is necessary for the two possible spin states the electrons can take.

Putting this in energy space:

$$\varepsilon = \frac{\hbar^2 k^2}{2m} \quad \therefore \quad k = \sqrt{\frac{2m\varepsilon}{\hbar^2}}, \quad \frac{d\varepsilon}{dk} = \frac{\hbar^2 k}{m} \quad \therefore \quad dk = \frac{md\varepsilon}{\hbar^2 k}$$

So:

$$\begin{aligned} 2 \left(\frac{4\pi k^2 dk}{(\frac{2\pi}{L})^3} \right) &= 2^3 \pi \hbar^{-2} 2m\varepsilon m \hbar^{-2} \hbar 2^{-\frac{1}{2}} m^{\frac{-1}{2}} \varepsilon^{\frac{-1}{2}} L^3 2^{-3} \pi^{-3} d\varepsilon = 2^{\frac{1}{2}} \pi^{-2} \hbar^{-3} m^{\frac{3}{2}} \varepsilon^{\frac{1}{2}} L^2 \\ &= \frac{2^{\frac{1}{2}} m^{\frac{3}{2}} V \varepsilon^{\frac{1}{2}}}{\pi^2 \hbar^3} \end{aligned}$$

where $V = L^3$. Note that ε is used as E is used for electric field.

So for the 3-dimensional case $G(\varepsilon) \propto \varepsilon^{\frac{1}{2}}$.

The Fermi energy, ε_F , is defined as the energy of the highest filled electron state at $T = 0K$. Using the fact that the total number of electrons should lie between zero and the Fermi energy, one can obtain an expression:

$$\text{Number of electrons} = N_e = \int_0^{\varepsilon_F} G(\varepsilon) d\varepsilon = \int_0^{\varepsilon_F} \frac{2^{\frac{1}{2}} m^{\frac{3}{2}} V \varepsilon^{\frac{1}{2}}}{\pi^2 \hbar^3} d\varepsilon = \left[\frac{2^{\frac{3}{2}} m^{\frac{3}{2}} V \varepsilon^{\frac{3}{2}}}{3\pi^2 \hbar^3} \right]_0^{\varepsilon_F}$$

$$\therefore \frac{2^{\frac{3}{2}} m^{\frac{3}{2}} V \varepsilon_F^{\frac{3}{2}}}{3\pi^2 \hbar^3} = N_e \quad \therefore \varepsilon_F^{\frac{3}{2}} = \frac{3N_e \pi^2 \hbar^3}{m^{\frac{3}{2}} V 2^{\frac{3}{2}}} \quad \therefore \varepsilon_F = \left(\frac{N_e 3\pi^2 \hbar^3}{m^{\frac{3}{2}} V 2^{\frac{3}{2}}} \right)^{\frac{2}{3}}$$

$$\therefore \varepsilon_F = \frac{\hbar^2}{2m} \left(3\pi^2 \frac{N_e}{V} \right)^{\frac{2}{3}}$$

By combining the equations for $G(\varepsilon)$ and ε_F we can write:

$$G(\varepsilon_F) = \frac{2^{\frac{1}{2}} m^{\frac{3}{2}} V \varepsilon_F^{\frac{1}{2}}}{\pi^2 \hbar^3} = \frac{3N_e}{2\varepsilon_F}$$

8.1 The Fermi-Dirac Distribution function

The occupation of electron states is governed by the Pauli exclusion principle and the probability of occupation of a state with energy ε at temperature T is:

$$f_{FD}(\varepsilon) = \frac{1}{\exp\left(\frac{\varepsilon - \mu}{kT} + 1\right)}$$

where $\mu(T)$ is the chemical potential - the amount by which the energy of the system would change if an additional particle were introduced, with the entropy and volume held fixed. At absolute zero, $\mu = \varepsilon_F$ as in the limit $T \rightarrow 0$, $f_{FD}(\varepsilon)$ changes discontinuously from 1 (filled) to 0 (empty).

The Fermi level is the level given to the state which has $f_{FD}(\varepsilon) = \frac{1}{2}$ (i.e. when $\varepsilon = \mu$)

When $kT \ll \varepsilon_F$ (which includes room temperature), we ignore the temperature dependence of μ and replace it with ε_F .

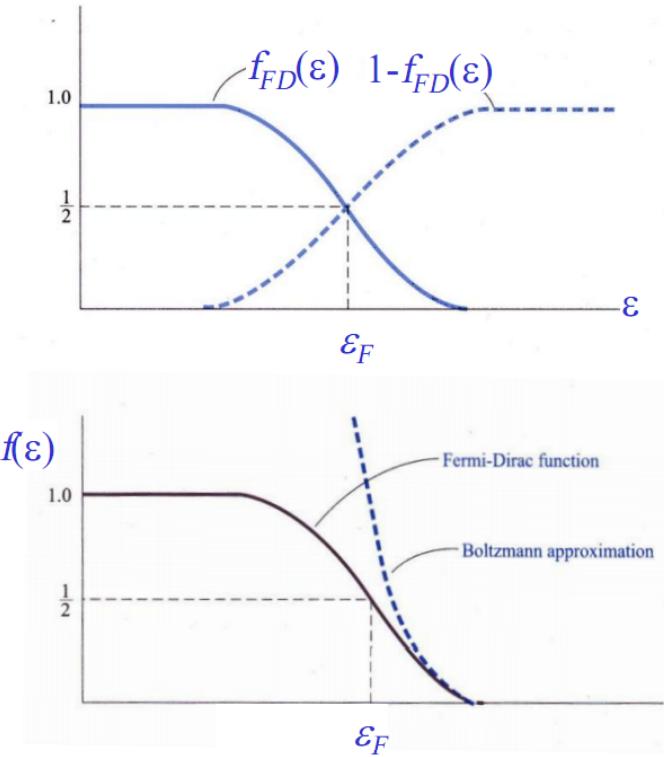
So usually:

$$f_{FD}(\varepsilon) = \frac{1}{\exp[(\varepsilon - \varepsilon_F)/kT] + 1}$$

Note the symmetry of the distribution, i.e. the probability of a state a distance $d\varepsilon$ above ε_F being occupied is the same probability of a state $d\varepsilon$ below ε_F being unoccupied (empty).

Also note the simplification, that when $\varepsilon - \varepsilon_F \gg kT$, the exponential term in the denominator of the Fermi-Dirac probability function is much greater than 1. And so the distribution can be approximated by the Maxwell-Boltzmann approximation:

$$f_{FD}(\varepsilon) \approx \exp\left[\frac{-(\varepsilon - \varepsilon_F)}{kT}\right]$$



For temperatures above T_F , the electron gas will act as a classical gas rather than being dominated by the Pauli Exclusion Principle (note that most metals will vapourise before T_F is reached): $varepsilon_F = kT_F$.

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

The occupied states of the FEG are described by k -values in a sphere of radius k_F . The average velocity of an electron in an atom at absolute zero is the Fermi velocity:

$$v_F = \left(\frac{2\varepsilon_F}{m} \right)^{\frac{1}{2}}$$

8.2 Heat Capacity of the Quantum Free Electron Gas

The actual electron population distribution (density) is given by:

$$n(\varepsilon) = \text{probability distribution function} \times \text{density of states function}$$

So at temperature T , the total energy of electrons is:

$$\varepsilon_{\text{el}} = \int_0^\infty \varepsilon f(\varepsilon) G(\varepsilon) d\varepsilon = \int_0^\infty \frac{\varepsilon G(\varepsilon)}{\exp((\varepsilon - \varepsilon_F)/kT) + 1} d\varepsilon \quad \text{and} \quad C_{\text{el}} = \frac{d\varepsilon_{\text{el}}}{dT}$$

The exact answer to this integral is complicated, and not covered here, but it turns out to be:

$$C_{\text{el}} = \frac{\pi^2}{3} k^2 G(\varepsilon_F) T$$

so:

$$C_{\text{el}} \propto T$$

The quantum derivation gives a different answer to the classical derivation because we only consider electrons that can actually be excited, i.e. those within kT of ε_F .

8.3 Rough estimate of C_{el}

We can estimate C_{el} in a few lines to confirm the linear dependence on temperature. The number of electrons that can absorb thermal energy is:

$$\approx (2kt)G(\varepsilon_F)f(\varepsilon_F) = kTG(\varepsilon_F)$$

Using that $f(\varepsilon_F) = \frac{1}{2}$.

The total thermal energy of electrons at temperature T is:

$$\varepsilon(T) \approx \varepsilon_0 + \left(\frac{3}{2}kT\right)kTG(\varepsilon_F) = \varepsilon_0 + \frac{3}{2}k^2T^2G(\varepsilon_F)$$

As each contributes energy $\frac{3}{2}kT$.

So the FEG heat capacity at T is:

$$C_{\text{el}} = \frac{d\varepsilon_{\text{el}}}{dT} \approx 3k^2G(\varepsilon_F)T$$

So:

$$C_{\text{el}} = \frac{\pi^2}{3}k^2G(\varepsilon_F)T = \gamma T$$

However, this linear dependence cannot be measured directly as the heat capacity of a real metal has an additional contribution from the lattice vibrations (phonons):

$$C(T) = C_{\text{el}} + C_{\text{lattice}} = \gamma T + \alpha T^3 \quad \text{where } \gamma_{\text{QFEG}} = \frac{C_{\text{el}}}{T} = \frac{\pi^2}{3}k^2G(\varepsilon_F)$$

One can measure γ (the Sommerfeld parameter) by plotting $\frac{C}{T}$ against T^2 .

Experimental results give reasonable agreement, but the theoretical values are always too high. This discrepancy is fixed by defining an effective electron mass m^* that takes in to account the complex interactions of the electron with the lattice in an abstract nature.

8.4 Electrical conductivity in reciprocal space

In the absence of an electric field there are the same number of electrons moving in every direction, and so the net current is zero.

When an E-field is applied (along the $-x$ direction for example), the Fermi sphere is shifted by an amount related to the net change in momentum of the Free Electron Gas.

The resulting Force is:

$$\vec{F} = -e\vec{E} = \frac{d\vec{p}}{dt} = \frac{\delta\vec{k}}{\delta t} = \frac{\Delta\vec{k}}{\tau}$$

Every orbital has it's k -vector changed by $\delta\vec{k} = \frac{-e\vec{E}\tau}{\hbar}$.

This creates a net current flow since more electrons now move in the $+x$ direction than the $-x$ direction, so the incremental velocity is:

$$v = \frac{\Delta k}{m} = \frac{-eE\tau}{m}$$

In a constant electric field E (Ohm's Law):

$$J = N_{el}qv = N_{el}q \left(\frac{\Delta k}{m} \right) = \frac{N_{el}e^2\tau E}{m} = \sigma E$$

so:

$$\boxed{\sigma = \frac{N_{el}e^2\tau}{m}}$$

8.5 Quantum version of the Wiedemann-Franz Law and Lorenz number

From $G(\varepsilon_F) = \frac{3N_e}{2\varepsilon_F}$ (Quantum DOS) and $C_{el} = \frac{\pi^2}{3}k^2G(\varepsilon_F)T$ one obtains the heat capacity of the quantum FEG per unit volume:

$$C_{el} = \frac{\pi^2k^2N_{el}T}{2\varepsilon_F}$$

The thermal conductivity per unit volume is:

$$K' = \frac{\pi^2l^2N_{el}\tau}{3m}T$$

$$L_{QFEG} = \frac{K}{\sigma T} = \frac{\pi^2k^2}{3e^2} = 2.45 \times 10^{-8} \left(\frac{J}{CK} \right)^2$$

Note that this is independent on N_{el} , m and τ .

The agreement with experiment is good, except at low temperatures near 10K where the actual value of L is about a factor of 10 less. This temperature dependence will be explained later.

8.6 Limitations of the Quantum FEG model

Whilst the Quantum FEG model (Drude + QM) predicts many basic properties of metals (heat capacity, electrical and thermal conductivity, WF-law), strict quantitative agreement with experimental results is not achieved.

It also does not explain why some materials are metals, some are insulators and some are semiconductors.

Which leads to the question, how do the electrons respond to an applied E-field?

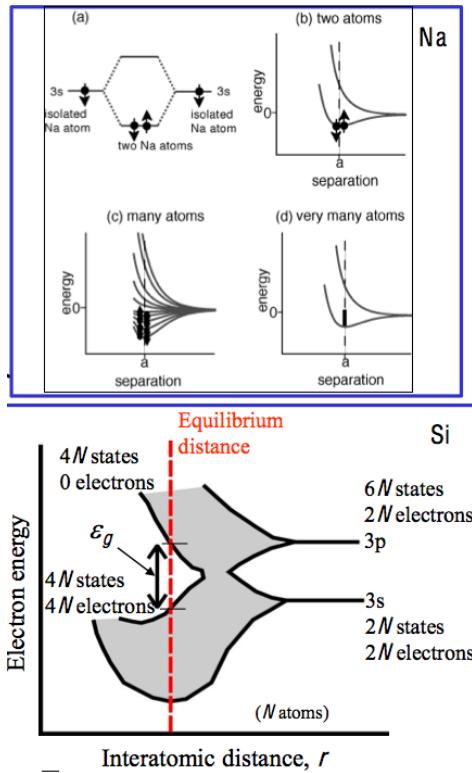
The electrons in the crystal are arranged in bands separated by gaps in which there are no allowed electron states. These energy/band gaps result from the interaction of the electron wavefunction with the periodic potential of the ion cores.

So we must take into account the periodic lattice, which leads to the concept of effective mass (although we continue to ignore electron-electron interactions so the single-electron Schrodinger equation can be used).

8.7 Idea of Energy bands and gaps

In a metal, the bonding and antibonding energy levels produce a quasi-continuum of states between the lowest and highest-filled bands. The Fermi level then lies so that the band is half-filled.

In Silicon we have different behaviour, as it is a non-metal with a band gap.



Expand this after reading

8.8 Nearly free electron gas model

The lattice periodicity has the effect that the electron wavefunctions are no longer sinusoidal travelling waves of constant amplitude, but instead exhibit the lattice periodicity in their amplitude.

The electrons can be scattered by the lattice. When the de-Broglie wavelength of the electron corresponds to a periodicity in the spacing of the ions, the electron interacts particularly strongly with the lattice Bragg Scattering.

Bragg reflection of electron waves is the cause of the energy gaps, which in turn dictate the conductivity of the material.

The model we use makes the following approximation of the potential:

$$U(x) = U_0 + U_1 \cos\left(\frac{2\pi x}{a}\right)$$

Considering the following solutions to the 1D TISE:

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(r) + U\psi(r) = \varepsilon\psi(r)$$

$$U_1 = 0 : \psi = Ae^{i(kx-\omega t)}, \varepsilon = \frac{\hbar^2 k^2}{2m}$$

Wavefunctions are plane waves and energy bands are parabolic.

$$U_1 \neq 0 : k \ll \frac{\pi}{a}$$

Electron wavelengths are much larger than a so wave functions and energy bands are nearly the same as above.

$$U_1 \neq 0 : k \leq \frac{\pi}{a}, \psi_{\pm} = Ae^{-i(kx-\omega t)} \pm Be^{-i(kx-\omega t)} B < A$$

Electrons wavelengths approach a , so waves begin to be strongly back-scattered.

$$U_1 \neq 0 : k = \frac{\pi}{a}, \psi_{\pm} = C(e^{i(kx-\omega t)} \pm e^{-i(kx-\omega t)}) = \frac{1}{\sqrt{2}}A(e^{ikx} \pm e^{-ikx})e^{-i\omega t}$$

Electron waves are strongly back-scattered (Bragg scattering) so standing waves are formed.

9 Electron energies in a Periodic Potential

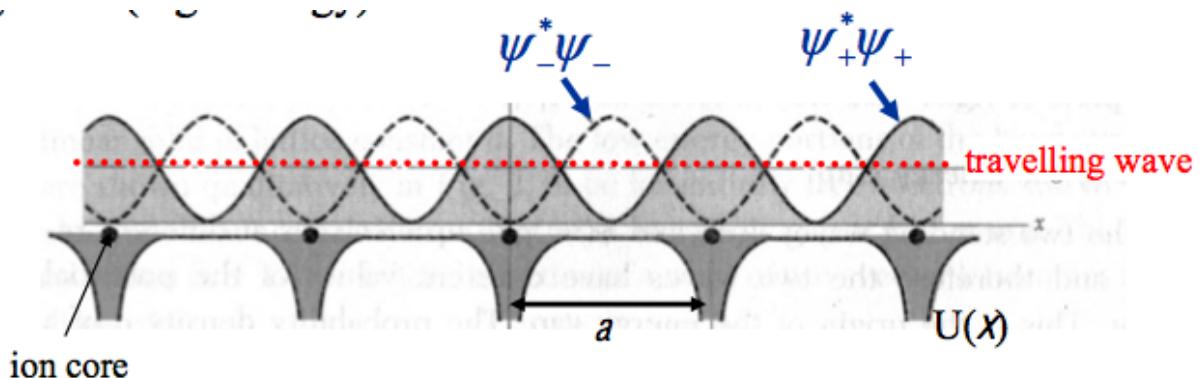
Two electron standing waves at $k = \frac{\pi}{a}$ so electrons accumulate at different regions with respect to the lattice, and so there are different potential energies with the same wavevector, so there is an energy gap.

$$\psi_+ = \frac{1}{\sqrt{2}}A(e^{ikx} + e^{-ikx})e^{-i\omega t} = \frac{1}{\sqrt{2}}2iA \cos(kx)e^{-i\omega t} \rightarrow \psi_+^* \psi_+ = \boxed{2A^2 \cos^2\left(\frac{\pi x}{a}\right)}$$

$$\psi_- = \frac{1}{\sqrt{2}}A(e^{ikx} - e^{-ikx})e^{-i\omega t} = \frac{1}{\sqrt{2}}2iA \sin(kx)e^{-i\omega t} \rightarrow \psi_-^* \psi_- = \boxed{2A^2 \sin^2\left(\frac{\pi x}{a}\right)}$$

Where the boxed equations are the electron probability densities.

$\psi_+^* \psi_+$ has its peaks at $x = 0, a, 2a, 3a, \dots$ at the positions of the atoms, where U is at its minimum (low energy wavefunction). The other solution $\psi_-^* \psi_-$ has its peaks at $x = \frac{a}{2}, \frac{3a}{2}, \frac{5a}{2}, \dots$ at positions in between atoms, thereby raising U above that seen by a travelling wave (high energy)



9.1 The energy gap

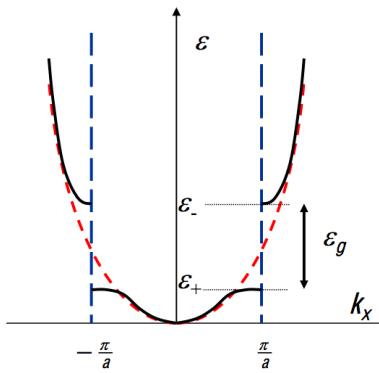
The magnitude of the gap (to first order):

$$\varepsilon_- - \varepsilon_+ = \varepsilon_g \approx \int_{x=0}^a U(x) (\psi_-^* \psi_- - \psi_+^* \psi_+) dx = -2A^2 \int_{x=0}^a U_1 \cos\left(\frac{2\pi x}{a}\right) \left(\cos^2\left(\frac{\pi x}{a}\right) - \sin^2\left(\frac{\pi x}{a}\right)\right) dx$$

where $A = \frac{1}{\sqrt{a}}$ found from normalising the wavefunction.

i.e. the gap is equal to the amplitude of the periodic potential.

In between the two energies there are no allowed energies. The periodic potential $U(x)$ splits the free-electron $\varepsilon(k)$ in to “energy bands” separated by gaps at each Brillouin Zone boundary.



9.2 Different representations of $\varepsilon(k)$

Applying periodic boundary conditions to the 1D crystal, the energy bands are invariant under a reciprocal lattice translation vector:

$$\varepsilon(k + G) = \varepsilon(k), \quad G = \frac{2\pi p_{\hat{i}}}{a}$$

This can be plotted in various schemes:

Periodic zone scheme redraw $\varepsilon(k)$ in each zone and superimpose (all curves)

Reduced zone scheme all states with $|k| > \frac{\pi}{a}$ are translated back in to the first Brillouin Zone (only curves in first BZ)

Extended zone scheme Plot $\varepsilon(k)$ from $k = 0$ through all possible BZs (thick)

Curves give phase velocity, group velocity and effective mass.

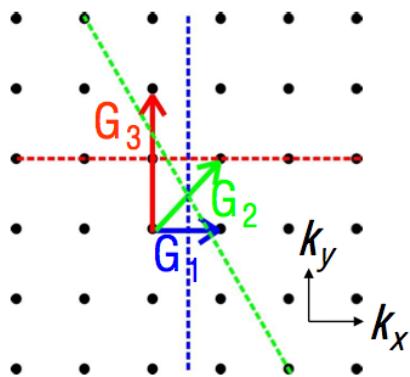
9.3 Metals, Insulators and Semiconductors

We saw previously that monovalent metals have half-filled bands, i.e. electrons can easily be excited from states just below the Fermi energy to states just above (as there are allowed states on both sides).

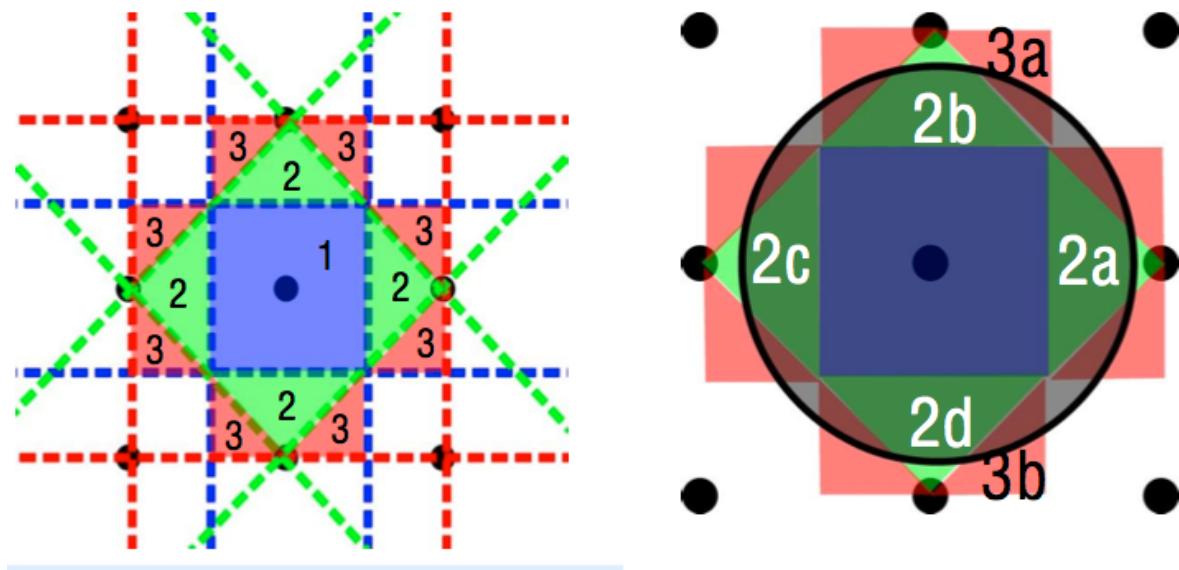
It will be shown later that semiconductors and insulators have completely filled or empty bands and an energy gap separating the highest filled and lowest unfilled band. This makes thermal excitation of electrons unlikely if the gap is large (insulators).

9.4 Fermi surfaces

- Surface of constant energy ε_F in k-space (spherical for 3D FEG)
- The material's electrical properties are determined by the shape of the Fermi surface since the currents are due to changes in the occupancy of states near the Fermi surface.
- Interactions between the electrons and the periodic potential of the lattice perturb the Fermi surface near the Brillouin Zone boundaries (QFEG)
- It is therefore necessary to first consider the construction of the Brillouin Zones in k-space. Considering a 2D square lattice: The three shortest forms of reciprocal lattice vectors (G) are shown, together with their associated perpendicular bisectors (i.e. Bragg condition)



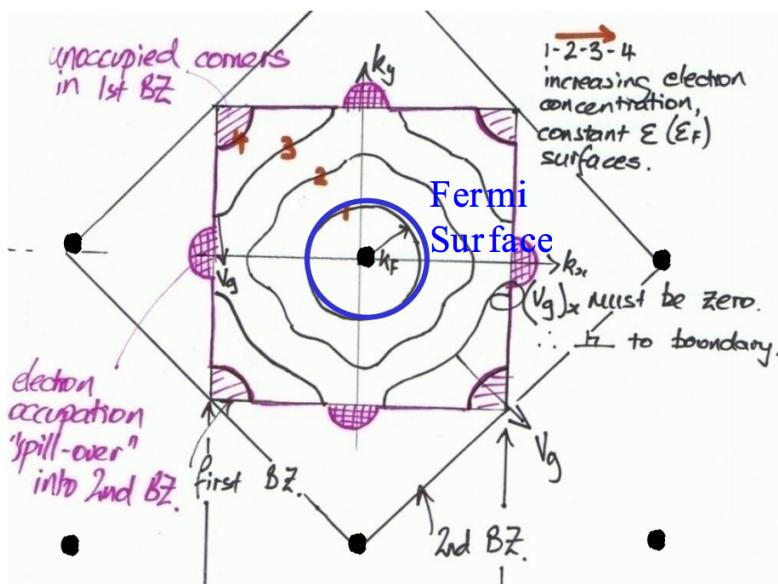
On constructing all lines equivalent by symmetry to those bisectors, we obtain the regions of k-space associated with the first three Brillouin Zones.



- On the bottom diagram the black circle represents the free electron Fermi surface for an arbitrary electron concentration.
- The total area depends on electron concentration.
- The shape of the Fermi surface depends on the interaction of electrons with the lattice.
- Plotting regions of the same BZ detached from one another is messy, so use reduced zone scheme.

k_F grows with the electron concentration n , so at higher electron concentrations the border to the Brillouin Zone perturbs the Fermi surface (contours of constant energy). Note the group velocity vector is normal to the constant energy surfaces. At the Brillouin Zone boundary the group velocity must be zero, and so the Fermi surface is perturbed as it must be perpendicular at the boundary.

In insulators there is large perturbation, so all the states in the first Brillouin Zone have an energy below any of those in the second Brillouin Zone.



10 Bloch's Theorem

Bloch's Theorem provides the mathematical form of the wavefunction in a periodic potential. For a periodic potential:

$$U(\mathbf{r} + \mathbf{R}) = U(\mathbf{r})$$

Bloch showed that the solutions to the SE are a product of the plane wave, and a function describing the periodicity of the lattice:

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

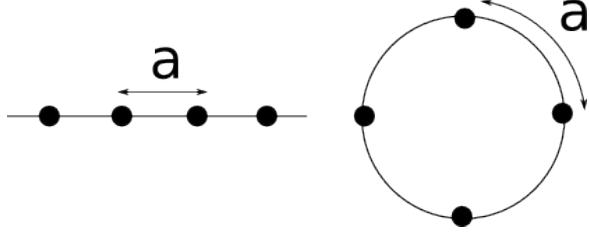
So Bloch's Theorem implies:

$$\psi_k(\mathbf{r} + \mathbf{R}) = u_k(\mathbf{r} + \mathbf{R})e^{i\mathbf{K}\cdot(\mathbf{r}+\mathbf{R})} = u_k(\mathbf{r} + \mathbf{R})e^{i\mathbf{k}\cdot\mathbf{r}}e^{i\mathbf{k}\cdot\mathbf{R}} = u_k(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}e^{i\mathbf{k}\cdot\mathbf{R}} = \psi_k(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{R}}$$

So to prove Bloch's Theorem we must prove:

$$\psi_k(\mathbf{r} + \mathbf{R}) = \psi_k(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{R}}$$

Periodic boundary conditions: Periodic boundary conditions imply:



$$\psi(x + Qa) = \psi(x)$$

Symmetry of the ring implies there is a solution:

$$\psi(x + a) = C\psi(x)$$

If there are Q lattice points, then applying this translation Q times will return to the initial position:

$$\psi(x + Qa) = C^Q\psi(x) = \psi(x)$$

So:

$$C^Q = 1, \text{ has solution: } C^Q = e^{2\pi pi} \text{ where } p = 0, \pm 1, \pm 2, \dots$$

$$C = e^{\frac{2\pi pi}{Q}} = e^{ika} \quad \therefore k = \frac{2\pi p}{Qa}$$

Now that we know C we can write:

$$\psi(x + a) = C\psi(x) = e^{ika}\psi(x)$$

This can be generalised to the 3D case:

$$\psi_k(\mathbf{r} + \mathbf{R}) = \psi_k(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{R}}$$

So by Bloch's theorem we can calculate the energy bands in a crystal if we know the form of the potential experienced by the electrons.

We define the **crystal momentum** as $\mathbf{p} = \hbar\mathbf{k}$. To calculate the energy bands we take the SE and substitute $\psi_k(\mathbf{r}) = u_k(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$:

$$\left[-\frac{\hbar^2}{2m} (\nabla + i\mathbf{k})^2 + U(\mathbf{r}) \right] u_k(\mathbf{r}) = \varepsilon_k u_k(\mathbf{r})$$

This is an eigenvalue equation and so the energy is a multivalued function of \mathbf{k} - leads to band gaps.

10.1 Bloch electrons in an E-field

We know that metals have partially filled bands, whilst insulators have completely filled bands - but what is the difference here?

For a single electron: $\mathbf{j} = -e\mathbf{v}$

For a collection of electrons: $\mathbf{J} = -e \sum_k \mathbf{v}(\mathbf{k})$

For each electron $\mathbf{v} = \frac{1}{\hbar} \nabla_{\mathbf{k}} \varepsilon(\mathbf{k})$

But symmetry of energy band requires: $\varepsilon(-\mathbf{k}) = \varepsilon(\mathbf{k})$

So $\mathbf{v}(-\mathbf{k}) = -\mathbf{v}(\mathbf{k})$

So for a filled band, which has an equal number of electrons with k positive and negative:

$$\mathbf{J} = -e \sum_{\text{1st BZ}} \mathbf{v}(\mathbf{k}) = 0$$

So filled energy bands carry no current, this is true even when an electric field is applied.

Note that the electrons are not stationary, but the net current is zero as there are an equal number moving in each direction.

In the absence of an electric field the same argument applies for partially filled bands (as the motion is random and cancels out). But with the application of an electric field the electrons all experience a force F that causes a change in their \mathbf{k} values:

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

If the electrons are in a partially filled band then the symmetry is broken and so there is a net current. Whereas in a filled band, all the \mathbf{k} vectors are changed symmetrically and so there is still no net current.

Note that the crystal momentum $\hbar\mathbf{k}$ is not the real momentum as it only takes in to account the external forces.

Looking at the effect of the external electric field in the more realistic case with scattering one obtains:

$$\frac{d\mathbf{k}}{dt} = \frac{-e\mathbf{E}}{\hbar} \approx \frac{\Delta\mathbf{k}}{\tau} \rightarrow \Delta k_x \approx \frac{-eE_x}{\hbar}\tau \quad /rightarrow \quad \Delta k_x \approx \frac{-eE_x}{\hbar}\tau$$

The steady state situation can be represented (for a nearly free electron metal) by a very small shift in the Fermi sphere.

This leads to a net current as there is no longer perfect cancellation of terms corresponding to $\pm k$ values as when $E = 0$.

$$J_x = -e \sum_{\text{1st BZ}} v_x(k) = -(g(\varepsilon_F)\Delta\varepsilon) \cdot \varepsilon \cdot v_{Fx}$$

Where $\Delta\varepsilon$ is the energy shift in the Fermi sphere, v_{Fx} is the x-component of the Fermi velocity, J_x is the x-component of the current density.

Can simplify this expression with the chain rule:

$$J_x = -(g(\varepsilon_F)\Delta\varepsilon) \cdot e \cdot v_{Fx} = -ev_{Fx}g(\varepsilon_F) \left(\frac{d\varepsilon}{dk_x} \right)_{\varepsilon_F} \Delta k_x$$

Now using the approximation of the free-electron band $\varepsilon(k)$:

$$v_g = \frac{d\omega}{dk} = \frac{1}{\hbar} \frac{d\varepsilon}{dk}, \quad \Delta k_x \approx \frac{-eE_x}{\hbar}\tau$$

$$J_x = -ev_Fxg(\varepsilon_F)(\hbar v_F) \Delta k_x = -ev_Fx^2 g(\varepsilon_F) \left(\frac{-eE_x}{\hbar} \tau_F \right) = e^2 v_F^2 \tau_F g(\varepsilon_F) E_x$$

For a spherical Fermi surface:

$$v_F^2 = v_Fx^2 + v_Fy^2 + v_Fz^2 = 3v_Fx^2 \rightarrow v_Fx^2 = \frac{1}{3}v_F^2$$

This yields:

$$J_x = \frac{1}{3}e^2 v_F^2 \tau_F g(\varepsilon_F) E_x = \sigma E_x \rightarrow \boxed{\sigma = \frac{1}{3}e^2 v_F^2 \tau_F g(\varepsilon_F)}$$

This result reduces to the free-electron expressio when the FEG values of $g(\varepsilon_F)$ and ε_F are substituted (**add this later**).

But this equation is much more general and highlights the importance of the density of states and the mean free parth rather than the electron concentration.

11 Effective Mass of an Electron

We can use our previous results to write the equation of a motion of a Bloch electron in 1D:

$$a_x = \frac{dv_x}{dt} = \frac{dv_x}{dk_x} \frac{dk_x}{dt} \frac{dv_x}{dk_x} = \frac{dv_g}{dk_x} = \frac{d}{dk_x} \left(\frac{1}{\hbar} \frac{d\varepsilon}{dk_x} \right) = \frac{1}{\hbar} \frac{d^2\varepsilon}{dk_x^2} \text{ and } \frac{dk_x}{dt} = \frac{F_x}{\hbar}$$

This gives:

$$a_x = \frac{1}{\hbar} \frac{d^2\varepsilon}{dk_x^2} \left(\frac{F_x}{\hbar} \right)$$

Or in the form of Newton's 2nd Law:

$$F_x = \frac{\hbar^2}{\frac{d^2\varepsilon}{dk_x^2}} a_x$$

This allows us to define an effective mass:

$$m^* = \frac{\hbar^2}{\frac{d^2\varepsilon}{dk_x^2}} \text{ or } \boxed{\frac{1}{m^*} = \frac{1}{\hbar^2} \frac{d^2\varepsilon}{dk_x^2}}$$

In contrast to the free-electron mass, the effective mass varies depending on the electron energy - i.e. it varies depending upon the location of the electron in the band.

The concept of the effectove mass allows us to retain the notion of a free-electron even when there is a periodic potential, as the effective mass m^* takes account of the effect of the lattice on the acceleration of the electron.

Note that the effective mass is inversely proportional to the curvature of the band.

11.1 How can an electron of mass m_0 respond to an applied field as if its mass were m^* ?

- At the bottom of the band, the electron is adequately described as a plane wave. On approach to the BZ boundary, momentum transfer from the lattice (Bragg reflection) becomes quite large. At the boundary it becomes equal to the forward component and standing waves are formed.
- Positive m^* occurs near the bottom the band (upward curvature). Negative m^* means that on going from k to $k + \Delta k$ the momentum transfer from the lattice is larger than the momentum transfer to the electron from the applied force. Although k is increased by Δk by the applied field, the approach to Bragg reflection gives an overall *decrease* in its forward momentum.
- Just below a BZ boundary the energy of the electron is too low to satisfy the Bragg reflection condition and the beam is transmitted. But applying a voltage across the crystal increases k so the Bragg condition is satisfied and the beam is reflected (described as negative m^*).

12 Electrons and Holes

Consider the lower band during the excitation of an electron from a filled lower band in to an unfilled upper band. When the lower band was filled the net electron wavevector was zero, that is:

$$\sum_i^{\pm N} k_i = 0 \text{ or } \left(\sum_{i \neq +j}^{\pm N} k_i \right) + k_j = 0 \text{ or } \sum_{i \neq +j}^{\pm N} k_i = -k_j = k_{-j}$$

As the band is symmetric so $k_{-j} = -k_j$.

Similarly we now analyse the current flow in the incomplete band under the influence of a field E .

For a complete band:

$$J = \sum_i^{\pm N} (-ev_i) = 0, \quad J = \sum_{i \neq +j}^{\pm N} (-ev_i) - ev_j = 0$$

This shows that an incomplete band (with state $+j$ empty) behaves just like a positive charge moving with the same velocity as an electron would have in that state.

So the properties of all of the remaining electrons in the incomplete band are equivalent to those of the vacant state j if the vacant state has:

- A k -vector $-k_j$ (i.e. $k_h = -k_e$)
- A velocity v_{+j} (i.e. $v_h = v_e$)
- A positive charge $+e$ (i.e. $q_h = +e$)

If this hole is accelerated in an applied electric field then:

$$m_h^* \frac{dv_h}{dt} = +eE$$

Whilst for an electron the equation is:

$$m_e^* \frac{dv_e}{dt} = -eE$$

But we've deduced that the hole velocity is the same as that of the missing electron:

$$v_h = v_e$$

By equating the derivatives:

$$\frac{eE}{m_h^*} = -\frac{eE}{m_e^*} \rightarrow \boxed{m_h^* = -m_e^*}$$

However, note that near the top of a band, where are most likely, the band curvature is negative so the effective electron mass is also negative. This means the corresponding hole mass is positive.

So the equation of motion of a hole in an electromagnetic field is:

$$F = \hbar \frac{dk_h}{dt} = e(\mathbf{E} + v_h \times \mathbf{B})$$

This explains why some metals have positive Hall coefficients.

To summarise:

- Effective mass, m^* , relates quantum mechanical results to classical force equations.
- In most instances, the electron in the bottom of a band can be thought of as a classical Newtonian particle provided the internal forces and QM properties are taken in to account through m^* .
- An electron near the top of an allowed energy band behaves as if it has negative mass.
- States of negative m^* occur near the top of the band on approach to the BZ boundary.
- We introduce the concept of holes to avoid negative m^* .
- **It has a positive effective mass m_h^* and a positive electronic charge so it moves in the same direction as the applied field.**

12.1 Charge carriers in Silicon

At $T = 0K$, each Si atom is surrounded by 8 valence electrons (4 of its own and one contributed from each of its 4 neighbours) **in their lowest state** and involved in covalent bond.

All the electrons are in their lowest state and so the lower band is filled and the upper band is empty, i.e. it is an **insulator**.

As T is increased, some electrons gain enough energy to break their bonds and jump from the highest-energy filled band to the lowest-energy unfilled band.

Note that the hole can move through the lattice just like an electron. **Electrons (in the CB) and holes (in the VB) are equally important charge carriers and both contribute towards the current.**

13 Insulators, Metals and Semiconductors

In **insulators**:

- The valence electrons completely fill one or more bands, leaving the higher bands empty.
- The filled band is separated by a large energy gap so that electrons cannot be thermally excited to the upper band to produce electron-hole pairs i.e. $\varepsilon_g \approx 5\text{eV}$
- There is no way to change the total momentum of the electrons if every accessible state is filled, so nothing changes even when an E-field is applied, **no current flows**.

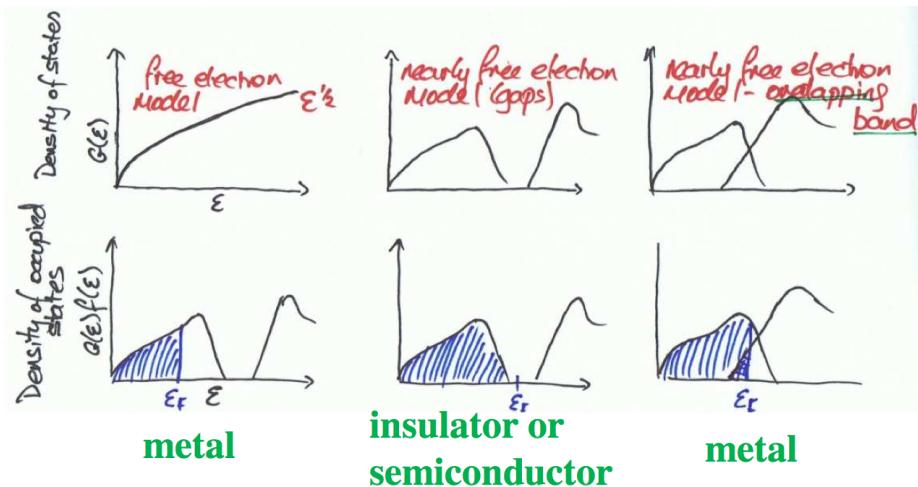
Note that crystals can **only** be insulators if the number of valence electrons in the primitive cell is an even integer. But note that this does not mean all divalent elements are insulators as they can have overlapping band gaps.

Metals have either:

- Partially filled bands: i.e. *Alkali* metals (group 1 e.g. Li, Na, K) and the *noble* metals (Cu, Ag, Au, Pd, Pt) have one valence electron per primitive cell and so the band is half-filled and can electrons are easily excited above the Fermi energy.
OR
- Overlapping bands: Alkaline earth metals (group 2 e.g. Be, Mg, Ca) have two valence electrons per primitive cell and could be insulators, but the bands overlap to give metals (though not very good ones).
- In both cases there are a large number of electrons and a large number of empty energy states into which the electrons can move. Very little energy is required to find new energies for the electrons to occupy, i.e. for them to move and create a current, this gives a low resistivity.

To summarise:

Semiconductors are classified by their electrical resistivity at room temperature, in the range $10^{-4} < \rho < 10^7 \Omega\text{m}$.



I.e. in Silicon and Germanium there are eight valence electrons per primitive cell and the bands do not overlap. So at $T = 0$ it acts as an insulator. As T is increased, electrons are thermally excited from the valence band to the conduction band, creating electron-hole pairs.

Note that both the electrons in the CB and holes in the VB contribute to the electrical conductivity (and so n and ρ are determined by ε_g/kT).

I.e. for Silicon $10^{10} < n < 10^{13} \text{ cm}^{-3}$ between 300K and 400K, and is **strongly temperature dependent** ($\varepsilon_g \approx 1 \text{ eV}$).

Note that in 3D crystals the electrons travelling in different directions in a crystal not only encounter different k -space boundaries but also different potential patterns. Therefore $\varepsilon - k$ dispersion diagrams are plotted as a function of the k -space direction.

13.1 Direct Bandgap Semiconductors (GaAs)

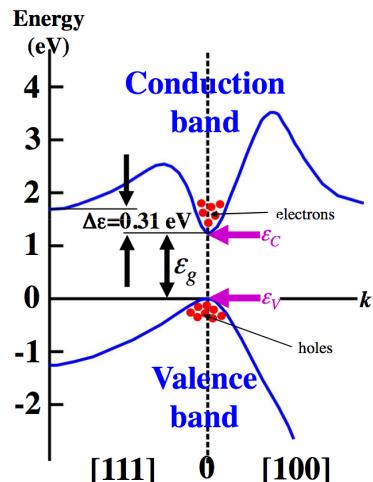


Figure 3: Simplified diagram of GaAs band structure

- The VB maximum (valence band edge, ε_V) and CB minimum (conduction band edge, ε_C) both occur at $k = 0$ for GaAs, so all charge carriers congregate at $k = 0$.

- Direct bandgap semiconductors have this property that transitions between the allowed bands can occur with no change in crystal momentum, that is **the holes congregate directly above where the electrons congregate** - the same k value.
- Makes these materials suited for semiconductor lasers.
- Note GaAs is a **compound (III-V) semiconductor** as it is made of two or more elements.

13.2 Indirect Bandgap Semiconductors (Si)

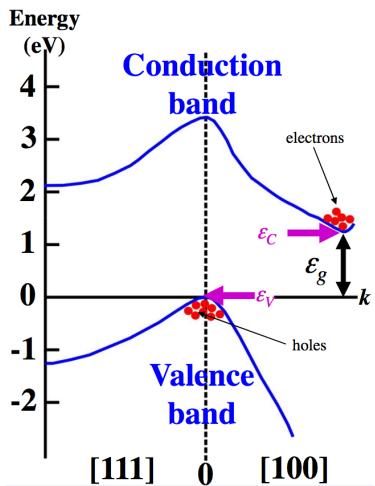


Figure 4: Simplified diagram of Si band structure

- Here the minimum of the CB does not align with the maximum of the VB.
- To make the transition an interaction with phonons must occur so crystal momentum is conserved.
- Note Si is an **elemental semiconductor**, as each atom is of the same type.

13.3 Using the effective mass concept

- Near the bottom of the CB and top of the VB, a parabolic approximation is effective.
- So at bottom of CB can write:

$$\varepsilon = \varepsilon_C + \frac{\hbar^2 k^2}{2m_e^*}$$

where the free electron mass is replaced by the effective mass.

- The corresponding density of allowed energy states in the conduction band per unit volume is:

$$g_C(\varepsilon) = \frac{4\pi (2m_e^*)^{3/2}}{h^3} \sqrt{\varepsilon - \varepsilon_c}$$

- Similarly, the DOS in the VB can be obtained by treating the hole as a “free” particle with effective mass:

$$\varepsilon = \varepsilon_V - \frac{\hbar^2 k^2}{2m_h^*} \xrightarrow{\varepsilon \leq \varepsilon_V} g_V(\varepsilon) = \frac{4\pi(2m_h^*)^{\frac{3}{2}}}{h^3} \sqrt{\varepsilon_V - \varepsilon}$$

13.4 Determination of effective mass

The effective mass can be experimentally determined using the cyclotron resonance technique: when placed in a static magnetic field the electrons move on circular or spiral orbits on a constant energy surface around the axis of the field with an angular frequency given by:

$$\omega_c = \frac{Be}{m_e^*}$$

We actually observe different m_e^* values in different directions but usually a statistical average is good for calculations.

14 The semiconductor in equilibrium

The difference between a metal and an insulator at $T = 0$ was defined as whether the DOS is zero or finite. To define semiconductors the idea is that the bandgap is sufficiently small such that thermally excited carriers are possible at reasonable temperatures.

We assume equilibrium, i.e. there are no voltages, fields or temperature gradients and combine the DOS with the Fermi-Dirac distribution function to yield the electron and hole concentrations.

With doping, the type of dopant dictates whether the dominant charge carrier is an electron or hole, and the Fermi Energy is a function of the type and concentration of dopant atoms.

14.1 Equilibrium distribution of electrons and holes in an Intrinsic Semiconductor

Electron population distribution in the conduction band:

$$n(\varepsilon) = g_C(\varepsilon) f_{FD}(\varepsilon)$$

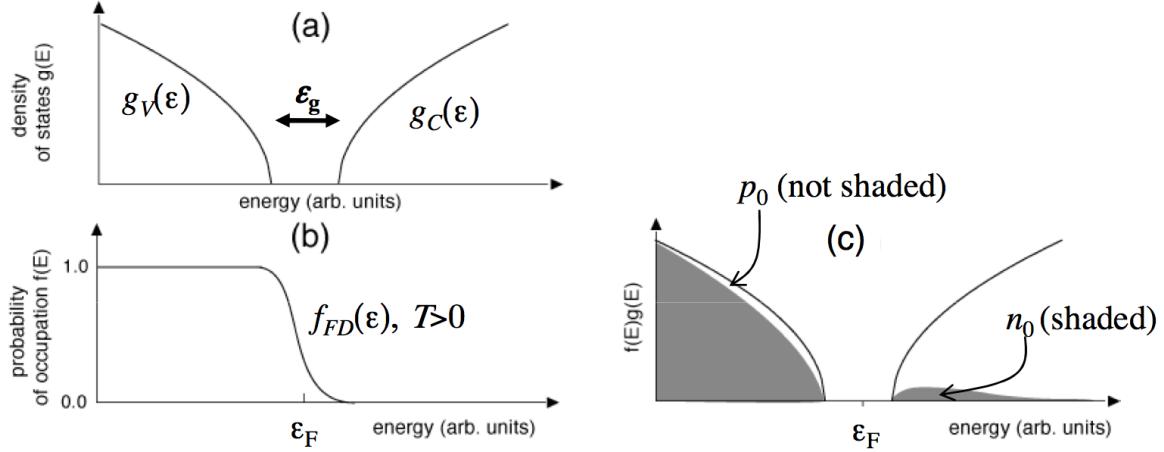
Hole population distribution in the valence band:

$$p(\varepsilon) = g_V(\varepsilon) [1 - f_{FD}(\varepsilon)]$$

Assuming m_e^* and M_h^* to be equal means that $g_C(\varepsilon)$ and $g_V(\varepsilon)$ are symmetrical about the mid-gap energy $\frac{\varepsilon_C + \varepsilon_V}{2}$. This means that the Fermi Energy ε_F lies at the middle of the gap to maintain net charge neutrality.

In intrinsic semiconductors, electrons and holes are created in pairs by thermal energy so the number of holes in the valence band is **always equal** to the number of electrons in the conduction band: $n_0 = p_0$

But if the effective masses of the electrons and holes are not equal then $g_C(\varepsilon)$ and $g_V(\varepsilon)$ are not symmetrical about the midgap energy. This means that the Fermi level of the intrinsic



semiconductor will shift from the midgap energy in order to maintain equal electron and hole concentrations $n_0 = p_0$.

The thermal equilibrium density of electrons is given by integrating over the conduction band:

$$n_0 = \int_{\varepsilon_C}^{\infty} g_C(\varepsilon) f_{FD}(\varepsilon) d\varepsilon$$

Here the energy limits are the bottom of the conduction band, ε_C and infinity.

Since $(\varepsilon_C - \varepsilon_F) \gg kT$ we can use the Maxwell Boltzmann approximation to the Fermi Dirac probability function (density of states in CB in blue, MB probability function in red):

$$n_0 = \int_{\varepsilon_C}^{\infty} \frac{4\pi (2m_e^*)^{\frac{3}{2}}}{h^3} \sqrt{\varepsilon - \varepsilon_C} \cdot \exp\left[\frac{-(\varepsilon - \varepsilon_F)}{kT}\right] d\varepsilon$$

We solve the integral by making a change of variable, $\eta = \frac{\varepsilon - \varepsilon_C}{kT}$ and using the standard integral (gamma function) of form:

$$\int_0^{\infty} \exp(-\eta) d\eta = \frac{\sqrt{\pi}}{2}$$

Obtaining:

$$n_0 = 2 \left(\frac{2\pi m_e^* k T}{h^2} \right)^{\frac{3}{2}} \exp\left[\frac{-(\varepsilon_C - \varepsilon_F)}{kT}\right]$$

We then write the thermal equilibrium electron concentration in the conduction band as:

$$n_0 = N_C \exp\left[\frac{-(\varepsilon_C - \varepsilon_F)}{kT}\right]$$

where the parameter $N_C = 2 \left(\frac{2\pi m_e^* k T}{h^2} \right)^{\frac{3}{2}}$ is called the **effective density of states** in the conduction band. Assuming $m_e^* = m_e$ then at $T = 300K$, $N_C = 2.5 \times 10^{19} \text{ cm}^{-3}$ which is the order of magnitude for most semiconductors.

Similarly, for holes:

$$p_0 = \int g_V(\varepsilon) [1 - f_{FD}(\varepsilon)] d\varepsilon = \int_{-\infty}^{\varepsilon_V} \frac{4\pi (2m_h^*)^{\frac{3}{2}}}{h^3} \sqrt{\varepsilon_V - \varepsilon} \exp\left[\frac{-(\varepsilon_F - \varepsilon)}{kT}\right] d\varepsilon$$

This time we make the change of variable $\eta' = \frac{\varepsilon_V - \varepsilon}{kT}$.

The thermal equilibrium concentration of holes in the valence band can then be written as:

$$p_0 = N_V \exp\left[\frac{-(\varepsilon_F - \varepsilon_V)}{kT}\right]$$

where N_V is the effective density of states in the valence band, and is:

$$N_V = 2 \left(\frac{2\pi m_h^* kT}{h^2} \right)^{\frac{3}{2}}$$

The magnitude of N_V is also $\approx 10^{19} \text{ cm}^{-3}$ at 300K for most semiconductors.

Note the $T^{\frac{3}{2}}$ dependence of N_V and N_C

So if a question gives you N at 300K but asks about what happens at 200K you must do:

$$N_{200K} = \left(\frac{200}{300} \right)^{\frac{3}{2}} N_{300K}$$

Note that N_C and N_V are constant for a given semiconductor at a fixed temperature.

In an intrinsic material n_i and p_i are used to denote the intrinsic electron and hole concentrations. The **intrinsic Fermi level** is denoted by ε_{Fi} .

So the earlier equations can be written for an intrinsic material as:

$$p_0 = p_i = n_i = N_V \exp\left[\frac{-(\varepsilon_{Fi} - \varepsilon_V)}{kT}\right], \quad n_0 = n_i = N_C \exp\left[\frac{-(\varepsilon_C - \varepsilon_{Fi})}{kT}\right]$$

So we can write:

$$n_i^2 = N_C N_V \exp\left[\frac{-(\varepsilon_C - \varepsilon_V)}{kT}\right] \rightarrow n_i^2 = N_C N_V \exp\left[\frac{-\varepsilon_g}{kT}\right]$$

For a given material at a constant temperature, the value of n_i is constant and independent of the Fermi energy

For an intrinsic semiconductor $n_i = p_i$ so setting the boxed equations equal to each other:

$$N_C \exp\left[\frac{-(\varepsilon_C - \varepsilon_{Fi})}{kT}\right] = N_V \exp\left[\frac{-(\varepsilon_{Fi} - \varepsilon_V)}{kT}\right]$$

Taking natural logs of both sides and solving for ε_{Fi} :

$$\varepsilon_{Fi} = \frac{1}{2}(\varepsilon_C + \varepsilon_V) + \frac{1}{2}kT \ln\left(\frac{N_V}{N_C}\right)$$

Using the definitions of N_C and N_V and substituting for the mid-gap energy, $\varepsilon_{\text{midgap}} = \frac{\varepsilon_C + \varepsilon_V}{2}$, obtains:

$$\boxed{\varepsilon_{Fi} - \varepsilon_{\text{midgap}} = \frac{3}{4}kT \ln \left(\frac{m_h^*}{m_e^*} \right)}$$

The important consequences of this equation are:

- If $m_h^* = m_e^*$ then $\varepsilon_{Fi} = \varepsilon_{\text{midgap}}$
- If $m_h^* > m_e^*$ then the Fermi level moves above the midgap energy and vice versa. This is because the DOS equation is directly related to the carrier effective mass - i.e. a larger m^* means a larger DOS function. And so the intrinsic Fermi level must move away from the band with the larger density of states in order to maintain an equal number of holes and electrons.
- The deviation of the Fermi level from the midgap is linear with T .

14.2 Dopants

A doped semiconductor is called an **extrinsic** semiconductor. There are two types of dopants, donors and acceptors.

14.3 Acceptors

Acceptor impurities accept an electron from the lattice. An example is a group 3 element Boron, where all 3 of the electrons are involved in covalent bonding and one bonding position is empty. An electron occupying this state would make the net charge of the B atom negative and so takes a small amount of energy (but far less than the conduction band energy), valence electrons gain a small amount of thermal energy and move about the crystal, this vacancy on the B atom is then occupied and other valence electron positions become empty. So the hole is effectively moving through the lattice, whilst the negatively charged B atom is fixed in the crystal.

- Acceptor impurities create holes in the valence band without producing electrons in the conduction band.
- A fixed negatively charged ion is created in the impurity band.
- An acceptor impurity atom has less valence electrons than its host material.
- Acceptor doped semiconductors are referred to as **p-type** semiconductors due to the domination of holes.

14.4 Donors

- Donor impurities donate an electron, as the energy required to elevate the donor electron in to the conduction band is much less than that of the electrons involved in covalent bonding.
- It can then move through the crystal and generate a current, whilst a fixed positively charged ion is left behind in the impurity band. E.g. Phosphorous group V

- Donor impurity atoms add electrons to the conduction band without creating holes in the valence band.
- Donor doped materials are referred to as **n-type** semiconductors.
- Donor impurity atoms have more electrons than their host material.

14.5 Ionization Energy

We want to calculate the distance of the electron from the donor impurity atom, and the energy required to elevate the donor electron into the conduction band. So $\varepsilon_C - \varepsilon_d$ is the ionization energy.

We use the Bohr model of the atom since we have an electron orbiting a donor ion. Set the Coulomb force of attraction equal to the centripetal force. κ is the permittivity of the semiconductor ($\kappa = \kappa_0 \kappa_r$), where κ_0 is the permittivity of free space and κ is the relative permittivity (as we are already using ε for energy) and r is the orbit radius:

$$\frac{e^2}{4\pi\kappa r^2} = \frac{m_e^* v^2}{r}$$

Assume the angular momentum is quantized, where n is a positive integer:

$$m_e^* r v = n \hbar$$

Solving for the radius, also quantised:

$$r_n = \frac{n^2 \hbar^2 4\pi \kappa}{m_e^* e^2}$$

The Bohr radius is defined as:

$$a_0 = \frac{4\pi \kappa_0 \hbar^2}{m_e e^2} = 0.053 \text{ nm}$$

Normalising the radius of the donor orbital to that of the Bohr radius:

$$\frac{r_n}{a_0} = n^2 \kappa_r \left(\frac{m_e}{m_e^*} \right)$$

Considering the lowest energy state, $n = 1$ for silicon ($\kappa_r = 11.7$ and $m_e^* = 0.26 m_e$):

$$\frac{r_1}{a_0} = 45 \approx 5 \text{ lattice constants}$$

so the orbiting donor electron encompasses many Si atoms and is not tightly bound. The Kinetic Energy of the electron is:

$$T = \frac{1}{2} m_e^* v^2$$

Substituting the velocity $m_e^*rv = n$ and for $r_n = \frac{n^2\hbar^2 4\pi\kappa}{m_e^* e^2}$:

$$T = \frac{m_e^* e^4}{2(n\hbar)^2(4\pi\kappa)^2}$$

The Potential energy is (Coulomb attraction):

$$V = \int_{R=r}^{\infty} \frac{e^2}{4\pi\kappa R^2} dR = \frac{-e^2}{4\pi\kappa r_n} = \frac{-m_e^* e^4}{(n\hbar)^2(4\pi\kappa)^2}$$

Total energy $T + V$ is therefore:

$$\varepsilon = \frac{-m_e^* e^4}{2(n\hbar)^2(4\pi\kappa)^2}$$

Group IV elements (i.e. silicon and germanium) are called *amphoteric* impurities since they can act as either donors or acceptors depending upon the material they are in.

14.6 Equilibrium distribution of charge carriers in the extrinsic semiconductor

The charge carrier distribution changes in an extrinsic semiconductor because you can add donor electrons without producing holes, or vice versa, changing the probability distribution dramatically. Adding donor impurities moves the Fermi level towards the conduction band, whilst adding acceptor impurities moves the Fermi level towards the valence band.

Must derive an alternate form of the thermal equilibrium concentration equations, add and subtract an *intrinsic* Fermi energy in the exponential of the equation for n_0 :

$$n_0 = N_C \exp \left[\frac{-(\varepsilon_C - \varepsilon_{Fi}) + (\varepsilon_F - \varepsilon_{Fi})}{kT} \right] = \textcolor{red}{N_C} \exp \left[\frac{-(\varepsilon_C - \varepsilon_{Fi})}{kT} \right] \exp \left[\frac{(\varepsilon_F - \varepsilon_{Fi})}{kT} \right]$$

But red term is *intrinsic* carrier concentration n_i so we write:

$$n_0 = n_i \exp \left[\frac{\varepsilon_F - \varepsilon_{Fi}}{kT} \right]$$

And similarly for holes:

$$p_0 = n_i \exp \left[\frac{\varepsilon_{Fi} - \varepsilon_F}{kT} \right]$$

Taking the product of the general expressions for n_0 and p_0 obtains:

$$n_0 p_0 = N_C N_V \exp \left[\frac{-(\varepsilon_C - \varepsilon_F)}{kT} \right] \exp \left[\frac{-(\varepsilon_F - \varepsilon_V)}{kT} \right]$$

$$n_0 p_0 = N_C N_V \exp \left[\frac{-\varepsilon_g}{kT} \right]$$

This equation is exactly the same as that derived for an intrinsic material and so we have:

$$n_0 p_0 = n_i^2$$

known as the **Law of mass action**

This equation states that the product $n_0 p_0$ is always constant for a given material at a given temperature and is a fundamental principle of semiconductors in thermal equilibrium. Note that it is only valid when the Boltzmann approximation is valid.

14.7 Degenerate and Nondegenerate semiconductors

- So far we have assumed that the concentration of dopant atoms is small compared to the density of host atoms - i.e. there is no interaction between donor electrons in n-type material.
- We've assumed that the impurities introduce discrete energy states, these are referred to as *nondegenerate* semiconductors.
- If the impurity concentration increases, then the donor electrons will eventually interact. The single discrete donor energy ε_d will split into a band of energies which may overlap with the bottom of the conduction band when the donor concentration N_d is on the order of the effective density of states N_C .
- When $N_d > N_C$ then ε_F lies within the conduction band and we have a **degenerate** n-type semiconductor.
- Similarly, acceptor doping can lead to a degenerate p-type material. Energy states below ε_F are mostly filled i.e. in the n(p)-type material most of the states between ε_C (ε_V) and ε_F are occupied (empty) so the electron (hole) concentration in the conduction (valence) band is very large.

15 Statistics of donors and acceptors

- The electron concentration in the donor states is given by:

$$n_d = N_d \times f_d(\varepsilon)$$

where N_d is the concentration of donor atoms and the probability distribution function based on Fermi-Dirac is:

$$f_d(\varepsilon_d) = \frac{1}{1 + \frac{1}{2} \exp\left(\frac{\varepsilon_d - \varepsilon_F}{kT}\right)}$$

- Note the new factor of $\frac{1}{2}$ this is because each donor level has two possible spin orientations for the donor electron, i.e. every donor has two quantum states.
- However the insertion of an electron into one of the states satisfies the vacancy requirement of the atom, and the insertion of a second electron is forbidden.
- This factor is sometimes written as $\frac{1}{g}$ where g is the degeneracy factor.
- So the concentration of electrons occupying the donor level (as opposed to being excited into the CB) can be written:

$$n_d = \frac{N_d}{1 + \frac{1}{2} \exp\left(\frac{\varepsilon_d - \varepsilon_F}{kT}\right)} = N_d - N_d^+$$

- and for holes:

$$p_a = \frac{N_a}{1 + \frac{1}{4} \exp\left(\frac{\varepsilon_F - \varepsilon_a}{kT}\right)} = N_a - N_a^+$$

where N_a^+ is the concentration of ionised donors, N_a is the concentration of acceptor atoms, ε_a is the acceptor energy level, p_a is the concentration of holes in the acceptor states and N_a^- is the concentration of ionised acceptors.

- i.e. a hole in the acceptor state corresponds to an acceptor atom that is neutrally charged and still has an empty bonding position.
- Note the factor of $\frac{1}{4}$ in the expression for p_a (applying to Si and GaAs) due to the complex valence band structure.

15.1 Complete Ionisation

If $(\varepsilon_d - \varepsilon_F) \gg kT$ (i.e. for non-degenerate semiconductors we use the MB approximation) then:

$$n_d \approx \frac{N_d}{\frac{1}{2} \exp\left(\frac{\varepsilon_d - \varepsilon_F}{kT}\right)} = 2N_d \exp\left[\frac{-(\varepsilon_d - \varepsilon_F)}{kT}\right]$$

The Boltzmann approximation is also valid for the electrons in the conduction band so that:

$$n_0 = N_C \exp\left[\frac{-(\varepsilon_C - \varepsilon_F)}{kT}\right]$$

So:

$$\frac{n_d}{n_d + n_0} = \frac{1}{1 + \frac{N_C}{2N_d} \exp\left[\frac{-(\varepsilon_C - \varepsilon_d)}{kT}\right]}$$

Which is the relative number of electrons in the donors state compared to the total number of electrons - note it is independent of ε_F . Where $(\varepsilon_C - \varepsilon_d)$ is the ionisation of the donor electrons.

15.2 Complete Ionisation and Freeze-out

- Using this for Phosphorus doping ($N_D \approx 10^{16} \text{ cm}^{-3}$) in Silicon at 300K, there are very few electrons (0.4%) in the donor state compared with the conduction band.
- So the donor states are essentially completely ionised at room temperature, and so almost all donor impurity atoms have donated an electron to the conduction band.
- Similarly, at room temperature there is complete ionisation of acceptor atoms (i.e. each has accepted an electron from the valence band, so $p_a = 0$)
- At absolute zero, the complete opposite happens: all the electrons are in their lowest state, i.e. for an n-type material, each donor state contains an electron.

- So at $T = 0K$, $n_d = N_d$ or $N_d^+ = 0$ and so from the previous expression for n_d :

$$n_d = \frac{N_d}{1 + \frac{1}{2} \exp\left(\frac{\varepsilon_d - \varepsilon_F}{kT}\right)} = N_d \text{ at } T=0 \quad \therefore \quad \exp\left[\frac{(\varepsilon_d - \varepsilon_F)}{kT}\right] = 0$$

- Since $T=0K$, this occurs in an n-type material when:

$$\left[\frac{(\varepsilon_d - \varepsilon_F)}{kT}\right] = -\infty \quad \text{i.e. } \varepsilon_F > \varepsilon_d$$

- For p-types at $T=0K$, the impurity atoms will not contain any electrons so the Fermi energy is below the acceptor energy level.
- A detailed analysis shows that at $T=0K$, ε_F is positioned halfway between ε_C and ε_d for n-type and half-way between ε_V and ε_a for p-type materials.

- **Freeze-out:** when no electrons from the donor state are thermally elevated in to the conduction band *or* when no electrons from the valence band are elevated in to the acceptor state.
- Between absolute zero and room temperature there is partial ionisation.

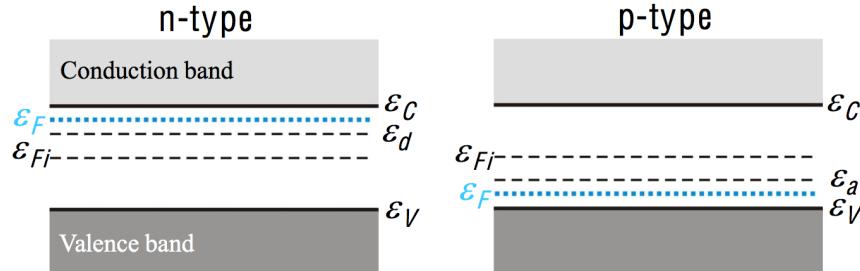
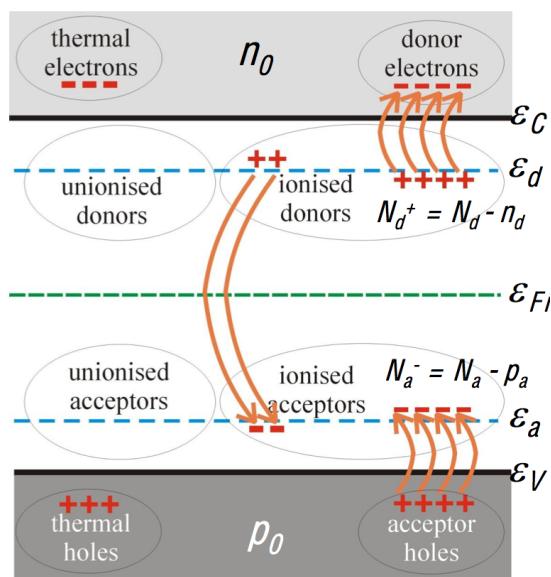


Figure 5: Energy level diagrams at T=0K

15.3 Compensated Semiconductors

- Compensated Semiconductors contain both donor and acceptor impurities in the same region i.e. formed by diffusing acceptor impurities in to an n-type material.
- The compensated conductor is n-type if $N_d > N_a$ and p-type if $N_a > N_d$.
- When $N_d = N_a$ the material is completely compensated and has the characteristics of an intrinsic semiconductor. Recall that n_d is the concentration of electrons in the donor state, p_a is the concentration of holes (or empty bonding positions) in the acceptor states.



15.4 Thermal equilibrium electron and hole concentrations

- In thermal equilibrium, the semiconductor crystal is electrically neutral. So for the compensated material:

$$n_0 + N_a^- = p_0 + N_d^+ \rightarrow n_0 + (N_a - p_a) = p_0 + (N_d - n_d)$$

- Assuming complete ionisation: $n_d = 0$, $p_a = 0$ and expressing p_0 as $\frac{n_i^2}{n_0}$ from the law of mass action:

$$n_0 + N_a = \frac{n_i^2}{n_0} + N_d \rightarrow n_0 = \frac{(N_d - N_a)}{2} + \sqrt{\left(\frac{N_d - N_a}{2}\right)^2 + n_i^2}$$

- Note we take the positive root since n_0 must be positive, valid for n-type semiconductor where $N_d > N_a$.
- Similarly for p-type semiconductors when $N_a > N_d$:

$$p_0 = \frac{(N_a - N_d)}{2} + \sqrt{\left(\frac{N_a - N_d}{2}\right)^2 + n_I^2}$$

- Note these equations **apply only for the majority charge carrier concentration** and are valid only for extrinsic materials i.e. at room temperature where we can assume complete ionisation. To calculate minority carrier concentrations first calculate the majority concentration using one of these equations and then use the law of mass action.
- If $N_d \gg n_i$ then $n_0 \approx N_d$ and the thermal equilibrium majority and minority carrier concentrations differ by many orders of magnitude.
- For an n-type semiconductor, the concentration of electrons in the conduction band (majority carrier) increases above the intrinsic value n_i as the donor impurity concentration is increased. At the same time the minority (hole) concentration decreases below n_i - why?

16 Redistribution of charge carriers

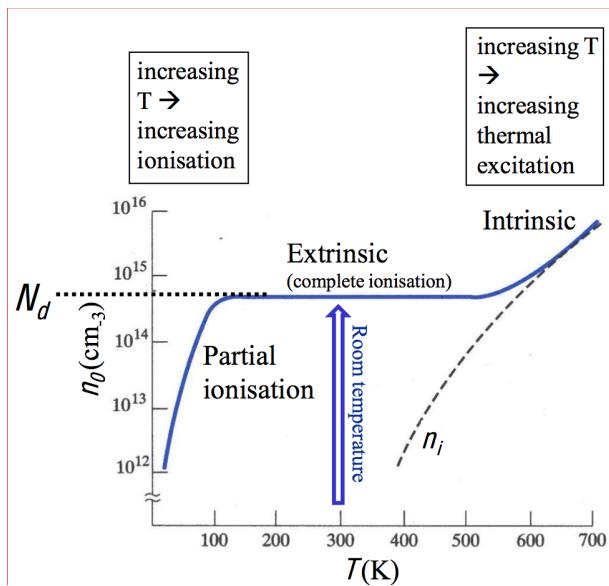
- As donors electrons are added there is a redistribution of electrons among the available energy states.
- A few of the donor electrons fall in to the empty states in the valence band and annihilate some of the intrinsic holes.
- And so the minority carrier concentration will fall.

- Because of this redistribution the number of electrons in the conduction band at complete ionisation is **not** simply equal to the donor concentration plus the intrinsic electron concentration.
- Similarly, adding acceptor impurities reduces the intrinsic electron concentration.
- Notice that if donor impurity concentration is not too different from intrinsic carrier concentration then the thermal equilibrium majority carrier electron concentration is influenced by the intrinsic concentration.

16.1 Electron concentration and temperature

- The intrinsic carrier concentration, n_i , is a strong function of temperature. So as T rises, n_i will begin to dominate.
- At high T , **the semiconductor will eventually lose its extrinsic characteristics**.
- At room temperature, all the impurities are ionised so $n_0 = N_d$
- Notice the onset of freeze-out at low T when none of the impurities are ionised so $n_d = N_d$ so the concentration of electrons in the impurity band is equal to the concentration of impurities.
- Plotting the equation for n_0 :

$$n_0 = \frac{(N_d - N_a)}{2} + \sqrt{\left(\frac{N_d - N_a}{2}\right)^2 + n_i^2}$$



16.2 Compensated Semiconductors

- If we assume complete ionisation, and if $(N_d N_a) \gg n_i$, then the majority carrier electron concentration is approximately the difference between the acceptor and donor concentrations and hence:

$$p_0 = \frac{n_i^2}{n_0} = \frac{n_i^2}{(N_d - N_a)}$$

- Note you must always calculate the minority carrier concentration from the law of mass action *after* having calculated the majority carrier concentration via one of the previous equations.

16.3 Position of Fermi level

- We know that the electron and hole concentrations change as Fermi Energy level moves across bandgap. We also know how the electron and holes concentrations depend on impurity concentrations. So we can determine the Fermi Energy level as a function of doping concentration and temperature.
- If we assume the Boltzmann approximation is valid then we can rearrange the equation for n_0 :

$$\varepsilon_C - \varepsilon_F = kT \ln \left(\frac{N_C}{n_0} \right)$$

- For an n-type semiconductor at room temperature (complete ionisation) which is doped with $N_d \gg n_i$, then $n_0 \approx N_d$:

$$\varepsilon_C - \varepsilon_F = kT \ln \left(\frac{N_C}{N_D} \right)$$

- i.e. as the donor concentration increases, the Fermi level moves closer to the CB and so the electron concentration in the CB increases.
- Note that if the semiconductor is compensated then N_d must be replaced with the net effective donor concentration $N_d - N_a$.
- We also know that $n_0 = n_i \exp \left[\frac{(\varepsilon_F - \varepsilon_{F_i})}{kT} \right]$ so we can write:

$$\varepsilon_F - \varepsilon_{F_i} = kT \ln \left(\frac{n_0}{n_i} \right)$$

- This equation can be used specifically for an n-type semiconductor where n_0 is given by:

$$n_0 = \frac{(N_d - N_a)}{2} + \sqrt{\left(\frac{N_d - N_a}{2} \right)^2 + n_i^2}$$

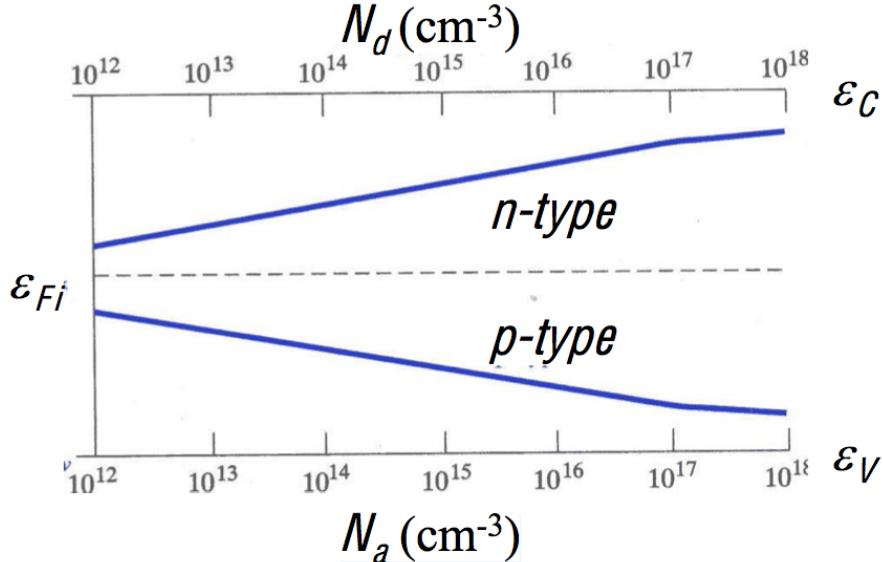
- Note if the net effective donor concentration is zero (i.e. it is completely compensated) then $N_d - N_a = 0$ and so $n_0 = n_i$ and $\varepsilon_F = \varepsilon_{Fi}$ so a completely compensated semiconductor has the properties of an intrinsic material.
- Similarly for a p-type semiconductor:

$$\varepsilon_F - \varepsilon_V = kT \ln \left(\frac{N_V}{p_0} \right) \xrightarrow[N_a >> n_i]{\text{Assuming 300K and}} \boxed{\varepsilon_F - \varepsilon_V = kT \ln \left(\frac{N_V}{N_a} \right)}$$

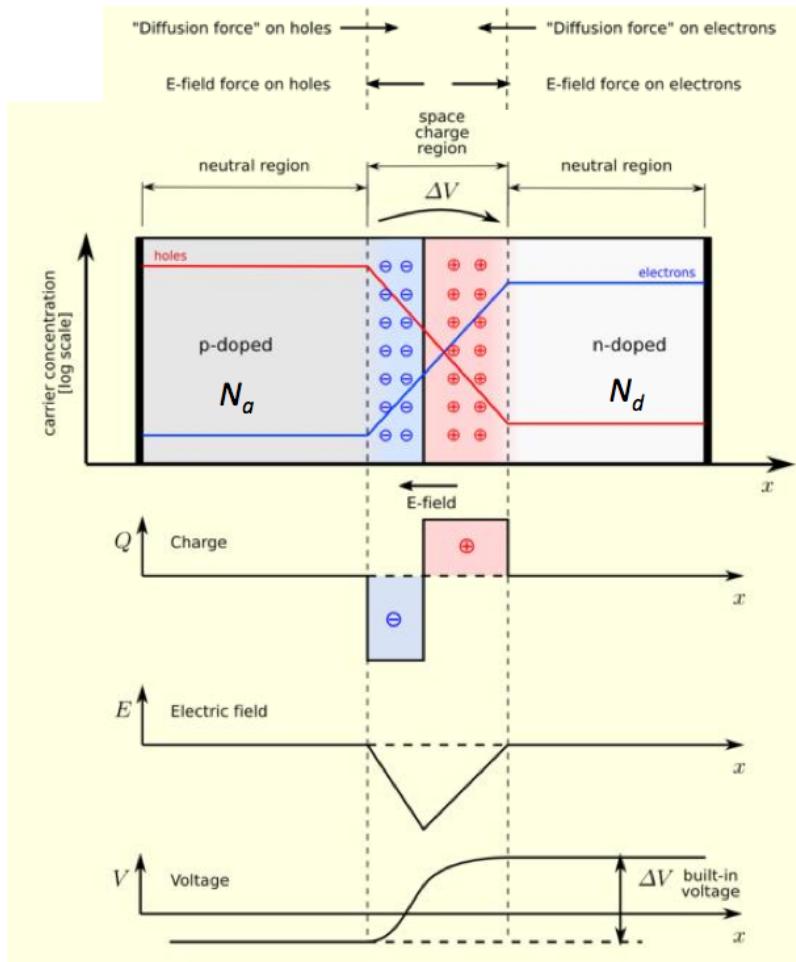
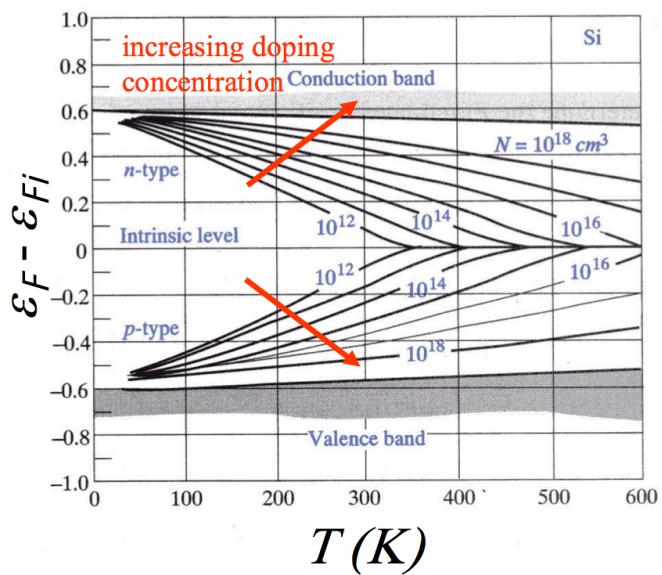
- And the same as for electrons, we can derive an expression in terms of the intrinsic Fermi energy level:

$$\boxed{\varepsilon_{Fi} - \varepsilon_F = kT \ln \left(\frac{p_0}{n_i} \right)}$$

- i.e. for an n-type material $n_0 > n_i$ and the Fermi level is above ε_{Fi} , and for a p-type material $p_0 > n_i$ and the Fermi level is below ε_{Fi}
- We can plot the Fermi level as a function of doping concentration.



- The intrinsic carrier concentration is a strong function of temperature, and therefore so is the Fermi energy level.
- As the temperature increases, n_i increases and ε_F moves closer to the intrinsic level.
- At high temperatures the semiconductor becomes intrinsic like.
- At very low temperatures, freeze-out occurs, the Fermi-level goes above the donor level for n-types and below the acceptor level for p-types - the Boltzmann approximation is no longer valid and many of the derived equations no longer apply.



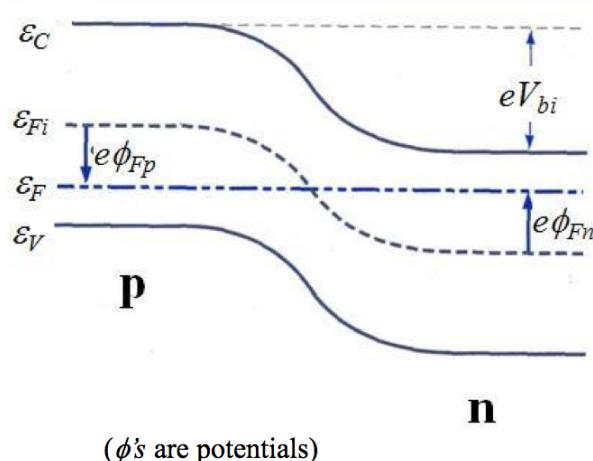
17 p-n junctions

- In thermal equilibrium, the Fermi level is constant through-out a system. If 2 systems with different Fermi levels are brought in to contact, electrons will diffuse from the

higher Fermi level system, in to the lower one to fill empty states.

- A p-n junction can be formed from a crystal with 2 sides doped differently.
- Consider a step junction between the two regions that are uniformly doped - initially there is a very large change in the density gradient of both electron and hole concentrations.
- Majority charge carriers from the n-region will diffuse in to the p-region and recombine with holes, and majority carrier holes will diffuse in to the n-region and recombine.
- Positively charged donor atoms and negatively charged acceptor atoms are left behind - this net charge induces an electric field around the metallurgical junction.
- The region around the junction becomes charged and depleted of charge carriers: “space charge” or “depletion” region - the region acts as an insulator.
- Density gradients still exist in the majority carrier concentrations at the edge of the space charge region, this produces a “diffusion force” that acts on the majority carriers.
- The junction is in thermal equilibrium when, with no voltage applied, the “diffusion force” and E-field force balance each other out.
- Therefore the Fermi level is constant throughout.

17.1 Energy bands of p-n junctions: zero applied bias



- The n region is left positively charged and the p region is left negatively charged. This results in the lowering of the electron energy levels on the n side and the raising on the p side, which causes the ϵ_F to be position-independent as required i.e. the CB and VB bend.

- Electrons in the conduction band of the n-region see a *built-in potential barrier*, V_{bi} , in trying to move in to the p-region. It maintains equilibrium between the majority carriers in one region and the minority carriers in the other (i.e. electrons in the n-region and the electrons in the p-region, and holes in the n-region and holes in the p-region).
- Since ε_{Fi} is equidistant from the conduction band edge through-out then:

$$V_{bi} = |\phi_{Fn} + \phi_{Fp}|$$

Where we define the potentials:

$$e\phi_{Fn} = \varepsilon_{Fi} - \varepsilon_F \text{ in the n-region , } e\phi_{Fp} = \varepsilon_{Fi} - \varepsilon_F \text{ in the p-region}$$

- In the n-type region: Assuming complete ionisation $N_d \gg n_i$ using previous definitions

$$\begin{aligned} n_0 = N_d &= n_i \exp \left[\frac{\varepsilon_F - \varepsilon_{Fi}}{kT} \right] = n_i \exp \left[\frac{-e\phi_{Fn}}{kT} \right] \\ \phi_{Fn} &= \frac{-kT}{e} \ln \left(\frac{N_d}{n_i} \right) \end{aligned}$$

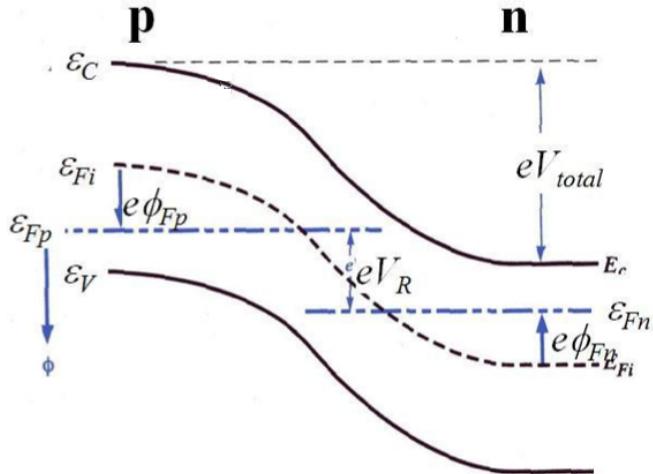
- Similarly, in the p-type region:

$$\begin{aligned} \phi_{Fp} &= \frac{+kT}{e} \ln \left(\frac{N_a}{n_i} \right) \\ \therefore V_{bi} &= \frac{kT}{e} \ln \left(\frac{N_a N_d}{n_i^2} \right) = V_t \ln \left(\frac{N_a N_d}{n_i^2} \right) \end{aligned}$$

where V_t is the **thermal voltage**, note the change of notation: N_d is now the *net* donor concentration *in the n-region* and N_a is the *net* acceptor concentration *in the p-region*.

17.2 Energy bands of p-n junctions: reverse bias

- Diagram shows the energy band diagram when a voltage is applied to the n-region with respect to the p-region.
- We are no longer in an equilibrium condition and the Fermi energy will not be constant
- as “positive” potential is downwards, then the Fermi level on the n-side is below that on p-side.
- The holes in the p-region are pulled away from the junction, causing the width of the depletion zone to increase. Similarly, the electrons will also be pulled away from the junction. Therefore the space charge region widens, and does so increasingly with increasing reverse bias voltage. This increases the voltage barrier causing a high resistance to the flow of charge carriers thus allowing minimal current to cross the p-n junction.



- $$V_{total} = V_R + V_{bi}$$

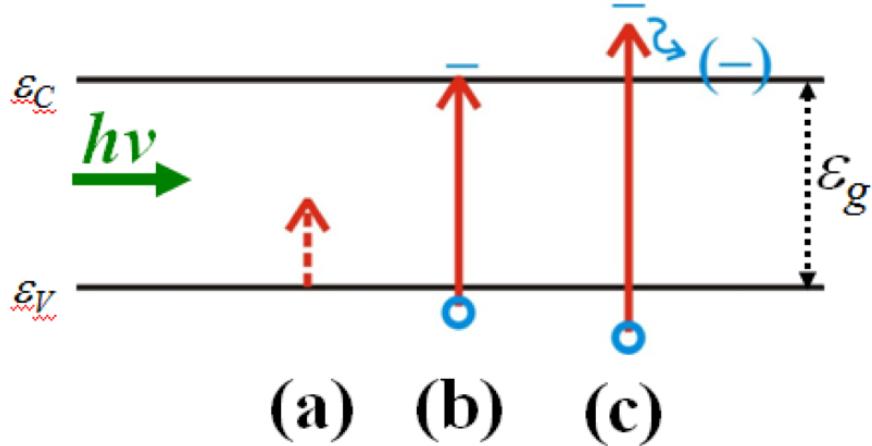
where V_R is the applied reverse bias voltage and V_{bi} is the thermal equilibrium built-in potential.

18 Elementary optical properties of semiconductors

- There are possible photon-semiconductor interactions with the lattice, impurities, defects, and valence electrons.
- If enough energy is imparted to elevate electrons from the valence band to the conduction band. There is **optical absorption** to generate electron-hole pairs, producing “excess carrier concentrations”.
- Recombination leads to the emission of a photon - **luminescence**
- Optical energy (hf) absorbed in a semiconductor generates *excess electron-hole* pairs producing “photocurrents”. The output terminals of a solar cell are connected to a resistive load which produces electrical power.
- Photodetectors also convert optical signals in to electronic signals. Photo-generated excess electron-hole pairs change the semiconductors conductivity (measured as a change in current). A *photodiode* is a p-n junction with reverse bias voltage applied: excess electron-hole pairs created in the space charge region are separated very quickly by electric field producing a photocurrent (proportional to photon flux).
- Inverse of photon flux is *electroluminescence*: this is the process of generating photon emission when the excitation of excess electron-hole pairs is a result of an electric current produced by an applied E-field. If a “forward bias” voltage is applied across a p-n junction then electrons and holes are injected across the space charge region, where they become excess minority carriers. They then diffuse in to the neutral semiconductor and recombine with the majority carriers, if it is a direct bandgap material then a photon is emitted (i.e. junction acts as a Light Emitting Diode)

18.1 Fundamental (optical) absorption and absorption coefficient

If we illuminate the semiconductor with light, then photons either propagate through the semiconductor or are absorbed - depending on the photon energy and magnitude of the band gap.



- (a) If photon energy $hf < \varepsilon_g$ then photons are not readily absorbed and the light is transmitted.
- (b) and (c): If $hf = \varepsilon_g$ or $hf > \varepsilon_g$ then the photon can interact with the electron and may elevate it into the conduction band. Electron-hole pair is produced and photon is absorbed. Any excess energy will give electron/hole additional kinetic energy which will be dissipated as heat.
- Absorption at a barrier:
- I_ν : Intensity of photon flux (ν is frequency) (units of energy $\text{cm}^{-2} \text{ s}^{-1}$)
- Energy absorbed per unit time in distance dx is $\alpha I_\nu(x)dx$ where α is the absorption coefficient: the relative number of photons absorbed per unit distance (cm^{-1})

$$I_\nu(x + dx) - I_\nu(x) = \frac{dI_\nu(x)}{dx}dx = \alpha I_\nu(x)dx \therefore \frac{dI_\nu(x)}{dx} = -\alpha I_\nu(x)$$

So if $I_\nu(0) = I_{\nu 0}$ then:

$$I_\nu(x) = I_{\nu 0} \exp(-\alpha x)$$

19 Electron-hole pair generation rate

- The rate at which energy is absorbed per unit volume is $\alpha I_\nu(x)$. If we assume one absorbed photon at an energy $h\nu$ creates one electron-hole pair then the generation rate of electron-hole pairs is:

$$G = \frac{\alpha I_\nu(x)}{h\nu}$$

Units are number per cm^3 per s

- If on average, one absorbed photon produces less than one electron-hole pair, then this equation must also be multiplied by an efficiency factor η
- Absorption coefficient α is a very strong function of the photon energy and band gap energy.
- Probe band structure by measuring **absorption spectrum** i.e. use beam of photons and study changes in transmitted radiation.
- **Fundamental Absorption:** i.e. excitation of an electron from valence band to conduction band, manifested as a rapid rise in absorption: estimation of energy gap can be made using absorption edge.

20 Absorption Coefficient

$$\alpha(h\nu) = A \sum_{i,f} f_{if} g_i g_f$$

It is summed for all possible transitions between states separated by an energy difference of $h\nu$. Where f_{if} is the probability of transition from the initial state to the final state, g_i is the density of electrons in the initial state and g_f is the density of available (empty) final states.

- We consider absorption transitions between 2 direct valleys, assuming all valence states are filled, and all conduction states are empty (true for undoped semiconductors at T=0K), and that all momentum-conserving transitions are allowed (i.e. the probability function f is independent of energy).
- Photon energy must be equal to the energy gap or larger so $\nu \geq \frac{\varepsilon_g}{h}$.
- Momentum of photon, $\frac{h}{\lambda} \ll$ crystal momentum, $\frac{h}{a}$. so the excited electron state has effectively the same k -value as the initial state.
- A photon is annihilated and an electron-hole pair created.
- Energy and momentum are conserved and electron and hole have opposite k (neglecting momentum of light).

21 Allowed direct transitions: Absorption

- i.e. every initial state is connected with a final state $\varepsilon_f = h\nu + \varepsilon_i$
- In parabolic bands:

$$\varepsilon_f - \varepsilon_g = \frac{\hbar^2 k^2}{2m_e^*}, \quad \varepsilon_i = \frac{-\hbar^2 k^2}{2m_h^*}$$

- So:

$$h\nu - \varepsilon_g = \frac{\hbar^2 k^2}{2} \left(\frac{1}{m_e^*} + \frac{1}{m_h^*} \right) = \frac{\hbar^2 k^2}{2} \frac{1}{m_r}$$

where m_r is reduced mass so:

$$\frac{1}{m_r} = \frac{1}{m_e^*} + \frac{1}{m_h^*}$$

- Density of **directly associated states**:

$$g(h\nu) = \frac{4\pi (2m_r)^{\frac{3}{2}}}{h^3} \sqrt{h\nu - \varepsilon_g}$$

- So:

$$\boxed{\alpha(h\nu) = A^* \sqrt{h\nu - \varepsilon_g}}$$

- So the photon ensures energy is conserved $k_f = k_i$

22 Indirect transitions

- Transition between indirect valleys
- Photon cannot provide a change in momentum (i.e. horizontal shift) so a two-step process is required.
- Momentum is conserved via a **phonon** interaction (quantum of lattice vibration)
- A phonon is either emitted (if the photon energy raises above the dip level), or absorbed if it less and so requires the phonon energy boost. So either:

$$h\nu_{\text{abs}} = \varepsilon_f - \varepsilon_i - \varepsilon_p$$

$$h\nu_{\text{em}} = \varepsilon_f - \varepsilon_i + \varepsilon_p$$

- All occupied states of valence band can connect to all empty states of the conduction band

- Density of initial states:

$$g(\varepsilon_i) = \frac{4\pi (2m_h^*)^{\frac{3}{2}}}{h^3} \sqrt{|\varepsilon_i|}$$

- Density of final states:

$$g(\varepsilon_f) = \frac{4\pi (2m_e^*)^{\frac{3}{2}}}{h^3} \sqrt{\varepsilon_f - \varepsilon_g}$$

- Substituting previous equations in to this:

$$g(\varepsilon_f) = \frac{4\pi (2m_e^*)^{\frac{3}{2}}}{h^3} \sqrt{h\nu - \varepsilon_g \mp \varepsilon_p + \varepsilon_i}$$

- For indirect transitions, absorption coefficient α is proportional to product of $g(\varepsilon_f)$ and $g(\varepsilon_i)$, integrated over all possible combinations of states separated by $h\nu \pm \varepsilon_p$.
- But, α is also proportional to the probability of interaction with phonons, which is itself a function $f(n_p)$ of the number of phonons. Phonons obey Bose-Einstein statistics, so the number of phonons with energy ε_p is given by:

$$n_p(\varepsilon_p) = \frac{1}{\exp\left(\frac{\varepsilon_p}{kT} - 1\right)}$$

- So:

$$\alpha(h\nu) = A f(n_p) \int_0^{-(h\nu - \varepsilon_g \varepsilon_p)} \sqrt{|\varepsilon_i|} \sqrt{h\nu - \varepsilon_g \varepsilon_p + \varepsilon_i} d\varepsilon_i$$

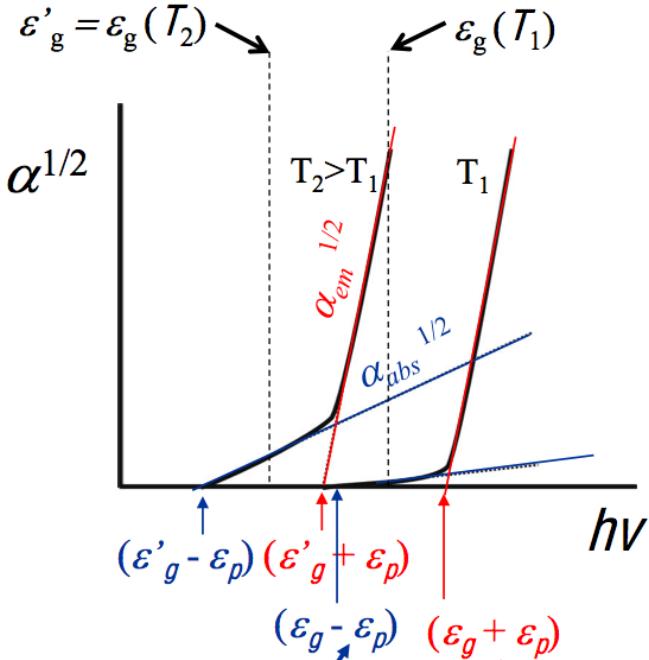
- After integration and substitution of the Bose-Einstein function:

$$\alpha_{\text{abs}}(h\nu) = \frac{A(h\nu - \varepsilon_g + \varepsilon_p)^2}{\exp(\varepsilon_p/kT) - 1}$$

for phonon absorption i.e. $h\nu > \varepsilon_g - \varepsilon_p$

$$\alpha_{\text{em}}(h\nu) = \frac{A(h\nu - \varepsilon_g - \varepsilon_p)^2}{1 - \exp(-\varepsilon_p/kT)}$$

for phonon emission i.e. $h\nu > \varepsilon_g + \varepsilon_p$ note the different factor on the bottom since the probability of phonon emission is proportional to $n_{ph} + 1$.



- Both transitions are possible when $h\nu > \varepsilon_g + \varepsilon_p$, but the phonon concentration is very small at low temperatures and so α_{abs} is also small at low T.

- Note ε_g is dependent on temperature, because as the temperature increases the band gap energy decreases as the crystal lattice expands and the interatomic bonds are weakened, and so less energy is needed to break the bonds and get electrons to the conduction band.

23 Burstein-Moss shift

- If a semiconductor is degenerately (heavily) doped then the Fermi level can be inside the band (i.e. the conduction band for an n-type semiconductor) by a quantity ξ_n
- Since states between ε_C and ε_F are already filled, fundamental transitions to states below $\varepsilon_g + \xi_n$ are forbidden.
- So absorption edge shift to higher energies, i.e. more energy is required as the electron must get above the already filled states.

23.1 Indirect transitions

- In heavily doped indirect bandgap semiconductors, it is also possible to conserve momentum by a scattering without interaction with a phonon: by electron-electron scattering or impurity scattering.
- Now the scattering probability is proportional to the number of scatters N_{sc} :

$$\alpha(h\nu) = AN_{sc}(h\nu - \varepsilon_g - \xi_n)^2$$

- Indirect transitions between direct valleys are also possible as momentum can be conserved by second order processes (phonons or scattering) and the absorption coefficient is the same as previous indirect transitions.
- But multistep processes will have a much lower probability than the equivalent direct transition.

24 Absorption

- In discussing fundamental absorption, we've assumed that the electron and hole created are both "free". However they will attract one another and may form a bound state, in which they revolve around their combined centre of mass, this state is called an exciton.
- Its binding energy is very small (just below the CB edge).

24.1 Impurity Photon Absorption

There are four cases of impurity absorption, none resulting in the creation of an electron-hole pair:

1. Neutral donor accepts photon and electron excited in to CB (or higher impurity level).
2. Valence band accepts photon and electron is excited in to acceptor level.

3. VB to ionised donor, or form ionised acceptor to CB.
4. Ionised acceptor to ionised donor.

24.2 Free Carrier absorption

- Absorption of radiation by electrons or holes to change the state in which they exist, but without becoming excited into other bands
- These are labelled **intraband transitions** and they dominate the absorption spectrum below the fundamental edge.

24.3 Recombination of electron-hole pairs and photon emission

- Basic interband transitions: transitions directly from band to band, if there are energetic holes/electrons i.e. with energies greater than the band gap energy then recombination will produce an emission spectrum with an associated bandwidth.
- Involving impurity sites: Transitions from band to impurity, impurity to band, or impurity to impurity. A deep trap can also create recombination by trapping holes and electrons simultaneously, this is non-radiative (no photon produced).
- Auger recombination processes: These are all non-radiative and occur in heavily doped materials. The energy is transferred to the other charge carriers and eventually lost as heat.
- Note that emission of a photon is not necessarily at a single, discrete energy.
- Spontaneous emission rate has the form:

$$I(\nu) = \nu^2 \sqrt{h\nu - \varepsilon_g} \exp\left[\frac{-(h\nu - \varepsilon_g)}{kT}\right]$$

- The peak photon energy decreases with temperature because bandgap energy decreases with temperature.
- Not all recombination processes are radiative. An efficient luminescent material is one where the radiative transitions are dominant.
- Quantum efficiency is defined as the ratio of the radiative recombination rate to the total recombination rate for all processes:

$$\eta = \frac{R_r}{R} = \frac{R_r}{R_r + R_{nr}}$$

24.4 Luminescent efficiency

- We define the radiative and non-radiative recombination rates as:

$$R_r = \frac{N}{\tau_r}, \quad R_{nr} = \frac{N}{\tau_{nr}}$$

Where N is the number of electron-hole pairs, τ_r is the radiative recombination lifetime and τ_{nr} is the non-radiative recombination rate.

- Recombination rate is inversely proportional to lifetime, then:

$$\eta = \frac{\frac{1}{\tau_r}}{\frac{1}{\tau_{nr}} + \frac{1}{\tau_r}} = \frac{\tau_{nr}}{\tau_{nr} + \tau_r}$$

- So for high luminescent efficiency, the non-radiative lifetime must be large.
- The interband recombination rate is directly proportional to the number of electrons available, and the number of empty states.

$$R = Bnp$$

Where B is a constant of proportionality. Note B is about 10^6 times smaller for indirect bandgap materials than direct bandgap materials!

- One problem with direct bandgap materials is the reabsorption of emitted photons.

$$\frac{dN}{dt} = G - R \text{ so for steady state solution: } \frac{dN}{dt} = 0 \therefore N_0 = G\tau$$

25 Transport properties of solids

- Net flow of electrons and holes produces currents - there are two such transport mechanisms: **drift** and **diffusion**. Carrier transport phenomena are the foundation for determining the semiconductor's current-voltage characteristics.
- Temperature gradients may also produce carrier currents, but since semiconductors are small, we neglect this.

25.1 Drift Current

- An applied E-field will produce a force on electrons and holes so that they will experience net movement provided there are available states to which they can move.
- The resulting **drift current** is a function of concentration of mobile electrons and holes and the net drift velocity of these charge carriers.
- Average drift velocity is related to the electric field by a parameter called **mobility**.
- Mobility gives an indication of how well a carrier moves through a semiconductor (i.e. charge carriers do not move unimpeded, as they are involved in collisions with semiconductor and impurity atoms).

25.2 Drift Current and mobility

- Consider a volume of positive charge density, R , moving at an average drift velocity, v_d . Then the drift current density is given by:

$$J_{\text{drift}} = Rv_d$$

- If the volume charge is due to holes, then:

$$J_{\text{h, drift}} = (ep)v_{d,h}$$

- Equation of motion of a positively charged hole in an E-field is:

$$F = m_h^* a = eE$$

- Due to collisions with impurities and lattice atoms, the velocity of charged particles does not increase linearly with time.
- Hole accelerates due to E-field, collides, and loses most of its energy. It again accelerates, gains energy and is involved in another scattering process, etc.
- Overall, the particle will gain an average drift velocity, which for low electric fields is directly proportional to E :

$$v_{d,h} = \mu_h E$$

Where μ is the proportionality factor, called the **mobility** - it describes how well a particle will move in an E-field.

- By combining this with the previous expression:

$$J_{\text{h, drift}} = e\mu_h p E$$

i.e. drift due to holes is in **same** direction as an applied E-field

- Similarly, for electrons:

$$J_{\text{e, drift}} = Rv_{d,e} = (-en)v_{d,e}$$

$$v_{d,e} = -\mu_e E$$

i.e. net motion of electron is opposite to E-field direction.

$$J_{\text{e, drift}} = (-en)(-\mu_e E) = e\mu_e n E$$

- Hence **conventional** drift current due to electrons is in the **same** direction as the applied E-field, even though the actual movement is in the opposite direction.
- Both electrons and holes contribute to the drift current, hence the **total** drift current is:

$$J_{\text{drift}} = e(\mu_e n + \mu_h p)E$$

25.3 Mobility

- A DC electric field does not produce a steadily increasing current because of collisions with the lattice vibrations (phonons) and impurities, so we write:

$$m_h^* \left(\frac{dv}{dt} + \frac{v}{\tau_{ch}} \right) = eE$$

The effect of the second term in the bracket is to cause v to decay exponentially to zero with a time constant τ when the field is removed.

- v is the drift velocity, i.e. the additional velocity associated with departure from thermal equilibrium given by the Fermi distribution function.
- For a DC electric field only, the above equation has the steady-state solution (for holes):

$$v_h = \frac{e\tau_{ch}}{m_h^*} E$$

- Imagine the random motion of a hole, with a mean time between collisions of τ_{ch}
- When a small E-field is applied, there is a net drift of the hole in the direction of the field.
- So the Hole mobility is given by:

$$\boxed{\mu_h = \frac{v_{dh}}{E} = \frac{e\tau_{ch}}{m_h^*}}$$

- And for electrons:

$$\boxed{\mu_e = \frac{e\tau_{ce}}{m_e^*}}$$
- There are two collision mechanisms that dominate in semiconductors:
 - Phonon or lattice scattering
 - Ionised impurity scattering
- For $T > 0$ K, lattice atoms have thermal energy that causes them to vibrate, this disrupts the perfectly periodic potential used in the earlier analysis. Such a perfect periodicity would otherwise allow electrons to move unimpeded.
- Vibrations in the lattice result in interactions between electron/hole and atoms - **phonon scattering**. The theory states, to first order, that:

$$\mu_L \propto T^{-3}$$

- Impurity atoms are all ionised at room temperature, so a Coulomb interaction exists between electrons/holes and ionised impurities. It results in scattering, and also alters the velocity characteristics of the charge carriers. To first order:

$$\mu_I \propto \frac{T^{\frac{3}{2}}}{N_I}$$

where $N_I = N_d^+ + N_a^-$ is the total ionised impurity concentration)

- i.e. for lattice scattering, the probability of scattering occurring increases as temperature increases and therefore mobility decreases.
- For ionised impurity scattering, if T goes up, then the random thermal velocity of a carrier increases, reducing the time the carrier spends in the vicinity of the Coulomb force, reducing the scattering effect and increasing the mobility. If N_I goes up then the probability of scattering also goes up and the value of μ_I goes down.
- If τ_L is the mean time between collisions due to lattice scattering, then $\frac{dt}{\tau_L}$ is the probability of a lattice scattering event occurring in time dt . Similarly for ionised impurity scattering τ_I .
- If the two scattering processes are independent, then the total probability of a scattering event occurring in time dt is the sum of the individual events i.e.

$$\frac{dt}{\tau} = \frac{dt}{\tau_I} + \frac{dt}{\tau_L}$$

This is known as Matthiessen's Rule.

- Comparing with the previous equations for μ_e , μ_h :

$$\frac{1}{\mu} = \frac{1}{\mu_I} + \frac{1}{\mu_L}$$

where μ is the net mobility.

26 Conductivity and resistivity

- From above, and Ohm's Law we can write:

$$J_{\text{drift}} = e(\mu_e n + \mu_h p)E = \sigma E = \frac{E}{\rho}$$

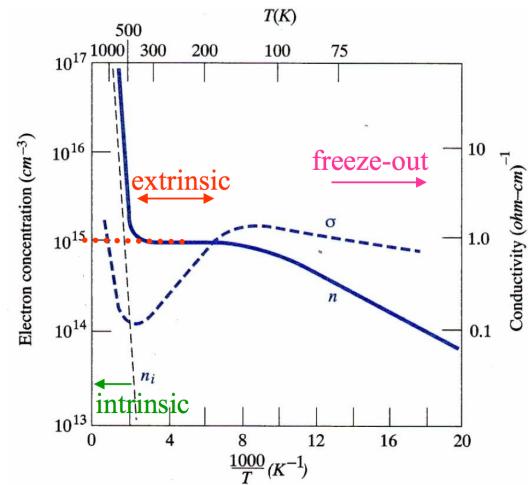
- Since mobilities are a function of impurity concentration then so is conductivity.
- E.g. consider a p-type semiconductor with acceptor doping N_a (so $N_d = 0$) in which $N_a \gg n_i$. If we assume electron and hole mobilities are of the same order of magnitude then:

$$\sigma = e(\mu_e n + \mu_h p) \approx e\mu_h p$$

i.e. the conductivity and resistivity of an extrinsic semiconductor are a function primarily of the majority charge carrier.

26.1 Conductivity

Plotting the carrier concentration and conductivity against temperature yields:



Mid-temperature (extrinsic) range: Complete ionisation and electron concentration remains essentially constant. The variation in conductivity is because mobility is a function of temperature - i.e. there is competition between μ_L and μ_I

At higher temperatures: intrinsic carrier concentration, n_i begins to dominate the electron concentration and conductivity. $\sigma_i = e(\mu_e + \mu_h)n_i$ since $n = p = n_i$

At the lowest temperatures: Freeze-out begins to occur: electron concentration and conductivity decrease with decreasing temperature.

Note at room temperature, phonon scattering dominates over ionised impurity scattering.

26.2 Velocity saturation

- We've assumed $v_d = \mu E$ i.e. that drift velocity increases linearly with applied electric field.
- $v_{\text{tot}} = v_{\text{th}} + v_d$ where v_{th} os the random thermal velocity.
- At room temperature $v_{\text{th}} \approx 10^7 \text{ cm s}^{-1}$ for electrons in Si. (From $\frac{1}{2}mv_{\text{th}}^2 = \frac{3}{2}kT$).
- By comparison, $\mu_e = 1350 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ apply $75 \text{ V cm}^{-1} \rightarrow v_d = 10^5 \text{ cm s}^{-1}$ which is 1% of the thermal velocity. So the applied electric field does not appreciably alter the energy of the electron.
- The linear relationship $v_d = \mu E$ breaks down when high fields are applied. When the carrier enery increases beyond the optical phonon energy, the probability of emitting an optical phonon increases abruptly.
- This mechanism causes the carrier velocity to saturate with increasing electric field. For carriers in silicon and other materials, which do not contain accessible higher bands, the velocity versus field relation increases monotonically.

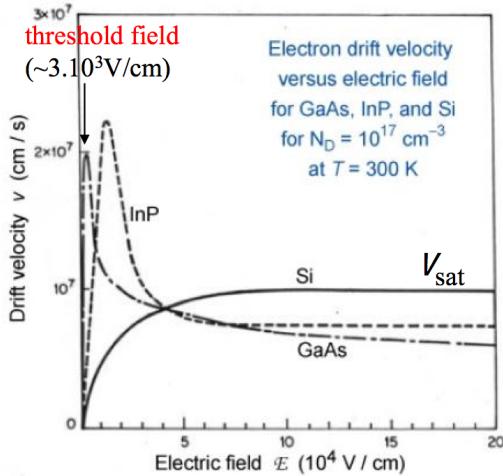
- The analysis is more complex for materials such as GaAs and InP, which contain multiple closely spaced conduction band minima.

26.3 Negative Differential Resistance

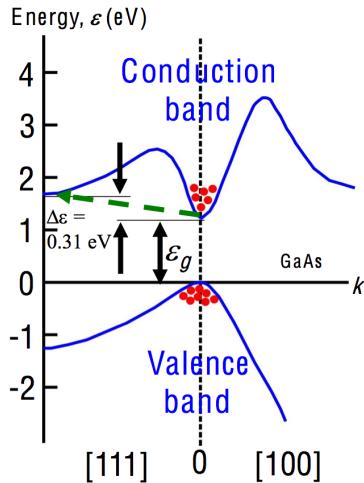
- As a first approximation, considering Silicon:

$$v_{\text{sat}} = \sqrt{\frac{2\varepsilon_{\text{phonon}}}{m^*}}$$

- i.e. Materials with a small m^* and high optical phonon energy are more likely to have a high v_{sat} . Materials with multiple band minima can have a rather low v_{sat} relative to the peak velocity, if the carriers in the higher minima have a larger m^* .



- Now consider GaAs: at low E , the relationship between v_d and E is linear. As the field increases, v_d peaks and then decreases, the negative slope represents a **negative differential mobility**.
- i.e. over certain voltage ranges, current is a decreasing function of the voltage. Note that the resistance is still positive, it is the slope which is negative, and so Ohm's law is not obeyed.
- At the CB minimum there is a low effective mass and so a large mobility.
- Because of their high mobility, electrons in the CB minimum are easily accelerated by a strong E-field and scattered in to the satellite valley
- Here the effective mass and effective DOS is much higher
- So the mobility is much smaller and the intra-valley transfer results in a decreasing average drift velocity with electric field (negative differential mobility).
- This is called the **Gunn effect** and is useful for producing oscillators (signal generators and power supplies).



27 Diffusion

- Particles flow from regions of high concentrations to regions of low concentrations - producing a **diffusion current**.
- For a simplified analysis assume only 1-dimensional electron concentration variation.
- We will calculate the current density, J , by considering the net flow of electrons per unit time passing across the $x = 0$ plane.
- l is the mean-free path of an electron, i.e. the average distance travelled between collisions, $l = v_{\text{th}}\tau_{\text{ce}}$
- So one half of the electrons at $x = -l$ will be travelling to the right, and one half of the electrons at $x = +l$ will be travelling to the left, and both groups will cross the $x = 0$ plane.
- So

$$J = -e \times \text{flux of electrons} = -e \left(\frac{1}{2}n(-l)v_{\text{th}} - \frac{1}{2}n(+l)v_{\text{th}} \right) = \frac{-ev_{\text{th}}}{2} (n(-l) - n(+l))$$

- Using a Taylor expansion around $x = 0$, keeping only first two terms:

$$J = \frac{-ev_{\text{th}}}{2} \left(\left[n(0) - l \frac{dn}{dx} \right] - \left[n(0) + \frac{dn}{dx} \right] \right) = +ev_{\text{th}}l \frac{dn}{dx}$$

- i.e. electron diffusion current is proportional to density gradient of electron concentration where the conventional current density is in the x-direction
- The diffusion current is also written as (Fick's Law):

$$J_{e,x,dif} = eD_e \frac{dn}{dx}$$

where D_e is the **electron diffusion coefficient** (in $\text{cm}^2 \text{ s}^{-1}$)

$$D = v_{\text{th}} l$$

- Since holes are positively charged, then the conventional diffusion current density is in the negative x-direction:

$$J_{\text{hx, dif}} = -eD_h \frac{dp}{dx}$$

where D_h is the **hole diffusion coefficient**.

- Therefore the *total current density* is made up of two drift and two diffusion terms, so in 1D:

$$J = en\mu_e E_x + ep\mu_h E_x + eD_e \frac{dn}{dx} - eD_h \frac{dp}{dx}$$

- The mobility gives an indication of how well a carrier moves in a semiconductor due to the force of an electric field.
- The diffusion coefficient gives an indication of how well a carrier moves in a semiconductor as a result of a density gradient.
- But these are not independent parameters, they are related by the Einstein relation.

28 Graded Impurity Distribution

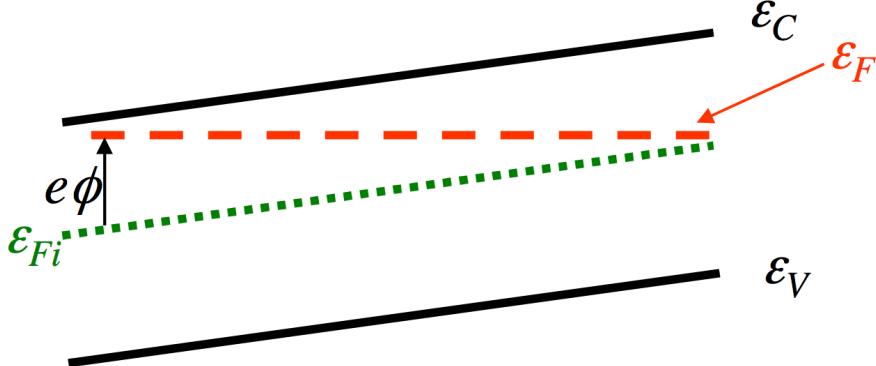
- Consider a non-uniformly doped n-type semiconductor (i.e. it has a graded impurity distribution) such that the doping concentration **decreases** as x increases.
- If the semiconductor is in equilibrium, then the Fermi energy level is constant throughout.
- There will be a diffusion of majority carrier electrons from the region of high to low concentration (the +ve x-direction), leaving behind positively charged donor ions.
- The separation of charge induces an electric field which opposes the diffusion process.
- When equilibrium is reached, the mobile carrier concentration is not exactly equal to the fixed impurity concentration, and the induced electric field prevents any further separation of charge.
- However, usually the space-charge induced by the diffusion process is a small fraction of the impurity concentration, thus the mobile carrier concentration is not too different from the impurity dopant density.
- Electric potential is related to the electron potential energy by the electron charge:

$$\phi = -\frac{1}{e} (\varepsilon_{F_i} - \varepsilon_F)$$

- So in 1D:

$$E_x = -\frac{d\phi}{dx} = \frac{1}{e} \frac{d\varepsilon_{Fi}}{dx}$$

i.e. if the intrinsic Fermi level changes as a function of distance throughout the semiconductor in thermal equilibrium, then an electric field exists.



- Assuming a quasi-neutrality condition, where the electron concentration is almost equal to the donor impurity concentration, then we can still write:

$$n_0 = n_i \exp \left[\frac{\varepsilon_F - \varepsilon_{Fi}}{kT} \right]$$

$$\varepsilon_F - \varepsilon_{Fi} = kT \ln \left(\frac{N_d(x)}{n_i} \right)$$

- The Fermi level is constant for thermal equilibrium so we obtain:

$$\frac{-d\varepsilon_{Fi}}{dx} = \frac{kT}{N_d(x)} \frac{dN_d(x)}{dx}$$

- Combining this with the previous equation for the Electric field in terms of the differential of the potential:

$$E_x = \frac{-kT}{e} \frac{1}{N_d(x)} \frac{dN_d(x)}{dx}$$

28.1 The Einstein relation

- Considering the non-uniformly doped semiconductor, and assuming there are no electrical connections, so the semiconductor is in thermal equilibrium.
- So the individual electron and hole currents must be zero.
- So for electrons:

$$J_e = 0 = en\mu_e E_x + eD_e \frac{dn}{dx}$$

- Assuming quasi-neutrality, so that $n(x) \approx N_d(x)$ and substituting in the previous equation for E_x obtains:

$$-\mu_e \left(\frac{kT}{e} \right) + D_e = 0$$

- The equation is valid for the condition:

$$\frac{D_e}{\mu_e} = \frac{kT}{e}$$

- The hole current must also be zero, so:

$$\frac{D_h}{\mu_h} = \frac{kT}{e}$$

- Combining these two equations yields the Einstein relation:

$$\frac{D_e}{\mu_e} = \frac{D_h}{\mu_h} = \frac{kT}{e}$$

- i.e. the diffusion coefficient and mobility are not independent parameters. Their relation is known as the **Einstein relation**
- Note mobilities are a strong function of temperature, as are the diffusion coefficients. Hence the specific temperature dependence in the Einstein relation is only a small fraction of the real temperature characteristic.

28.2 Thermal Properties of Solids

- We need to understand **phonons** (a quantization of the elastic energy of sound waves) to discuss the thermal properties of solids:
 - Insulators and semiconductors: Debye theory - heat capacity, thermal conduction.
 - Metals (again) but including the lattice effects in addition to electrons. The heat capacity leads to the thermal conduction, electrical resistance and Weidemann-Franz law.
- The thermal equilibrium number of phonons with frequency ω , is given by the Planck distribution:

$$n_{\text{ph}}(\hbar\omega) = \frac{1}{\exp\left[\frac{\hbar\omega}{kT}\right] - 1}$$

this is a special case of **Bose-Einstein statistics** which determines the statistical distribution of identical indistinguishable bosons over the energy states in thermal equilibrium.

- Bosons, unlike fermions, are not subject to the Pauli exclusion principle: an unlimited number of particles may occupy the same state at the same time (Bose-Einstein condensate).

- For $\hbar\omega \gg kT$:

$$n_{\text{ph}} \approx \exp\left(\frac{-\hbar\omega}{kT}\right)$$

- For $\hbar\omega \ll kT$:

$$n_{\text{ph}} \approx \frac{kT}{\hbar\omega}$$

28.3 Phonons: Debye approximation

- This treats the vibrations of the atomic lattice as phonons in a box.
- Assumes identical oscillators and neglects the dispersion of acoustic waves. This continuum approximation is valid because the interatomic spacing is much smaller than the wavelength.
- i.e. we assume a linear dependence: $\omega = v_s k$
- The number of allowed states in the range $k \rightarrow k + dk$ is given by:

$$G(k)dx = 3 \cdot \frac{4\pi k^2 dk}{\left(\frac{2\pi}{L}\right)^3} = \frac{12\pi\omega^2 d\omega}{\left(\frac{2\pi}{L}\right)^3 v_s^3}$$

where the factor of 3 is from the 3 acoustic modes.

$$G(\omega) = \frac{G(k)dk}{d\omega} = \frac{3V}{2\pi^2} \frac{\omega^2}{v_s^3}$$

- We define the Debye frequency, ω_D , such that above this cut-off there are no modes. We consider the total number of states (recall there are $3N$ vibrational modes of a 3D solid of N atoms):

$$3N = \int_0^{\omega_D} G(\omega)d\omega = \frac{V}{2\pi^2} \frac{\omega_D^3}{v_s^3}$$

28.4 Debye Temperature

- We can define the Debye temperature, θ_D with:

$$\hbar\omega_D = k_B\theta_D$$

- Rearranging the expression for ω_D and substituting it in, yields:

$$\theta_D = \frac{\hbar v_s}{k_B} \left(\frac{6\pi^2 N}{V} \right)^{\frac{1}{3}}$$

- i.e. $\theta_D \sim v_s(N/V)^{1/3} \sim (Y/\rho)^{1/2}(N/V)^{1/3} \sim (Y/M)^{1/2}$ where Y is the elastic constant, Young's Modulus, and M is the atomic mass.
- Therefore the stiffer the crystal, the smaller the M , the higher the θ_D
- It is the temperature of a crystal's highest mode of vibration, i.e. the highest temperature that can be achieved due to a single normal vibration.

29 Debye heat capacity

- The total energy is:

$$\begin{aligned}\varepsilon &= V \int_0^{\omega_D} \hbar\omega n_{\text{ph}} g(\omega) d\omega \\ \varepsilon &= V \int_0^{\omega_D} \frac{3\omega^2}{2\pi^2 v_s^3} \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} d\omega\end{aligned}$$

- Therefore

$$C_V = \frac{\partial \varepsilon}{\partial T} = \frac{3V\hbar}{2\pi^2 v_s^3} \frac{\partial}{\partial T} \int_0^{\omega_D} \left(\frac{\omega^3}{\exp\left(\frac{\hbar\omega}{kT}\right) - 1} \right)$$

- By substituting $x = \frac{\hbar\omega}{kT}$ and $x_D = \frac{\theta_D}{T}$ and recalling that $\theta_D = \frac{\hbar v_s}{k} \left(\frac{6\pi^2 N}{V} \right)^{\frac{1}{3}}$:

$$C_V = 9Nk_B \left(\frac{T}{\theta_D} \right)^3 \int_0^{x_D} \left(\frac{x^4 e^x}{(e^x - 1)^2} \right)$$

- This is the Debye Heat Capacity, associated with the lattice and defines a “universal” function scaled only by the Debye temperature θ_D

29.1 Debye heat capacity: High temperature limit

- At high temperature limit, $T \gg \theta_D$, the Debye heat capacity equation becomes:

$$C_V = 9Nk_B \left(\frac{T}{\theta_D} \right)^3 \int_0^{x_D} \left(\frac{x^4}{(1 + x - 1)^2} \right) dx = \boxed{3Nk_B}$$

- i.e. this is the Dulong-Petite **classical** result, in which the crystal is simply described by $3N$ classical harmonic oscillators. Each degree of freedom contributes Nk_B to C_V . **This does not work for low T!**
- This can be understood because at high temperatures, in excess of their Debye temperature, $n_{\text{ph}}(\omega) \approx \frac{kT}{\hbar\omega}$, and the average energy of the mode ω is:

$$\varepsilon_\omega = n_{\text{ph}}(\omega) \hbar\omega \approx kT$$

29.2 Debye heat capacity: Low temperature limit

- Now consider the low temperature limit i.e. $T \ll \theta_D$
- First we must write the total energy of the phonons in the system as:

$$\varepsilon = 9NkT \left(\frac{T}{\theta_D} \right)^3 \int_0^{x_D} \left(\frac{x^3}{e^x - 1} \right) dx$$

by using the same substitutions in to the expression for energy as was used to derive the heat capacity.

- Then, because x is large, we can make the approximation to carry out the integral to infinity (instead of x_D). Then the integral has a value of $\frac{\pi^4}{15}$ and we obtain the Debye T^3 law:

$$C_V = \left(\frac{12\pi^4}{5} \right) N k_B \left(\frac{T}{\theta_D} \right)^3$$

This is valid for $T < 0.1\theta_D$.

- The T^3 law is best seen in materials with no free electrons to absorb heat (like Ar, as opposed to metals), and it works well because it provides a good description of vibrational modes at low energies (long wavelengths) where only the lowest energy modes are excited.

29.2.1 Why is it a T^3 law

- Only those phonons having $\hbar\omega < k_B T$ will be excited to any appreciable extent at low temperatures
- $\hbar\omega_{\max} \sim \hbar\omega_T = k_B T$ and the fraction of occupied modes is then (using that $g(\omega) \approx \omega^2$):

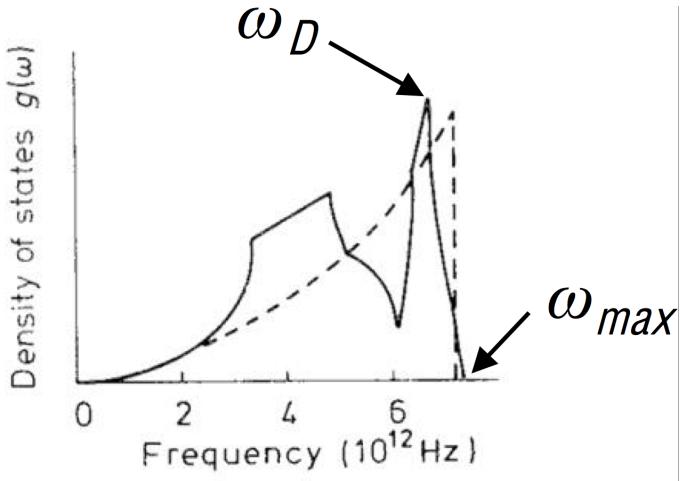
$$\frac{\int_0^{\omega_T} g(\omega) d\omega}{\int_0^{\omega_D} g(\omega) d\omega} = \left(\frac{\omega_T}{\omega_D} \right)^3 = \left(\frac{k_T}{k_D} \right)^3 = \left(\frac{T}{\theta_D} \right)^3$$

- The excitation of these modes will be approximately classical, with an energy close to $k_B T$
- There are of the order $3N \left(\frac{T}{\theta_D} \right)^3$ excited modes, each having energy $\sim k_B T$
- Hence the total energy $\sim 3Nk_B T \left(\frac{T}{\theta_D} \right)^3$ and the heat capacity:

$$C_V = \frac{d\varepsilon}{dT} \propto k_B N \left(\frac{T}{\theta_D} \right)^3$$

29.3 Real Systems

- We've shown that the non-dispersive Debye model gives a good approximation to the heat capacity of most solids even though the actual DOS may vary significantly from the Debye assumption:
- The differences are due to:
 - The actual dispersion of sound ($\frac{d\omega}{dk}$ reduces to zero) lowers the cutoff frequency and causes a rise in the density of states just below it.
 - The variation in the cut-off wavenumber with crystal orientation blurs the sharp cut-off in the DOS, and raises the maximum frequency.
- The consequence is that the main rise in heat capacity comes at a lower T than one would expect from the Debye theory.



- In real systems, the experimentally measured θ_D is not constant.
- In conclusion, the one parameter (θ_D) model is too simplistic to account for all the information about atomic structure and forces.

29.4 Anharmonic effects

- We have so far neglected *anharmonic effects* i.e. positive and negative stresses don't produce strains of equal magnitude (a deviation from Hooke's Law).
 - For the symmetric harmonic approximation (parabola) to the interatomic potential, the mean separation of the atoms does not change with increasing amplitude of vibration (temperature): atoms are independent of one another. This approximation is necessary for the separation of the lattice motions into independent normal modes.
 - For large atomic displacements we must include anharmonic (higher than quadratic) terms in the Taylor series expansion of the interaction between nearest neighbours: This leads to coupling of modes that can be perceived to be collisions between phonons associated with the modes.
 - As a result of these collisions, the phonon mean free path ; $\propto \frac{1}{T}$ limits the thermal conductivity (this dependence assumes no geometrical or impurity scattering).
 - At $T < 10K$, $l \approx$ specimen size (a constant) due to the lack of phonon availability for significant phonon-phonon scattering, and the inefficiency of (small) impurities to scatter long wavelength modes.

29.5 Phonon-phonon collisions (coupling)

- Coupling of normal modes by anharmonic terms of interatomic potential is equivalent to collisions between the phonons.
- Energy must be conserved so $\omega_3 = \omega_1 + \omega_2$

- Normal process: $\vec{k}_3 = \vec{k}_1 + \vec{k}_2$ where $|\vec{k}_3| < \frac{\pi}{a}$ i.e. momentum is conserved and there is no change in the flow of phonons.
- Umklapp process: $|\vec{k}_3| > \frac{\pi}{a}$ and is equivalent to a wrapped around phonon with wavevector $|\vec{k}_{3'}| < \frac{\pi}{a}$ i.e.

$$\vec{k}_3 \pm n \frac{2\pi}{a} = \vec{k}_1 + \vec{k}_2$$

where $G = \frac{2\pi}{a}$.

- Such a non-zero ($= G$) change in total wavevector is allowed in periodic crystals, as the lattice absorbs the momentum.
- The U-processes are highly effective at changing the momentum of the phonon as a collision of two phonons with a positive k_x can create a phonon with negative k_x .
- The U-processes limit the mean free path of phonons at high temperatures.
- The difference in momentum associated with the U-process is transferred to the centre of mass of the lattice.

30 Thermal conduction: Insulators, semiconductors and metals

- A temperature gradient in a solid leads to a flow of heat from hot to cold.
- In insulators and semiconductors, the most important contribution to thermal conduction is associated with the flow of phonons (lattice vibrations). Applying the results from the Debye theory (of a phonon gas).
- Recall that thermal conduction is limited by (i) interactions with other phonons and (ii) geometrical scattering (imperfections, boundaries).
- In metals, the largest contribution comes from electrons. The thermal energy flux j is given by:

$$j = -\kappa \frac{dT}{dx}$$

30.1 Metals vs. Insulators

$$j = \sum_k n_k k$$

where n_k is the number of phonons having wavevector k . For distributions with heat flow ($j \neq 0$), N-processes are not able to establish thermal equilibrium since they leave j unchanged.

Phonon gas (insulators): speed of particles is essentially constant, but both the number density and energy density are greater at the hot end. Heat flow is primarily due to phonon flow, with phonons being created at the hot end and destroyed at the cold end. Note that if only Normal (N)-processes occur, the phonon flux is unchanged in momentum on collision and some phonon flux will persist across the sample length i.e. the resistivity is zero. U-processes are large momentum changes, and cause the initial phonon flux to decay rapidly (hot to cold).

Real Gas (also electron gas i.e. ideal metal): No flow of particles. The average kinetic energy per particle is greater at the hot end, but the number density is greater at the cold end. The energy density is therefore uniform because the pressure is uniform. Heat flow is only due to transfer of kinetic energy from one particle to another during collisions (a minor effect in the phonon case).

30.2 Thermal conduction using kinetic theory

Works with phonon gas and real gas as particle conservation is not assumed.

- $\frac{dT}{dx}$ implies that the process of thermal energy transfer is a random process i.e. energy *diffuses* from one end to the other (otherwise j would be instead proportional to ΔT if the energy was transferred directly - there would be no dependence on the sample length)
- Hence the random nature brings temperature gradient, and, as we shall see, a mean free path in to the expression for the thermal flux.
- Assuming a uniform temperature gradient in the x-direction.
- The net flow of energy across the plane ($x=0$):

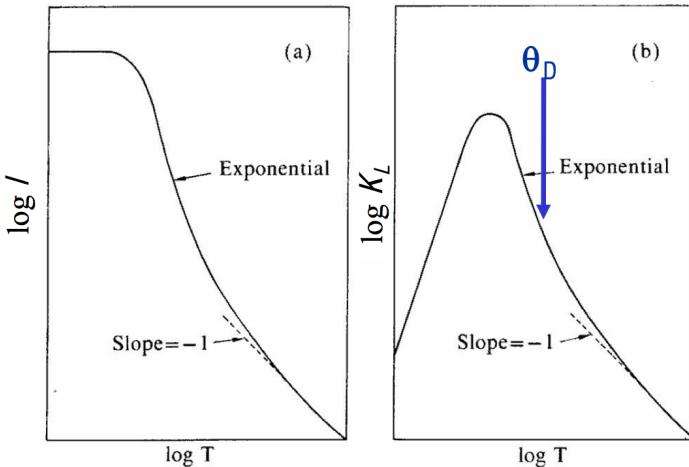
$$\begin{aligned} j &= \frac{1}{2}nv\varepsilon(x = l) - \frac{1}{2}nv\varepsilon(x = -l) \\ -j &= \frac{1}{2}nv \frac{d\varepsilon(x = 0)}{dx} 2l = nvl \frac{d\varepsilon}{dT} \frac{dT}{dx} \\ -j &= v l C_V \frac{dT}{dx} \end{aligned}$$

note the negative sign since on j since the energy flows in the opposite direction to ΔT . n is the concentration of particles, v is the average velocity of the particles, ε is the energy per particle, τ is the time between collisions, l is the mean free path, $C_V = n \frac{d\varepsilon}{dT}$ is the heat capacity per unit volume (specific heat).

30.3 Thermal conductivity

- Thus in 1D, $\kappa = C_V v l$ so in 3D, $\kappa = \frac{1}{3}C_V v l$, the factor of $\frac{1}{3}$ arises due to averaging.
- Note in heat transfer, the phonons (lattice vibrations) and electrons (and holes) take part in heat transfer $\kappa = \kappa_e + \kappa_L$

- In semiconductors, $\kappa_e \ll \kappa_L$
- In order to understand $\kappa(T)$ we must consider the temperature dependence of l and C_V
- l is determined by geometrical scattering (static imperfections, surfaces) and scattering with other phonons (anharmonic interactions)
- At very low T ($\approx 10K$), atomic displacements are very small and atoms only vibrate the bottom of their potential wells and hence anharmonic terms are negligible. l becomes large and is determined by surfaces and impurities. Because the distances are fixed, l is temperature independent (and on the order of the specimen size).
- At high T, l is dependent on the number of phonons that can interact, and therefore varies inversely with the number of phonons $\sim \frac{k_B T}{\hbar \omega}$ i.e. $\frac{1}{T}$ dependence.
- Consider κ_L , where $\kappa = \frac{1}{3} C_V v l$
- At low T, l is constant and the temperature dependence is determined by C_V which goes as T^3 (the Debye result).
- At high T, C_V has very little temperature dependence, and so K_L will vary as $\frac{1}{T}$
- Intermediate T range, (T below about θ_D) κ_L rises more steeply with falling T **despite** C falling in this region. This is because the contribution to κ from collisions involving the Umklapp process still dominate here, limiting l . Two phonons with $\varepsilon \sim \frac{1}{2} k_B \theta_D$ are required to create a scattered phonon with $k > /frac{\pi}{a}$. Since the number of these high energy phonons decreases exponentially with temperature, we may expect l to increase exponentially with T .



30.4 Metals: Electron scattering

- In metals only electrons close to the Fermi sphere can participate in scattering - other collisions are forbidden as they would scatter electrons into occupied states.
- The scattered electrons have energy ε_F , velocity v_F and a mean free path l
- The value of τ_{el} is determined by collisions. This is dominated by the influence of static defects (impurities and imperfections) at low temperatures (T-independent) and by thermal phonons and thermal electrons at high temperatures (where τ decreases with increasing T)
- And also remember that these collisions mean that a dc electric field does not produce a steadily increasing current. We define the drift velocity as the additional velocity associated with the departure from thermal equilibrium given by the Fermi distribution function - all the electrons gain the drift velocity and so the whole Fermi sphere is displaced. $v_d = \frac{-e\tau}{m_e^*} E$

30.5 Mobility and resistivity in metals

- Electric current density:

$$j = n(-e)v_d = \frac{ne^2\tau}{m_e^*} E = \sigma E \rightarrow \boxed{\sigma = \frac{1}{\rho} = \frac{ne^2\tau}{m_e^*} = ne\mu_e}$$

- As we've already discussed (Mattheissen's rule) the rates of different scattering mechanisms are independent. So $\frac{1}{\tau} = \sum_j \frac{1}{\tau_j}$ or $\rho = \sum_i \rho_i$
- So $\rho = \rho_i + \rho_L(T)$ (resistivity due to scattering by static impurities plus the scattering from phonons)
- The **residual resistance** ratio $\frac{\rho_{T=300K}}{\rho_{T=4.2K}}$ is a measure of the *sample purity*. Note that $\rho_{L,T=0K} = 0$ and $\rho_{L,T>\theta_D} \propto T$ while in the intermediate range $\rho_{|,0 < T < \theta_D} \propto T^5$

30.6 Heat capacity of metals

- Recall that at temperatures much below the Debye temperature and the Fermi temperature, the heat capacity of metals may be written as the sum of electron and phonon contributions: $C_V = \gamma T + \alpha T^3$
- Because of the heat transport by conduction electrons, the thermal conductivity of metals is normally much larger than that of non-metals. (Note that the thermal conductivity due to phonons in metals is generally less than that in non-metals due to the scattering of phonons by the electrons). Therefore, the flow of heat is by the *same* mechanism as in a real gas.
- So using kinetic theory: $\kappa_e = \frac{1}{3} C_V v_F l$

- Using $l = v_F\tau$, $\varepsilon_F = \frac{1}{2}mv_F^2$ and $C_{V,el}$ (the contribution from electrons):

$$\kappa = \frac{1}{3} \frac{1}{2} \pi^2 n k_B^2 T \frac{v_F^2 \tau}{\frac{1}{2} m_e^* v_F^2} = \frac{\pi^2 n k_B^2 T \tau}{3 m_e^*}$$

30.7 The Wiedemann-Franz Law

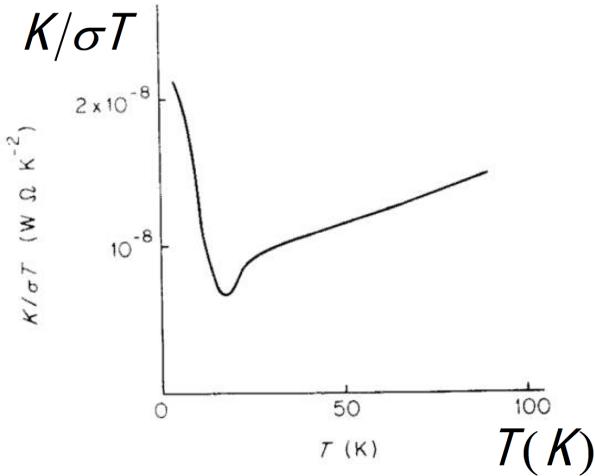
Recall:

$$\kappa = \frac{\pi^2 n k_B^2 T \tau}{3 m_e^* e}, \quad \sigma = \frac{n e^2 \tau}{m_e^*}$$

- The ratio of $\frac{\kappa}{\sigma}$ is independent of the electron gas parameters:

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

- Consider the graph below. Note that the Lorenz number L is significantly below prediction, but tends towards predicted values as T tends to zero, where collisions with impurities are the dominant scattering mechanism.



- **Breakdown of the Wiedemann-Franz Law** happens in the range from a few Kelvin up to θ_D , and is a result of the failure of the assumption that the collision times limiting the flow of electric and heat currents are the same.
- An electric current is produced by more electrons travelling one way than the other. To reduce it, electrons have to change directions (scatter must produce large changes in momentum, \vec{k}).
- Conversely, a thermal current is produced by electrons moving in one direction being more energetic than electrons travelling in the other. To reduce it, only small changes in momentum are required.

31 Breakdown of the Wiedemann-Franz Law

31.1 Scattering events that limit electrical conductivity

- Displacement of the Fermi sphere due to E-field is given by:

$$\delta k = \frac{F\tau}{\hbar} = \frac{-e\tau}{\hbar} E$$

- The displacement is very small.
- For the current to decay to zero, the Fermi sphere must relax (i.e. collisions must occur that remove an electron from the right (assuming rightward displacement) and add one to the left side)
- These transitions involve a change of the electron k-vector of magnitude $\sim 2k_F$

31.2 Scattering events that limit thermal conductivity

- In a T-gradient, the finite temperatures mean there are some vacant states below the Fermi surface and some occupied states above.
- If the material is hotter at the left of a sample, then electrons moving from left to right across the sample (i.e. those with $k_x > 0$) will have a distribution corresponding to a higher temperature than those coming from right to left ($k_x < 0$) (as shown in k-space, the Fermi surface is more blurred for positive values of k).
- The difference in blurring due to the T-gradient corresponds to the temperature difference in a mean free path ($\sim 100\text{nm}$ at 300K)
- The scattering processes effective in limiting thermal conductivity are ones which even out the blurring of the Fermi surface. This can involve large or small momentum changes due to the wrap-around from U-processes.
- Since more scattering processes are effective in limiting thermal conduction, the relaxation time for thermal conductivity is shorter and the Lorenz number falls below the Wiedemann-Franz value.
- To explain this further and investigate the temperature dependence of thermal and electrical conductivities, we must study the scattering of electrons by phonons.
- The maximum energy change of an electron is therefore the maximum phonon energy $\sim k_B\theta_D$
- It is therefore much less than the energy of an electron on the Fermi surface $k_B T_F$
- Therefore since a colliding electron must scatter in to a vacant state, and vacant states only occur close to the Fermi surface, then only electrons close to the Fermi surface can be scattered by phonons.

- At $T \gg \theta_D$ phonons have energy $k_B\theta_D$ and have enough momentum to cause large scale momentum transfer collisions and so electrical resistance is produced.
- And so at high temperatures the relaxation of thermal and electrical resistances are similar and the Wiedemann-Franz law is obeyed well.
- Since $\tau \propto \frac{1}{T}$ then $\rho \propto T$ (recall $\sigma = \frac{1}{\rho} = \frac{ne^2\tau}{m_e^*}$)
- Inserting $\tau \propto \frac{1}{T}$ in to the equation for κ ($\kappa_e = \frac{\pi^2 n k_B^2 T \tau}{3 m_e^*}$) Predicts that the thermal conductivity is independent of temperature at high temperatures.
- So at high temperatures the WF law works since $\kappa\rho \propto T$
- At the lowest temperatures electron-impurity collisions are dominant. They are elastic and unable to produce small momentum changing collisions but can produce large momentum changing collisions.
- So the effective scattering times for electrical and thermal conductivities are identical and the WF law is obeyed.
- The electron mean free path due to electron-impurity scattering is independent of temperature, leading to a temperature independent electrical conductivity and a thermal conductivity proportional to T .
- So the WF Law works for low T as $\kappa\rho \propto \rho$
- **At intermediate temperatures** (still $T \ll \theta_D$) the average phonon energy $\sim k_B T$, and the total energy of phonons is proportional to T^4 and so the phonon number (the total lattice energy divided by the average phonon energy) is proportional to T^3 .
- Phonons of energy $k_B T$ have just the energy required to take part in scattering events to produce thermal resistance.
- The corresponding mean free path and scattering time determining the thermal conductivity are therefore inversely proportional to the phonon number: $\tau_{\text{th}} \propto T^{-3}$
- So the phonon-limited thermal conductivity due to electrons will exhibit $\kappa \propto T^{-2}$ for low T.
- Now considering the electrical resistance at $T \ll \theta_D$
- A phonon of energy $k_B T$ has too little momentum to cause the large scale momentum changes to produce electrical resistance. Large momentum changes will only be able to occur via the addition of many small changes.
- Because the initial and final electron states must be close to the Fermi surface, then a small momentum change implies scattering through a small angle.

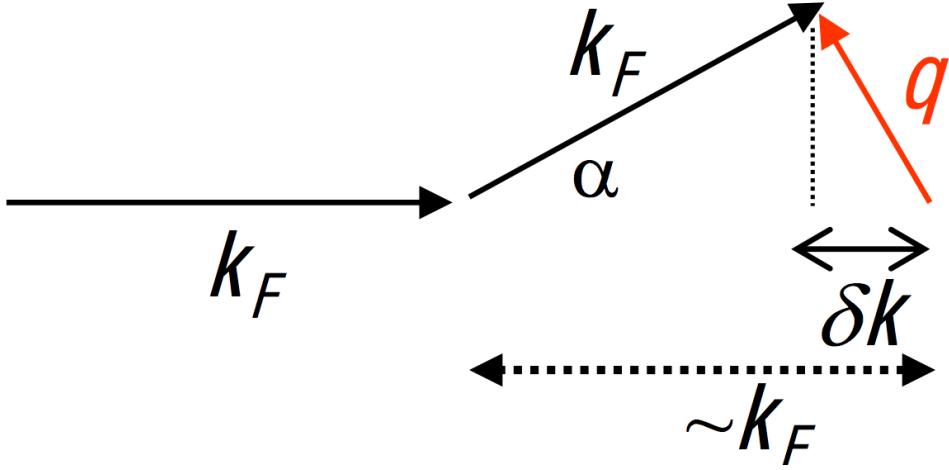


Figure 6: The change in momentum of an electron when it is scattered through an angle α by the absorption of a phonon. Note that the magnitude of the electron momentum does not change, as it remains close to the Fermi surface.

- The loss of momentum $\hbar\vec{\delta k}$ of the electron in its original direction tells us about the effectiveness of the collision in producing electrical resistance, which is equal to $\frac{\hbar q^2}{2k_F}$ for small angle α .
- $\frac{\hbar q^2}{2k_F}$ is proportional to T^2 for typical phonons at low temperatures.
- Hence the scattering rate $\frac{1}{\tau_{\text{el}}}$ to be used in calculating the resistivity must be weighted by an effectiveness factor of order $\left(\frac{T}{\theta_D}\right)^2$ so that:

$$\frac{1}{\tau_{\text{el}}} \sim \left(\frac{T}{\theta_D}\right)^2 \frac{1}{\tau_{\text{th}}} \propto T^5$$

so:

$$\rho_i(T) \propto T^5$$

- So the WF law fails in the mid-temperature region.

