PH 409: Introduction to Condensed Matter Physics

Course instructor : Kantimay Das Gupta
Lectures are held: Mon (10.35-11.30), Tue (11.35-12.30), Thu (8.30-9.25) [IC 1]

Course Contents:

1. Describing condensed matter microscopically
    - structure factor
    - lattice and fourier transforming of densities.
    - Unit cell concept
    - Xray scattering
2. The free electron gas in metals
    - Drude theory of transport
    - Thermodynamic quantities
    - response to electromagnetic fields
3. Bloch's theorem
    - Gap at zone boundaries and reduced zone scheme. Draw it with free electrons, show that in the second zone, filled states are not near the zone center.
    - Band structure as a perturbation problem.
    - Classification of metals/insulators/semiconductors
    - Kronig-Penny model of a solid.
4. Tight binding method
    - Linear Combination of Atomic Orbitals as the way to build a molecule from atoms.
    - ID example, 2D Graphene, 3D, some lattices like cubic/BCC/FCC.
    - What are the shortcomings. APW and then e-e interaction.
    - Mott insulators (NiO) if possible.

    - Revision of Brillouin zone. Drawing the free electron Fermi surface. The Fermi level in Alkali metals, divalent/trivalent metals.
    - Why some metals can give hole-like Hall voltage.
5. semiclassical dynamics of electrons in a band.
    - Introduce the equations, write Bloch oscillation condition, group velocity near zone boundary, why k → k-G flipping does not give unphysical results?
    - Concept of holes: what is the correct k vector of a hole, equation of motion, group velocity etc.
    - Concept of orbits in a magnetic field (electron like, hole like and open)
    - Concept of holes
    - de-Haas van Alphen oscillation mechanism.
    - Optical absorption and dHvA as two key methods of band structure determination
6. Phonons
    - specific heat (lattice + free electron) ,
    - thermal conductivity Bloch-Gruniessen if possible.

- One example of inelastic neutron scattering (perhaps from liq. Helium)

7. Chapter on bonding: electrovalent, covalent solids

8. Magnetism

    - Pauli paramagnetism in the electron gas
    - writing the free energy of an electron gas in magnetic field
    - diamagnetism,
    - ferromagnetism
    - exchange interaction , ferro-antiferro, hysteresis,
    - spin wave magnon,

9. Special topics (may be modified...)

    - Conservation laws in a crystal
    - fluctuation, dimensionality and order
    - Collective modes/excitations

10. disordered material

    - glass, colloids etc.
    - Introduce g(r) and its significance

BOOKS: There are a (very!) large number of books on introductory Condensed matter physics. Two of the most commonly used ones at this level are probably

1. C. Kittel, *Introduction to Solid State Physics*

2. N.W. Ashcroft and N.D. Mermin, *Solid State Physics*

There are many more texts like:

1. J. M. Ziman, *Principles of the Theory of Solids*

2. P. M. Chaikin and T. Lubensky, *Principles of the Condensed Matter Physics.* (This has a focus on statistical and ”soft” matter)

3. N.F. Mott and H. Jones, *The Theory of the Properties of Metals and Alloys* (The early development of the subject with very close attention to interpreting experimental data.)

From time to time I will suggest other references during the lectures.

Evaluation :

1. Class Quiz : 15%

2. Mid sem : 30%

3. Class Quiz : 15%

4. End sem : 40%

This material is not a book or the complete course. It gives you some guidelines only. You should always refer to the recommended books to see the standard development and data relating to a chapter.

Contents

Chapter 1

Describing a collection of atoms: Structure factor and lattice

Lecture notes for PH409 (Jul-Dec 2013): K. Das Gupta

In elementary quantum mechanics, we usually work with a single particle. In condensed matter physics - we need to describe lots of particles together. The very first thing we need to do is to develop a language to do so. We will first develop some methods to describe a *static* collection of atoms.

1.1 Fourier transform and Structure factor

It so turns out that the way to do so is to use the concept of Fourier transform. We first assume that each particle (*i.e.* atom) is like a point mass and can be described by a delta function. So the density of matter around any point $\mathbf{r}$ can be written as

$$\rho(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \tag{1.1}$$

where the sum runs over all particles in the system. What is the fourier transform of this density?

$$\begin{aligned} \rho(\mathbf{k}) &= \int\limits_{\substack{\text{all}\\ \text{space}}} \mathrm{d}^3\mathbf{r}\; e^{i\mathbf{k}.\mathbf{r}} \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \\ &= \sum_i e^{i\mathbf{k}.\mathbf{r}_i} \end{aligned} \tag{1.2}$$

For $\mathbf{k} = 0$ eqn 1.2 predicts a large peak whose magnitude is equal to the number of particles in the system. We are more interested in the modulus of the (complex) function $\rho(\mathbf{k})$, which is :

$$|\rho(\mathbf{k})|^2 = \sum_{i,j} e^{-i\mathbf{k}.\mathbf{r}_i}\; e^{i\mathbf{k}.\mathbf{r}_j} \tag{1.3}$$

Once again, if we define it this way we will get a huge peak at $\mathbf{k} = 0$, whose magnitude would be N^2, so we redefine the expression in a slightly different way to take care of this peak

$$S(\mathbf{k}) = \frac{1}{N}\left(\left| \sum_i e^{i\mathbf{k}.\mathbf{r}_i} \right|^2 - N^2\delta(\mathbf{k}) \right) \tag{1.4}$$

The function $S(\mathbf{k})$ contains a lot of information about the ordering present in the material and is also normalised by the total density.

How does the quantity look for a gas, liquid, crystalline solid, glass/ amorphous solid (discuss...)? We see a lot more of amorphous solids and gases and liquids around us, compared to crystalline solids. It turns out though, that the crystalline solids are mathematically easiest to describe.

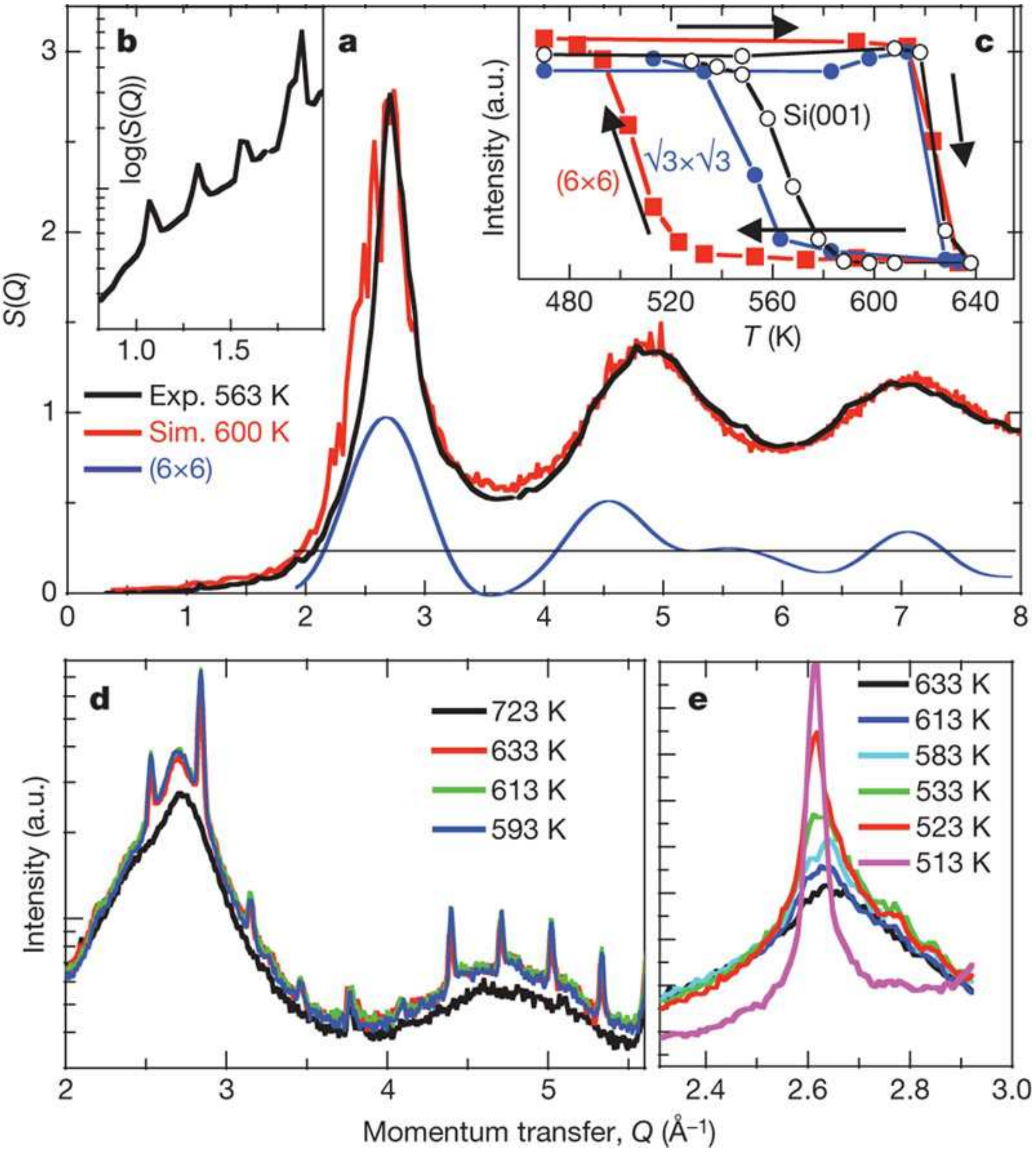


Figure 1.1: Evolution of the structure factor during cooling and solidification of Pb. The data is taken from U Schülli *et al.* *Nature* **464**, 1174-1177 (2010) doi:10.1038/nature08986. Don't worry too much about the details, but notice that the liquid is not a completely structureless thing, the same is true for glasses also. We will see soon that a crystalline solid will have many prominent peaks in its $S(k)$, rather than just a few.

1.1.1 What is a crystal?

A crystalline solid means that the positions of the particles are given by

$$\begin{aligned} (\text{ 1D }) \qquad \mathbf{r}_\mathrm{i} &= n_1\mathbf{a}_1 \qquad (1.5)\\ (\text{ 2D }) \qquad \mathbf{r}_\mathrm{i} &= n_1\mathbf{a}_1 + n_2\mathbf{a}_2 \qquad (1.6)\\ (\text{ 3D }) \qquad \mathbf{r}_\mathrm{i} &= n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \qquad (1.7)\end{aligned}$$

where $\mathbf{a}_1$,$\mathbf{a}_2$,$\mathbf{a}_3$ are *fixed* vectors and n_1, n_2, n_3 are *any* integers. An important point is that for a collection of points to qualify as a crystal all points of the form in the sets of equation must belong to the collection. We cannot say that only *some* points of the specified form are there and some are not there! Notice that this idealised definition (due to Bravais) implies that the crystal must be infinite. This of course cannot be true.

The set of vectors $\mathbf{a}_1$,$\mathbf{a}_2$,$\mathbf{a}_3$ are called the primitive (translation) vectors of the lattice. They are not necessarily orthogonal but must be linearly independent. The line segment (in 1D), parallelogram (in 2D) or parallelopiped (in 3D) formed by the primitive vectors is called the unit cell of the crystal.

It follows from the definitions 1.5, 1.6 and 1.7 that the addition (or subtraction) of two lattice vectors *must* produce another lattice vectors. If we can locate two lattice vectors, which when added leads to a point that does not belong to the original set of points, then the set of points do not form a Bravais lattice. See the example in Fig. 1.2 :

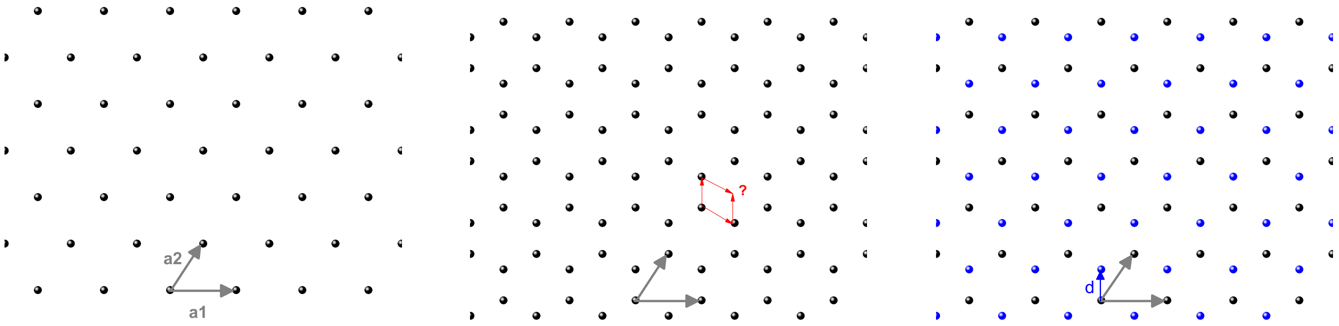


Figure 1.2: (left)The lattice has unit vectors $\mathbf{a}_1 = (1, 0)$ and $\mathbf{a}_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$. (middle)But the set of points in the middle is not a Bravais lattice because addition of the two lattice vectors marked in red, leads to a point that is not in the set. (right) But then the same set of points can be treated as a lattice with a basis set of points at $(0, 0)$ and $(0, \frac{1}{\sqrt{3}})$

PROBLEM: Consider two points in a lattice given by:

$$\begin{aligned}\mathbf{r}_m &= m_1\mathbf{a}_1 + m_2\mathbf{a}_2 + m_3\mathbf{a}_3\\ \mathbf{r}_n &= n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3\end{aligned}$$

Show that the distance between them is given by the ”metric” like expression:

$$d^2 = \mathbf{s}^T M \mathbf{s} \qquad (1.8)$$

where

$$M = \begin{pmatrix} \mathbf{a}_1.\mathbf{a}_1 & \mathbf{a}_1.\mathbf{a}_2 & \mathbf{a}_1.\mathbf{a}_3\\ \mathbf{a}_2.\mathbf{a}_1 & \mathbf{a}_2.\mathbf{a}_2 & \mathbf{a}_2.\mathbf{a}_3\\ \mathbf{a}_3.\mathbf{a}_1 & \mathbf{a}_3.\mathbf{a}_2 & \mathbf{a}_3.\mathbf{a}_3 \end{pmatrix} \qquad (1.9)$$

and $\mathbf{s} = (m_1 - n_1, m_2 - n_2, m_3 - n_3)$. Then show that:

$$\det M = (\mathbf{a}_1.\mathbf{a}_2{\times}\mathbf{a}_3)^2 \qquad (1.10)$$

Lattice and basis

To every lattice point, we can attach whatever we want - provided it is the *same* thing that we attach to every point. Whatever is attached to each lattice point is called the basis. The simplest basis is of course just the lattice point itself. But it can be as complex as we want to make it!

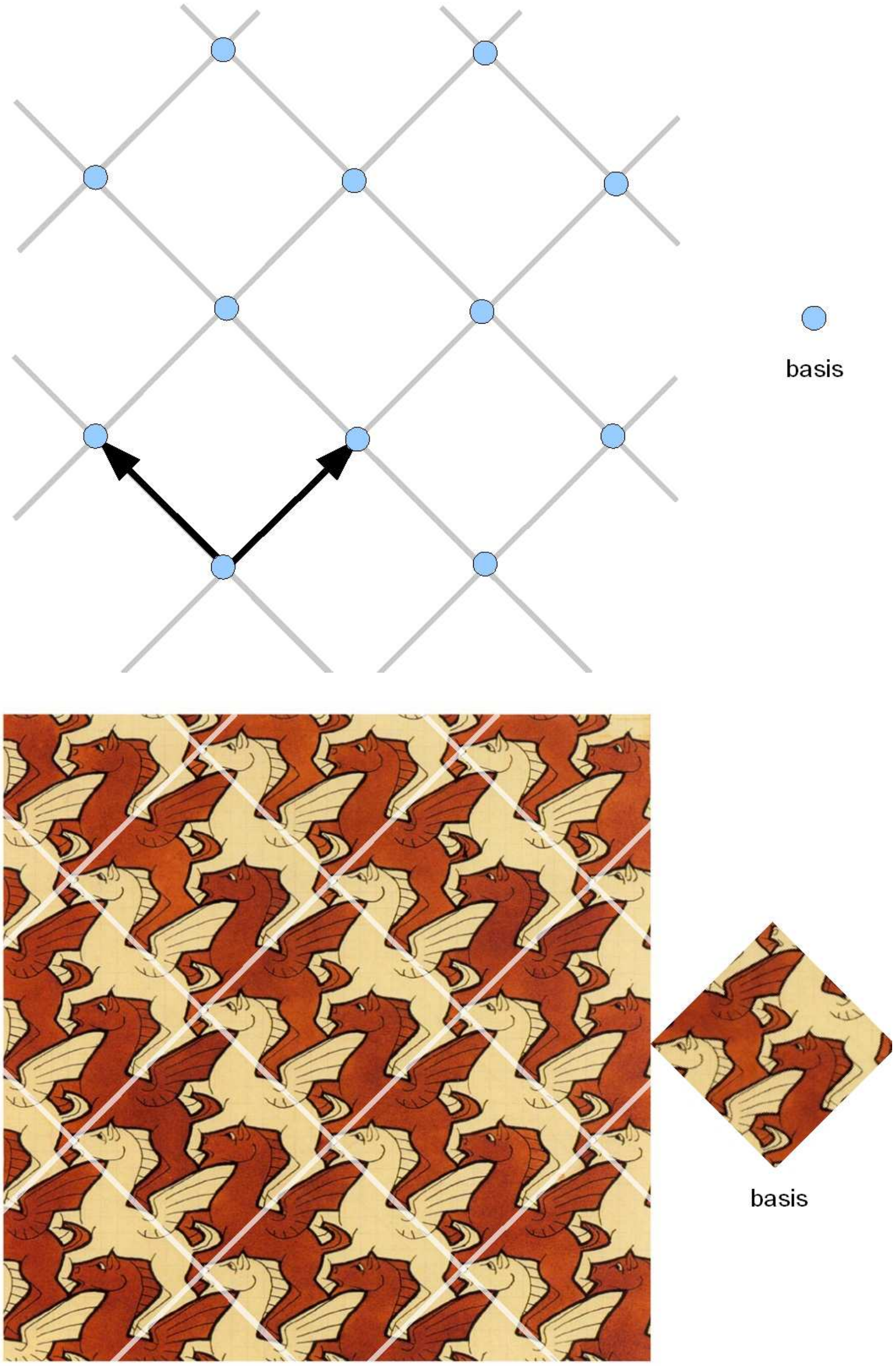


Figure 1.3: The basis can be trivial or complex. Convince yourself that the choice of the basis is not unique, though in many cases there will be a most sensible choice!

1.1.2 Fourier transform of the lattice points

How would $\rho(\mathbf{k})$ look for the lattice?

$$\begin{aligned}
\rho(\mathbf{k}) &= \sum_i e^{i\mathbf{k}.\mathbf{r}_i} \\
&= \sum_{n_1,n_2,n_3} e^{i\mathbf{k}.(n_1\mathbf{a}_1+n_2\mathbf{a}_2+n_3\mathbf{a}_3)} && (1.11) \\
&= \sum_{n_1} e^{i\mathbf{k}.\mathbf{a}_1} \sum_{n_2} e^{i\mathbf{k}.\mathbf{a}_2} \sum_{n_3} e^{i\mathbf{k}.\mathbf{a}_3} && (1.12) \\
& && (1.13)
\end{aligned}$$

Now suppose that we have some $\mathbf{k}$ vectors for which the following hold:

$$\begin{aligned}
\mathbf{k}.\mathbf{a}_1 &= 2\pi \times \text{some integer} \\
\mathbf{k}.\mathbf{a}_2 &= 2\pi \times \text{some integer} \\
\mathbf{k}.\mathbf{a}_3 &= 2\pi \times \text{some integer} && (1.14)
\end{aligned}$$

Clearly if $\mathbf{k}$ satisfies this, then any integer multiple of $\mathbf{k}$ would also satisfy this. The magnitude of $\rho(\mathbf{k})$ will be large whenever this happens because all the exponentials will add up in phase.

Now comes a very important point : *these vectors are sharply defined*. Let's look at this using the 1D case, which is easy to plot out without having to bother about vector components. In this case we can write

$$\begin{aligned}
\rho(k) &= \sum_n e^{ikr_n} \\
&= \sum_n e^{ika_0n} \\
&= N \quad \text{when} \quad ka_0 = 2\pi
\end{aligned}$$

But how much is $\rho(k+\delta k)$, when we move away a little bit from the peak?

$$\begin{aligned}
\rho(k+\delta k) &= \sum_n e^{i(k+\delta k)r_n} \\
&= \sum_n e^{ika_0n} .e^{i\delta ka_0n} \\
&= \sum_n e^{i\delta ka_0n}
\end{aligned}$$

Let's plot $\left|\frac{\rho(\frac{2m\pi}{a_0}+\delta ka_0)}{N}\right|$ for $N = 100, 500, 1000, 5000$ etc. Notice how the phase of each term starts changing rapidly and cancelling on average as N becomes large. (Problem: The sum can be done analytically also.....)

The structure factor of a crystalline solid thus has a series of sharp peaks. Liquids have some partially developed broad peaks, gases have none.

PROBLEM: If the lattice has a basis attached to each point, the "density" would be:

$$\rho(\boldsymbol{r}) = \sum_{\boldsymbol{R}_i} f(\boldsymbol{r}-\boldsymbol{R}_i) \tag{1.15}$$

$$\text{Then show that :} \quad \rho(\boldsymbol{k}) = \left(\sum_i e^{i\boldsymbol{k}.\boldsymbol{R}_i}\right) f(\boldsymbol{k}) \tag{1.16}$$

where the summation is something we have seen before and $f(\boldsymbol{k})$ is the fourier transform of $f(\boldsymbol{r})$ Can you comment on the significance of the result?

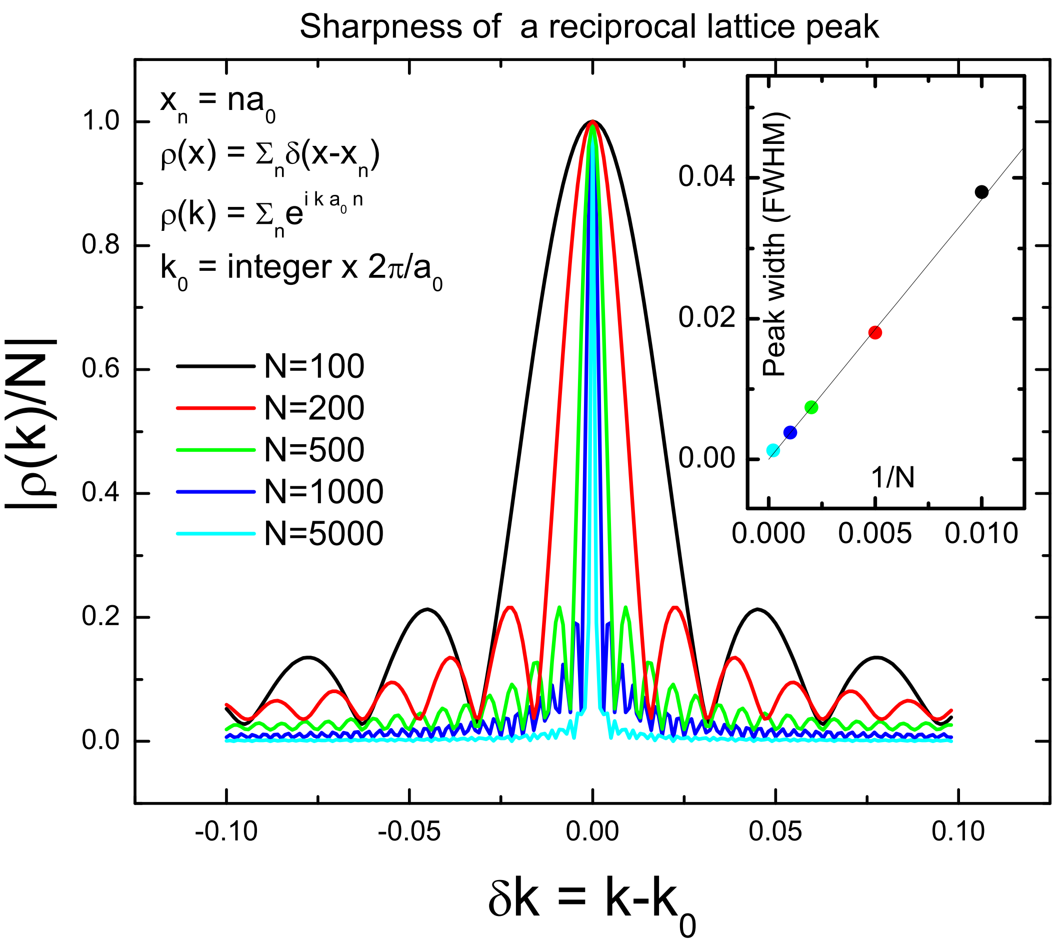


Figure 1.4: Notice how the sharpness of the peak grows with the size of the sample. We will see later that the width of such peaks, which can be experimentally measured by X-ray diffraction, can be a measure of the grain-size of the material.

1.2 The reciprocal lattice

Now, back to the question of Fourier components which contribute to the lattice. We shall denote them with **G**. Clearly they need to satisfy for *any* lattice vector **r**

$$\mathbf{G}.\mathbf{r} = 2\pi \times \text{ any integer} \tag{1.17}$$

There is a very elegant way of writing down the set of all vectors which satisfy the condition 1.17. Define

$$\begin{aligned} \mathbf{b}_1 &= 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1.\mathbf{a}_2 \times \mathbf{a}_3} \\ \mathbf{b}_2 &= 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1.\mathbf{a}_2 \times \mathbf{a}_3} \\ \mathbf{b}_3 &= 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1.\mathbf{a}_2 \times \mathbf{a}_3} \end{aligned} \tag{1.18}$$

Neither ($\mathbf{a}_1$,$\mathbf{a}_2$,$\mathbf{a}_3$) nor ($\mathbf{b}_1$,$\mathbf{b}_2$,$\mathbf{b}_3$) need to form an orthogonal set. But they satisfy

$$\mathbf{a}_i.\mathbf{b}_j = 2\pi\delta_{ij} \tag{1.19}$$

It is straightforward to see that any vector of the form

$$G_{hkl} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3 \tag{1.20}$$

would satisfy the condition 1.17. The set of points with integer *h,k,l,* as defined clearly form a lattice. This lattice is called the reciprocal lattice of the original lattice.

PROBLEM : Show that the reciprocal of the reciprocal lattice is again the direct lattice, as one might expect.

Some mathematics texts define these vectors without the factor of 2π. In which case the condition 1.17, should also not contain the 2π factor. It is then (usually) called the dual lattice.

1.2.1 Specifying directions and planes in a lattice

We need to know two more geometrical aspects of the lattice. We have this infinite collection of points. We need to be able to specify a direction and specify a plane (take a cut) through these points. How does one do these?

Defining hkl planes: Miller indices

Recall condition 1.17. This definition allows us to classify all points in the direct lattice into non-intersecting sets, in a way that no point will be left out at the end. We fix a reciprocal lattice vector $\mathbf{G}_{hkl}$, and an integer n_0 then form the set of all points for which $\mathbf{G}_{hkl}.\mathbf{r}_j = 2\pi \times n_0$.

The set of points belong to a plane and the vector $\mathbf{G}_{hkl}$ is perpendicular to the plane. This is very simple to show,

$$\begin{aligned} \mathbf{G}_{hkl}.\mathbf{r}_j &= 2\pi \times n_0 \\ \therefore G_x x + G_y y + G_z z &= 2\pi \times n_0 \end{aligned} \tag{1.21}$$

which is the equation of a plane in cartesian co-ordinates. Taking the gradient of the LHS of eqn 1.21 yields the direction perpendicular to the plane, which is clearly the vector with components (G_x, G_y, G_z).

Now we can choose some other n_0 and we will get another plane parallel to the original plane. These set of planes are called the *hkl* planes. The perpendicular to this family of planes is $\mathbf{G}_{hkl}$.

We also ensure that *hkl* ar relatively prime and reduced to the smallest possible set of integers. For example 123 and 246 define the same family, we will always use 123 to denote the set. This is called the Miller index of a plane.

The convention is to reduce the set of numbers to the smallest possible integers and then denote the family by (hkl). It is possible that some of the integers will be negative - such intercepts are denoted with a bar on top. for example a plane may cut $\mathbf{a}_1$ at -1, $\mathbf{a}_2$ at 2 and $\mathbf{a}_3$ at 3. Then we will denote this family as $(\bar{6}, 3, 2)$

No point in the direct lattice has been left out in the process, because condition 1.17 holds for all points of the lattice, so any point must belong to some plane in any *hkl* family.

PROBLEM : Show that an *hkl* plane intersects the primitive axes formed by $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ making intercepts in the ration $1/h : 1/k : 1/l$

Spacing of the hkl planes

We know that $e^{i\mathbf{G}.\mathbf{r}}$ is also the equation of a plane wave, with wavevector $\mathbf{G}$ propagating in the direct lattice. The surfaces over which $\mathbf{G}.\mathbf{r}$ is a constant are the wavefronts. The discussion in the preceeding section implies that the hkl planes are precisely such wavefronts. Therefore the spacing between two successive hkl planes, for which the constant n_0 in eqn 1.21 differs exactly by 1, must be the wavelength associated with $\mathbf{G}$. We thus get a very important result -if d_{hkl} denotes the distance between two successive planes, then

$$d_{hkl} = \frac{2\pi}{|\mathbf{G}|} \tag{1.22}$$

Density of points on an hkl plane

The density of points (per unit area) in an hkl plane is

$$n_{hkl} = \frac{d_{hkl}}{V} \tag{1.23}$$

where $V = \mathbf{a}_1.\mathbf{a}_2{\times}\mathbf{a}_3$ is the volume of the primitive unit cell. The proof is left as an exercise.

PROBLEM : Prove eqn. 1.23

These two results will come in very handy when we discuss diffraction from a lattice.

Direction in a lattice

If a vector points along

$$\mathbf{r} \quad = \quad n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$$

then we denote this direction as $[n_1, n_2, n_3]$. Note the use of different brackets to avoid confusion with Miller indices.

Chapter 2

Diffraction and basics of crystal structure

Lecture notes for PH409 (Jul-Dec2013): K. Das Gupta

References:

1. Chapter 5 *Solid State Physics*, N. W. Ashcroft and N.D. Mermin
2. Chapter 1 & 2 *Introduction to Solid State Physics, C. Kittel*
3. Useful (free) software for visualising crystal structures:
    - XCrysden : http://www.xcrysden.org
    - POV-ray : http://www.povray.org
    - Crystosim : http://home.iitk.ac.in/ sangals/crystosim/finalsolution/startpage.html

---

How does one *know* about the structure of a solid or a liquid or a glass? There is a generic answer to this question. We take a small piece of the material. We then send a beam of X-ray or neutrons or something else at the material. We measure what comes out. Because electromagnetic waves and a beam of particles can be treated in a similar way the experiment essentially involves a single question. If we throw a wave with wavevector $\mathbf{k}_i$ at the material what is the probability that the same beam/particle will come out (make a transition to) with wavevector $\mathbf{k}_f$? We will give a ver general answer to this question and show that the answer always involves the structure factor of the material which scattered the object thrown at it. But this we will do at the end of this chapter. We will do a few specific cases first and then come back to this.

You may also ask whether diffraction tells us something about the structure that a probe like Scanning microscope cannot? The answer is that all scanning probe techniques sample the surface (or at best the first couple of layers) whereas diffraction *can* (but not necessarily does) sample the whole volume. The structure of the few layers near the surface is often very different from the bulk structure. This (called surface reconstruction) is not an anomaly but has clear reasons behind it. Though we will not get into this, it is worth remembering that the structure seen by an Scanning tunneling microscope (STM) or Atomic Force microscope (AFM) can be different from what a volume-sampling X-ray diffraction would suggest.

2.1 Bragg diffraction

Historically X-ray diffraction was first used by Bragg to determine the structure of crystals. Recall what we learnt about the (hkl) planes in chapter 1.2.

- *All* the atoms can be grouped into a set of parallel planes labelled by the perpendicular to the family of the planes (eqn. 1.20)

$$G_{hkl} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3$$

- No atom would be left out and the planes would be separated by (eqn. 1.22)

$$d_{hkl} = \frac{2\pi}{|\mathbf{G}|}$$

- If the numbers (hkl) get too big, then the density of points on each plane would be low and the separation would be very small as well. So we restrict ourselves to (hkl) numbers not more than $\sim$4 (typically).

Thus a beam of radiation can see the sample as a set of planes, and bounce off these planes. But would the reflections from the successively deeper planes be in phase and interfere constructively?

The simple answer is that if the path difference (CX_2C') in Fig. 2.1 is an integer multiple of the wavelength. This is satisfied, if

$$2d_{hkl}\sin\theta = n\lambda \tag{2.1}$$

We choose X-rays because they have wavelengths comparable to atomic spacing (typically few Å. If the spacing was much larger, we would have chosen longer wavelengths. Similar phenomena would occur if we shine a (visible) laser on "colloidal crystals".

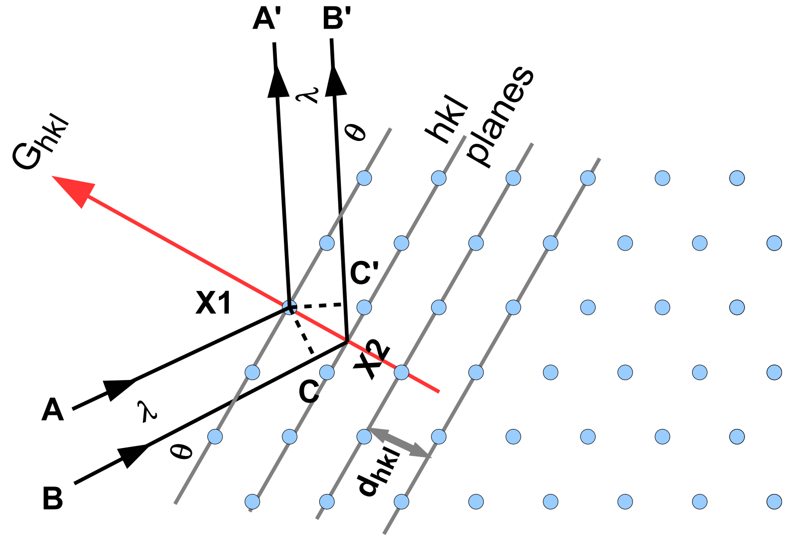


Figure 2.1: When would the successive reflections interfere constructively? This requires that each reflection from the successively deeper planes differ in "optical path" by a multiple of the wavelength of the radiation. In the figure X1C is ⊥ BX2 and X1C' is ⊥ on X2B'. Note that the outgoing beam suffers a deviation (from its original direction) equal to 2θ.

2.1.1 Powder diffraction : Debye Scherer "camera"

The Bragg condition eqn. 2.1 contains a sharp insight about when a crystal gives specular reflection. Doing the experiment in reality with a 3D crystal would give rise to spots corresponding to the reflections. However what if a single crystal is not in hand? Very often one gets poly crystalline sample or a powder, which means that the angles at which small bits of the crystal are arranged are random.

The direction of the incoming beam remains the same. Now to get a certain *hkl* reflection we can imagine rotating the $\mathbf{G}_{hkl}$ vector around the incoming beam, keeping θ constant. The reflections would then spread out on a cone as shown in the figure 2.2. The film would intercept these cones, from the film (detector signal peak) we can measure the half angle of the cone, which would be 2θ. Note that for this method to work the X-ray source must be monochromatic. A "monochromator" would usually filter the radiation coming into the sample, (not shown in the figure). This could simply be a cleanly cut single crystal (*e.g.* Germanium), giving a sharp reflection at a known angle. The broad circular patches in the filmstrip are the entry and exit points of the beam.

A plot of the intensity against the deviation (2θ) and we get results like:

- Knowing the angles and λ we calculate the d.
- Now if we have a guess about the structure, we try to match the reflections with certain (hkl) values.
- If our model accounts for all the reflections we consider the problem to be solved.

The sharpness of the peaks will depend on crystal "quality", grain size etc. See the example in Fig 2.4. It is possible to estimate the grain size of particles (if they are very small) from the FWHM ($\Delta\theta$) of the peaks. This is given by:

$$\Delta\theta \approx \frac{K\lambda}{L\cos\theta} \tag{2.2}$$

where K is a geometrical factor that depends on the shape of the grain ($K = 0.9$ for spheres), λ is the x-ray wavelength, L is grain diameter, θ is the (Bragg) angle at which the peak was seen. Recall the size effect we have seen before in summing the contribution to the fourier components of a 1D chain in chapter 1.

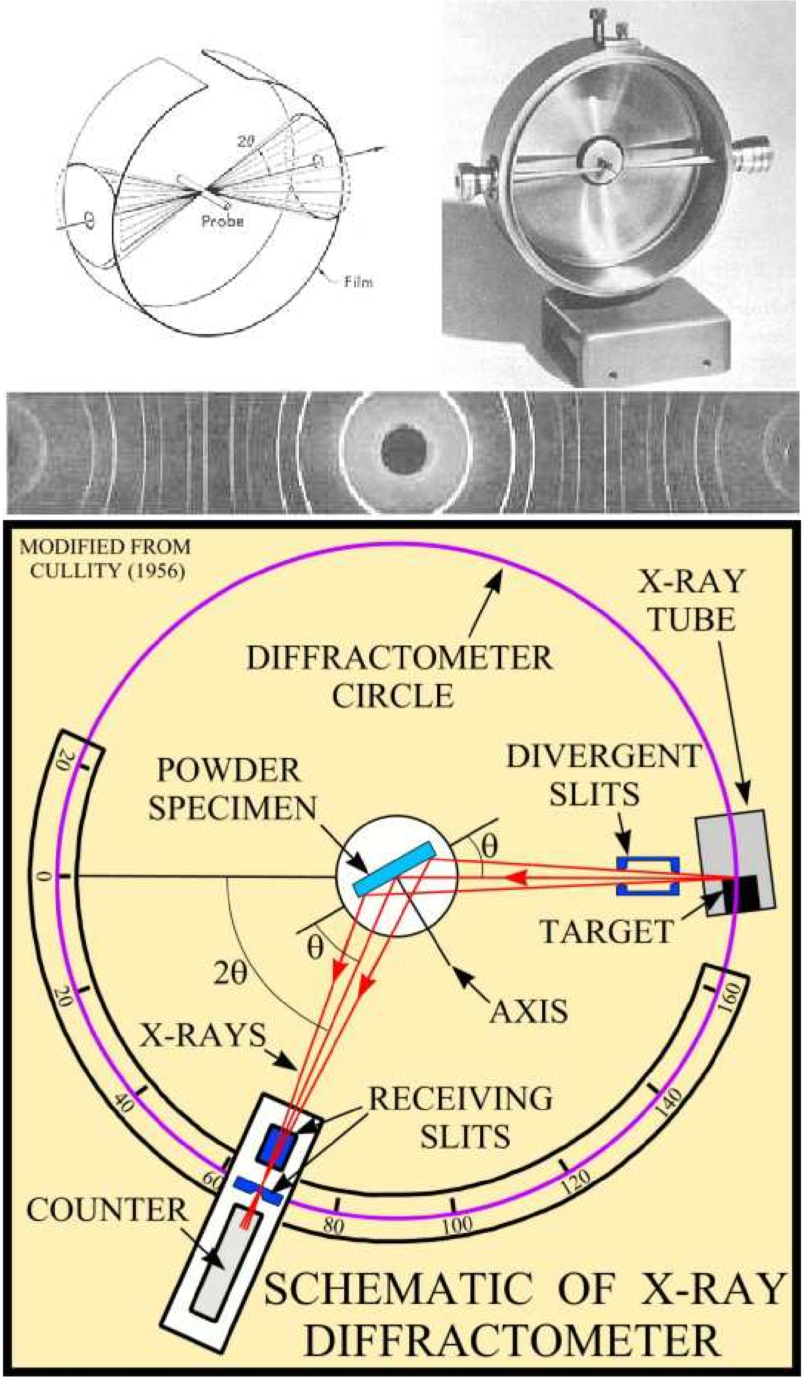


Figure 2.2: Debye-Scherer camera, the film would have been fitted in the inside of the circle, when they were first made. Now of course it would be an electronic detector scanning through the angle as shown in the next schematic. The photo taken form <http://www.tf.uni-kiel.de/matwis/amat/mw2_ge/kap_3/backbone/r3_5_1.html>

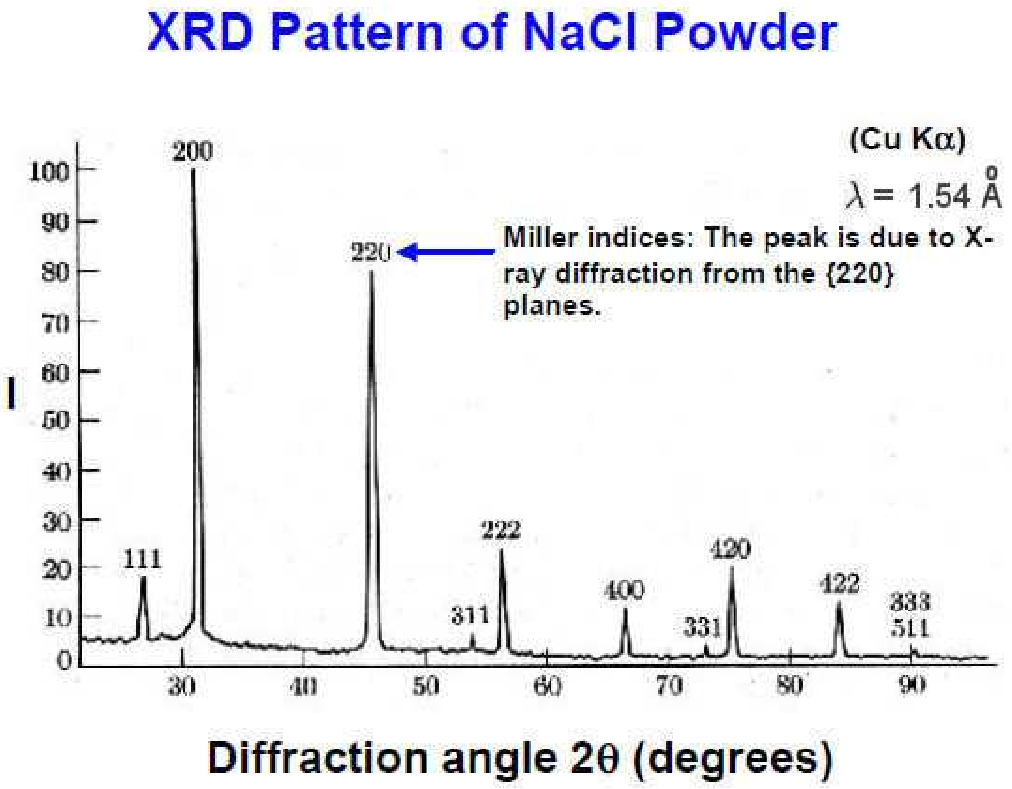


Figure 2.3: Bragg diffraction from NaCl. The figure is taken from <http://www.universe.review.ca>. The problem is to figure out the structure, given a data like this and some reasonable guesses.

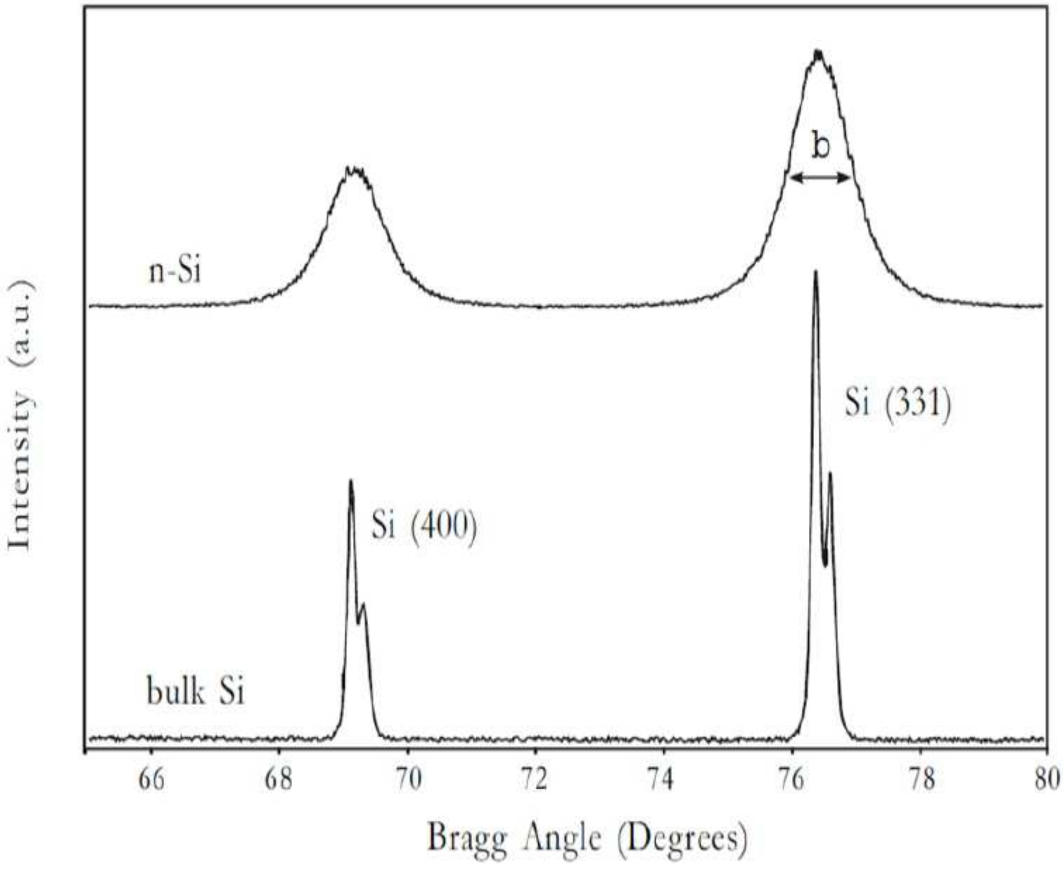


Figure 2.4: Bragg diffraction from bulk Silicon and nanocrystalline Silicon. Notice the broadening of the peaks. figure is taken from http://www.vanbokhoven.ethz.ch/education/XRD_excercises.

2.1.2 Handling the basis in the scattering problem

The Bragg condition, as we remarked, offers a sharp insight as to how a crystal reflects the X-ray. It is however, not sufficient. We said nothing about the basis, sort of implicitly assuming a monoatomic lattice in the previous section. We now consider the scattering problem a bit more formally.

A standard scattering problem in quantum mechanics is depicted in Fig. 2.5. An incoming plane wave hits a scattering center. The scattering process creates an outgoing spherical wave (here treated as a scalar wave). The sum total that the observer will see is the sum of the original and the scattered wave. The scattering is elastic so the magnitude of the wavevector does not change.
The observer (at $\mathbf{r}$) sees (we are ignoring normalisation at this point)

$$\psi(\mathbf{r}) \quad \sim \quad e^{i\mathbf{k_0}.\mathbf{r}} + S(\mathbf{k_0}, \mathbf{k})\frac{e^{ik|\mathbf{r}|}}{|\mathbf{r}|} \tag{2.3}$$

Elastic scattering implies

$$|\mathbf{k_0}| \quad = \quad |\mathbf{k}| \tag{2.4}$$

$$\mathbf{k} \quad = \quad k_0\frac{\mathbf{r}}{r} \tag{2.5}$$

We place the observer at large distances : $|\mathbf{r}| >> |\mathbf{r}'|$. Then in Born approximation:

$$S(\mathbf{k_0}, \mathbf{k}) \approx \int_{\text{all crystal}} \mathrm{d}^3\mathbf{r}' \; e^{-i\mathbf{k}.\mathbf{r}'}V(\mathbf{r}')e^{i\mathbf{k_0}.\mathbf{r}'} \tag{2.6}$$

This form will appear for all cases whenever the initial and final states can be represented by plane waves. So this will work for X-rays, neutron beams etc. The scattering potential can be written as a sum over each lattice point (unit cell):

$$V(\mathbf{r}') = \sum_n v(\mathbf{r}' - \mathbf{R}_n) \tag{2.7}$$

where each term is taken to be non-zero in the n^{th} unit cell. Due to the symmetry of the lattice ($\mathbf{R}_n$ are lattice vectors)they are identical in functional form. The interaction potential between the lattice and the x-ray is proportional to the electron density at the point.

Hence the Born approximation implies:

$$S(\mathbf{k_0}, \mathbf{k}) \quad = \quad \sum_n e^{i(\mathbf{k_0}-\mathbf{k}).\mathbf{R}_n} \int_{\text{unit cell}} \mathrm{d}^3\boldsymbol{\xi} \; e^{i(\mathbf{k_0}-\mathbf{k}).\boldsymbol{\xi}}v(\boldsymbol{\xi}) \tag{2.8}$$

$$\quad = \quad f(q)\sum_n e^{i\mathbf{q}.\mathbf{R}_n} \tag{2.9}$$

where ξ runs over the *unit cell* only and $\mathbf{q} = \mathbf{k_0} - \mathbf{k}$ is the change in wavevector due to the scattering process. The sum (eqn 2.9)would be significantly different from zero only if $\mathbf{q}$ is a reciprocal lattice vector. The measured intensity would be proportional to $|S(\mathbf{k_0}, \mathbf{k})|^2$ and hence the structure factor would come into the expression naturally. The relation between scattered intensity and structure factor is thus very general and forms the fundamental basis of many experimental methods.

PROBLEM : Show that the condition $\mathbf{k}-\mathbf{k}_0 = \mathbf{G}$ leads to the Brag condition deduced earlier : $2d\sin\theta = \lambda$

The unit cell has a few atoms at positions $\mathbf{d}_i$, then the scattered amplitude can be written as

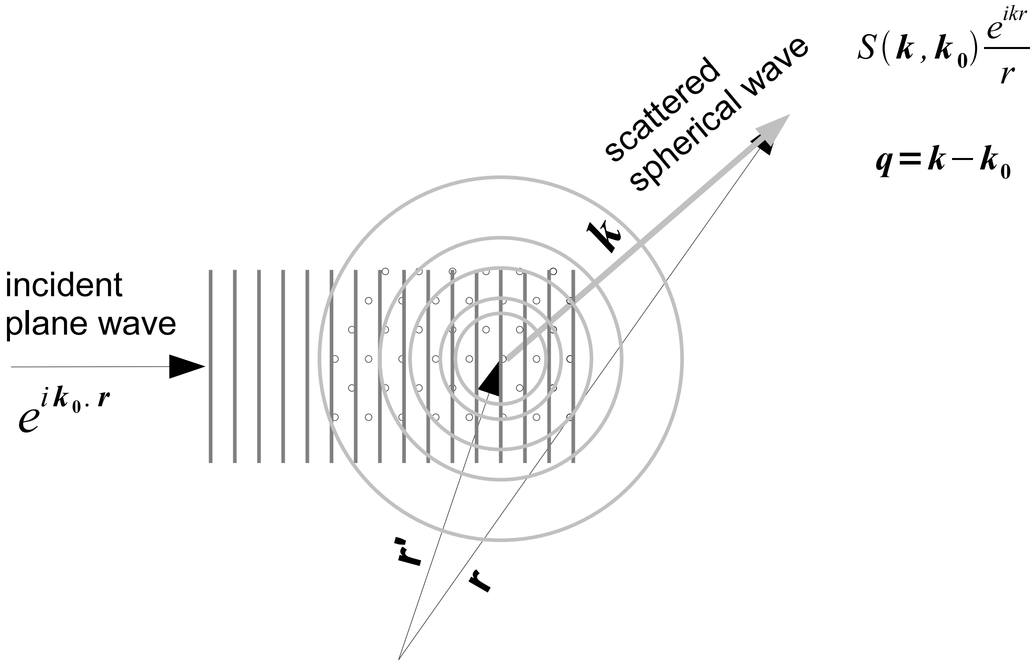


Figure 2.5: We write the scattering amplitude from each site and then sum over all sites to get the total amplitude.

$$f(\mathbf{q}) = \sum_{\text{unit cell}} e^{i\mathbf{q}.\mathbf{d}_i} f_i^{\text{atom}}(q) \tag{2.10}$$

The atomic form factors can be shown to be integrals over the electronic charge distribution $\rho(\mathbf{r})$. An implicit assumption is made that the electron distribution of the isolated atom (ion) can be used to calculate the scattering from the lattice.

$$f(\mathbf{q}) = \int_{atom} \mathrm{d}^3\mathbf{r}\rho(\mathbf{r})e^{i\mathbf{q}.\mathbf{r}} \tag{2.11}$$

Now note a few points :

- In general the integral (eqn. 2.11)will fall quite rapidly as $|\mathbf{q}|$ increases.
- The larger the number of electrons (Z) in the atom, the more visible it would be to x-rays.
- eqn. 2.9 and 2.10 together imply that a certain reflection allowed by the lattice may get cut off by the sum over the basis.

2.2 A few common lattice types

When we talk about two distinct lattice types, we mean that one lattice (with no basis) cannot be superposed on the other by simply expanding or contracting it. Thus the question of how many lattice types there can be, is related to enumerating the possible rotational and other symmetries (like 3 fold, 4 fold or inversion etc.) We will not go into this enumeration, but just describe the most common ones. (See chap1 of Kittel for example for a full listing)

2.2.1 Simple Cubic

The simplest possible lattice, where there the lattice points are distributed on the corners of a cube. It is not very common. Only one element is known to crystallise in this form - Polonium.

However there are many compounds which can be described as simple cubic with a basis. A representative example is Cesium Chloride.

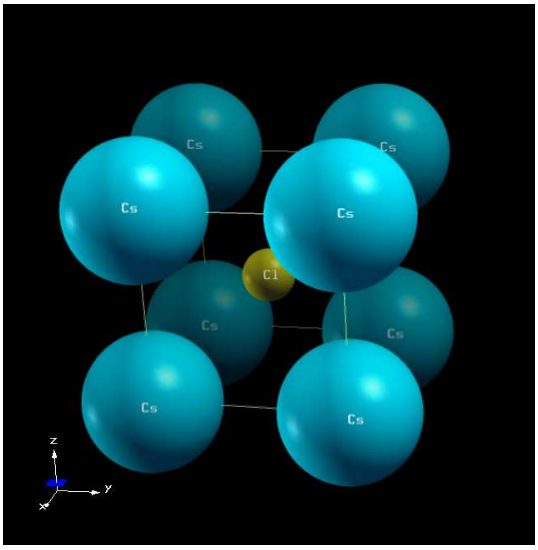


Figure 2.6: The Cesium chloride lattice, the lattice constant $a = 4.11$Å. Other materials which crystallize like this are NH_4Cl, $CuPd$, $CuZn$, $AlNi$, $BeCu$ etc.

The lattice vectors are (trivially):

$$\begin{aligned} \mathbf{a}_1 &= a\hat{\mathbf{x}} \\ \mathbf{a}_2 &= a\hat{\mathbf{y}} \\ \mathbf{a}_3 &= a\hat{\mathbf{z}} \end{aligned} \tag{2.12}$$

where a is the length of the sides of the cube. The basis at (See Fig. 2.6):

$$\begin{aligned} \mathbf{d}_1 &= 0 \\ \mathbf{d}_2 &= \frac{1}{2}\mathbf{a}_1 + \frac{1}{2}\mathbf{a}_2 + \frac{1}{2}\mathbf{a}_3 \end{aligned} \tag{2.13}$$

PROBLEM : Which *hkl* reflections would be cut off by this basis, following eqn. 2.10? if the sites were occupied by the same atom? However this is a bit hypothetical for Cesium chloride, because the atomic form factors of Cs and Cl are not the same.

2.2.2 Body-centered-cubic lattice

If we place an identical atoms at the center of a cube and on all the corners, we get a Body-centered-cubic lattice.

The conventional unit cell would be simple cubic with the basis vectors given by $(0, 0, 0)$ and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. But there is also another choice, where the unit cell is defined in a different way, so that there is only one

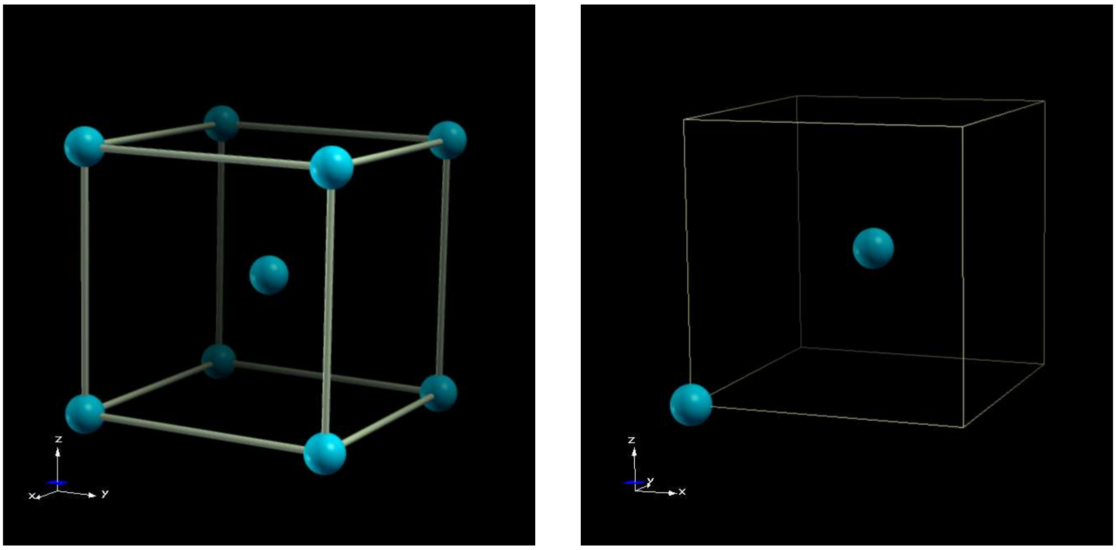


Figure 2.7: The lattice and the basis for a BCC. Alkali metals like Li,Na,K,Rb are BCC at room temperature, so is Fe.

lattice point per cell.

$$\begin{aligned}
\mathbf{a}_1 &= \frac{a}{2}\left(\hat{\mathbf{x}}+\hat{\mathbf{y}}-\hat{\mathbf{z}}\right)\\
\mathbf{a}_2 &= \frac{a}{2}\left(\hat{\mathbf{x}}-\hat{\mathbf{y}}+\hat{\mathbf{z}}\right)\\
\mathbf{a}_3 &= \frac{a}{2}\left(-\hat{\mathbf{x}}+\hat{\mathbf{y}}+\hat{\mathbf{z}}\right)
\end{aligned} \tag{2.14}$$

2.2.3 Face-centered-cubic lattice

If we place an identical atoms at the centers of each face of a cube and on all the corners, we get a Face-centered-cubic lattice.

If we choose the cubic unit vectors (eqn 2.12 then the four basis atoms in the *cubic* unit cell(see Fig. 2.8) should be at

$$\begin{aligned}
\mathbf{d}_1 &= 0\\
\mathbf{d}_2 &= \frac{1}{2}\mathbf{a}_1+\frac{1}{2}\mathbf{a}_2\\
\mathbf{d}_3 &= \frac{1}{2}\mathbf{a}_2+\frac{1}{2}\mathbf{a}_3\\
\mathbf{d}_4 &= \frac{1}{2}\mathbf{a}_3+\frac{1}{2}\mathbf{a}_1
\end{aligned} \tag{2.15}$$

However there is a way of choosing the lattice vectors so that there is only one basis atom in the unit cell. These primitive vectors are :

$$\begin{aligned}
\mathbf{a}_1 &= \frac{a}{2}\left(\hat{\mathbf{x}}+\hat{\mathbf{y}}\right)\\
\mathbf{a}_2 &= \frac{a}{2}\left(\hat{\mathbf{y}}+\hat{\mathbf{z}}\right)\\
\mathbf{a}_3 &= \frac{a}{2}\left(\hat{\mathbf{z}}+\hat{\mathbf{x}}\right)
\end{aligned} \tag{2.16}$$

Notice that these primitive lattice vectors (eqn. 2.16 are simply obtained by joining the basis atoms with the origin.

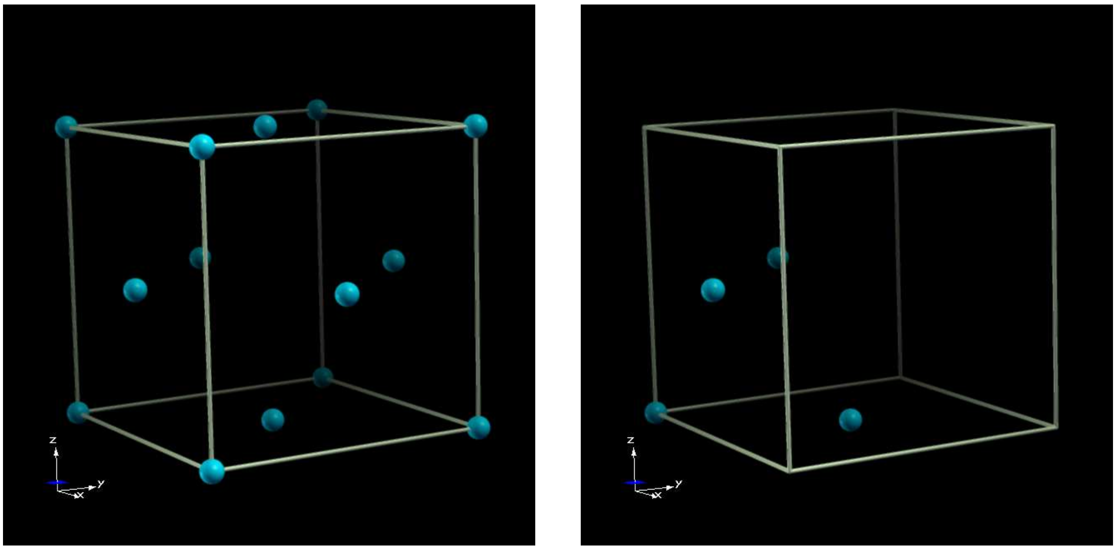


Figure 2.8: The lattice and the basis for a FCC. Many common metals like *Pb*, *Cu*, *Pt*, *Au* crystallize as FCC

PROBLEM : One of the standard problems in crystal structures: Show that the FCC and BCC lattices as described by the primitive vectors (eqn. 2.16 and eqn. 2.14) are reciprocal lattices of each other.

NaCl lattice

NaCl crystallizes into a lattice, which has alternate Na and Cl atoms along all the edges of the cube as well as through the perpendicular lines running along the middle of the faces. If the atoms were identical then it would be a simple cubic with half the lattice constant.
The basis atoms are at (referred to the cubic unit cell):

$$\begin{aligned}
\text{Na}: \quad \mathbf{d}_1 &= (0,0,0) \\
\mathbf{d}_2 &= \left(\frac{1}{2},\frac{1}{2},0\right) \\
\mathbf{d}_3 &= \left(\frac{1}{2},0,\frac{1}{2}\right) \\
\mathbf{d}_4 &= \left(0,\frac{1}{2},\frac{1}{2}\right) \\
\text{Cl}: \quad \mathbf{d}_5 &= \left(\frac{1}{2},\frac{1}{2},\frac{1}{2}\right) \\
\mathbf{d}_6 &= \left(0,0,\frac{1}{2}\right) \\
\mathbf{d}_6 &= \left(0,\frac{1}{2},0\right) \\
\mathbf{d}_6 &= \left(\frac{1}{2},0,0\right)
\end{aligned}$$

2.2.4 Diamond lattice

Carbon, Silicon, Germanium and compounds like GaAs, InSb, ZnS, SiC crytallize in a lattice that is fcc with a two atom basis.
The cubic unit cell has 8 atoms in it, each of the four fcc lattice points contributing two atoms. The second atom is displaced along the diagonal of the cube by $\frac{1}{4}^{\text{th}}$ of the cubes diagonal. So we have for the basis in terms of the cubic lattice vectors $\mathbf{a}_i$

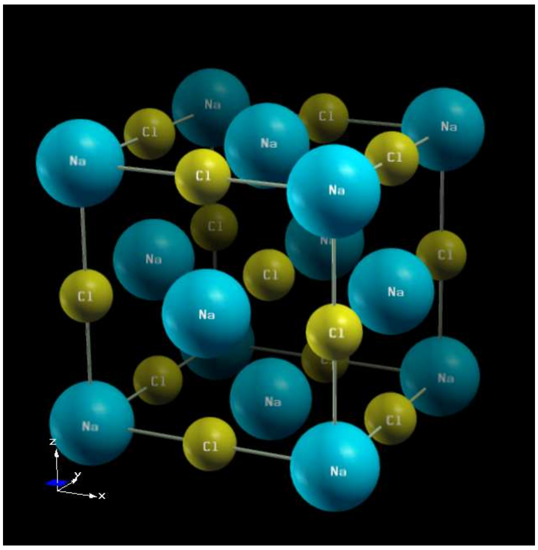


Figure 2.9: The $NaCl$ lattice. The cubic lattice constant $a = 5.6$Å and for KCl, KBr, LiH, $AgBr$, PbS all have similar lattices.

$$\begin{aligned}
\mathbf{d}_1 &= (0,0,0) \\
\mathbf{d}_2 &= \left(\frac{1}{2},\frac{1}{2},0\right) \\
\mathbf{d}_3 &= \left(0,\frac{1}{2},\frac{1}{2}\right) \\
\mathbf{d}_4 &= \left(\frac{1}{2},0,\frac{1}{2}\right) \\
\mathbf{d}_5 &= \left(\frac{1}{4},\frac{1}{4},\frac{1}{4}\right) \\
\mathbf{d}_6 &= \left(\frac{3}{4},\frac{3}{4},\frac{1}{4}\right) \\
\mathbf{d}_7 &= \left(\frac{1}{4},\frac{3}{4},\frac{3}{4}\right) \\
\mathbf{d}_8 &= \left(\frac{3}{4},\frac{1}{4},\frac{3}{4}\right) \qquad (2.17)
\end{aligned}$$

Another way to describe the diamond lattice is to use the FCC unit vectors, (still using the cubic cell to define our unit of length), so that the atoms are

$$\begin{aligned}
\mathbf{d}_1 &= (0,0,0) \\
\mathbf{d}_2 &= a\left(\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)
\end{aligned} \tag{2.18}$$

Then the reciprocal lattice vectors would be

$$\begin{aligned}
\mathbf{b}_1 &= \frac{2\pi}{a}(-\hat{\mathbf{x}}+\hat{\mathbf{y}}+\hat{\mathbf{z}}) \\
\mathbf{b}_2 &= \frac{2\pi}{a}(\hat{\mathbf{x}}-\hat{\mathbf{y}}+\hat{\mathbf{z}}) \\
\mathbf{b}_3 &= \frac{2\pi}{a}(\hat{\mathbf{x}}+\hat{\mathbf{y}}-\hat{\mathbf{z}})
\end{aligned} \tag{2.19}$$

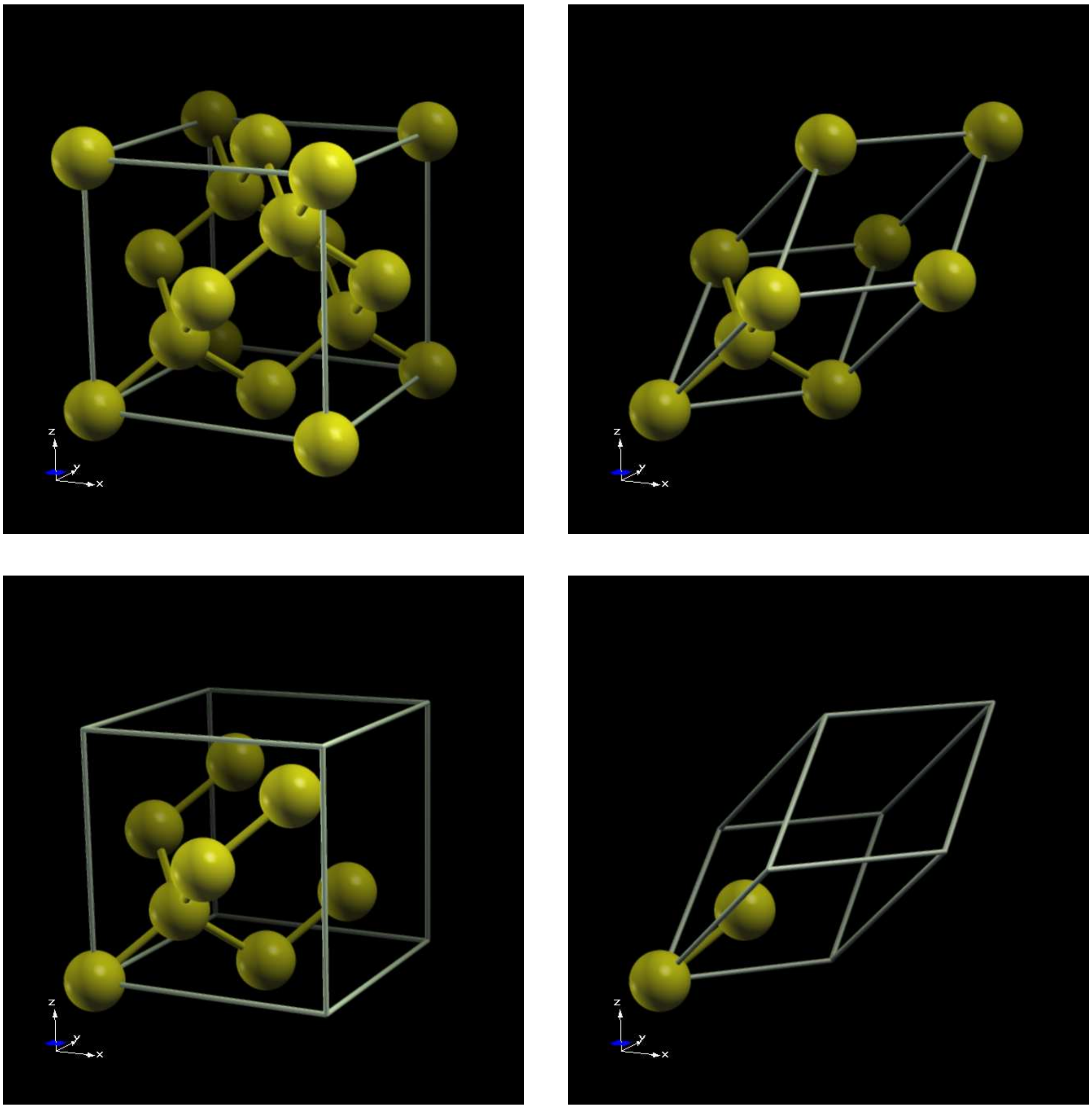


Figure 2.10: The diamond lattice. On the left column, we have used the cubic unit cell, on the right we use the FCC primitive vectors. The cubic lattice constant $a = 3.57$Å and for Si $a = 5.43$Å

PROBLEM : Which *hkl* reflections would be cut off by this basis, following eqns. 2.18 and 2.19? Consider the general reciprocal lattice vector to be of the form

$$\mathbf{G} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3$$

Can you see the advantage of making the basis as small as possible?

ZnS and GaAs lattice

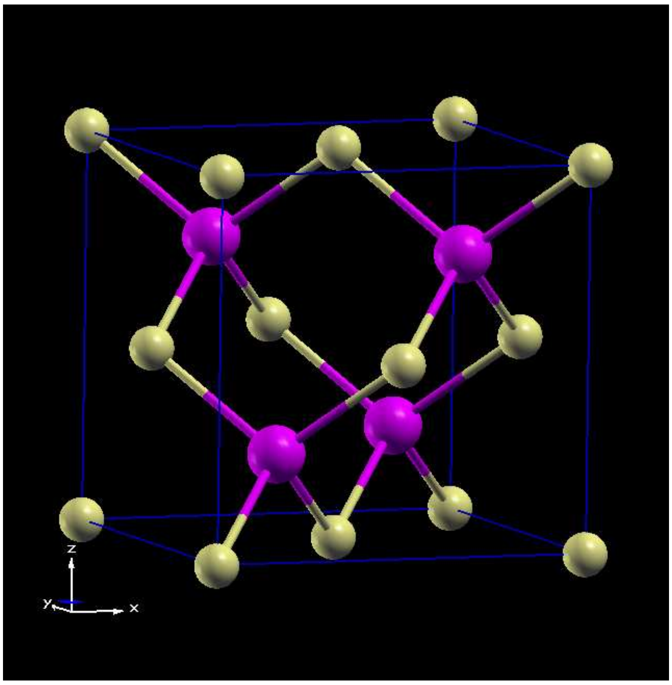


Figure 2.11: ZnS lattice, similar to diamond but the atoms are different.

The ABO3 Perovskite lattice

A large number of compounds that show interesting properties like piezoelectricity, ferroelectricity, giant magnetoresistance, superconductivity at high temperatures, have a composition of the type ABO_3, where A and B are two different positively charged ions of different size and O is oxygen. The structure can undergo deformations and will not be necessarily have lattice vectors of equal size, though the one shown here has.

A summary of various names

All the examples given so far could be described using a cubic system. This need not be so.

In the context of crystal structures one comes across a lot of names. Here's a brief survey of what they mean. Notice how the symmetry of the structure is given up in steps.

- **Cubic**: Simple cubic, BCC, FCC lattices. We can describe them using three lattice vectors of equal length, which are also orthogonal to each other.
- **Tetragonal**: The three lattice vectors are orthogonal, but two (not three) of them are of equal length. We can transform the cube to this structure by pulling one face.
- **Orthorhomic**: The three lattice vectors are orthogonal, but no two vectors are of equal length.

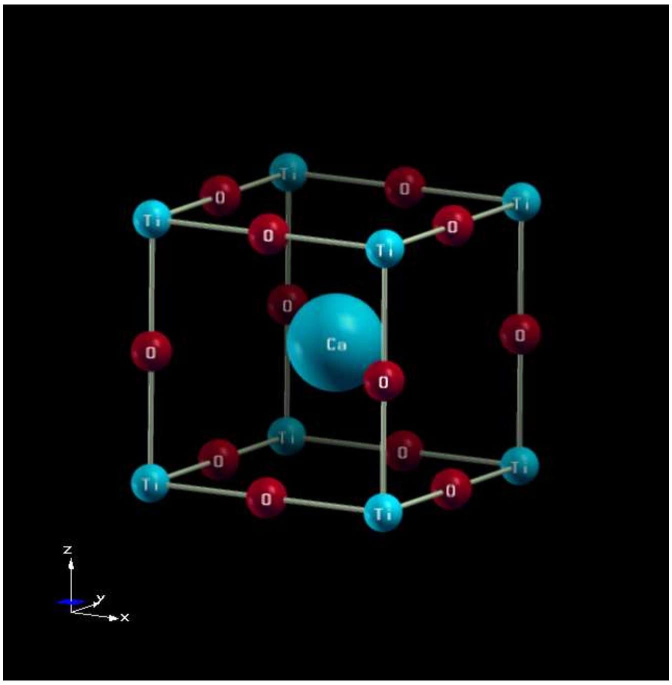


Figure 2.12: $CaTiO_3$ lattice, The lattice constant is 3.8Å.

- **Rhombohedral**: The three lattice vectors are of equal length but no two lattice vectors are orthogonal to each other, this will happen if a cube is pulled along its body diagonal.
- **Tricilinic**: The three lattice vectors are of unequal length and no two lattice vectors are orthogonal to each other.

2.3 A Geometrical excursion: Wigner Seitz cell, Brilloiun zones etc

In this section we are going to introduce a new concept, the significance of this is impossible to describe right at the beginning. For example it is impossible to justify or describe why Fourier transforms are important, at the time of introducing it. So we would not attempt it!

2.3.1 Wigner Seitz cell

Let's consider a collection of points in any dimension. They are not necessarily in any regular pattern. Then divide the entire pattern into zones, around each point , according to the following rule

Collect all points, that are closer to a certain lattice point, than to any other lattice point.

Notice how the pattern (Fig. 2.13) looks for a random collection of points.

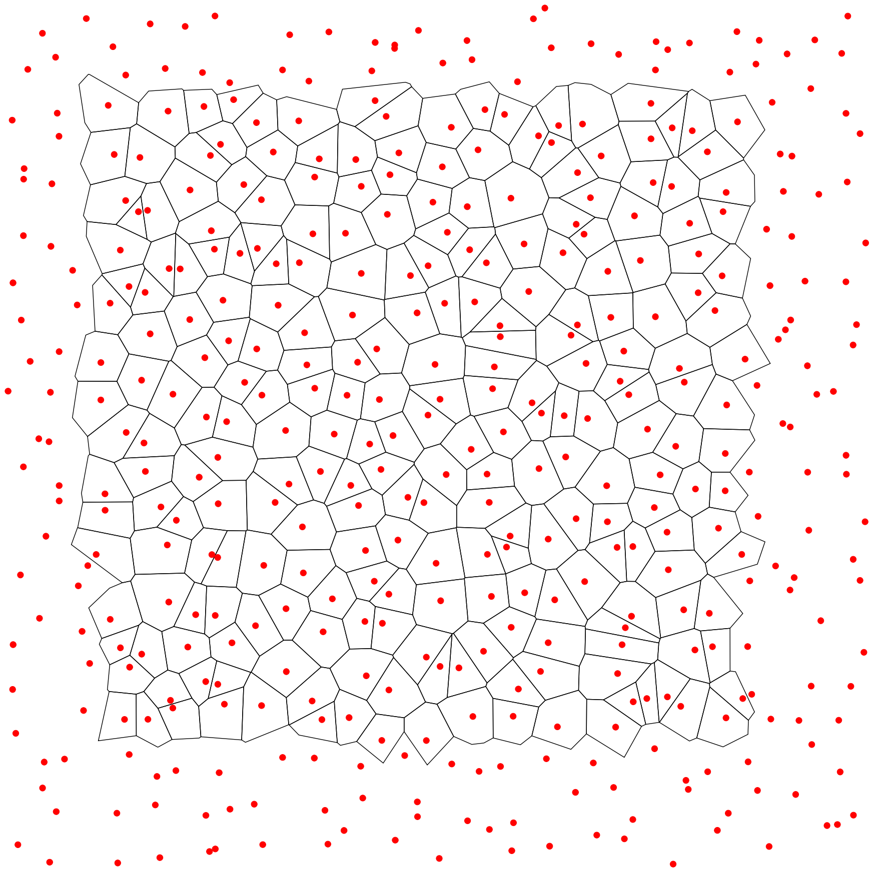


Figure 2.13: The area around a lattice point, that is closer to *that* point, than to any other lattice point.

The figure can be constructed according to the following rules

1. Join a particular point to a few (all others if you want) neighbouring points.
2. Construct the perpendicular bisector of each of these lines.
3. The smallest area around a point that is enclosed by the perpendicular bisectors form the Wigner-Seitz cell around that point.

Of course, if the lattice is regular then the cells will also be regular and is infact a possible way to form the unit cell of the lattice, containing exactly one lattice point at its "center".

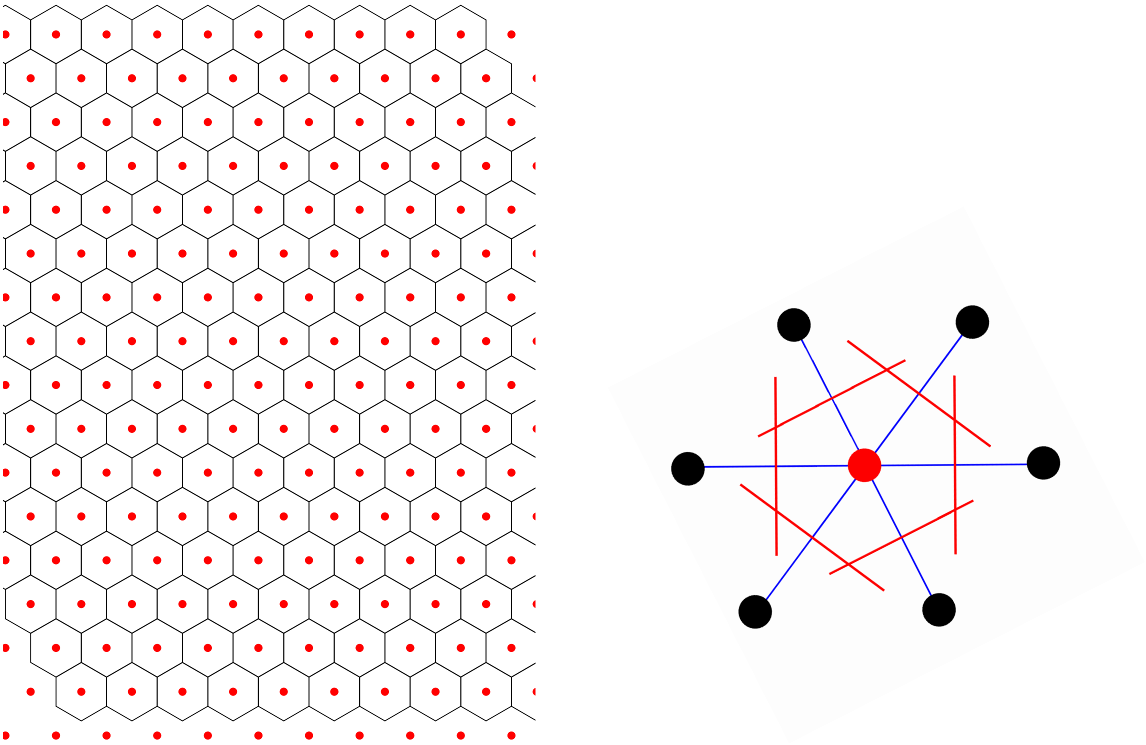


Figure 2.14: The Wigner-Seitz cell for the triangular lattice with lattice vectors (1,0) and $(\frac{1}{2},\frac{\sqrt{3}}{2})$. Part of the picture is taken from Wikipedia.

PROBLEM : Now you can convince yourself that irrespective of whether the lattice is ordered or disordered, the Wigner-Seitz cells would cover the area leaving no "gaps" or "overlaps". In mathematical language it *tiles* the plane. Show that a Wigner-Seitz cell is always a convex polygon.

Brillouin zones

Obviously these cells can be drawn in any lattice. The convention is that when Wigner-Seitz cells are drawn in the reciprocal lattice they are called Brillouin zones. However the concept of the Brillouin zone is a little more involved than the Wigner-Seitz cell itself.

Higher order Brillouin zones

2.3.2 In 3D....

We can extend the idea of the Wigner-Seitz cell to 3D. We are already familiar with some of the lattice types in 3D. The rule for constructing it similar to what we have stated, except *now the bisectors of the lines joining the points are no longer lines, but planes.* See how it would look for the two most frequently encountered lattice types BCC and FCC.

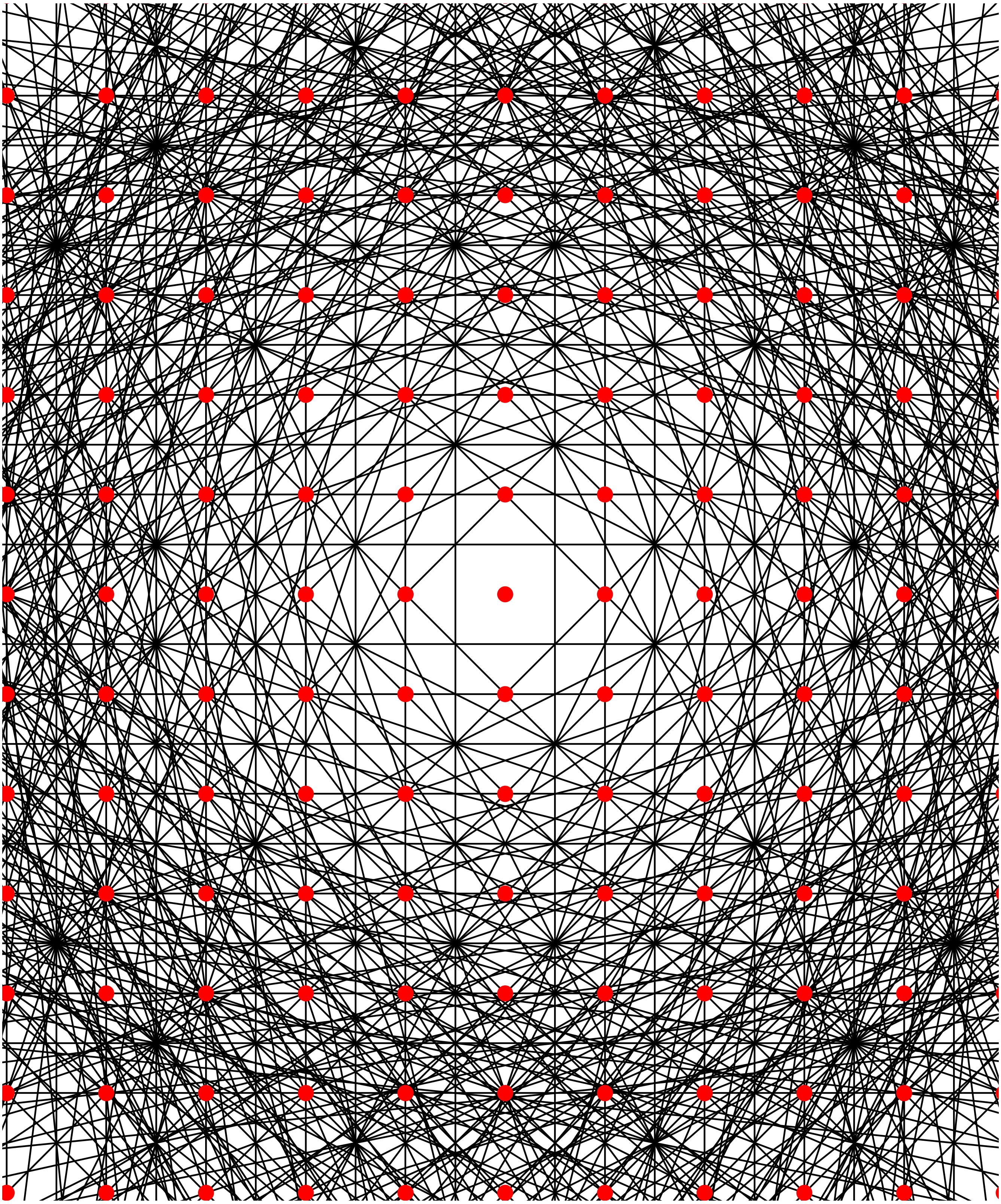


Figure 2.15: For a SQUARE lattice : notice how the bisectors of all the lines connecting the point at the center divide up the plane into many disjoint parts. The central part is the Wigner Seitz cell, but the other parts have their significance too.

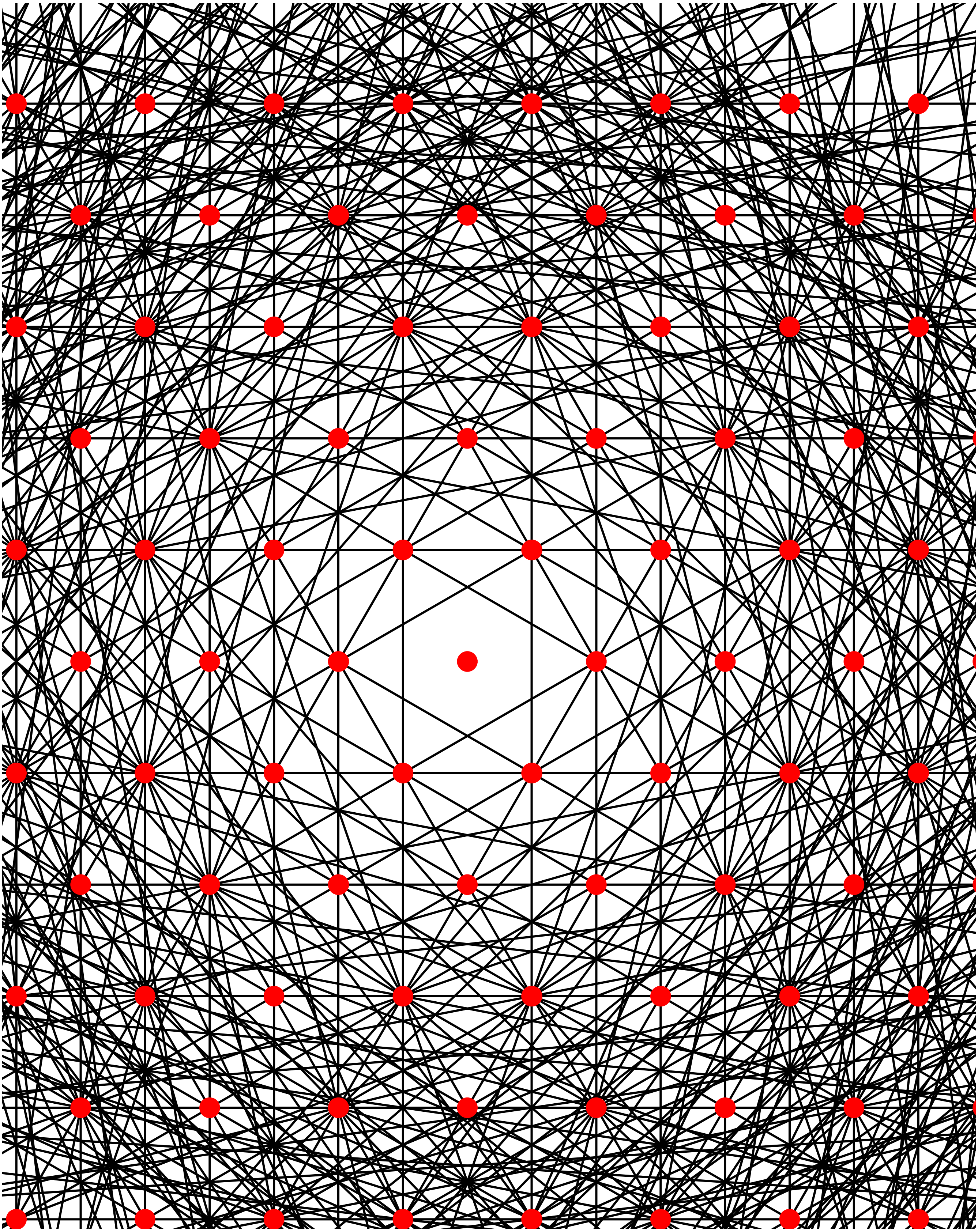


Figure 2.16: Same for a TRIANGULAR lattice...

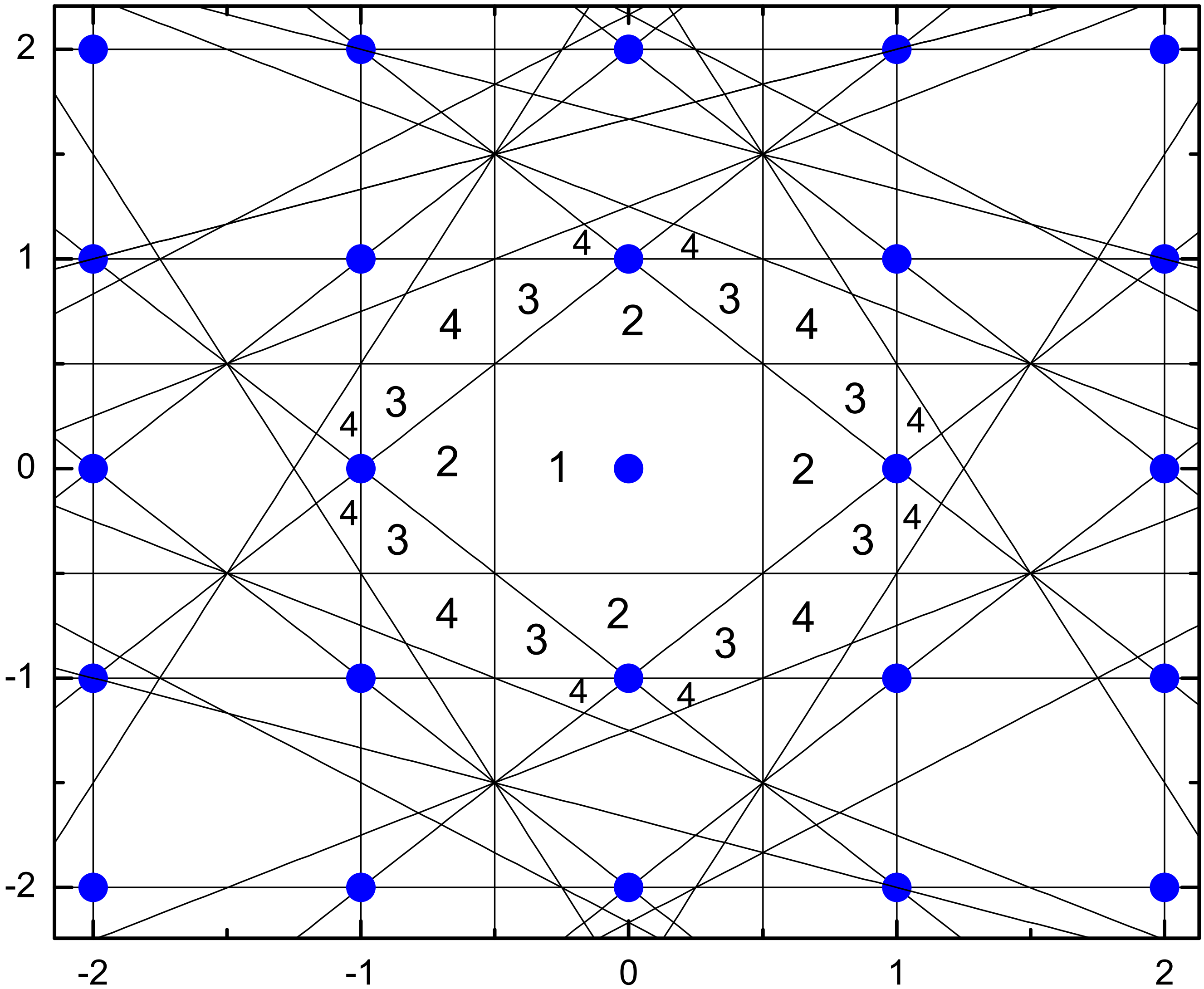


Figure 2.17: The higher order Brillouin zones in a square lattice. The significance of this will appear when we get into band structure. For the time being it is just a bit of interesting geometry

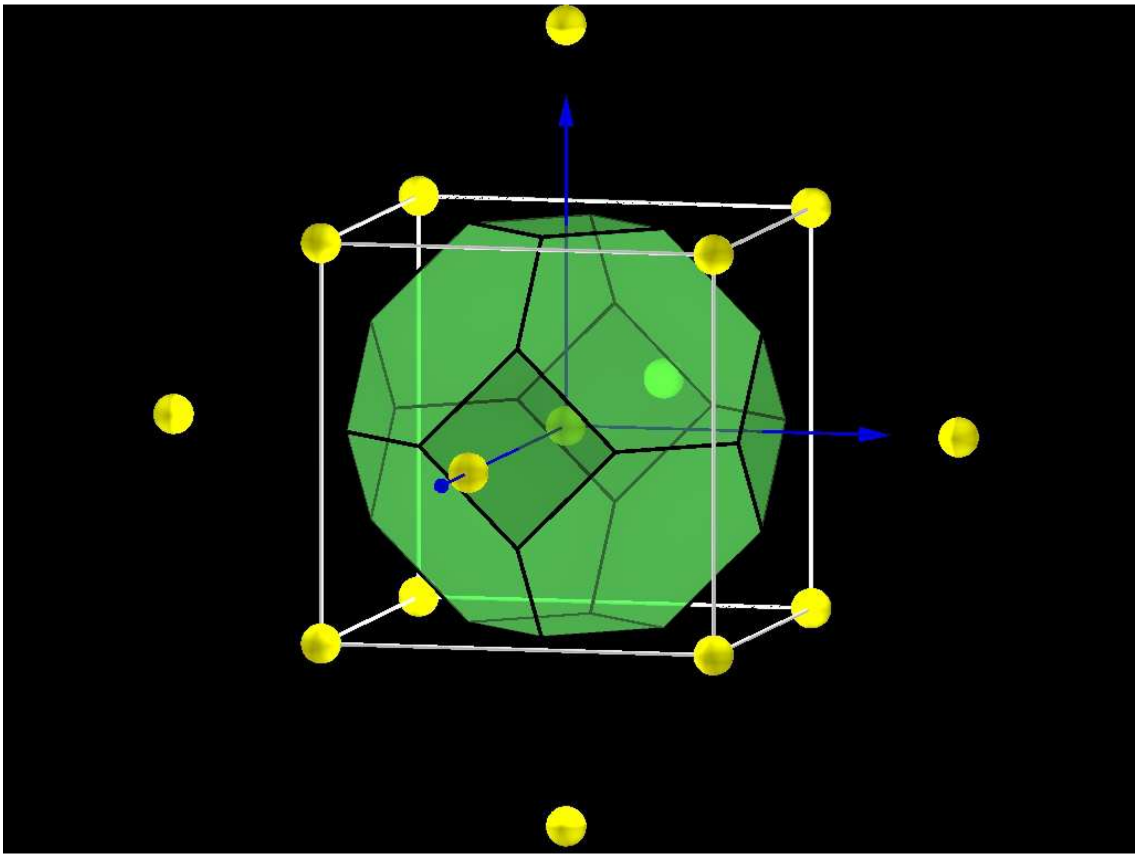


Figure 2.18: BCC Wigner-Seitz cell, shown in ”transparent” way..

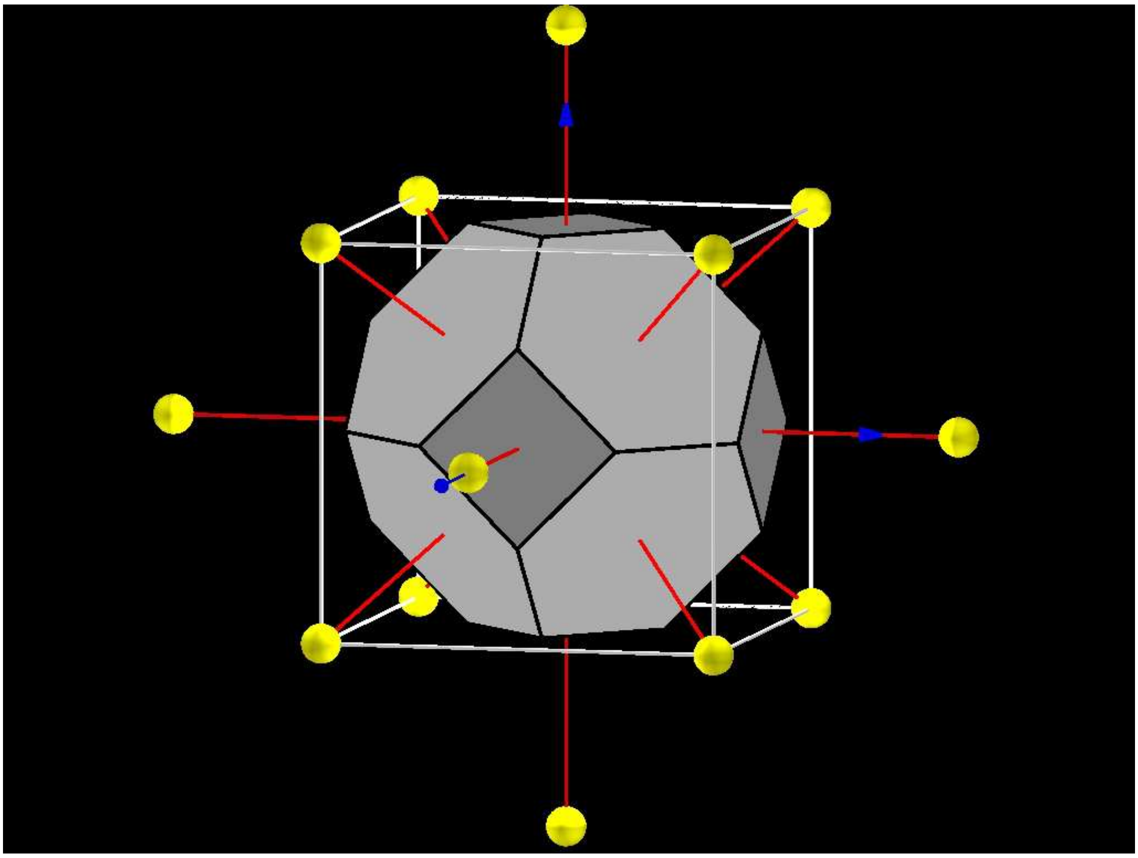


Figure 2.19: BCC Wigner-Seitz cell: another view showing which neighbours are involved in forming the planes.

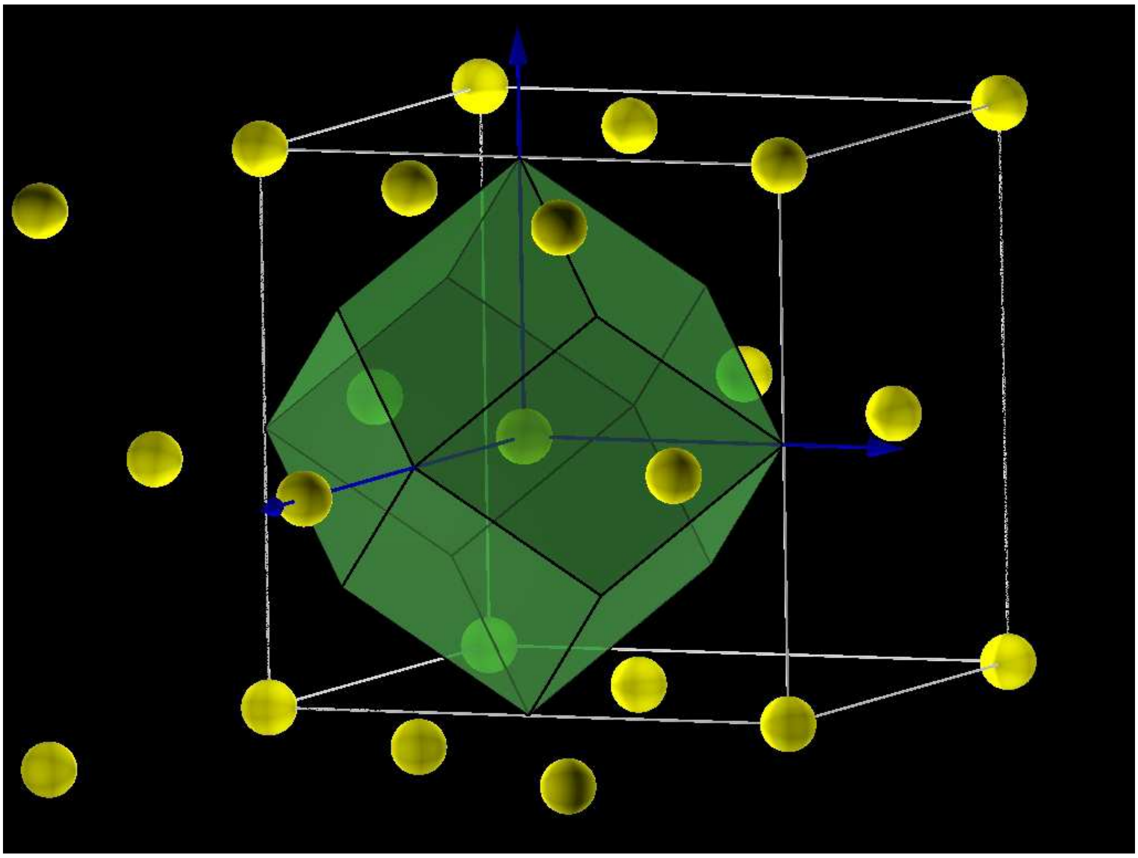


Figure 2.20: FCC Wigner-Seitz cell, shown in "transparent" way..

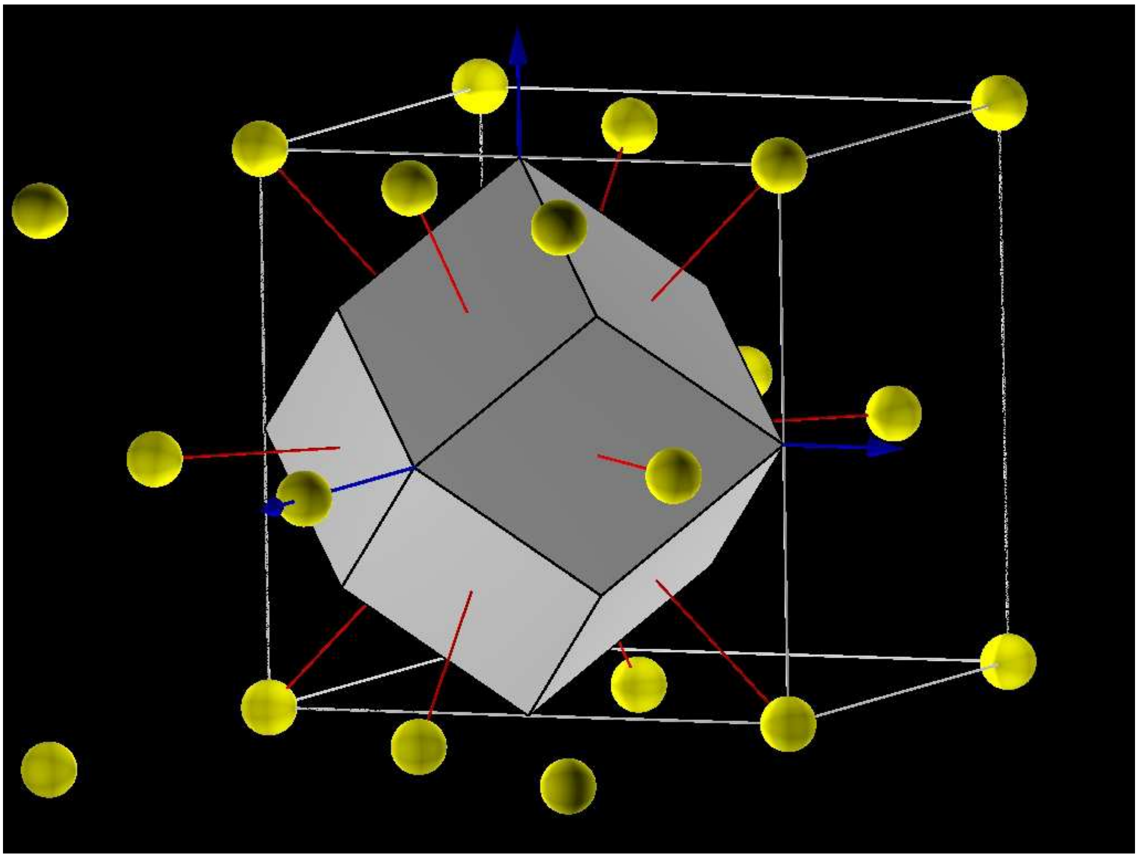


Figure 2.21: FCC Wigner-Seitz cell: another view showing which neighbours are involved in forming the planes.

Chapter 3

Free electrons in metals: basic thermodynamics, transport and electrodynamics

Lecture notes for PH409 (Jul 2012): K. Das Gupta

References:

1. Chapter 1 & 2 *Solid State Physics*, N. W. Ashcroft and N.D. Mermin
2. Chapter 12 & 16, *Solid State Physics*, N. W. Ashcroft and N.D. Mermin (for the Boltzmann transport equation)
3. page 61-67, *Quantum Heterostructures*, V. V. Mitin, V. A. Kochelap and M. A. Stroscio (for the Boltzmann transport equation)
4. Chapter 6 *Introduction to Solid State Physics, C. Kittel*
5.
    - An useful article that you can read as an aside, about resonant responses: *Reversing Light: Negative Refraction*, J. B. Pendry and D.R. Smith, *Physics Today*, December 2003.
    - A very interesting article on a similar subject: *The electrodynamics of substances with simultaneously negative values of ϵ and μ.* V.G. Veselago, *Soviet Physics Uspekhi*, **10**, 4 p509-515, January-February 1968.

To a first approximation we can think of metals as a gas of electrons in a box, behaving very much like ”particle(s) in a box”, independent of each other. There is a specific meaning of the word ”gas” in this context. When we say electron gas, we usually imply that the Coulomb interaction between the individual particles have been ignored. When we take the interaction into account, we usually refer to the system as a (fermi) liquid. The interaction between the electrons are not weak, but the approach works to a surprising extent. Metals have some basic properties - they are good conductors of heat and electricity, good (shiny) reflector of light etc. These have fairly decent explanations within the ”free electron” model. But how did we get so lucky here? The answer to that would have to wait for some time. Let's see what we can get after making this simplifying assumption as far as thermodynamics, transport of electric current and electrodynamics are concerned.

3.1 The basic model

Fig. 3.1 shows our picture of a (very!) simplified metal. The highest electron energy is 4-5 eV below the vacuum level. This is the typical energy that the electron must gain (in a photemission/thermal emission process, for example) to be able to escape from the ”box”. This is consistent with the fact that ultraviolet light is usually needed for photoemission from metals and electron emitting hot filaments need

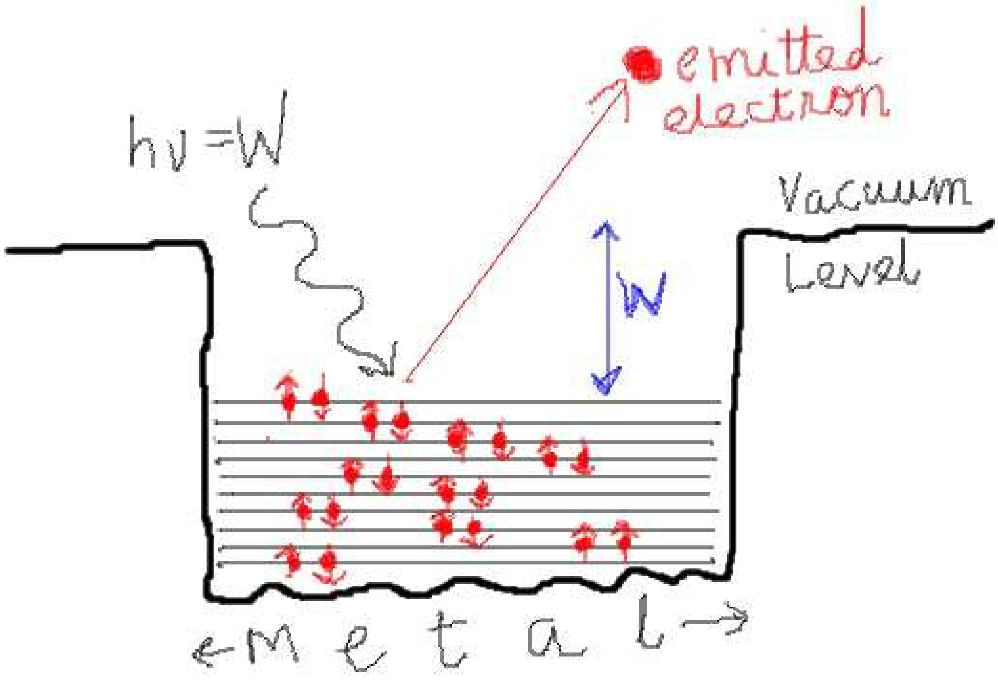


Figure 3.1:

to be heated to (typically) $> 1500^o K$

Recall the following consequences of fermi statistics and their "freedom" to move about inside the metal, which the non-interacting electrons must obey.

1. 2 fermions cannot occupy the same state.
2. Spin $\frac{1}{2}$ objects like electrons have two possible spin states.
3. The wavefunction of the electrons may be written as :

$$\psi(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}.\mathbf{r}} \tag{3.1}$$

4. The energy-momentum relation is like a free particle

$$E(\mathbf{k}) = \frac{\hbar^2}{2m}\left(k_x^2 + k_y^2 + k_z^2\right) \tag{3.2}$$

5. The energy level populations are given by:

$$f(E,\mu) = \frac{1}{1 + e^{(E-\mu)/k_B T}} \tag{3.3}$$

   where μ is called the chemical potential of the system. It is important to remember that the probability of an energy level being occupied drops sharply as one crosses μ from below.

3.1.1 How does the Fermi distribution look?

Look at Fig. 3.2, the distribution drops sharply near the chemical potential. At $T = 0$ it is a step function. The region over which it is significantly different from $T = 0$ characteristic is the spread of the distribution. This is $\sim k_B T$. k_B is the Boltzmann constant which has a value 1.38×10^{-23} Joule/K. At

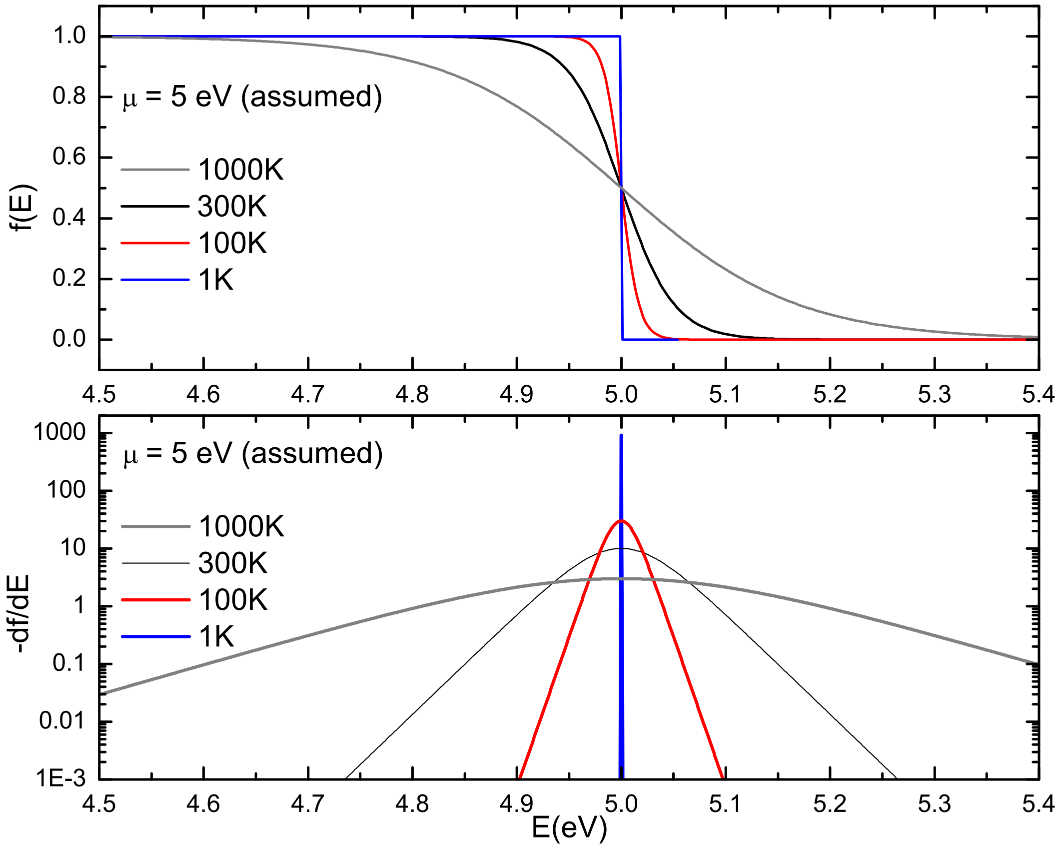


Figure 3.2: Note how the Fermi distribution becomes sharper at small T. We have chosen the chemical potential to be a realistic number for normal metals. How does the derivative $df(E)/dE$ look?

room temperature (300K), $k_B T \approx 25$meV in energy units.

PROBLEM : Certain combinations of the Fermi function, occur very frequently in expressions that involve scattering or transitions. It is useful to be familiar with the combination $f^0(1-f^0)$

Make a rough sketch of how $f^0(1-f^0)$ would look as a function of energy. How does the area under the curve of $f^0(1-f^0)$ vary with temperature?

3.1.2 What fixes the chemical potential (at $T = 0$)?

We have kept μ as an undetermined parameter. You can see that since

$$N = \sum_{\text{all states}} \frac{1}{1 + e^{(E-\mu)/kT}} \tag{3.4}$$

must hold, chemical potential and the number of particles in a system are implicitly related. To be able to evaluate the sum we need one more information. Thermodynamics gives the probability of a level being occupied, but does not tell us how many levels there are (density of states) between E to $E + dE$.

The density of states

We formulate our argument in terms of the **k** vector.

- The allowed values of **k** for ”particles in a box” is determined by the condition that the wavefunction must vanish at the boundary. This imposes a ”quantization”.
- One gets similar results for standing waves, electromagnetic modes in a cavity etc.

- Another way of getting the same result was introduced by Born and von Karman. We assume that the whole structure is periodic, so it is like a loop in 1D, torus in 2D etc. Periodic boundary condition means :

  $$\psi(x,y,z) = \psi(x+L,y,z) = \psi(x,y+L,z) = \psi(x,y,z+L) = \psi(x,y,z) \tag{3.5}$$

  Hence

  $$e^{ik_xL} = e^{ik_yL} = e^{ik_zL} = 1 = e^{i2n\pi} \tag{3.6}$$

  This means that two allowed points in **k**-space are separated by

  $$\Delta k_x = \Delta k_y = \Delta k_z = \frac{2\pi}{L} \tag{3.7}$$

  Hence the very important result (with $V = L^3$)

  $$\sum_{\text{all states}} \ldots \rightarrow \frac{V}{(2\pi)^3}\int \mathrm{d}^3\mathbf{k}\ldots \tag{3.8}$$

- We can then use the $E(k_x,k_y,k_z)$ relation to convert the density of states $D(\mathbf{k}) = \frac{V}{(2\pi)^3}$ to $D(E)$.
- Since the momentum is $\mathbf{p} = \hbar\mathbf{k}$, this also implies that the volume occupied by a point in *phase space* is $\sim h^3$

Calculating $D(E)$ from $D(\mathbf{k})$

For free particles we can write for *each spin state*:

$$\begin{aligned} D(E)\mathrm{d}E &= \frac{V}{(2\pi)^3}4\pi k^2\mathrm{d}k \\ D(E)/V &= \frac{1}{(2\pi)^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{E} \end{aligned} \tag{3.9}$$

The last step is written to get the density of states per unit volume.

PROBLEM : What will be form of eqn 3.9 in 1D and 2D?

A slightly different way of getting $D(E)$

Suppose we didn't assume the travelling wave like states for the particle. Recall that a particle in a 2D "box" of size $L\times L$ would have un-normalised wavefunctions given by

$$\Psi(x,y) = \begin{cases} \sin\frac{n_x\pi}{L}x\ \sin\frac{n_y\pi}{L}y & 0<x,y<L \\ 0 & \text{otherwise} \end{cases} \tag{3.10}$$

The energy of these states would be given by the usual free particle expression

$$E = \frac{\hbar^2\pi^2}{2mL^2}({n_x}^2 + {n_y}^2) \tag{3.11}$$

If that is the case then how many states are there between E to $E + dE$? Recall that in the travelling wave case each allowed value was spaced by $2\pi/L$ along each direction. In this case the allowed values are clearly spaced by π/L. Shouldn't it end up making $D(E)$ double? The answer is no.

It is true that the area per state is now less by a factor of four (in 2D). But in the n_x, n_y space *only positive values are allowed.* This is because $\sin n\pi x/L$ and $\sin -n\pi x/L$ are not distinct states, but e^{ikx} and e^{-ikx} are distinct states. To see why this is so, we can imagine measuring the momentum of these states. The state e^{ikx} would give $\hbar k$ and the state e^{-ikx} would give $-\hbar k$. These two states correspond to two distinct eignevalues of an operator and are distinct states. Both must be counted. Whereas the state

$\sin -n\pi x/L$ is simply -$\sin n\pi x/L$. They should not be counted twice.

Thus to count the number of states with energy less than E we need to know how many allowed (n_x, n_y) pairs in the first quadrant, we get such that

$$n_x{}^2 + n_y{}^2 < \frac{2mL^2}{\hbar^2\pi^2}E \tag{3.12}$$

The answer is

$$\begin{aligned} \Omega(E) &= \frac{1}{4}\pi\frac{2mL^2}{\hbar^2\pi^2}E & (3.13)\\ \therefore D(E) &= \frac{d\Omega(E)}{dE} = \frac{m}{2\pi\hbar^2}L^2 & (3.14)\end{aligned}$$

Compare this with the results of problems in the earlier section.

A general expression for density of states

So far we just considered the density of states of essentially free particle. Another way of stating this is that we just counted the eigenvalues of the Hamiltonian $H = \frac{p^2}{2m}$. We can then ask - what if the Hamiltonian is more complex? For example one can have a magnetic field or some potential. Obviously the eigenvalues of such a system would be very different from a free particle. One might expect that the density of states $D(E)$ should have a connection with the Hamiltonian. This is indeed correct, but the mathematical form of that relation requires some work. We will need to introduce the Green's function (operator) corresponding to the Hamiltonian defined as

$$G(E + i\varepsilon) = \frac{1}{E + i\varepsilon - H} \tag{3.15}$$

Notice that the operator G has singularities. The denominator will become zero when it acts on an eigenstate of H, unless we introduce the vanishingly small imaginary part, $i\varepsilon$. We will take the limit of the expression at the end. The non-trivial role of this quantity will become clear when we do a contour integration. A similar method is also used calculation of scattering amplitudes (See Sakurai, Modern Quantum Mechanics, for more details).
H has eigenvalues and normalised eignefunctions - E_i and ψ_i respectively. We also consider another complete set of normalised functions f_α, these do not have to be eigenfucntions of H. To get the density of states we need to evaluate (as always)

$$D(E) = \sum_i \delta(E - E_i) \tag{3.16}$$

Consider an intermediate quantity called the "projected density of states" on f_α as

$$n_\alpha(E) = \sum_i |\langle f_\alpha|\psi_i\rangle|^2 \delta(E - E_i) \tag{3.17}$$

You should be able to show the two results

1. $\int dE \; n_\alpha(E) = 1$
2. $\sum\limits_\alpha \; n_\alpha(E) = \sum\limits_i \delta(E - E_i) = D(E)$

Now consider the diagonal elements of the Greens' function in the α basis:

$$\begin{aligned} G_{\alpha\alpha}(E + i\varepsilon) &= \langle f_\alpha|\frac{1}{E + i\varepsilon - H}|f_\alpha\rangle \\ &= \sum_{i,j}\langle f_\alpha \psi_i\rangle\langle\psi_i|\frac{1}{E + i\varepsilon - H}|\psi_j\rangle\langle\psi_j|f_\alpha\rangle \\ &= \sum_i |\langle f_\alpha|\psi_i\rangle|^2\frac{1}{E + i\varepsilon - E_i} \end{aligned} \tag{3.18}$$

Notice how the operator H has been replaced by the number E_i because H was acting on its eigenfunctions. Rationalise the expression and pick out the imaginary part.

$$\text{Im } G_{\alpha\alpha} = \sum_i |\langle f_\alpha|\psi_i\rangle|^2 \frac{\varepsilon}{(E-E_i)^2+\varepsilon^2} \tag{3.19}$$

It is a standard exercise in contour integration to show the following:

$$\lim_{\varepsilon\to 0^+} \frac{\varepsilon}{(E-E_i)^2+\varepsilon^2} = \pi\delta(E-E_i) \tag{3.20}$$

eqns 3.19 and 3.17 now give us the important results:

$$n_\alpha(E) = -\frac{1}{\pi}\lim_{\varepsilon\to 0^+} \text{Im } G_{\alpha\alpha}(E+i\varepsilon) \tag{3.21}$$

$$\begin{aligned} D(E) &= \sum_\alpha n_\alpha(E) \\ &= -\frac{1}{\pi}\lim_{\varepsilon\to 0^+} \text{Im Tr } G_{\alpha\alpha}(E+i\varepsilon) \end{aligned} \tag{3.22}$$

Notice that the density of states have been related to the Hamiltonian via the Green's function construct. This is called the "spectral theorem".

$T=0$ **case**

The total number of particles and the highest occupied energy level are easily related at $T=0$. We start filling the states from the bottom one by one and stop when we run out of particles. The factor of 2 in the later steps accounting for two spin states.

$$N = \int_0^\mu \frac{V}{(2\pi)^3}\text{d}^3\mathbf{k} \tag{3.23}$$

$$N/V = \int_0^\mu \text{d}E\,\frac{1}{(2\pi)^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{E} \tag{3.24}$$

$$\mu(T=0) = \left(\frac{\hbar^2}{2m}\right)(3\pi^2 n)^{2/3} \tag{3.25}$$

We will call the chemical potential at $T=0$ the Fermi level E_F. It turns out that the difference between the values of μ at $T=0$ and $T=300$K for normal metals is very small. This we will prove soon.

The radius of the spherical surface which separates the occupied and unoccupied states is called k_F. In 3D, with $n=N/V$

$$k_F = \left(3\pi^2 n\right)^{1/3} \tag{3.26}$$

In real metals, $n \sim 10^{22}\text{cm}^{-3}$, so $k_F \sim 10^8\text{cm}^{-1}$ from 3.26. The value of $1/k_F$ is around 10^{-8} cm, corresponding to atomic distance. Using this value gives $\mu(0) \sim 4-5$eV. Since 1 eV is the energy corresponding to 10^4K, the room temperature, 300 K, is sufficiently low compared with the Fermi temperature, $T_F = \mu/k_B$. This fact is very important in understanding the electronic properties. The level splitting of one-electron energy near the Fermi surface is given by

$$\mu/N \simeq 10^{-22}\text{eV} \tag{3.27}$$

As a result, excitations near the Fermi energy can be created with vanishingly small energy, the discreteness of energy levels doesn't show up.

Energy of the system

The highest energy level occupied by the system is clearly $>> k_BT$, this is a consequence of fermi statistics. We now ask, what is the total (kinetic) energy of N electrons confined to a volume V? To answer this question we must calculate

$$U(T) = \int_0^\infty \mathrm{d}E\ ED(E)f(E,\mu) \tag{3.28}$$

$$U(T=0) = \int_0^\mu \mathrm{d}E\ ED(E)f(E,\mu) \tag{3.29}$$

$$= \frac{3}{5}N\mu(0) \tag{3.30}$$

The algebra is left as an simple exercise. The internal energy is far greater than that of a classical ideal gas with the same density. The energy per electron is of the order of 10^4K for normal metals.

3.1.3 At finite T

We want to be able to calculate response functions of the system like specific heat, susceptibility etc. For that we must learn to compute the integrals without assuming a sharp cut off at $T = 0$.

Sommerfeld expansion of Fermi integrals

While dealing with Fermi systems we often encounter integrals of the form

$$I = \int_0^\infty \mathrm{d}E\ g(E)f(E,\mu) \tag{3.31}$$

Often $g(E)$ is a fairly simple function of energy and the density of states and is well behaved near $E = \mu$. We also assume that we can integrate $g(E)$ such that

$$\frac{\mathrm{d}G(E)}{\mathrm{d}E} = g(E) \tag{3.32}$$

or equivalently

$$G(E) = \int_0^E \mathrm{d}x\ g(x) \tag{3.33}$$

We then integrate by parts to obtain

$$I = G(E)f(E,\mu)|_0^\infty + \int_0^\infty \mathrm{d}E\ G(E)\left(-\frac{\mathrm{d}f(E,\mu)}{\mathrm{d}E}\right) \tag{3.34}$$

$$= -G(0) + \int_0^\infty \mathrm{d}E\ \left[G(\mu) + (E-\mu)G'(\mu) + \frac{1}{2!}(E-\mu)^2G''(\mu) + \ldots\right]\left(\frac{\mathrm{d}f(E,\mu)}{\mathrm{d}E}\right) \tag{3.35}$$

$$I + G(0) = G(\mu)\int_0^\infty \mathrm{d}E\ \left(-\frac{\mathrm{d}f(E,\mu)}{\mathrm{d}E}\right) + G'(\mu)\int_0^\infty \mathrm{d}E\ (E-\mu)\left(-\frac{\mathrm{d}f(E,\mu)}{\mathrm{d}E}\right)$$

$$+ \frac{G''(\mu)}{2!}\int_0^\infty \mathrm{d}E\ (E-\mu)^2\left(-\frac{\mathrm{d}f(E,\mu)}{\mathrm{d}E}\right) + \ldots \tag{3.36}$$

Now, recall how $-\frac{\mathrm{d}f(E,\mu)}{\mathrm{d}E}$ behaves.

$$-\frac{\mathrm{d}f(E,\mu)}{\mathrm{d}E} = \frac{\beta}{(e^{\beta(E-\mu)}+1)(1+e^{-\beta(E-\mu)})} \tag{3.37}$$

Note the following:

- In the limit $T \to 0$ it gives $\delta(E-\mu)$, but at other temperatures also it is sharply peaked at $E = \mu$ and decays *exponentially* as we move away. This makes the contribution of any polynomial insignificant.
- We can change the limit of the integration from $[0, \infty]$, to $[-\infty, \infty]$, without introducing any significant error.
- It is easy to see that all terms with an odd power of $(E - \mu)$ vanish. You can fill in the steps to get the important result that allows us to express the integral as power series in T.
- We can ignore $G(0)$ on the LHS, because in almost all cases there will be a density of states factor in $G(E)$ which will make $G(0) = 0$.

So,

$$\begin{aligned} I &= G(\mu) + \sum_n \frac{G^{2n}(\mu)}{(2n)!}(k_B T)^{2n} \int_{-\infty}^{\infty} \mathrm{d}x\, \frac{x^{2n}}{(e^x+1)(e^{-x}+1)} \\ &= G(\mu) + \sum_n G^{2n}(\mu)(k_B T)^{2n} 2c_{2n} \qquad (3.38) \end{aligned}$$

Where the numbers c_{2n} are available as a standard integral

$$2c_{2n} = \frac{1}{(2n)!} \int_{-\infty}^{\infty} \mathrm{d}x\, \frac{x^{2n}}{(e^x+1)(e^{-x}+1)} \qquad (3.39)$$

$$c_2 = \frac{\pi^2}{12} \qquad (3.40)$$

$$c_4 = \frac{7\pi^4}{720} \qquad (3.41)$$

Variation of the chemical potential $\mu(T)$

Now we want to calculate the variation of the chemical potential with T. We (thermodynamically) distinguish between situations where N is kept fixed (microcanonical/canonical ensemble) and a grand canonical ensemble where μ is held fixed. In a canonical ensemble μ will change to keep N constant. In a grand canonical ensemble, that allows exchange of particles, some particles will come in or go out (N changes) to keep μ the same as T is changed. We assume fixed N.

We have using the Sommerfeld expansion

$$N = \int_0^{\infty} \mathrm{d}E\, D(E) f(E, \mu) \qquad (3.42)$$

$$= \int_0^{\mu} \mathrm{d}E\, D(E) + \frac{\pi^2}{6} \left.\frac{\mathrm{d}D(E)}{\mathrm{d}E}\right|_{E=\mu} (k_B T)^2 + \mathcal{O}(T^4) \qquad (3.43)$$

$$N = \int_0^{E_F} \mathrm{d}E\, D(E) + \Delta\mu D(E_F) + \frac{\pi^2}{6} \left.\frac{\mathrm{d}D(E)}{\mathrm{d}E}\right|_{E=E_F} (k_B T)^2 \qquad (3.44)$$

If N is to stay fixed then we must have

$$\begin{aligned} \Delta\mu &= -\frac{\pi^2}{6} \left.\frac{\mathrm{d}\ln D(E)}{\mathrm{d}E}\right|_{E=E_F} (k_B T)^2 \\ \therefore \mu(T) &\simeq \mu(0) - \frac{\pi^2}{12}\left(\frac{k_B T}{E_F}\right) k_B T \qquad (3.45) \end{aligned}$$

Specific heat

$$\begin{aligned}
U(T) &= \int_0^\infty \mathrm{d}E\; ED(E)f(E,\mu) && (3.46)\\
&= \int_0^\mu \mathrm{d}E\; ED(E) + \frac{\pi^2}{6}\left.\frac{\mathrm{d}ED(E)}{\mathrm{d}E}\right|_{E=\mu}(k_BT)^2 + \mathcal{O}(T^4)\\
\frac{\partial U}{\partial T} &= \mu D(\mu)\frac{\mathrm{d}\mu}{\mathrm{d}T} + \frac{\pi^2}{3}k_B^2T\left[D(E) + E\frac{\mathrm{d}D(E)}{\mathrm{d}E}\right]_{E=E_F} + \mathcal{O}(T^2)\\
&= \frac{\pi^2}{3}k_B^2TD(E_F) + \mu D(\mu)\left[\frac{\mathrm{d}\mu}{\mathrm{d}T} + \frac{\pi^2}{3}k_B^2T\frac{1}{D(E)}\frac{\mathrm{d}D(E)}{\mathrm{d}E}\right]\\
&= \frac{\pi^2}{3}k_B^2TD(E_F) + \mathcal{O}(T^2) && (3.47)\\
&= \left(\frac{\pi k_B^2T}{3\hbar^2}\right)mk_F && (3.48)
\end{aligned}$$

We have (in the last but one step) used eqn 3.45. The linear dependence of specific heat and its proportionality to the mass of the particle are strong predictions. The coefficient of temperature is usually called γ. Putting in approximate values of the free electron density allows shows that $\gamma \sim$ mJmole^{-1}K^{-2}, which is the right order of magnitude for many common metals like copper, silver, gold, aluminium etc. The classical prediction would have been simply a constant $\frac{3}{2}k_B$ per electron, which is clearly wrong. The quantum theory thus correctly accounts for one of the basic thermodynamic responses of metals.

PROBLEM : The following is a very useful exercise called "differentiation under the integral sign". As a mathematical exercise prove the following:

$$\begin{aligned}
\text{If}\quad I(u) &= \int_{a(u)}^{b(u)} F(x,u)dx\\
\therefore \frac{dI}{du} &= \int_{a(u)}^{b(u)} \frac{\partial F(x,u)}{\partial u}dx \quad + F(b,u)\frac{db}{du} \quad - F(a,u)\frac{da}{du} && (3.49)
\end{aligned}$$

where $a(u)$ and $b(u)$ are functions of u. The relation is used but often not stated very clearly.

Mass of the free electrons

We have claimed that we can measure the electronic specific heat - at this point we have not explained how to isolate this contribution from the contribution of the lattice. This we will do later - for the time being let's just assume that it can be done.

Eqn. 3.48 suggest that from a measurement of the specific heat we should be able to infer the mass of the electrons, if we know the density ($k_F^3 = 3\pi^2 n$) as well. for many metals it works within a factor of ∼2 or so. For Silver the γ term agrees with the observed value exactly. (See Kittel's book for a comprehensive list.)

If you look into the middle of the periodic table (f and d block elements), the agreement is not so good. There is a class of compounds for which the observed value suggests that the carriers are somehow very heavy - by factors of 1000. See the Figs. 3.4,3.3.

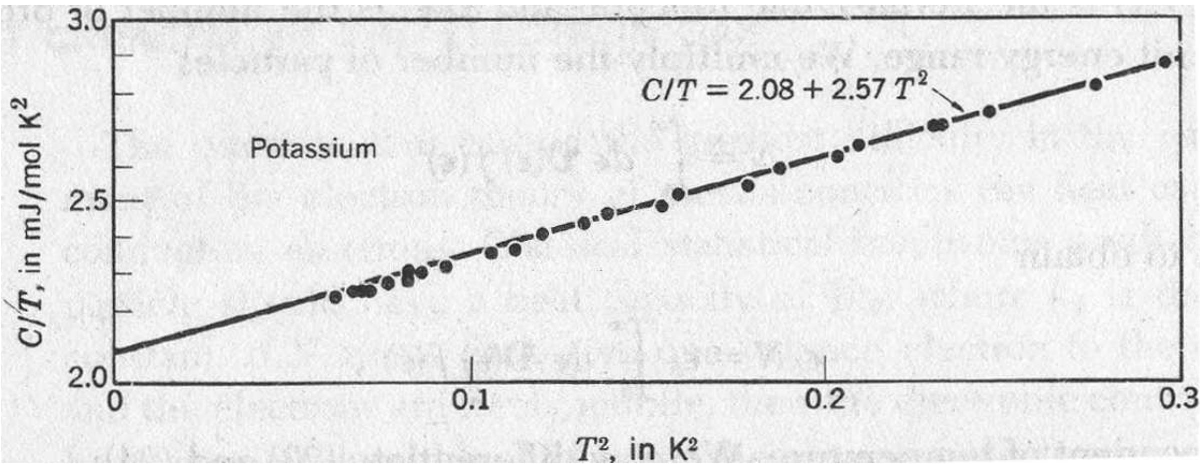


Figure 3.3: Electronic specific heat of Potassium (from Kittel). Notice that at similar temperatures the linear term is only of the order of $milliJoulemole^{-1}K^{-2}$, about 1000 times smaller than what you see in Fig. 3.4.

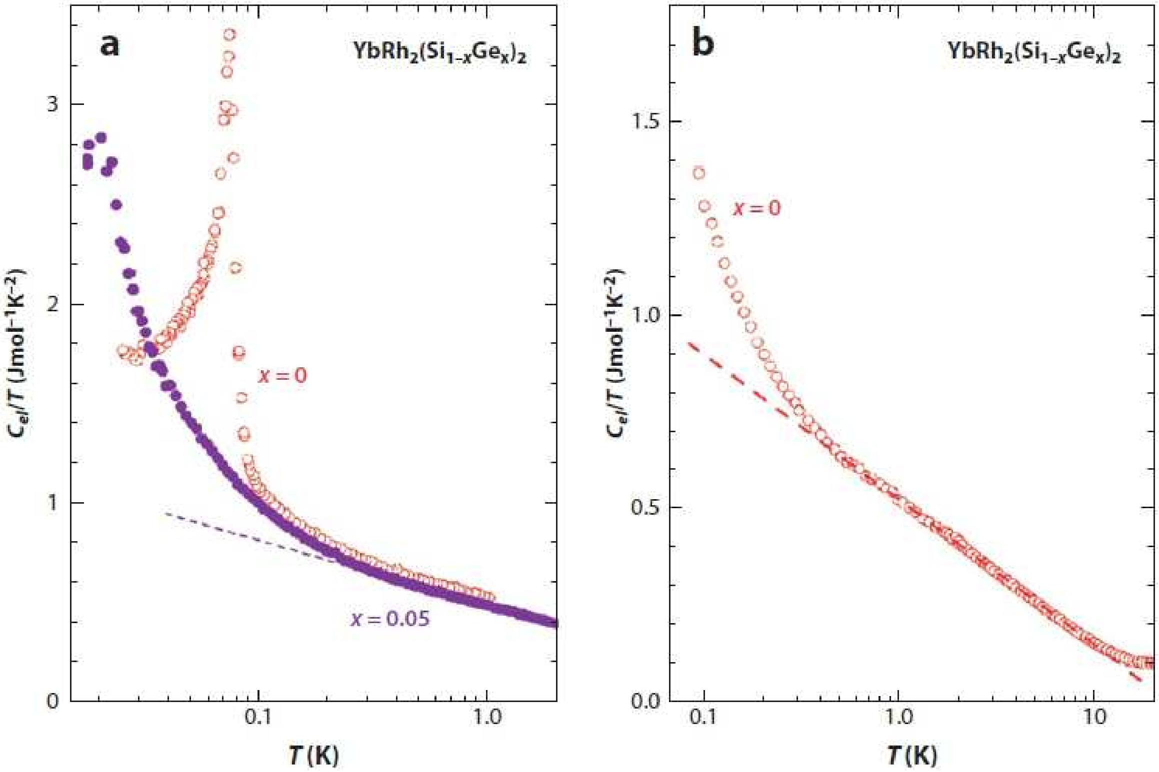


Figure 3.4: Electronic specific heat of "heavy fermion" compounds. These are *not* insulators and have enough free electrons. Their resistivity is $\rho \sim 10\mu\Omega{-}$cm and drops with temperature like metals. But the electrons are very "heavy" in the sense that they have a low-temperature specific heat whose linear term is up to 1000 times larger than the value expected from the free-electron theory. Heavy fermion materials are a specific type of intermetallic compounds, containing elements with 4f or 5f electrons. The figure is taken from "Unconventional Quantum Criticality in Heavy- Fermion Compounds" O. Stockert and F. Steglich, *Annual Review of Condensed Matter Physics* 2011. 2:7999.

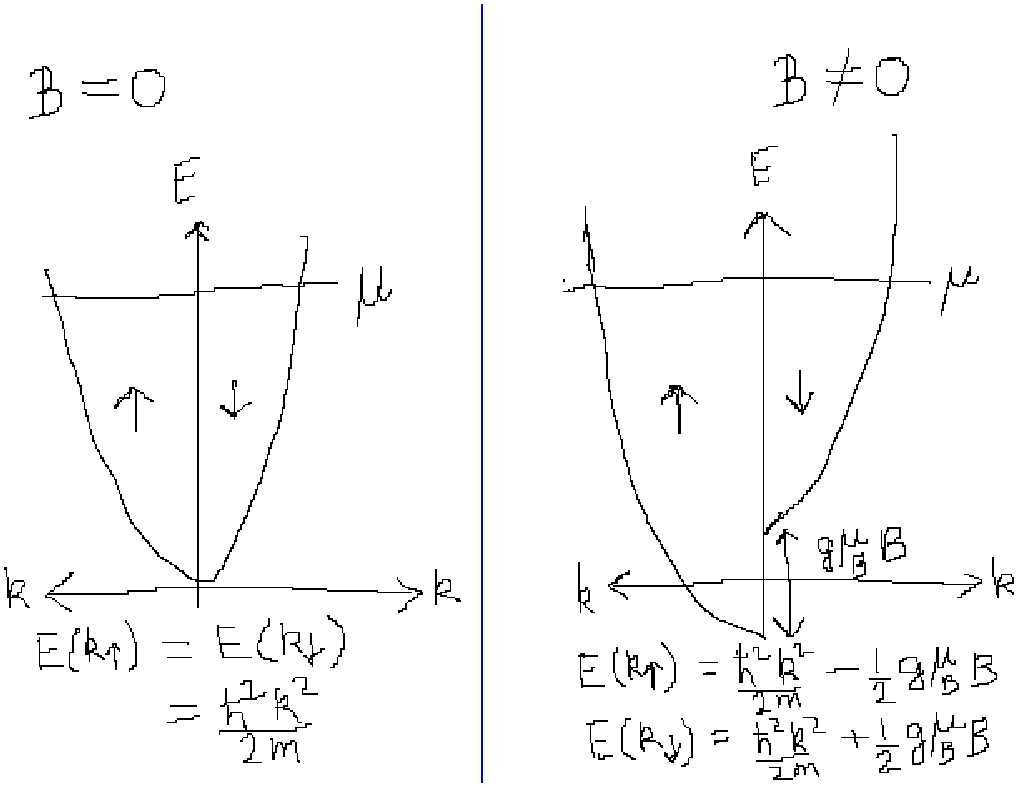


Figure 3.5: The figure shows what eqn 3.50 says. The relative shift of the distribution with respect to the chemical potential has been highly exaggerated. In reality $g\mu_B B/E_F \sim 10^{-3}$ at best

Spin susceptibility at T=0

If we apply a magnetic field then the spin-up and spin-down states at the same **k**, no longer have the same energy. We have at a magnetic field B

$$\begin{aligned} E(k_\uparrow) &= \frac{\hbar^2}{2m}{k_\uparrow}^2 - \sigma g \mu_B B \\ E(k_\downarrow) &= \frac{\hbar^2}{2m}{k_\downarrow}^2 + \sigma g \mu_B B \end{aligned} \qquad (3.50)$$

where recall that for free electrons $g = 2$ and $\sigma = \pm 1/2$, $\mu_B = 9.3 \times 10^{24}\mathrm{JT}^{-1}$. This means that even for fairly large magnetic fields (say B=10Tesla), the shift in energy die to the magnetic field is small compared to E_F, typically 4-5eV. Clearly $g\mu_B B/E_F \sim 10^{-3}$ at best. So we treat it as a small change.

The magnetisation is essentially the difference between the spin-up and spin-down populations. Notice that the chemical potential must be same for both species, but the highest filled k states will not be same ($k_{F\uparrow} > k_{F\downarrow}$).

Fig. 3.5 suggests how we should proceed with the calculation. Using the eqns 3.50 we obtain,

$$k_{F\uparrow} - k_{F\downarrow} = \frac{1}{2}\frac{g\mu_B B}{E_F} k_{F0} \qquad (3.51)$$

Where k_{F0} denotes the fermi wavevector at $B = 0$
Since for each species (up or down) the number of states contained within a sphere of radius k is given by $n = k^3/6\pi^2$, we find

$$\begin{aligned} M &= \mu_B(n_\uparrow - n_\downarrow) \\ &= \frac{k_{F0}^3}{2\pi^2}\frac{g\mu_B B}{E_F}\mu_B \\ &= {\mu_B}^2 D(E_F) B \end{aligned} \qquad (3.52)$$

The last step follows using the density of states relations discussed in previous sections. The "paramagnetic" susceptibility (M/B) then can be written as

$$\chi = {\mu_B}^2 D(E_F) \tag{3.53}$$

Now note:

- χ is independent of temperature. This is what is observed in non-ferromagnetic metals. Compare this with a paramagnetic "Curie" type susceptibility which has a $1/T$ dependence on temperature. The application of Fermi statistics tells us why we should not expect the electron gas to behave as if they are independent spins.
- This "response" too turns out to be proportional to $D(E_F)$,this is a very general feature and will keep appearing.
- Consider an electron far below the fermi level, which finds its spin aligned opposite to the field (spin down). This electron cannot turn over and align with the magnetic field - they can turn over only if there is an unoccupied spin-up state at similar energies. This can happen only close to E_F, this is what stops many electrons from responding to the field. It is a consequence of Fermi statistics/Pauli exclusion principle.

3.2 Transport in the free electron gas: Drude picture

Metals conduct heat and electricity very well. This distinguishes them from insulators. So it is natural to expect that the free electron gas must be responsible for that. Since we have been treating the electron gas in a metal like an ideal gas, we must be able to give a reasonable account of their transport (and optical/high frequency response) using this picture.

Now since we are talking about many moving particles, the obvious question is how they are distributed over energy, or equivalently over momentum/velocity. We assumed their energy is completely kinetic (like ideal gases). During Drude's time the only statistics that was known is the Maxwell-Boltzmann statistics, which would give for the probability of a particle having velocity between v to $v + \mathrm{d}v$

$$f_{MB}(v) = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi v^2 e^{-\frac{mv^2}{2kT}} \tag{3.54}$$

The rms velocity predicted by this is well known:

$$\begin{aligned} \frac{1}{2}mv^2 &= \frac{3}{2}kT \\ \therefore v &= \sqrt{\frac{3kT}{m}} \end{aligned} \tag{3.55}$$

But the Fermi-Dirac distribution would give:

$$f(v) = \frac{m^3}{\pi^2\hbar^3}\frac{v^2}{e^{-\frac{mv^2/2-\mu}{kT}}+1} \tag{3.56}$$

For electrons $m = 9.1 \times 10^{31}$kg, and $E_F \simeq 5$eV for a typical metal. So at $T = 300$K

$$\begin{aligned} \frac{1}{2}mv^2 &= \frac{3}{2}kT \simeq 40\text{meV} \\ \therefore v_{\text{classical}} &\simeq 10^5\text{ms}^{-1} \end{aligned} \tag{3.57}$$

$$\begin{aligned} \frac{1}{2}mv_F^2 &= E_F \simeq 5000\text{meV} \\ \therefore v_F &\simeq 10\times v_{\text{classical}} \simeq 10^6\text{ms}^{-1} \end{aligned} \tag{3.58}$$

This large error means that we underestimate the *energy* of the electrons by a factor of 100, to start with. Yet Drude's approach is useful, and forms a very good starting point. We will see how and why.

Getting Ohm's law

A raindrop falling through air and getting accelerated by gravity reaches a steady velocity. This happens when the energy dissipation is exactly balanced by the energy put in by gravity that accelerates the rain drop. This is an example where a uniform force produces a steady velocity. The case of electrons in a metal is quite similar.

We know that the current voltage relation for a piece of metal is very well described by Ohm's law. It implies that there must be a relaxation mechanism. The current density and its relation to the electric field is given by:

$$\begin{aligned} \mathbf{j} &= ne\mathbf{v}_d && (3.59) \\ \mathbf{j} &= \sigma\mathbf{E} && (3.60) \end{aligned}$$

If the electrons were continuously accelerated by $\mathbf{E}$ then the current could not have been proportional to the velocity, a relaxation and dissipation mechanism is needed to ensure that this does happen. We need

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -e\mathbf{E} - \frac{\mathbf{p}-\mathbf{p_0}}{\tau} \tag{3.61}$$

Here $\mathbf{p_0}$ denotes the equilibrium thermal velocity. To find the current we only need to get the small deviation from the (random) thermal velocity brought about by the electric field. *This change is the drift velocity* $\mathbf{v}_d$.
The steady state solution would then give

$$\begin{aligned} \mathbf{p}-\mathbf{p_0} &= -e\mathbf{E}\tau && (3.62) \\ \mathbf{j} &= ne\mathbf{v}_d = \frac{ne^2\tau}{m}\mathbf{E} && (3.63) \\ & && (3.64) \end{aligned}$$

We introduced the relaxation time τ in an ad-hoc way. If $\mathbf{E}$ is suddenly switched off, then $\mathbf{p}$ would relax back to $\mathbf{p_0}$ as

$$\mathbf{p}(t) = \mathbf{p}_0 + \Delta\mathbf{p}(0)e^{-t/\tau} \tag{3.65}$$

PROBLEM : Show that eqn.3.63 can be generalised for a time varying electric field $\mathbf{E}(\omega)e^{i\omega t}$to

$$\begin{aligned} \mathbf{j}(\omega) &= \sigma(\omega)\mathbf{E}(\omega) \qquad \text{where} \\ \sigma(\omega) &= \frac{\sigma(0)}{1+i\omega\tau} && (3.66) \end{aligned}$$

The relaxation time τ

What is the magnitude and meaning of the relaxation time that we introduced? The magnitude can be estimated from observed values of room temperature resistivity of simple metals, which is $\sim 10\mathrm{n\Omega m}$ and the known electron density $n \sim 10^{28}\mathrm{m}^{-3}$

$$\tau = \frac{m}{ne^2\rho} \simeq 10^{-14}\mathrm{s} \tag{3.67}$$

Drude assumed (with good reason!) that this time arises from the fact that the electrons bump off the fixed lattice ions as they drifted along a particular direction. Using our classical estimate of v from Maxwell-Boltzmann distribution, it would appear that the distance traveled by an electron in time τ would be $l \simeq v\tau = 10^{-9}\mathrm{m}$ (*i.e.* few nm). This appeared to be the right order of magnitude for the size of a unit cell, or lattice constant of typical metals. Although this seemed very plausible, this is not correct.

The collisions which give rise to the relaxation time, arises from two dominant sources - impurities and the lattice vibrations. We will analyse this later. Knowing what we know about the correct fermi distribution of the electrons we know that within this time τ, the electrons would travel 10-100 times this distance.

The ratio of thermal and electrical conductivity

The thermal conductivity of a gas of particles with mean free path l, rms velocity v and specific heat C_v can be written as

$$K = \frac{1}{3} C_v v l. \tag{3.68}$$

At room temperature metals have thermal conductivity 10-100 higher than insulators. You can hold the other end of a burning match-stick of wood but certainly not a metal wire of the same size heated on one end. So it is a reasonable assumption that the electrons *are* carrying the heat current. In pure metals the electronic contribution is dominant at all temperatures. In disordered alloys/ impure metals the lattice contribution (we have not discussed this yet) can be comparable.

PROBLEM: Show that the ratio:

$$\frac{K}{\sigma T} = \frac{\pi^2}{3} \left(\frac{k_B}{e}\right)^2 T \tag{3.69}$$

called the Lorenz number, is independent of all material parameters and as such should provide a robust confirmation of our assumptions, if the observed values agree with the prediction. This indeed happens, see Kittel/Ashcroft for a comprehensive list of observed values for many metals.

3.2.1 Hall effect & Drude electrons in a magnetic field

Hall effect

How do we know the number of electrons and (also the sign of the charge carrier) in a conducting solid. It is interesting to read the line of thinking that led to the "Hall effect" experiment (See Ashcroft/Mermin) but we will just give the summary here.

Consider a bar shaped metal (of known thickness) with crossed electric and mangetic fields as shown. As the electrons acquire a drift velocity due to the electric field, they must also feel the Lorenz force.

$$\mathbf{F} = -|e| \mathbf{v}_d \times \mathbf{B} \tag{3.70}$$

Since the current cannot flow out of the sides, we must assume that an electric field (E_y) arises in body of the sample that balances this force. This electric field (or the integral of this field, as a transverse voltage) is what we can measure. Since

$$\begin{aligned} V_H &= w.E_y = w.v_{dx} B \\ &= w.\frac{j_x}{ne} B \\ &= \frac{B}{ne} \frac{1}{t} I_x \end{aligned} \tag{3.71}$$

A remarkably simple result, with n as the only unknown because B, I_x, t and V_H are measured or set by the experimenter.

Doing it a little better...

Now it is clear that we need to concentrate on the deviation in momentum from the equilibrium, it is implied now that we are talking about momentum in the sense of an average deviation from equilibrium

$$\Delta \mathbf{p}(t) = \langle \mathbf{p}(t) - \mathbf{p}_0 \rangle = m \mathbf{v}_d \tag{3.72}$$

We write out the transport equation 3.61 with the magnetic field present as :

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -e \left(\mathbf{E} + \mathbf{v}_d \times \mathbf{B}\right) - \frac{\mathbf{p}}{\tau} \tag{3.73}$$

Let us assume that $\mathbf{E}$ is in the $x-y$ plane and $\mathbf{B}$ is along $\hat{\mathbf{z}}$, which is a very common configuration: Using the definition of current from eqn. 3.59 and writing out the components we get for the steady state

$$\begin{aligned} 0 &= -eE_x - \frac{eB}{m}p_y - \frac{p_x}{\tau} \\ 0 &= -eE_y + \frac{eB}{m}p_x - \frac{p_y}{\tau} \end{aligned} \tag{3.74}$$

Defining the cyclotron frequency

$$\omega_c = \frac{eB}{m} \tag{3.75}$$

and the zero field conductivity and resistivity as

$$\sigma_0 \equiv \frac{1}{\rho_0} = \frac{ne^2\tau}{m} \tag{3.76}$$

The set of eqns. 3.74 can be written as :

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \rho_0 \begin{pmatrix} 1 & \omega_c\tau \\ -\omega_c\tau & 1 \end{pmatrix} \begin{pmatrix} j_x \\ j_y \end{pmatrix} = \begin{pmatrix} \rho_0 & \frac{B}{ne} \\ -\frac{B}{ne} & \rho_0 \end{pmatrix} \begin{pmatrix} j_x \\ j_y \end{pmatrix} \tag{3.77}$$

PROBLEM : Eqn. 3.77 implies that $\mathbf{j}$ and $\mathbf{E}$ are not parallel to each other in presence of a magnetic field. The angle, δ, between these two vectors is called the "Hall angle". Show that

$$\tan\delta = \omega_c\tau \tag{3.78}$$

Notice that cleaner (large τ) the substance is, easier it is for the electrons to complete a rotation without suffering large collisions. The effect of the magnetic field will be significant if the charged particle can complete one full rotation without getting scattered.

Our analysis would be qualitatively correct as long as the $\mathbf{B}$ is not significantly altered by the material - this means that the metal is not ferromagnetic or superconducting. But everything is not quite correct at this point. If our theory was all right then, the R_H for all metals would have to be of the same sign (corresponding to the negative charge of the electron) and be consistent with the expected density of the free electrons.

3.2.2 Electrodynamics in the electron gas

We want to find out what an electromagnetic wave *with wavelength larger than the mean free path* does in the metal or anything containing a free electron gas. The constraint of large wavelength means that this analysis does not hold for X-rays, but holds for visible light, radio waves, microwaves etc.

Our solution would be complete if we can solve for the full dielectric function $\epsilon(q,\omega)$, that tells us how the electron gas would respond to an electromagnetic field varying as $\sim e^{i(\mathbf{q}\cdot\mathbf{r}-\omega t)}$. This is a central problem in condensed matter physics - but also a very complex one. We will only solve it in some special cases.

We need to take into account the contribution of the polarised lattice and the displacement of the free electrons. The usual dielectric susceptibility tells us about the response of the fixed lattice cores, we will denote the contribution of the lattice using χ_∞ and ϵ_∞. So we write the polarisation produced at finite frequencies but zero wavevector. This means that there is very little (or none) spatial variation of the electric field.

$$\mathbf{P} = (ne)\mathbf{u} + \epsilon_0\chi_\infty\mathbf{E} \tag{3.79}$$

The derivative of the displacement (of the electrons) $\mathbf{u}$ is the drift velocity $\mathbf{v}_d$. We take the equation to frequency domain, using the usual method of writing

$$\mathbf{E}(t) = \tilde{\mathbf{E}}(\omega)e^{-i\omega t} \tag{3.80}$$

Table 3.1: Table of observed Hall coefficients, the ratio in the third column should have been 1 for all, according to the simple free electron theory presented here (The list is taken from Ashcroft & Mermin's book. Notice that the agreement is quite good for alkali metals, but becomes poor as one moves across the periodic table. For trivalent Al, even the *sign* is wrong, *indicating that the charge carriers are behaving as positive particles*!

Metal	Valence	$-1/R_H ne$
Li	1	0.8
Na	1	1.2
K	1	1.2
Rb	1	1.0
Cs	1	0.9
Cu	1	1.5
Ag	1	1.3
Au	1	1.5
Be	2	-0.2
Mg	2	-0.4
In	3	-0.3
Al	3	-0.3

$$\begin{aligned}
\dot{\mathbf{P}} &= (ne)\dot{\mathbf{u}} + \epsilon_0\chi_\infty\dot{\mathbf{E}} \\
&= \mathbf{j} - \epsilon_0\chi_\infty\dot{\mathbf{E}} \\
-i\omega\tilde{\mathbf{P}} &= \sigma(\omega)\tilde{\mathbf{E}} - i\omega\epsilon_0\chi_\infty\tilde{\mathbf{E}} \\
-i\omega\epsilon_0\chi(\omega)\tilde{\mathbf{E}} &= \sigma(\omega)\tilde{\mathbf{E}} - i\omega\epsilon_0\chi_\infty\tilde{\mathbf{E}} \\
\therefore \sigma(\omega) &= -i\omega\epsilon_0\left(\chi(\omega) - \chi_\infty\right) \\
&= i\omega\epsilon_0\left(\epsilon_\infty - \epsilon(\omega)\right) \qquad (3.81)
\end{aligned}$$

Now using eqn 3.66 & eqn 3.81 we can write:

$$\begin{aligned}
\epsilon(\omega) &= \epsilon_\infty - \frac{\sigma(\omega)}{i\omega\epsilon_0} \\
&= \epsilon_\infty - \frac{ne^2}{m\epsilon_0}\frac{1}{i\omega\left(\frac{1}{\tau} - i\omega\right)} \qquad (3.82) \\
&= \epsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\omega/\tau} \qquad (3.83)
\end{aligned}$$

where we have introduced the plasma frequency of the electron gas

$$\omega_p^2 \equiv \frac{ne^2}{m\epsilon_0} \qquad (3.84)$$

for metallic electron densities $\omega_p \sim 10^{16}\mathrm{s}^{-1}$. The reason behind the subscript ∞ should be clear now, at very high frequencies the second term on the RHS of eqn. 3.83 would go to zero and hence $\epsilon(\omega) \to \epsilon_\infty$. For most substances ϵ_∞ is a number $\sim$ 2 or 3.

Real and imaginary parts of the response function

Let's absorb ϵ_∞ in the definition of $\epsilon(q,\omega)$, notice that the response will have real and imaginary parts, which vary a lot with frequency.

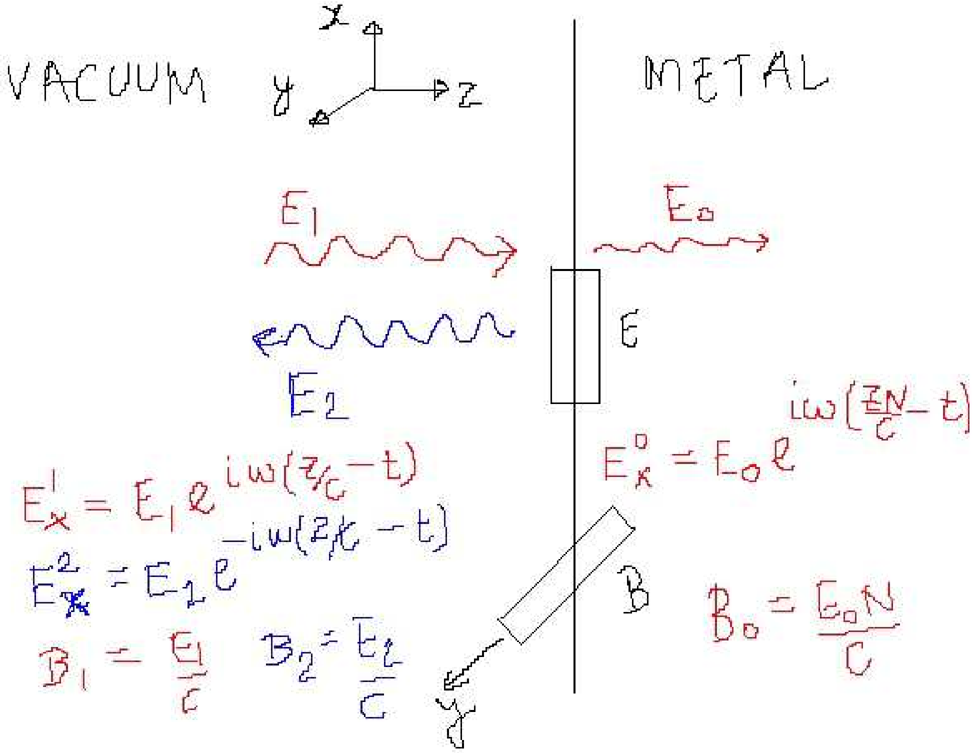


Figure 3.6: Reflection at the vacuum-metal interface, notice the boundary conditions.

Reflection of light

The phase velocity (v_p) of light would change if ϵ changes and the refractive index becomes

$$N = c/V_p = \sqrt{\epsilon(\omega)} \tag{3.85}$$

Here we have assumed that $\mu = 1$, hence the magnetic permeability remains unchanged from the vacuum value, $\mu_0 = 4\pi{\times}10^{-7}$Henry m^{-1}, remains the same, which is fine as long as we don't have ferromagnets, anti-ferromagnets, superconductors etc. Knowing the refractive index of the medium and vacuum we can calculate the reflectivity be using the standard boundary conditions for electromagnetic radiation. We do it for normal incidence. (The line integral of E and B along the two small loops shown must vanish.) Also E and B should be related The electric field vectors of the incident, reflected and transmitted light are:

$$\begin{aligned}
E_x^{inc} &= E_1 e^{i\omega(z/c-t)} && (3.86)\\
E_x^{refl} &= E_2 e^{-i\omega(z/c-t)} && (3.87)\\
E_x^{trans} &= E_0 e^{i\omega(zN/c-t)} && (3.88)\\
& && (3.89)
\end{aligned}$$

The magnetic field vector satisfies:

$$\begin{aligned}
\nabla{\times}\mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && (3.90)\\
\therefore \frac{\partial E_x}{\partial z} &= -\frac{\partial B_y}{\partial t} && (3.91)
\end{aligned}$$

You can show that applying the boundary condition leads to the set of equations:

$$\begin{aligned} E_0 &= E_1 + E_2 && (3.92) \\ B_0 &= B_1 - B_2 \\ \therefore NE_0 &= E_1 - E_2 && (3.93) \end{aligned}$$

Using eqns 3.92 and 3.93 we can eliminate E_0 and find the ratio of the transmitted and reflected amplitudes (E_2/E_1), and hence the reflectivity.
The reflectivity is plotted as a function of ω/ω_p in Fig. 3.7

$$R = \left|\frac{E_2}{E_1}\right|^2 = \left|\frac{1-N}{1+N}\right|^2 \tag{3.94}$$

An useful approximation (prove this as an exercise) when $|N|$ is large is

$$R = 1 - 4\,\mathrm{Re}\left(\frac{1}{N}\right) \tag{3.95}$$

By using some reasonable values for conductivity of metals and the dielectric constants, you can show that N is indeed very large for frequencies significantly below the plasma resonance, ω_p. We can then plot out this quantity (R), by calculating it numerically:

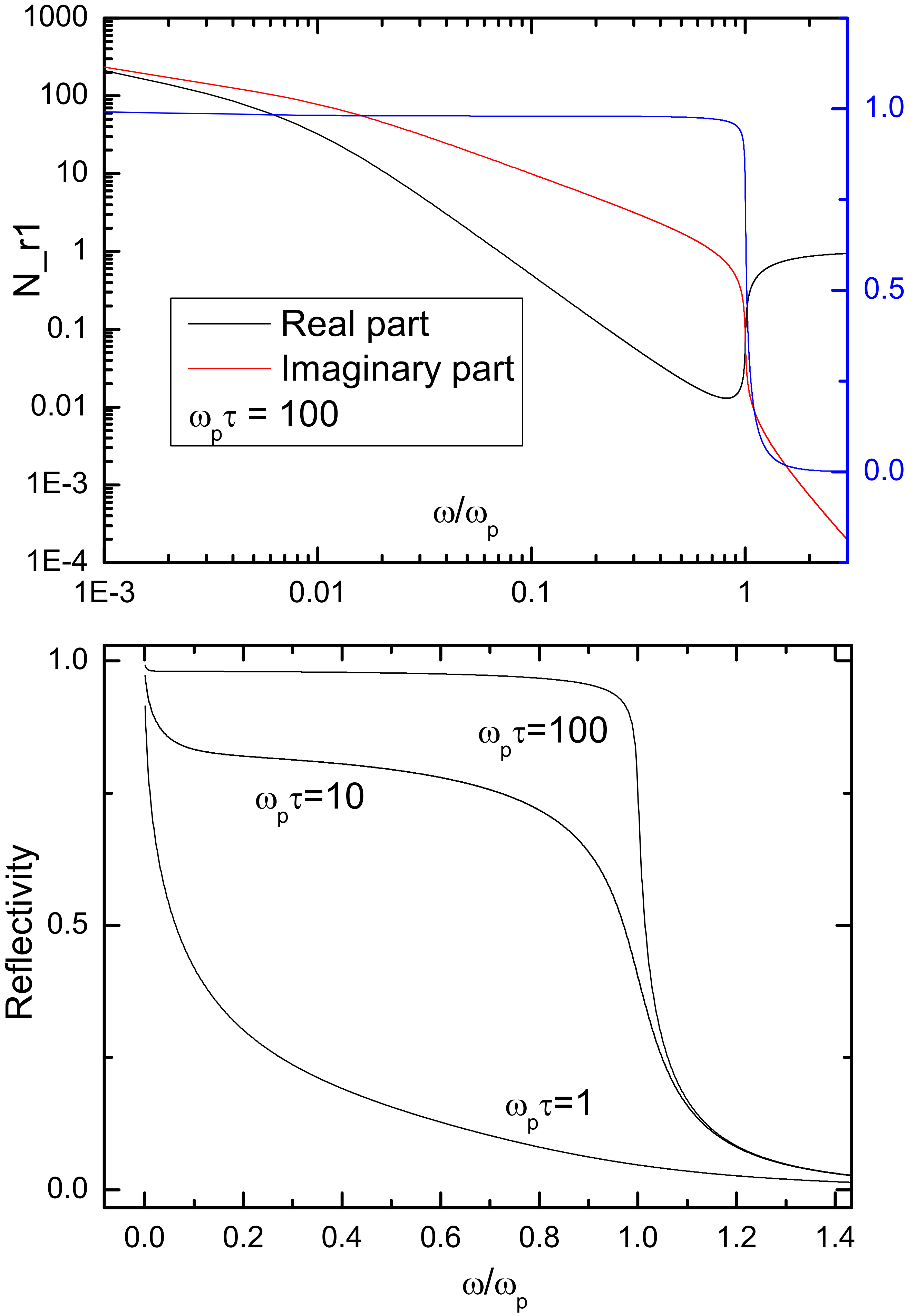


Figure 3.7: Notice how the quality of the metal affects the sharpness of reflection. The metal reflects almost everything below $\omega_p/\sqrt{\epsilon_\infty}$, here we have absorbed the $\sqrt{\epsilon_\infty}$ factor in the definition of the plasma frequency. This tells us why metals can be very shiny but still be transparent in the ultra-violet region.

Can a wave with (q,ω) propagate in this medium?

The usual way to get the wave equation in a medium that can support current flow is as follows. We begin with

$$\begin{aligned} \nabla\times\mathbf{E} &= -\frac{\partial\mathbf{B}}{\partial\mathbf{t}} \\ \nabla\times\mathbf{B} &= \mu_0\mathbf{j} + \epsilon_0\mu_0\frac{\partial\mathbf{E}}{\partial\mathbf{t}} \qquad (3.96) \\ \mathbf{j} &= \sigma\mathbf{E} \qquad (3.97) \end{aligned}$$

If this supports a wavelike solution

$$\mathbf{E}(\mathbf{r},t) = \tilde{\mathbf{E}}e^{i(\mathbf{k}.\mathbf{r}-\omega t)} \tag{3.98}$$

Then it must imply

$$\mathbf{k}^2 = \frac{\omega^2}{c^2}\left(1 - \frac{{\omega_p}^2}{\omega^2 + i\omega/\tau}\right) \tag{3.99}$$

Notice how ω_p has reappeared. If $\omega < \omega_p$, then k is complex and the waves are attenuated. Since $\omega_p\tau >> 1$ in a metal the imaginary part can be ignored at high $\omega > \omega_p$ for simplicity. We have hidden the magnetic permeability $(\mu\mu_0)$ in c^2. Recall that, $1/c^2 = \epsilon_0\epsilon\mu_0\mu$

PROBLEM : Deduce eqn. 3.99 from the set of eqns 3.97.

The sign of $\epsilon(\omega)$ and $\mu(\omega)$

Now notice that the real part of $\epsilon(\omega)$ changes sign near ω_p. A characteristic feature of resonance is that the response changes sign near resonance. Think of a harmonic oscillator with a small dissipation and a resonant frequency ω_0. The following is then well known:
If

$$\ddot{x} + \gamma\dot{x} + {\omega_0}^2 x = \frac{F(t)}{m} \tag{3.100}$$

then (using the standard complex notation)

$$\tilde{x}(\omega) = \frac{\tilde{F}(\omega)/m}{({\omega_0}^2 - \omega^2) - i\omega\gamma} \tag{3.101}$$

shows where the change of sign occurs. Now we may ask these questions:

1. What kind of resonance does eqn 3.99 indicating?
2. Is it possible that μ (the magnetic response) may also have similar features?
3. What if $\epsilon(\omega)$ and $\mu(\omega)$ go negative *simultaneously*? Notice that if that happens then k^2 can have positive solutions inspite of both ($\epsilon(\omega)$, $\mu(\omega)$) being negative over some range of ω.

Let's deal with the first question as a problem:

PROBLEM : Consider the electron system (N per unit volume) displaced from its equilibrium position by a distance d, this creates a capacitor like situation with a surface charge $\sigma = Ned$ per unit area. The block of metal does not carry any *net* charge, so the negative charge density due to the electrons is exactly balanced by the positive lattice at equilibrium, when $d = 0$. Calculate the electric field inside the metal and show that this leads to a restoring force proportional to the displacement and hence a harmonic oscillator. Calculate the frequency of the oscillation. You can assume that the lattice polarizability of the metal is ϵ_∞ at these frequencies for consistency. Calculate typical numbers for metallic densities. Can you see that the presence of *all* the electrons and the lattice is necessary for this phenomena?

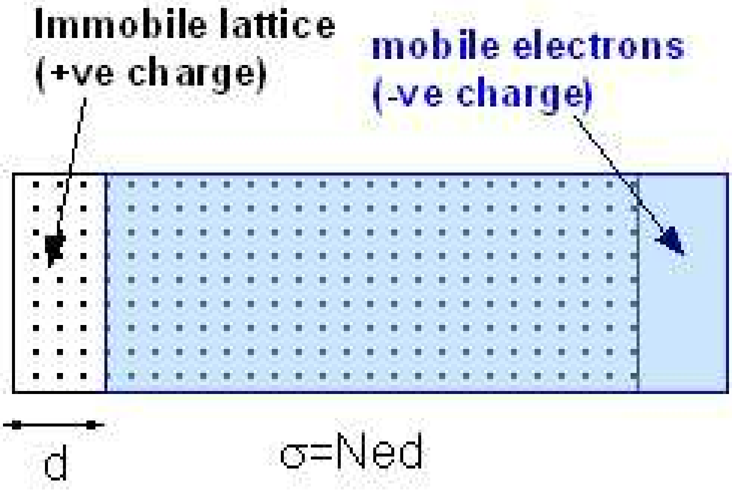


Figure 3.8: A physical interpretation of Plasma oscillation.

We know form experience that the ionosphere (that begins from about 50 km above the earth's surface) reflects radio waves of ~ 5Mhz and higher frequencies. Estimate the ion density in the ionosphere. Would you expect this to be same at all times of the day and night? Can you give a qualitative explanation of why distant radio stations are often better heard at night?

Let's now deal with the second and third points.

Since we don't have exact analogues of an electron (electric charge) in magnetism. The situation is slightly different. Magnetic response must be analyzed in terms of induced current loops. It is possible to artificially pattern some circuit like structure that would have a response very similar to eqn. 3.83. See the article by Pendry *et al* mentioned at the beginning. It is possible to engineer some split ring like circuits where over a small band of frequencies, ϵ and μ are both negative. See fig. 3.9 for a summary and some striking results.

PROBLEM: Electromagnetic field equations can be written in terms of **E**, **D** and **H**. They are:

$$\begin{aligned}
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\
\nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t} \\
\mathbf{B} &= \mu_0 \mu \mathbf{H} \\
\mathbf{D} &= \epsilon_0 \epsilon \mathbf{E}
\end{aligned}$$

For a plane monochromatic wave each quantity varies as $\sim e^{i(\mathbf{k}.\mathbf{r}-\omega t)}$. Show that if $\epsilon < 0$ *and* $\mu < 0$, then $\mathbf{E} \times \mathbf{H}$ points in the direction of $-\mathbf{k}$, in all other cases it points along $\mathbf{k}$. What significance of this result can you think of?

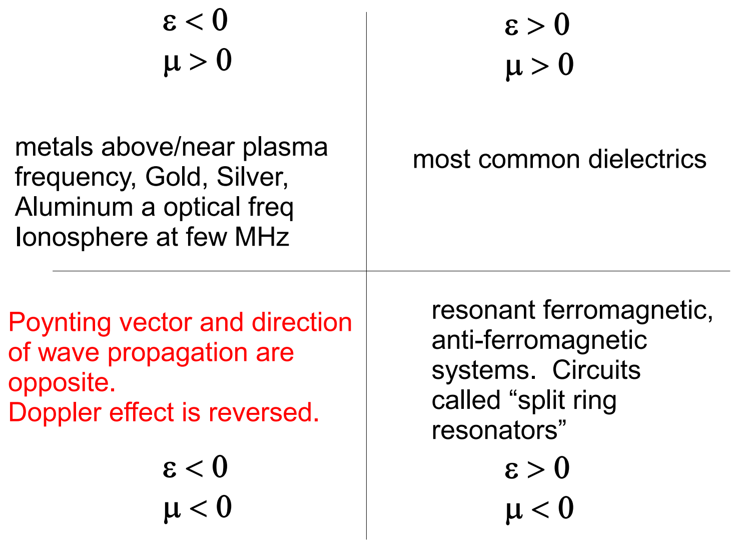


Figure 3.9: In a medium with $\epsilon < 0$ and $\mu < 0$ some remarkable phenomena can happen. We will not go into the details but you can read the two references cited at the beginning if you are interested.

3.2.3 Spatial variation of the potential, $\epsilon(q)$

You would have noticed that we neglected any spatial variation of the electric field. This means that we are assuming that **q** vector of the wave is close to zero (small q = very slow spatial variation). We now show how to handle this, but ignore the time variation of the field. This means that we are shying away from treating time and spatial variation simultaneously. The reason is that calculating $\epsilon(q, \omega)$ is a more complicated problem. It is one of the central problems in condensed matter, but there are some subtle aspects, that we don't want to face now....
Recall that in a dielectric there are no free electrons *i.e* nothing in the conduction band (CB). We haven't introduced the concepts of bands etc, for the time being just think of the band bottom as the lowest energy that the electrons can have. Each lattice site has atoms with bound electrons. When an external electric field (or potential) is applied, the electron wavefunction around the atom shifts a little bit. The result is that (in most cases) each atom becomes like a small dipole. The total dipole moment per unit volume is related to the polarisation (**P**) as

$$\nabla.\mathbf{P} = -\rho_{induced} \tag{3.102}$$

$$\mathbf{P} = \epsilon_0 \chi \mathbf{E} \tag{3.103}$$

If there are free charges then something more must happen, because the free charges in the CB would reorganise themselves more readily in response to a potential. We are going to consider a situation where the external potential is due to some fixed charges - like impurities or surface gates. The external field does not vary with time, it is a "static" response that we want to calculate.

Jellium

The bulk material is electrically neutral, because the charge of the free electron gas is exactly cancelled by the positively charged lattice. To simplify things we totally ignore the crystal structure or discreetness of the lattice at this point. We assume that there is an uniform background (like a block of jelly) in which the electrons move. The equivalent positive charge is spread out in an average structureless way.

Screened potential

Consider now the following basic problem. An external charge ρ_{ext} is placed in a sea of electrons. As a result some charge redistribution will take place creating an additional charge density ρ_{ind}. Let us assume that initially the potential was zero everywhere. The potential must now satisfy

$$\nabla^2 V = -\frac{\rho_{ext} + \rho_{ind}}{\epsilon_0 \epsilon_r} \tag{3.104}$$

where $\epsilon_0\epsilon_r$, denotes the contribution from the dipoles. There are many ϵs that we are talking about. Here we don't want to use the notation ϵ_∞ for the lattice contribution, to remind us of the fact that the field is not varying. Although for the lattice part only, this may not make a huge difference. If there are free electrons then the contribution from those at $\omega = 0$ would dominate.

Notice that we can solve this, if we relate ρ_{ind} to something known. Recall that in equilibrium with no battery connected to the system, the chemical potential (μ) must be constant. But if the electrostatic potential varies, then the bottom of the CB also varies, as shown in Fig 3.10. In a place where the bottom of the CB is high some electrons must be pushed out. How many?

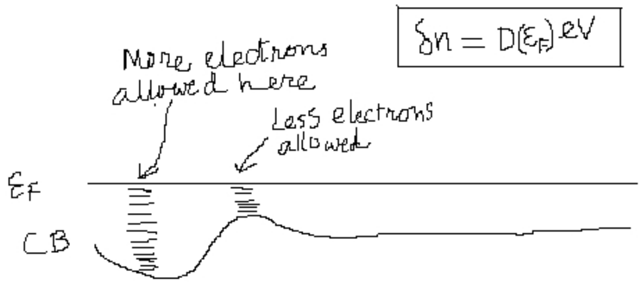


Figure 3.10: Whenever the conduction band rises (electrostatic potential is *lower*) some electrons are pushed out. If the electrostatic potential is higher, then more electrons tend to gather there. This is because the potential is defined with a positive charge in mind.

$$\begin{aligned} \delta n &= D(\mu)eV & (3.105)\\ \therefore\ \rho_{ind} &= -e^2 D(\mu) V & (3.106) \end{aligned}$$

Where $D(\mu)$ is the density of states at the chemical potential (Fermi level). Here the implicit assumption is that $eV << \mu$ and that the potential varies slowly, on the scale of $1/k_F$
We then write eqn 3.104 in a solvable form

$$\begin{aligned} \nabla^2 V - \frac{e^2 D(\mu) V}{\epsilon_0 \epsilon_r} &= -\frac{\rho_{ext}}{\epsilon_0 \epsilon_r} & (3.107)\\ \nabla^2 V - q_{TF}^2 V &= -\frac{\rho_{ext}}{\epsilon_0 \epsilon_r} & (3.108) \end{aligned}$$

Where we introduced the Thomas Fermi wavevector

$$q_{TF} = \left(\frac{e^2 D(\mu)}{\epsilon_0 \epsilon_r}\right)^{1/2} \tag{3.109}$$

Eqn 3.108 is solved by using Fourier transforms. In general the ∇^2 operator becomes multiplication by $-q^2$ in Fourier space.

$$\begin{aligned} V(\mathbf{r}) &= \frac{1}{(2\pi)^3} \int d^3\mathbf{q}\, \tilde{V}(\mathbf{q}) e^{-i\mathbf{q}.\mathbf{r}} & (3.110)\\ \rho(\mathbf{r}) &= \frac{1}{(2\pi)^3} \int d^3\mathbf{q}\, \tilde{\rho}(\mathbf{q}) e^{-i\mathbf{q}.\mathbf{r}} & (3.111) \end{aligned}$$

Using the fourier components we now write:

$$-q^2\tilde{V}(q) - q_{TF}^2\tilde{V}(q) = -\frac{\tilde{\rho}_{ext}(q)}{\epsilon_0\epsilon_r} \tag{3.112}$$

Hence

$$\begin{aligned}
\tilde{V}(q) &= \frac{\tilde{\rho}(q)}{\epsilon_0\epsilon_r\left(q^2+q_{TF}^2\right)} \\
&= \frac{1}{\epsilon_0\epsilon_r}\frac{\tilde{\rho}(q)}{q^2}\frac{1}{\left(1+\frac{q_{TF}^2}{q^2}\right)} \\
&= \frac{\tilde{V}_{ext}(q)}{1+\frac{q_{TF}^2}{q^2}} && (3.113) \\
\epsilon(q,\omega=0) &= 1+\frac{q_{TF}^2}{q^2} && (3.114)
\end{aligned}$$

We have thus related the external potential to the total potential and the induced charge density by eqns 3.114 and 3.106. The relation between the external potential and the final "screened" potential is essentially a relation between **E** and **D**.

PROBLEM : Due to the $D(\mu)$ factor, the Thomas-Fermi wavevector depends on the dimensionality of the system. Show that in 3D :

$$q_{TF}^2 = \frac{1}{\pi^2}\frac{me^2}{\epsilon_0\epsilon_r\hbar^2}k_f = \frac{4}{\pi}\frac{k_f}{a_B} \tag{3.115}$$

where a_B is the Bohr radius of the system.

Extent of validity of the Thomas Fermi approximation

The expression for screening that we deduced is valid for small q, as q approaches $2k_f$ the the method no longer gives correct results.

The screened Coulomb potential

What will be the total potential due to a single point charge placed in a sea of electrons? For a point charge (e) at the origin $\rho(r) = e\delta(r)$. Hence the fourier transform is just a constant $\tilde{\rho}(q) = 1$. Eqn. 3.114 give

$$\tilde{V}(q) = \frac{e}{\epsilon_0\epsilon_r(q^2+q_{TF}^2)} \tag{3.116}$$

The transform can be done by using a simple contour integral and gives

$$V(r) \sim \frac{e}{r}\, e^{-q_{TF}r} \tag{3.117}$$

The only "pole" that will contribute to the integral is at $z = +iq_{TF}$ Notice how fast the total "screened" potential falls off as a result of the presence of the free electron gas. In metals q_{TF} is large enough to make the total potential fall off within one or two lattice constants.

3.3 The basics of Boltzmann transport : shift in the distribution function

In a system of particles, not every particle has the same velocity (momentum). Any system of particles (be it a neutral gas or charged electron gas) is necessarily described by a distribution function. All physically measurable quantities, like electric current, heat current, particle density gradient, scattering rates (*not amplitudes*) must involve the correct thermodynamic distribution function. This distribution function may change over time or the extent of the sample itself.

What is the equation that governs the change of the distribution function? This question was addressed by Boltzmann. We will make no assumptions (to start with) about whether the distribution function is classical (Maxwell-Boltzmann) or Fermi-Dirac or Bose-Einstein, but try to deduce how it changes due to the presence of an external "force", like a electric field resulting from connecting a battery, or the thermal gradient appearing due to heating one end of the sample etc.

The "semiclassical" arguments do involve some subtle points, which we will not fully address, but emphasize the (*very useful*) final result.

- We try to follow a point in phase space $(\mathbf{r}, \mathbf{p})$, assuming that the classical concept of force as a time derivative of momentum is still useful. Later on we will make the association with quantum mechanics by saying that $\mathbf{p} = \hbar\mathbf{k}$. We also assume that velocity and momentum are related in a classical way. This application of quantum mechanics, without completely sacrificing the concept of a "trajectory" is a feature of i *semiclassical* arguments.
- So

$$\begin{aligned} \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} &= \mathbf{v} = \frac{\mathbf{p}}{m} \\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} &= \mathbf{F} \end{aligned} \tag{3.118}$$

- So a point in phase space that was at $(\mathbf{r},\mathbf{p})$ at time t, should find itself at $(\mathbf{r} + \mathbf{v}\delta t,\mathbf{p} + \mathbf{F}\delta t)$ unless it is scattered in the meantime due to collision with other particles.
- This implies

$$f(\mathbf{r} + \mathbf{v}\delta t, \mathbf{p} + \mathbf{F}\delta t, t + \delta t)\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{p}' - f(\mathbf{r}, \mathbf{p}, t)\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{p} = \left.\frac{\delta f}{\delta t}\right|_{\text{collision}} \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{p}\mathrm{d}t \tag{3.119}$$

- We need to prove that the "volume element" deforms but maintains its volume such that

$$\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{p}' = \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{p} \tag{3.120}$$

  But we will assume this and skip the proof.
- This then leads to the transport equation

$$\frac{\partial f}{\partial t} + \mathbf{v}.\nabla_r f + \mathbf{F}.\nabla_p f = \left.\frac{\delta f}{\delta t}\right|_{\text{collision}} \tag{3.121}$$

- We will ignore the first time derivative since we do not expect the distribution function to have an intrinsic time dependence. The collision term must be such that it will pull the system back towards equilibrium if it deviates from the equilibrium distribution, denoted by f^0. Note the two gradients taken w.r.t to position and momentum variables.
- The assumption that the collision is able to restore a disturbed system to equilibrium over a typical time scale τ, such that the equation 3.121 becomes

$$\mathbf{v}.\nabla_r f + \mathbf{F}.\nabla_p f = -\frac{f - f^0}{\tau} \tag{3.122}$$

- equivalently we may write, keeping the utility of **k** vectors in mind:

$$\frac{d\mathbf{r}}{dt}.\nabla_r f + \frac{d\mathbf{k}}{dt}.\nabla_k f = -\frac{f - f^0}{\tau} \tag{3.123}$$

This is called the relaxation time approximation which is extensively used in describing transport.

3.3.1 Effect of E and B on the distribution function

If there are electric and magnetic fields in the system, the "semiclassical" equation of motion would be:

$$\hbar\frac{d\mathbf{k}}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{3.124}$$

We will also assume that f has no explicit time dependence and that $\nabla_r f = 0$ as well - which in general means that there is no density gradient of particles in the system. This assumption is correct if we are dealing with a piece of copper wire at constant temperature, but not necessarily correct for a semiconductor or a even a piece of metal with a thermal gradient. Let us assume that the charge of each particle is "q". For the most common case of electrons in the conduction band we would need to put $q = -|e|$ to get the correct sign of the terms.

Electric field only

With these assumptions, equation 3.123 in presence of an electric field only reduces to

$$\frac{q}{\hbar}\mathbf{E}.\nabla_k f = -\frac{f - f^0}{\tau} \tag{3.125}$$

Then we make the first order approximation by taking the derivative around the equilibrium value

$$\begin{aligned} f(\mathbf{k}) &= f^0(\mathbf{k}) - \frac{q\tau}{\hbar}\mathbf{E}.\nabla_k f^0 && (3.126) \\ &= f^0(\mathbf{k} - \frac{q\tau}{\hbar}\mathbf{E}) && (3.127) \end{aligned}$$

This means that the equilibrium distribution function has retained its functional form but just got shifted by a certain amount. Think of how the graph of a function $f(x)$ would be related to $f(x - a)$. In the figure we have drawn it for a Fermi distribution in 2 dimensions. Note that if the relaxation mechanism is strong then τ would be small. On the other hand if the particle suffers very little scattering then τ would be large and the displacement of the Fermi circle (or sphere) would also be large.

PROBLEM: The free electron density in Copper is n=$8.5 \times 10^{28}\mathrm{m}^{-3}$ and near room temperature the relaxation time of most metals is of the order of $10^{-15} - 10^{-14}$ sec. From this data estimate the fractional shift of the distribution on the scale of the Fermi wavevector (k_F) for an electric field of 10V/m,(*i.e.* calculate $\Delta k/k_F$).

Our target is to calculate the current produced by this state:

$$\begin{aligned} \mathbf{j} &= q\sum_{\mathbf{k}} \mathbf{v}\delta f \\ &= \frac{2q}{(2\pi)^3}\int \mathrm{d}^3\mathbf{k}\ \ \mathbf{v}\delta f && (3.128) \end{aligned}$$

To proceed we need to evaluate eqn 3.127 for the case of Fermi distribution.

$$f^0 = \frac{1}{e^{\beta(E-E_f)} + 1} \tag{3.129}$$

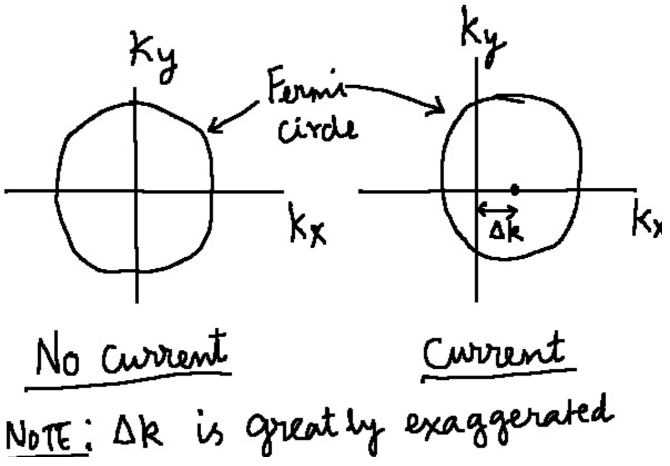


Figure 3.11: Displacement of the Fermi circle results in current flow.

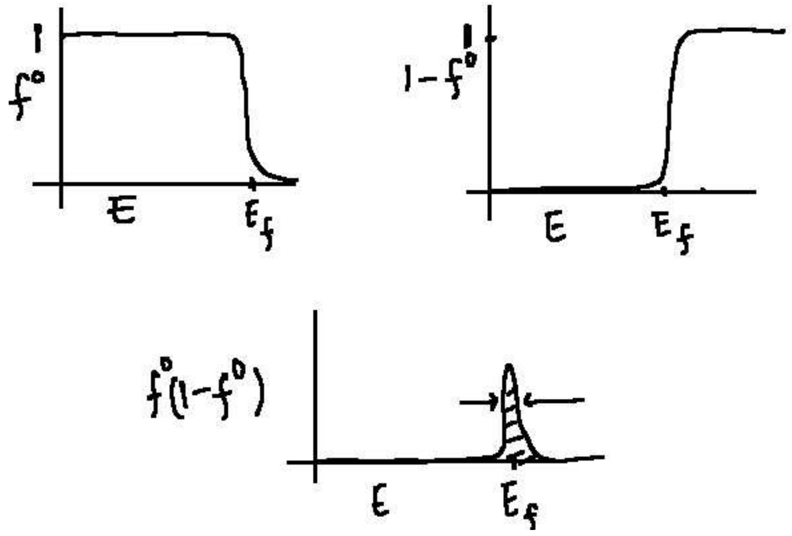


Figure 3.12: How does the product $f^0(1-f^0)$ behave?

$$
\begin{aligned}
\nabla_k f^0 &= -\left(\frac{1}{e^{\beta(E-E_f)}+1}\right)^2 e^{\beta(E-E_f)}\nabla_k\beta(E-E_f) \\
&= -\beta f^0(1-f^0)\nabla_k E \\
&= -\beta f^0(1-f^0)\hbar\mathbf{v_g} \qquad (3.130)
\end{aligned}
$$

Notice that the Fermi level is not a function of $\mathbf{k}$. The end result of 3.130 can also be written as :

$$\nabla_k f^0 = \frac{\partial f^0}{\partial E}\hbar\mathbf{v_g} \qquad (3.131)$$

Equations 3.130 and 3.131 are important results as these derivatives occur frequently in transport related physics. How does the product $f^0(1-f^0)$ behave ? Since f^0 drops sharply around E_f, $(1-f^0)$ must rise sharply around E_f, producing a sharp peak.

Using eqn 3.130 and eqn 3.127 we get

$$\delta f = q\tau\beta f^0(1-f^0)\mathbf{E}.\mathbf{v_g} \qquad (3.132)$$

Notice that the change occurs only near the Fermi surface. This is the generic reason phenomena like electrical or heat conduction are often referred to as a "Fermi surface property". Now we calculate the current as defined in eqn 3.128

$$\begin{aligned} \mathbf{j} &= \frac{q}{4\pi^3}\int \mathrm{d}^3\mathbf{k}\;\; \mathbf{v_g}\left(q\tau\beta f^0(1-f^0)\mathbf{E}.\mathbf{v_g}\right) \\ &= nq\left(\frac{q}{4\pi^3 n}\int \mathrm{d}^3\mathbf{k}\;\; \tau\mathbf{v_g}\otimes\mathbf{v_g}\left(-\frac{\partial f^0}{\partial E}\right)\right).\mathbf{E} \qquad (3.133) \end{aligned}$$

Notice that the part within the large brackets is determined by equilibrium properties of the system only. The outer product ($\otimes$) of two vectors is an object with two indices and can be written out like a matrix. For example

$$\begin{aligned} \mathbf{C} &= \mathbf{A}\otimes\mathbf{B} \quad \text{implies} \qquad (3.134) \\ C_{ij} &= A_i B_j \qquad (3.135) \end{aligned}$$

We will call the quantity inside the bracket as mobility. But it is often not necessary to evaluate this is full generality. We assume that the dispersion relation is spherically symmetric and evaluate the expression for low temperature. Low temperature implies that the Fermi distribution has a sharp drop near E_f and behaves like a step function at that point. The derivative of a step function is a (Dirac) delta function which would pick out the contribution of the integrand around its peak. So we can write

$$\lim_{T\to 0} -\frac{\partial f^0}{\partial E} = \delta(E-E_f) \qquad (3.136)$$

Let's go through the steps for evaluating the mobility integral:

$$\begin{aligned} \overleftrightarrow{\mu} &= \frac{q}{4\pi^3 n}\int \mathrm{d}^3\mathbf{k}\;\; \tau\mathbf{v_g}\otimes\mathbf{v_g}\left(-\frac{\partial f^0}{\partial E}\right) \qquad (3.137) \\ &= \frac{q}{n}\int \mathrm{d}E\;\; D(E)\tau\mathbf{v_g}\otimes\mathbf{v_g}\left(-\frac{\partial f^0}{\partial E}\right) \\ &= \frac{q}{n}\int \mathrm{d}E\;\; D(E)\tau\mathbf{v_g}\otimes\mathbf{v_g}\delta(E-E_f) \quad \text{as}\;\; \mathrm{T}\to 0 \qquad (3.138) \\ & \qquad (3.139) \end{aligned}$$

Now since $\mathbf{v_g} = \hbar\mathbf{k}/m$, we can write:

$$\mu_{ij} = \frac{q}{n}\int \mathrm{d}E\;\; D(E)\;\tau\left(\frac{\hbar}{m}\right)^2 k_i k_j\;\; \delta(E-E_f) \qquad (3.140)$$

This from works in all dimensions, provided the density n is interpreted correctly. Now μ_{ij} will average to zero if $i \neq j$, due to symmetry. If we fix k_i, we can find corresponding pairs of points at k_j and $-k_j$, which will add up to zero. So we need to calculate only the diagonal terms. Since there is nothing to distinguish the x, y or z directions, all the diagonal components must be equal. This allows us to write:

$$\begin{aligned} \mu_{ii} &= \frac{q}{n}\int \mathrm{d}E\;\; D(E)\;\tau\left(\frac{\hbar}{m}\right)^2 \frac{k_x^2+k_y^2+k_z^2}{3}\;\delta(E-E_f) \\ &= \frac{q}{3n}\int \mathrm{d}E\;\; D(E)\;\tau\frac{2E}{m}\;\delta(E-E_f) \qquad (3.141) \end{aligned}$$

Using the expression for density of states in 3D, eqn 3.141 reduces to:

$$\begin{aligned} \mu_{ii} &= \frac{q}{3\pi^2 mn}\int \mathrm{d}E\; k^3\;\tau\;\delta(E-E_f) \\ &= \frac{q\tau}{m} \quad \text{since}\;\; k_F^3 = 3\pi^2 n \qquad (3.142) \end{aligned}$$

Electric and magnetic field together

Now let's recall eqn 3.124 and allow a magnetic field. Eqn 3.126 that described the deviation of the distribution function from equilibrium should now read :

$$f(\mathbf{k}) = f^0(\mathbf{k}) - \frac{q\tau}{\hbar}(\mathbf{E} + \mathbf{v} \times \mathbf{B}).\nabla_k f^0 \tag{3.143}$$

Now because the force term has explicit dependence on $\mathbf{k}$ we can no longer write down the solution by inspection, as we did in eqn 3.127. However we now try a solution of the same form, with an unknown vector $\mathbf{Z}$. Our target is to write $\mathbf{Z}$ as a function of $\mathbf{E}$ and $\mathbf{B}$, but free of $\mathbf{k}$ and $\mathbf{v_g}$. Thus we want $\mathbf{Z}$, such that

$$f(\mathbf{k}) \;=\; f^0(\mathbf{k} - \frac{q\tau}{\hbar}\mathbf{Z}) \tag{3.144}$$

Hence,

$$\delta f \;=\; -\frac{q\tau}{\hbar}\mathbf{Z}.\nabla_k f^0 \tag{3.145}$$

We now use the assumed form (eqn 3.145) with eqns 3.123 and 3.124. This gives:

$$\frac{q}{\hbar}(\mathbf{v} \times \mathbf{B}).\left(\nabla_k f^0 + \nabla_k \delta f\right) + \frac{q}{\hbar}\mathbf{E}.\nabla_k f^0 = -\frac{\delta f}{\tau} \tag{3.146}$$

We already know that $\nabla_k f^0$ points along $\mathbf{v_g}$ and hence the first term in eqn 3.146 gives zero. This leaves us with

$$\frac{q}{\hbar}(\mathbf{v} \times \mathbf{B}).\nabla_k \delta f + \frac{q}{\hbar}\mathbf{E}.\nabla_k f^0 = \frac{q}{\hbar}\mathbf{Z}.\nabla_k f^0 \tag{3.147}$$

Now we need to calculate $\nabla_k \delta f$.

$$\begin{aligned}
\nabla_k \delta f &= \nabla_k \frac{q\tau}{\hbar}\mathbf{Z}.\nabla_k f^0 \\
&= \frac{q\tau}{\hbar}\nabla_k \left(-\beta f^0(1-f^0)\mathbf{Z}.\hbar\mathbf{v_g}\right) \\
&= -\beta q\tau \left((1-f^0)(\mathbf{Z}.\mathbf{v_g})\nabla_k f^0 + f^0(\mathbf{Z}.\mathbf{v_g})\nabla_k(1-f^0) + f^0(1-f^0)\nabla_k(\mathbf{Z}.\mathbf{v_g})\right)
\end{aligned} \tag{3.148}$$

Once again the first two terms in the RHS of 3.148 will give zero when dotted with $\mathbf{v} \times \mathbf{B}$ as they are $\propto \mathbf{v_g}$. The only term left is

$$\nabla_k \mathbf{Z}.\mathbf{v_g} = \nabla_k \mathbf{Z}.\frac{\hbar\mathbf{k} - q\mathbf{A}}{m} = \frac{\hbar}{m}\mathbf{Z} \tag{3.149}$$

In eqn 3.149, $\mathbf{A}$ denotes the vector potential of the magnetic field, $\mathbf{v_g}$ is related to the canonical momentum in presence of a magnetic field in the usual way. Combining eqns 3.148 and 3.149 we can write:

$$(\mathbf{v} \times \mathbf{B}).\nabla_k \delta f = -\beta q\tau f^0(1-f^0)(\mathbf{v} \times \mathbf{B}).\frac{\hbar}{m}\mathbf{Z} \tag{3.150}$$

So eqn 3.147 now simplifies to:

$$\begin{aligned}
&-\frac{\hbar}{m}\beta q\tau f^0(1-f^0)(\mathbf{v} \times \mathbf{B}).\mathbf{Z} + (\mathbf{E} - \mathbf{Z}).\nabla_k f^0 = 0 \\
\therefore\ &-\frac{\hbar}{m}\beta q\tau f^0(1-f^0)(\mathbf{v} \times \mathbf{B}).\mathbf{Z} + (\mathbf{E} - \mathbf{Z})\beta f^0(1-f^0)\hbar\mathbf{v_g} = 0 \\
\therefore\ &\frac{q\tau}{m}(\mathbf{v_g} \times \mathbf{B}).\mathbf{Z} + (\mathbf{E} - \mathbf{Z}).\mathbf{v_g} = 0 \\
\therefore\ &\frac{q\tau}{m}(\mathbf{B} \times \mathbf{Z}).\mathbf{v_g} + (\mathbf{E} - \mathbf{Z}).\mathbf{v_g} = 0 \\
\therefore\ &\mathbf{E} = \mathbf{Z} - \frac{q\tau}{m}\mathbf{B} \times \mathbf{Z}
\end{aligned} \tag{3.151}$$

We call $\mathbf{Z}$ as the Hall vector. When both $\mathbf{E}$ and $\mathbf{B}$ fields are present, this quantity in some way, "replaces" the electric field in the transport equation. But we still need to express $\mathbf{Z}$ explicitly in terms of $\mathbf{E}$ and $\mathbf{B}$, with $\mu = q\tau/m$. The proof is left as an exercise.

$$\mathbf{Z} = \frac{\mathbf{E} + \mu\mathbf{B} \times \mathbf{E} + \mu^2(\mathbf{B}.\mathbf{E})\mathbf{B}}{1 + \mu^2 B^2} \tag{3.152}$$

PROBLEM: If $\mathbf{E} = \mathbf{Z} - \mathbf{A} \times \mathbf{Z}$, then show that

$$\mathbf{Z} = \frac{\mathbf{E} + \mathbf{A} \times \mathbf{E} + (\mathbf{A}.\mathbf{E})\mathbf{A}}{1 + A^2}$$

Hint : Try $\mathbf{A} \times \mathbf{E}$ and $\mathbf{A}.\mathbf{E}$

Chapter 4

Electrons in a periodic potential: Bloch's theorem

We know that the lattice sites form a repetitive array, so the potential felt by an electron moving in it will be periodic as well, this means (as we have discussed during X-ray scattering) we can write potential as a Fourier sum over all reciprocal lattice vectors

4.1 Derivation of the theorem

$$\begin{aligned} H &= T+V \\ &= T+\sum_{\mathbf{G}} v_{\mathbf{G}} e^{i\mathbf{G}.\mathbf{r}} \end{aligned} \tag{4.1}$$

Let's assume we know all the components $v_{\mathbf{G}}$ and want to solve for the the eigenvalues and eigenfunctions. We know that if the potential was zero, then the solutions would have been free particle like, now as a consequence of the existence of the potential, many plane waves states $|\mathbf{k}\rangle$ must be mixed, so

$$|\psi\rangle = \sum_{\mathbf{k}} c_{\mathbf{k}} |\mathbf{k}\rangle \tag{4.2}$$

where

$$\langle \mathbf{r}|\mathbf{k}\rangle = \frac{1}{\sqrt{L^3}} e^{i\mathbf{k}.\mathbf{r}} \tag{4.3}$$

To solve for $c_{\mathbf{k}}$ is the obvious target. We let H act on the *ket* $|\psi\rangle$, which we assume to be an eigenstate of H, and then left multiply with the state *bra* $\langle \mathbf{k}'|$.

$$\begin{aligned} \langle \mathbf{k}'|H|\psi\rangle &= \sum_{\mathbf{k}} c_{\mathbf{k}} \langle \mathbf{k}'|T|\mathbf{k}\rangle + \sum_{\mathbf{k}} c_{\mathbf{k}} \langle \mathbf{k}'|V|\mathbf{k}\rangle \\ \therefore \qquad E c_{\mathbf{k}'} &= \frac{\hbar^2}{2m} k'^2 c_{\mathbf{k}'} + \sum_{\mathbf{k}} \sum_{\mathbf{G}} c_{\mathbf{k}} v_{\mathbf{G}} \delta_{\mathbf{k}',\mathbf{k}+\mathbf{G}} \\ \therefore \qquad 0 &= \left(\frac{\hbar^2}{2m} k'^2 - E\right) c_{\mathbf{k}'} + \sum_{\mathbf{G}} c_{\mathbf{k}'-\mathbf{G}} v_{\mathbf{G}} \end{aligned} \tag{4.4}$$

The sum is over the reciprocal lattice vectors $\mathbf{G}$. This has the form of an eigenvalue equation. But we have only one equation for a large number of unknowns -all the $c_{\mathbf{k}-\mathbf{G}}$. In principle there can be an infinite number of them.
Now notice a few important points

1. Eqn 4.4 connects a state $\mathbf{k}$ with states that can be reached by reciprocal lattice translations. *i.e.* The potential only mixes the states $|\mathbf{k}\rangle$, $|\mathbf{k}-\mathbf{G_1}\rangle$, $|\mathbf{k}-\mathbf{G_2}\rangle$ and so on.

2. But there are some states which cannot be taken to one another by reciprocal lattice translations: take any two states in the first Brillouin zone, that are not on the zone boundary.
3. The previous two points together imply that **k** from the first Brillouin zone can be used to label a wavefunction that will contain $|\mathbf{k}\rangle$, $|\mathbf{k}-\mathbf{G_1}\rangle$, $|\mathbf{k}-\mathbf{G_2}\rangle$.... These wavefunctions are the Bloch wavefunctions.

Now we complete the solution. In the process of getting to eqn. 4.4 we could have left multiplied with another state $\langle\mathbf{k}'-\mathbf{K}|$ where **K** is any reciprocal lattice vector. Then we would get the equation

$$\left(\frac{\hbar^2}{2m}(\mathbf{k}'-\mathbf{K})^2 - E\right)c_{\mathbf{k}'-\mathbf{K}} + \sum_{\mathbf{G}} c_{\mathbf{k}'-\mathbf{K}-\mathbf{G}}v_{\mathbf{G}} = 0 \tag{4.5}$$

Hence

$$\left(\frac{\hbar^2}{2m}(\mathbf{k}'-\mathbf{K})^2 - E\right)c_{\mathbf{k}'-\mathbf{K}} + \sum_{\mathbf{G}'} c_{\mathbf{k}'-\mathbf{G}'}v_{\mathbf{G}'-\mathbf{K}} \quad = \quad 0$$

where we have written $\mathbf{G}' = \mathbf{K}+\mathbf{G}$, remembering that $\mathbf{k}'$ and **K** are fixed vectors for a particular row. The primes are now unnecessary in the summation indices.

$$\left(\frac{\hbar^2}{2m}(\mathbf{k}-\mathbf{K})^2 - E\right)c_{\mathbf{k}-\mathbf{K}} + \sum_{\mathbf{G}} c_{\mathbf{k}-\mathbf{G}}v_{\mathbf{G}-\mathbf{K}} = 0$$

G and **K** run over the same set of (reciprocal lattice) vectors. Now we see the following:

1. There will be as many solutions as the number of reciprocal lattice vectors we keep.
2. Each **k** vector (not a reciprocal lattice vector) in the first Brillouin zone gives a matrix for us to solve. The $E(\mathbf{k})$ relation thus has as many branches as the dimension of the matrix.

We can see that the wavefunction must have the form

$$\begin{aligned}
|\psi_{\mathbf{k}}\rangle &= \sum_{\mathbf{G}} c_{\mathbf{G}}|\mathbf{k}-\mathbf{G}\rangle \\
\langle\mathbf{r}|\psi_{\mathbf{k}}\rangle \equiv \psi_{\mathbf{k}}(\mathbf{r}) &= \sum_{\mathbf{G}} c_{\mathbf{G}}e^{i(\mathbf{k}-\mathbf{G}).\mathbf{r}} \\
&= e^{i\mathbf{k}.\mathbf{r}}\sum_{\mathbf{G}} c_{\mathbf{G}}e^{-i\mathbf{G}.\mathbf{r}} \\
&= e^{i\mathbf{k}.\mathbf{r}}u(\mathbf{r}) \qquad (4.6) \\
\therefore \quad \psi_{\mathbf{k}}(\mathbf{r}+\mathbf{R}) &= e^{i\mathbf{k}.(\mathbf{r}+\mathbf{R})}u(\mathbf{r}+\mathbf{R}) \\
&= e^{i\mathbf{k}.(\mathbf{r}+\mathbf{R})}\sum_{\mathbf{G}} c_{\mathbf{G}}e^{-i\mathbf{G}.\mathbf{r}}e^{-i\mathbf{G}.\mathbf{R}} \qquad (4.7) \\
&= e^{i\mathbf{k}.\mathbf{R}}\psi(\mathbf{r}) \qquad \left(\because e^{-i\mathbf{G}.\mathbf{R}} = 1\right)
\end{aligned}$$

Since where **k** is in the first Brillouin zone. This is Bloch's theorem in many equivalent forms. The function $u(\mathbf{r})$ has the symmetry of the direct lattice.

4.1.1 Translation invariance and Bloch's theorem

The Hamiltonian of the lattice is invariant under a translation by any lattice vector **R**. We construct the translation operator in terms of the momentum operator **p** as

$$\begin{aligned}
T(\mathbf{R}) &= e^{i\mathbf{R}.\mathbf{p}/\hbar} \qquad (4.8) \\
T(\mathbf{R})^{-1}HT(\mathbf{R}) &= H \\
\therefore [T(\mathbf{R}), H] &= 0 \qquad (4.9)
\end{aligned}$$

Since $\mathbf{p}$ is hermitian, $T(\mathbf{R})$ is unitary. The commutator tells us that $T(\mathbf{R})$ and H has simultaneous eigenfunctions. So if $|\Psi\rangle$ is an eigenfunction of H, it should also satisfy the two following conditions

$$\begin{aligned} T(\mathbf{R})\psi(\mathbf{r}) &= c(\mathbf{R})\psi(\mathbf{r}) && (4.10) \\ T(\mathbf{R})\psi(\mathbf{r}) &= \psi(\mathbf{r}+\mathbf{R}) && (4.11) \end{aligned}$$

Now what is $c(\mathbf{R})$?

$$\begin{aligned} \langle\mathbf{k}|T(\mathbf{R})|\psi\rangle &= \langle\psi|T^\dagger(\mathbf{R})|\mathbf{k}\rangle^* \\ &= \left(\int d^3\mathbf{r}\ \psi^*(\mathbf{r}).e^{-i\mathbf{R}.\mathbf{p}/\hbar}.e^{i\mathbf{k}.\mathbf{r}}\right)^* \\ &= \left(\int d^3\mathbf{r}\ \psi^*(\mathbf{r}).e^{i\mathbf{k}.(\mathbf{r}-\mathbf{R})}\right)^* \\ &= e^{i\mathbf{k}.\mathbf{R}}\langle\mathbf{k}|\psi\rangle && (4.12) \end{aligned}$$

Then left multiply 4.10 with $\langle\mathbf{k}|$ and put together the result with 4.12. This leads to

$$\left(c(\mathbf{R}) - e^{i\mathbf{k}.\mathbf{R}}\right)\langle\mathbf{k}|\psi\rangle = 0 \tag{4.13}$$

Using this and eqn 4.11 you should be able to reproduce the properties of the Bloch functions.

4.1.2 Significance of k

The vector $\mathbf{k}$ looks like momentum ($\hbar\mathbf{k}$), but it is not. To see this, let us calculate the momentum of a particle in a Bloch state. We need the expectation value of the momentum operator $\mathbf{p} = -i\hbar\nabla$

$$\langle\mathbf{p}\rangle = \langle\psi_\mathbf{k}| - i\hbar\nabla|\psi_\mathbf{k}\rangle = \hbar\mathbf{k} + \hbar\sum_\mathbf{G}\mathbf{G}|c_{\mathbf{k}+\mathbf{G}}|^2 \tag{4.14}$$

PROBLEM : Prove eqn. 4.14.

To distinguish $\hbar\mathbf{k}$ from momentum, we will call it the *crystal momentum.* The significance of this will be clear when we write the equations of semiclassical dynamics of the electron.

Response of a Bloch electron to an Electric field

Consider a chain of atoms in the form of a loop (1D) so that periodic boundary conditions are easy to apply. However it is precisely because of the periodic boundary condition an extra electric field is difficult to add to the potential $V(x)$. The extra field (resulting from connecting a voltage source to measure electric current for example) is hard to accommodate if we write it as $V = -eEx$. We use another way of introducing an extra field. Recall

$$\mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t} - \nabla V \tag{4.15}$$

We introduce the extra E using a time dependent $A = -Et$. So the Hamiltonian including $A(t)$ becomes:

$$H = \frac{(p+eA)^2}{2m} + V(x) \tag{4.16}$$

This has *time dependent* eigenfunctions (of the Bloch form) and eigenenergies, which we call $|\Psi(x,t)\rangle$ and $\varepsilon(t)$ to distinguish it from the electric field E.

$$\left(\frac{(p+eA)^2}{2m} + V(x)\right)|\Psi(x,t)\rangle = \varepsilon(t)|\Psi(x,t)\rangle \tag{4.17}$$

We now introduce an unknown parameter Λ and write

$$\begin{aligned} |\Psi(x,t)\rangle &= e^{i\Lambda x}|\Phi(x,t)\rangle && (4.18)\\ \therefore (p+eA)|\,\Psi(x,t)\rangle &= \hbar\Lambda e^{i\Lambda x}|\Phi(x,t)\rangle - i\hbar e^{i\Lambda x}\frac{\partial}{\partial x}|\Phi(x,t)\rangle + eAe^{i\Lambda x}|\Phi(x,t)\rangle && (4.19) \end{aligned}$$

The first and last terms will cancel if we choose

$$\hbar\Lambda + eA = 0 \tag{4.20}$$

This choice results in the simplification

$$(p+eA)|\,\Psi(x,t)\rangle = -i\hbar e^{i\Lambda x}\frac{\partial}{\partial x}|\Phi(x,t)\rangle = e^{i\Lambda x}p|\Phi(x,t)\rangle \tag{4.21}$$

Eqn. 4.17 can now be written as (a few intermediate steps are left for you to work out...)

$$\left(\frac{p^2}{2m} + V\right)|\Phi\rangle = \varepsilon(t)|\Phi\rangle \tag{4.22}$$

$|\Phi\rangle$ must have solutions of the Bloch form such that

$$\Phi(x) = e^{ikx}u_k(x) \tag{4.23}$$

Using this and the periodic boundary conditions we get:

$$\begin{aligned} \Psi(x+L) &= \Psi(x)\\ \therefore e^{-ieA(x+L)/\hbar}e^{ik(x+L)}u_k(x+L) &= e^{-ieA(x)/\hbar}e^{ik(x)}u_k(x)\\ \therefore e^{-ieAL/\hbar}e^{ikL} &= 1\\ \therefore eEt/\hbar + k &= 2n\pi/L\\ \therefore \hbar\frac{dk}{dt} &= -eE && (4.24) \end{aligned}$$

The label k *behaves like* momentum. The result is a striking consequence of periodicity and elementary quantum mechanics put together! Later on we will treat this problem with an electric and magnetic field as well.

4.1.3 Origin of the band gap: solving the matrix equation

Let us for the moment consider a simplified "toy" case to appreciate an (perhaps the most) important consequence of a periodic potential.

1. We consider a 1-dimensional case (chain with a period a)
2. We consider a periodic potential

$$V(x) = 2V_0\cos\frac{2\pi x}{a} = V_0e^{i\frac{2\pi}{a}x} + V_0e^{-i\frac{2\pi}{a}x} \tag{4.25}$$

which means that only two (reciprocal lattice) components of the potential contribute:

$$V_{\frac{2\pi}{a}} = V_{-\frac{2\pi}{a}} = V_0 \tag{4.26}$$

So as to write out the matrix of eqn 4.5 we need to order all the reciprocal vectors according to some rule (which we can frame for ourselves). In the 1D case we can write all the RLVs as:

$$V_n = n.\frac{2\pi}{a} \qquad (\text{n} = ...-2,-1,0,1,2,....) \tag{4.27}$$

So the matrix resulting from the eqn. 4.6 is a tridiagonal matrix, a *part* of which looks like (assuming $\hbar^2/2m = 1$) for simplicity

$$\begin{vmatrix} (k+3\frac{2\pi}{a})^2-E & V_{-\frac{2\pi}{a}} & 0 & 0 & 0 & 0 & 0 \\ V_{\frac{2\pi}{a}} & (k+2\frac{2\pi}{a})^2-E & V_{-\frac{2\pi}{a}} & 0 & 0 & 0 & 0 \\ 0 & V_{\frac{2\pi}{a}} & (k+\frac{2\pi}{a})^2-E & V_{-\frac{2\pi}{a}} & 0 & 0 & 0 \\ 0 & 0 & V_{\frac{2\pi}{a}} & k^2-E & V_{-\frac{2\pi}{a}} & 0 & 0 \\ 0 & 0 & 0 & V_{\frac{2\pi}{a}} & (k-\frac{2\pi}{a})^2-E & V_{-\frac{2\pi}{a}} & 0 \\ 0 & 0 & 0 & 0 & V_{\frac{2\pi}{a}} & (k-2\frac{2\pi}{a})^2-E & V_{-\frac{2\pi}{a}} \\ 0 & 0 & 0 & 0 & 0 & V_{\frac{2\pi}{a}} & (k-3\frac{2\pi}{a})^2-E \end{vmatrix} = 0 \tag{4.28}$$

In the notation of eqn. 4.6, **G** varies along a row, **K** is kept fixed. For every **K** we have a new row. If we had more components of the potential ($e.g. V_{\frac{4\pi}{a}}, V_{-\frac{4\pi}{a}}$) then another band would appear symmetrically about the diagonal line. The matrix is necessary symmetrical and "band diagonal".

Zero potential case : the "empty lattice"

Now let's consider the apparently trivial case (but there is purpose!) where all the potential components are zero. Then the eignevalues must be the free electron values

$$E = \frac{\hbar^2}{2m}\left(k - V_{n\frac{2\pi}{a}}\right)^2 \qquad \text{with} \quad -\frac{\pi}{a} < k < \frac{\pi}{a} \tag{4.29}$$

It might be a little counter-intuitive to see how the plot with shifted parabolas look. We are used to seeing one continuous parabola ($E \propto k^2$) as the dispersion, because in the free particle case there is no lattice. But if we now shift any vector from outside the first Brillouin zone into the first zone by reciprocal vector translations then this is how it will look. This way of plotting it is called the "reduced zone" scheme and is the most commonly used way of plotting bandstructure. If the lattice is 2d or 3d (Kittel gives an example) then these plots can look quite striking, but it is perfectly logical way of plotting it. We will see soon that the effect of a small non-zero potential would be to smooth the many sharp corners and crossings in the plot as seen in Fig. 4.1.

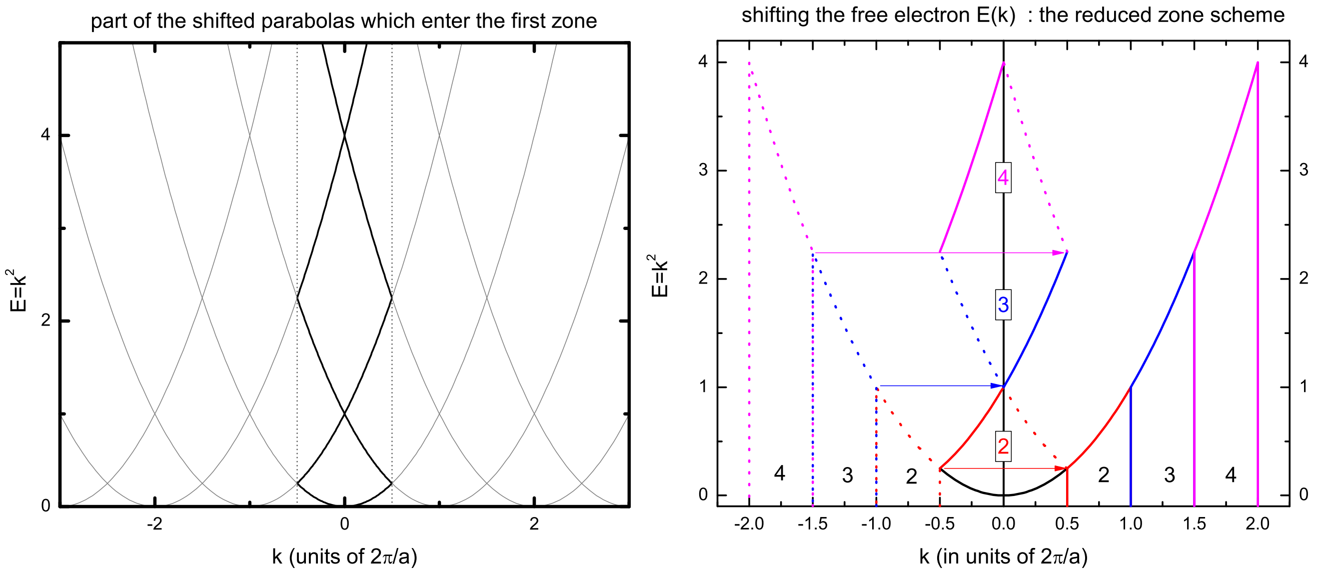


Figure 4.1: (left)Free electron $E(k)$ in the reduced zone scheme. The equations $E = \left(k - n.\frac{2\pi}{a}\right)^2$, and the part which enters the first Brillouin zone. The part in the first zone is in black, the outside parts are grayed out. (right) Another similar plot , showing how the same can be obtained from $E = k^2$ by shifting different parts by different reciprocal lattice translations.

Turning on a small potential

We can see that at the point where the shifted parabolas cross, two eignevalues are degenerate. The first case happens at $k = \frac{\pi}{a}$. Here the $E = k^2$ and $E = \left(k - \frac{2\pi}{a}\right)^2$ branches give the same value if $V = 0$. What happens if we turn on a small V? We can treat the problem with perturbation (and we will later) but first we just solve for the eigenvalues of the matrix equation 4.28. The plot shows a crucial prediction of the Bloch equation. If we have a potential with a Fourier component ($v_{\mathbf{G}}$) then a gap opens at $\mathbf{G}/2$ These are precisely the zone boundaries. Note that the sharp crossings of Fig. 4.1 have been "rounded off". It turns out that this is a generic feature of turning on a small interaction potential and you will see this in many situations.

PROBLEM : A quick (though bit handwaving) way of seeing what happens when a small potential is "turned on" is to solve the 2×2part of the matrix eqn 4.28:

$$\begin{vmatrix} k^2 - E & V_0 \\ V_0 & \left(k - \frac{2\pi}{a}\right)^2 - E \end{vmatrix} = 0 \tag{4.30}$$

where $V_{-\frac{2\pi}{a}} = V_{\frac{2\pi}{a}} = V_0$. Now solve for the eignevalues, this should give the behaviour near $k = \pi/a$. You should be able to show that the two eigenvalues at $k = \pi/a$ differ by $2V_0$.

Extended and repeated zone schemes

Exactly the same data can be plotted in some different ways. The repeated zone scheme simply means that we emphasize the periodicity of the solution that $E(\mathbf{k}) = E(\mathbf{k} + \mathbf{G})$. The extended zone scheme shows very clearly the deviation from the free electron parabola. The data used to plot Fig. 4.2 has been plotted in these two schemes in the next figure, Fig .4.3.

4.1.4 Band structure as a perturbation problem

The potential is being turned on slowly and we traced the evolution of the energy eignevalues starting with the free electron case. Surely, we could have treated it as a perturbation problem. The answer is "yes", but there are some degeneracies to be aware of. This fact and the connection with Bloch's theory will now become clear.

Time-independent perturbation tells us that the unperturbed state kets $|\mathbf{k}\rangle$ will now be mixed. The point to note is that only those states of the form $|\mathbf{k} + \mathbf{G}\rangle$ will mix with $|\mathbf{k}\rangle$. We write the perturbed state kets to *first* order and perturbed energies to *second* order.

$$\begin{aligned} |\psi_{\mathbf{k}}\rangle &= |\mathbf{k}\rangle + \sum_{\mathbf{G}\neq\mathbf{0}} \frac{\langle\mathbf{k}|V|\mathbf{k}+\mathbf{G}\rangle}{E^0(\mathbf{k}) - E^0(\mathbf{k}+\mathbf{G})}|\mathbf{k}+\mathbf{G}\rangle \\ &= \sum_{\mathbf{G}\neq\mathbf{0}} \frac{v_{\mathbf{G}}}{E^0(\mathbf{k}) - E^0(\mathbf{k}+\mathbf{G})}|\mathbf{k}+\mathbf{G}\rangle \end{aligned} \tag{4.31}$$

$$\begin{aligned} E(\mathbf{k}) &= E^0(\mathbf{k}) + \cancel{\langle\mathbf{k}|V|\mathbf{k}\rangle} + \sum_{\mathbf{G}\neq\mathbf{0}} \frac{|\langle\mathbf{k}|V|\mathbf{k}+\mathbf{G}\rangle|^2}{E^0(\mathbf{k}) - E^0(\mathbf{k}+\mathbf{G})} \\ &= E^0(\mathbf{k}) + \sum_{\mathbf{G}\neq\mathbf{0}} \frac{|v_{\mathbf{G}}|^2}{E^0(\mathbf{k}) - E^0(\mathbf{k}+\mathbf{G})} \end{aligned} \tag{4.32}$$

PROBLEM : We have dropped the first order energy shift in eqn. 4.32. Why are we justified in doing this? When will this be incorrect?

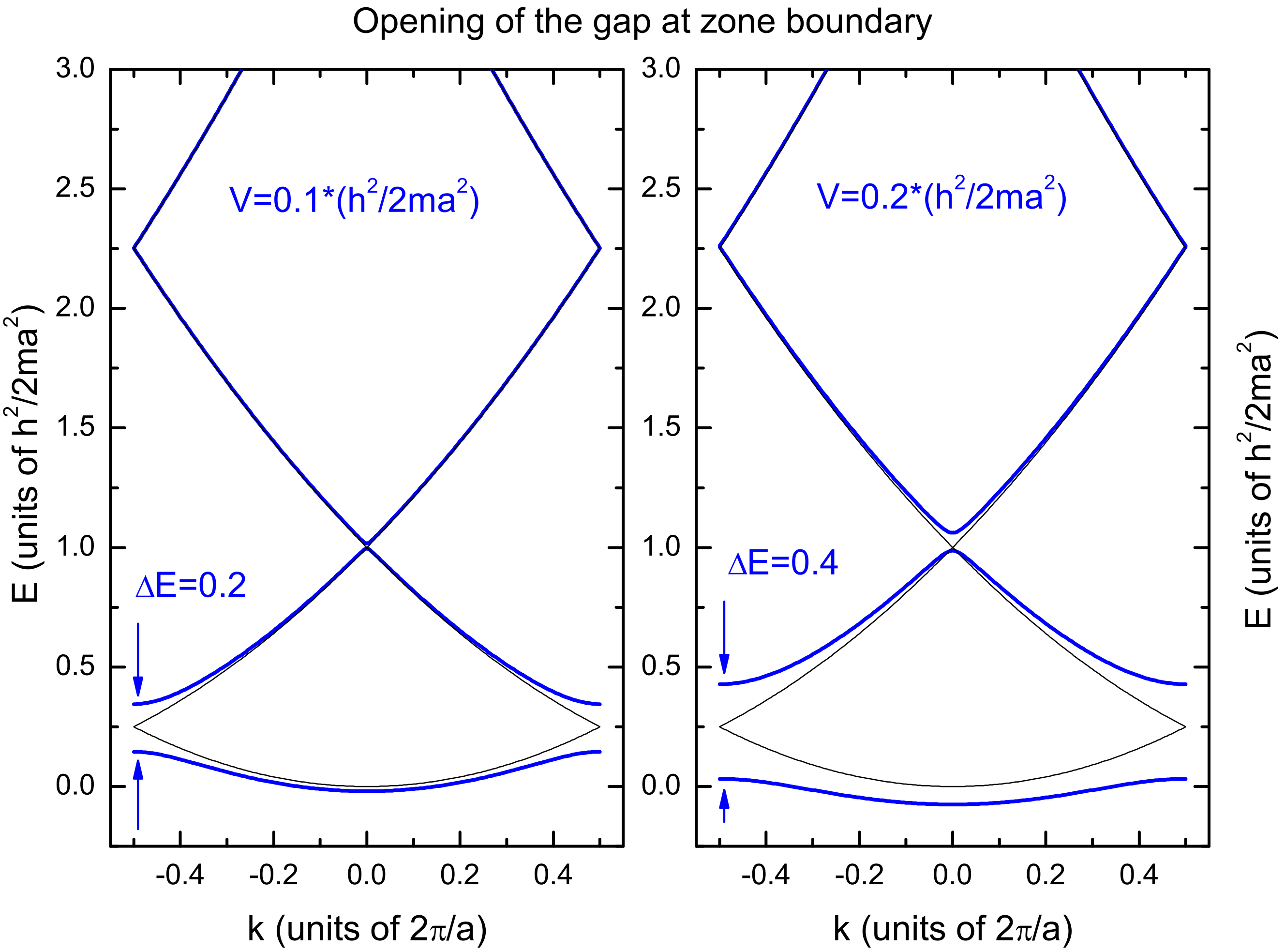


Figure 4.2: $E(k)$ in the reduced zone scheme, with a single non-zero component of the potential as in the equation 4.28. The energies have been scaled with $E = h^2/2ma^2$, which is 4 times the kinetic energy at the zone boundary, to make the plot. Similarly k has been scaled by $2\pi/a$. The black lines show the free electron result with no potential turned on. Note that the potential also affects higher bands but much less..

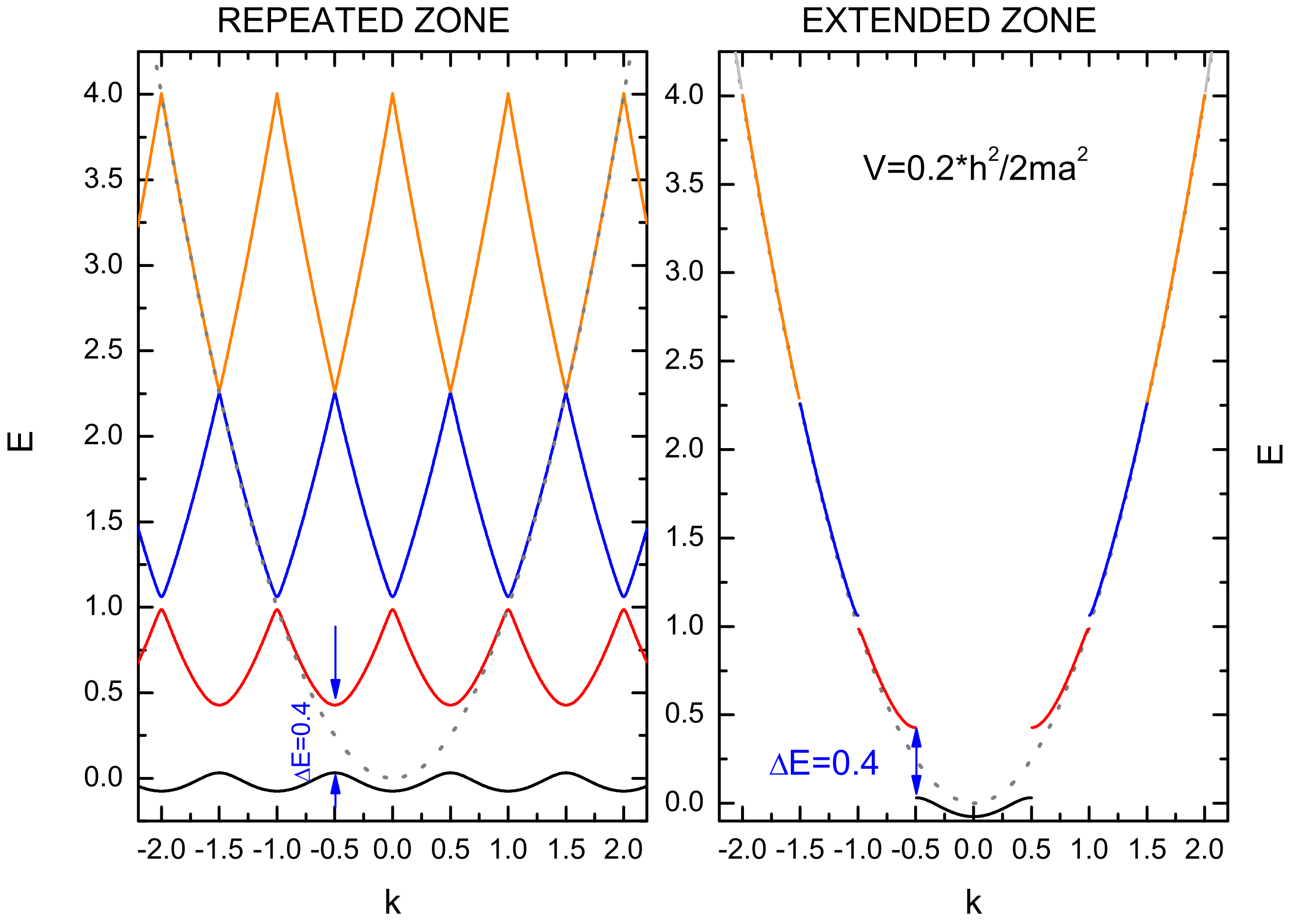


Figure 4.3: $E(k)$ in the repeated and extended zone schemes. The data used is exactly the same as in Fig. 4.2.

1. We do not need to make the sum in eqn. 4.31 run over all $\mathbf{k}' \neq \mathbf{k}$ but only some of them because $\langle \mathbf{k}|V|\mathbf{k}'\rangle = 0$ unless $\mathbf{k}' = \mathbf{k} + \mathbf{G}$. The consequence of all this is that the first order result gives us something of the Bloch form.

2. But if $E^0(\mathbf{k}) - E^0(\mathbf{k}+\mathbf{G}) = 0$ then the denominator will vanish and this will not work. This happens when two states are degenerate. This is precisely where two parabolas in Fig. 4.1 intersect. These are the zone boundaries.

3. At the zone boundaries, the energy shift is no longer of second order, this is then a degenerate first order problem. The solution (see any quantum mechanics text) for the energy shifts are to be obtained by taking the eigenvalues of the matrix V_{ij} between the degenerate states. This is precisely what we did when we solved for (in a previous problem):

$$\begin{vmatrix} E^0(\mathbf{k}) - E & \langle \mathbf{k}|V|\mathbf{k}+\mathbf{G}\rangle \\ \langle \mathbf{k}+\mathbf{G}|V|\mathbf{k}\rangle & E^0(\mathbf{k}+\mathbf{G}) - E \end{vmatrix} = 0 \tag{4.33}$$

4.1.5 The Kronig-Penny model

A simple model of a periodic potential is shown in the figure. The well-barrier-well sequence is more realistic than a single fourier component. It is a very useful "toy model" for understanding how the energy bands and band gaps in a solid arises.

PROBLEM : Revise the single electron in a finite quantum well. Consider a potential given by

$$\begin{aligned} V &= 0 \quad \text{for} \quad -w < x < 0 \\ &= V_0 \qquad \text{otherwise} \end{aligned}$$

Write the wavefunctions for bound states, piecewise as follows

$$\begin{aligned} \Psi_1 &= Ae^{i\alpha x} + Be^{-i\alpha x} \quad \text{for} \quad -w < x < 0 \\ \Psi_2 &= Ce^{-\beta x} \qquad \text{for} \quad x > 0 \\ \Psi_3 &= De^{\beta x} \qquad \text{for} \quad x < -w \end{aligned} \tag{4.34}$$

where α and β are given by:

$$\begin{aligned} \alpha^2 &= \frac{2mE}{\hbar^2} \\ \beta^2 &= \frac{2m(V_0 - E)}{\hbar^2} \end{aligned} \tag{4.35}$$

The wavefn and its derivative must be continuous at the two boundaries of the potential well. Show that the solutions are obtained by solving the set of linear equations:

$$\begin{pmatrix} 1 & 1 & -1 & 0 \\ i\alpha & -i\alpha & \beta & 0 \\ e^{-i\alpha w} & e^{i\alpha w} & 0 & -e^{-\beta w} \\ i\alpha e^{-i\alpha w} & -i\alpha e^{i\alpha w} & 0 & -\beta e^{-\beta w} \end{pmatrix} \begin{pmatrix} A \\ B \\ C \\ D \end{pmatrix} = 0 \tag{4.36}$$

There are 5 unknowns, but one of the coefficients can be chosen arbitrarily and normalisation will take care of it. Setting the determinant to zero you can show that the wavevector α can be obtained from one of the two conditions

$$\begin{aligned} \tan\frac{\alpha w}{2} &= \frac{\beta}{\alpha} \\ \tan\frac{\alpha w}{2} &= -\frac{\alpha}{\beta} \end{aligned} \tag{4.37}$$

Once k is determined calculate E and then the coefficients A, B, C, D. By arbitrarily setting $C = 1$ The *unnormalised* wavefunction can be written as

$$\begin{aligned}
\Psi_1 &= \cos\alpha x - \frac{\beta}{\alpha}\sin\alpha x && \text{for } [-w, 0] \\
\Psi_2 &= e^{-\beta x} && \text{for } [0, \infty] \\
\Psi_3 &= \left(\cos\alpha w + \frac{\beta}{\alpha}\sin\alpha w\right) e^{\beta(w+x)} && \text{for } [-\infty, -w]
\end{aligned} \tag{4.38}$$

We now put many such potential wells in an array separated by b units. So we have a repeating structure with period $a = b + w$. In the barrier regions the wavevector cannot be real, because $E < V_0$. However the exponential can have both decaying and growing parts, because the barrier region does not tend to infinity. There is no chance for one exponential to blow up. We can now write the wavefunction in well (Ψ_w) and barrier (Ψ_b) as :

$$\begin{aligned}
\Psi_w &= Ae^{i\alpha x} + Be^{-i\alpha x} \\
\Psi_b &= Ce^{\beta x} + De^{-\beta x} && (4.39) \\
& && (4.40)
\end{aligned}$$

where α and β are defined by 4.35 as before. Because of the periodicity of the structure, the wavefunction must satisfy the Bloch condition. Hence we can write , with k as crystal momentum:

$$\begin{aligned}
\Psi_w(0) &= \Psi_b(0) \\
\left.\frac{d\Psi_w}{dx}\right|_{x=0} &= \left.\frac{d\Psi_b}{dx}\right|_{x=0} \\
\Psi_b(b) &= e^{ika}\Psi_w(-w) \\
\left.\frac{d\Psi_b}{dx}\right|_{x=b} &= e^{ika}\left.\frac{d\Psi_w}{dx}\right|_{x=-w}
\end{aligned} \tag{4.41}$$

Notice how the Bloch condition helped us to write the last two equations of the set. We now have four equations connecting A, B, C, D and k.
The four equations are :

$$\begin{aligned}
A + B &= C + D \\
Ai\alpha - Bi\alpha &= C\beta - D\beta \\
Ce^{\beta b} + De^{-\beta b} &= e^{ika}\left[Ae^{-i\alpha w} + Be^{i\alpha w}\right] \\
C\beta e^{\beta b} - D\beta e^{-\beta b} &= e^{ika}\left[Ai\alpha e^{-i\alpha w} - Bi\alpha e^{i\alpha w}\right]
\end{aligned} \tag{4.42}$$

This leads to the following consistency condition:

$$\left(\frac{\beta^2 - \alpha^2}{2\alpha\beta}\right)\sinh\beta b\sin\alpha w + \cosh\beta b\cos\alpha w = \cos ka \tag{4.43}$$

What are the allowed values of k

Let's assume the structure is a loop with length L such that $L = Na$ Because we impose periodic boundary conditions we must have

$$\Psi(x + L) = \Psi(x + Na) = \Psi(x) \tag{4.44}$$

But Bloch condition requires

$$\Psi(x + L) = e^{ikL}\Psi(x) \tag{4.45}$$

These two together imply

$$\begin{aligned} e^{ikL} &= 1 \\ \therefore kL &= 2n\pi \\ \therefore k &= \frac{2n\pi}{L} \\ \therefore k &= \frac{n}{N}\frac{2\pi}{a} \end{aligned} \tag{4.46}$$

In any interval $2\pi/a$ there are thus N states, where N is the number of unit cells in the structure. Later on we will prove it for 2D and 3D also for any lattice.
The equation 4.43 can be solved numerically, following the process:

- Notice that the solution $E(k)$ is periodic in k. So we restrict k within one period.
- Choose a k, in the range $-\pi/a < k < \pi/a$. This fixes the RHS.
- The only unknown in the LHS is the energy E, because both α and β are determined by E.
- Allow α to vary from 0 to $\sqrt{2mV_0/\hbar^2}$. Pick up all the solutions. There are a finite number of them only.
- Calculate the corresponding E values and plot them vs k (not α).
- This is the band structure.

Here's an example
And a couple more, where we have increased the separation, keeping all other parameters same.
And then:

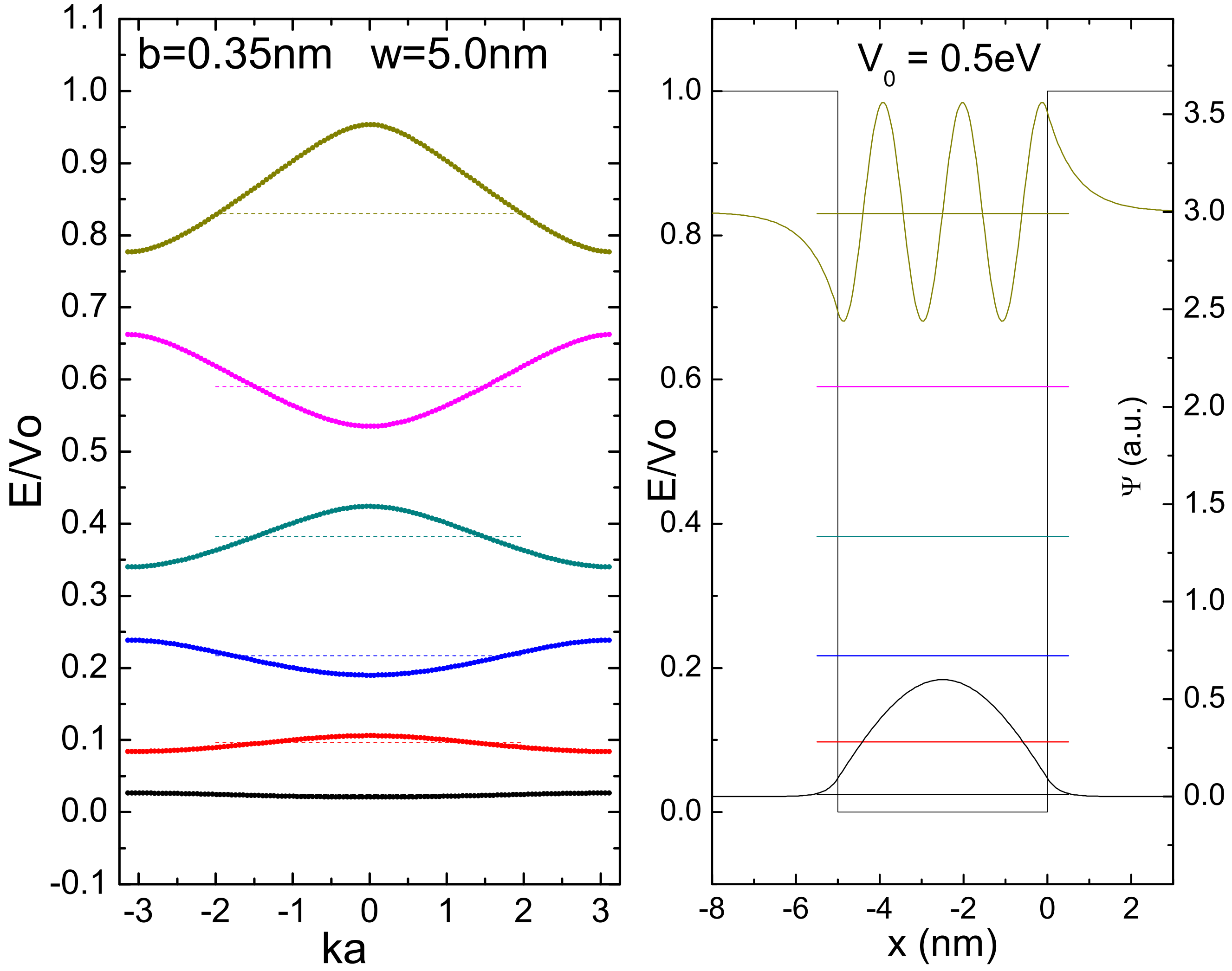


Figure 4.4: Notice how a single energy level has spread out into a band. The deeper levels are not very spread out, the higher (loosely bound) levels are spread more. This indeed is the crux of band structure. The more the possible overlap between wavefunctions at neighbouring sites, more will be the spread of the band. The right hand figure shows the single potential solutions for eigenvalues and eignefunctions. The left hand figure is a plot of the eigenvalues (allowed energies) when a series of them are laid out on a chain.

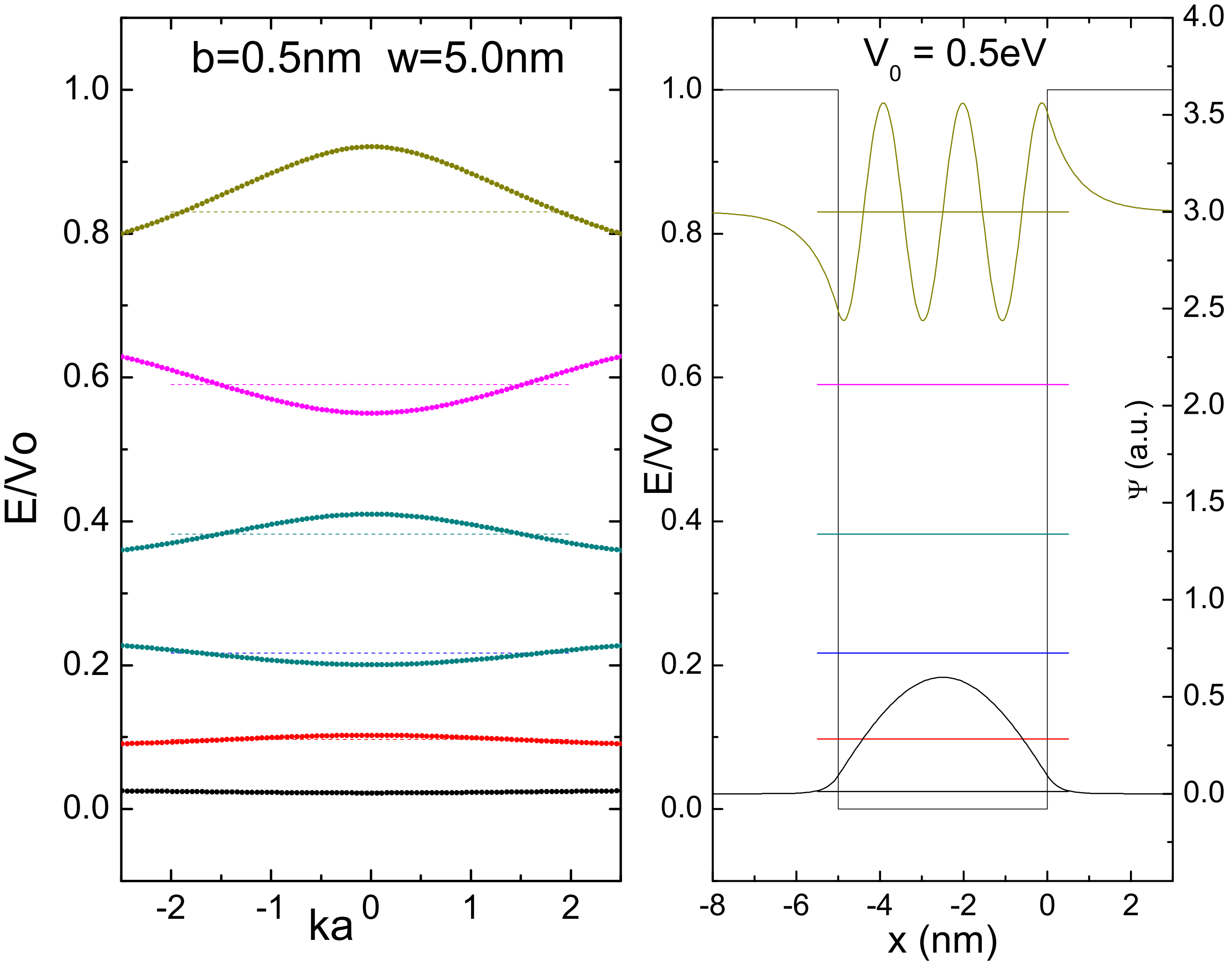


Figure 4.5: b=0.5nm, w= 5nm, V_0=0.5eV. Notice that the bands have become narrower

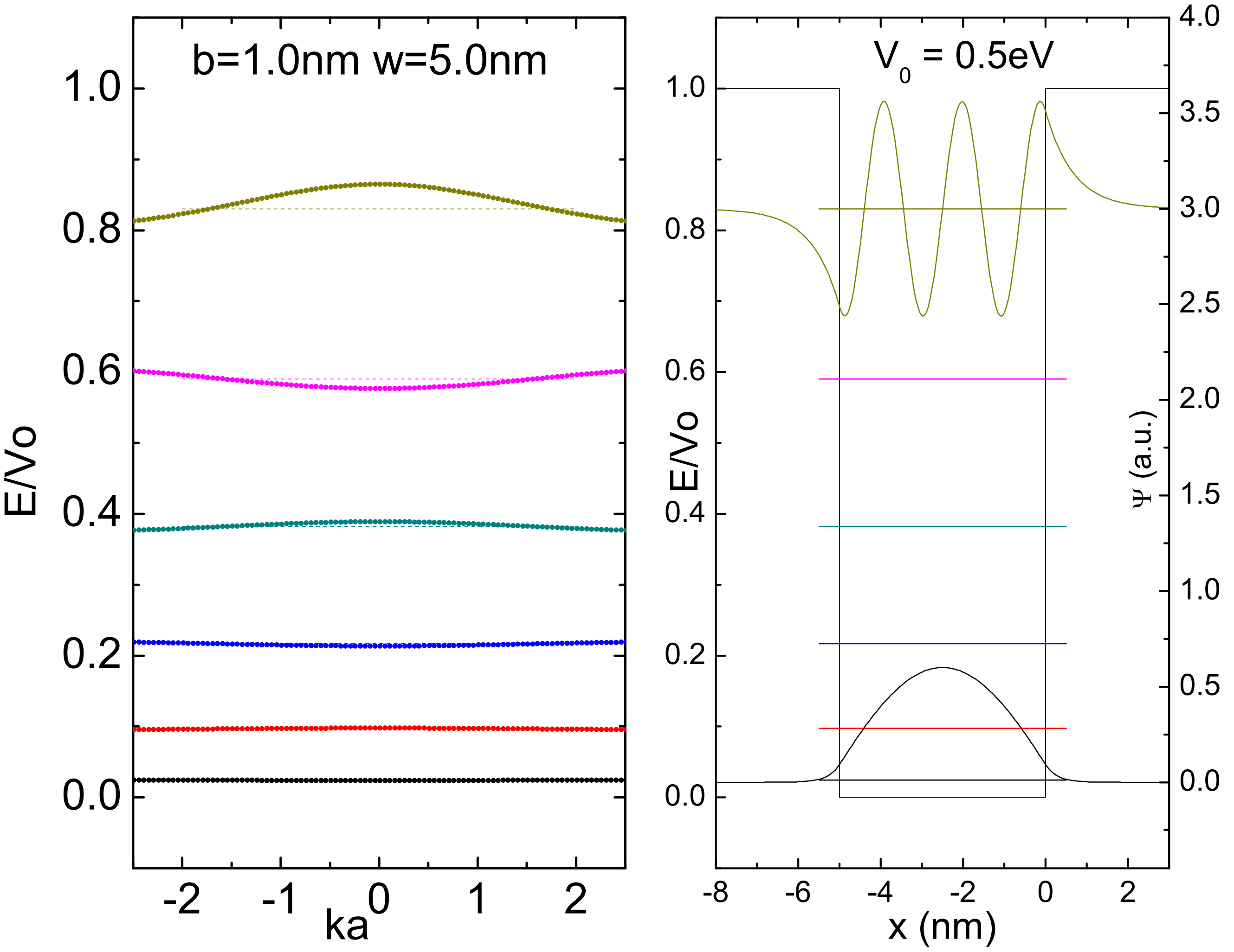


Figure 4.6: b=1nm, w= 5nm, V_0=0.5eV. The bands are getting even more narrow.

4.1.6 The Band gap: classification of conducting, non-conducting and semiconducting substances

We have spoken about the importance of the density of states at the Fermi level before, while calculating quantities like specific heat, susceptibility, conductivity etc. Now since the energy spectrum has gaps in it, it is possible to have the Fermi level in qualitatively different positions.

Fermi level in a continuous band : Metals

If the electrons fill a band partially, then the $D(E_F)$ will in general be large. It will be easy to accelerate an electron with a small electric field. Responses which depend on $D(E_F)$ - like thermal and electrical conductivities will also be large.

Fermi level in a large bandgap: Insulators

If the band gap is large (few electron volts) and the available electrons fill up the lower band fully, then the Fermi level is forced to lie in a gap. $D(E_F) = 0$ and the material cannot respond to small electric fields. This is an insulator.

Fermi level in a small bandgap: Semiconductors

By small we mean typically not more than ~ 2 eV. The distinction between semiconductors and insulators is a qualitative one. In semiconductors like Silicon, Germanium, Gallium Arsenide there are some carriers in the upper band due to thermal excitations at room temperature. Large bandgap materials can behave like semiconductors at high temperatures.

4.2 Motion of an electron in a band: Bloch oscillation, group velocity and effective mass

How do the concepts of position, velocity etc carry over to the Bloch electrons? Recall that a Bloch state $|\Psi\rangle = e^{ikx}|u_k\rangle$ is not in general an eigenfunction of the momentum operator : $p = -i\hbar\frac{\partial}{\partial x}$.

The group velocity of Bloch electrons:

$$\begin{aligned}
p|\Psi\rangle &= \hbar k|\Psi\rangle + e^{ikx}(-i\hbar)\frac{\mathrm{d}}{\mathrm{d}x}|u_k\rangle \\
\therefore (p-\hbar k)\,|\Psi\rangle &= e^{ikx}p|u_k\rangle \\
\therefore p|\Psi\rangle &= e^{ikx}\,(p+\hbar k)\,|u_k\rangle
\end{aligned} \tag{4.47}$$

It is left as a little exercise to show that this implies the following result

$$\begin{aligned}
\left(\frac{p^2}{2m}+V\right)|\Psi\rangle &= E(k)|\Psi\rangle \\
H(k)|u_k\rangle = \left(\frac{(p+\hbar k)^2}{2m}+V\right)|u_k\rangle &= E(k)|u_k\rangle \\
\frac{\mathrm{d}}{\mathrm{d}k}\left(\frac{(p+\hbar k)^2}{2m}+V\right)|u_k\rangle &= \frac{\mathrm{d}}{\mathrm{d}k}E(k)|u_k\rangle \\
\frac{\hbar}{m}\,(p+\hbar k)\,|u_k\rangle + H(k)\frac{\mathrm{d}}{\mathrm{d}k}|u_k\rangle &= \frac{\mathrm{d}E}{\mathrm{d}k}|u_k\rangle + E\frac{\mathrm{d}}{\mathrm{d}k}|u_k\rangle \\
\frac{\hbar}{m}\langle u_k|p+\hbar k|u_k\rangle + \cancel{\langle u_k|H(k)\frac{\mathrm{d}}{\mathrm{d}k}|u_k\rangle} &= \frac{\mathrm{d}E}{\mathrm{d}k}\langle u_k|u_k\rangle + \cancel{E\langle u_k|\frac{\mathrm{d}}{\mathrm{d}k}|u_k\rangle} \qquad (4.48) \\
\therefore \langle\Psi|\frac{p}{m}|\Psi\rangle &= \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k} \qquad (4.49)
\end{aligned}$$

Bloch oscillation

We now know the following

- The $E(k)$ relation is periodic in k. Let's consider a case where $E = E_0 - 2\gamma \cos ka$.
- An electric field (F) changes the k vector such that $dk/dt = -eF/\hbar$. (eqn 4.24).
- Now suppose there is no scattering, what is the equation of motion of the electron?

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k} \\ &= \frac{2\gamma a}{\hbar}\sin k(t)a \\ &= \frac{2\gamma a}{\hbar}\sin\left(k(0)a - \frac{eaFt}{\hbar}\right) \\ \therefore x(t) &= \frac{2\gamma a}{\hbar}\int_0^t \mathrm{d}t\ \sin\left(k(0)a - \frac{eaFt}{\hbar}\right) \end{aligned} \tag{4.50}$$

We therefore see that the electron will *oscillate* with a frequency $\omega_{Bloch} = eaF/\hbar$, a very striking prediction of quantum mechanics that is believed to be correct but has only been partially verified. Putting $a = 1$nm, $F = 10$ V/cm we find that the frequency would be $\sim 10^9$Hz.

PROBLEM : How is the amplitude of Bloch oscillation related to the other quantities in the problem? For a bandwidth of 1 ev, how does this length compare to typical mean free paths in a very clean metal at low temperature - say 500 microns?

Effective mass

Mass is the constant of proportionality between the applied force, **F**, and the acceleration produced **a**. If we carry over this idea to the motion of electrons in a band, we get

$$\begin{aligned}\mathbf{a} &= \frac{d}{dt}\mathbf{v} \\ &= \frac{d}{dt}\frac{1}{\hbar}\nabla_{\mathbf{k}}E(\mathbf{k})\end{aligned}$$

Writing it out componentwise and remembering that v_x can be a function of *all* components k_x, k_y and k_z we get:

$$\begin{aligned}a_i &= \frac{\mathrm{d}}{\mathrm{d}t}v_i(k_j) \\ &= \frac{\partial v_i}{\partial k_j}\frac{\mathrm{d}k_j}{\mathrm{d}t} \\ &= \frac{1}{\hbar}\frac{\partial^2 E}{\partial k_i \partial k_j}\frac{F_j}{\hbar}\end{aligned} \tag{4.51}$$

The sum over the repeated index j is implied. The structure of the eqn. 4.51 tells us that we can write an inverse effective mass matrix

$$\mathrm{M}_{\mathrm{ij}}^{-1} = \frac{1}{\hbar^2}\frac{\partial^2 E}{\partial k_i \partial k_j} \tag{4.52}$$

Calculating the effective mass in a band: $k.p$ method

Using what we did for calculating the group velocity of Bloch electrons, you can show the following

---

PROBLEM : A Bloch state is a solution of the Schrodinger equation in a periodic potential, $V(\mathbf{r})$, of the crystal. It can be written as

$$\Psi_{n\mathbf{k}}(\mathbf{r}) \quad = \quad e^{i\mathbf{k}.\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r})$$

where $u_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r}+\mathbf{R})$ for any direct lattice vector $\mathbf{R}$. Given that $\Psi_{n\mathbf{k}}(\mathbf{r})$ is a solution of

$$\left[\frac{1}{2m}\mathbf{p}^2 + V(\mathbf{r})\right]\Psi_{n,\mathbf{k}}(\mathbf{r}) \quad = \quad E_{n,\mathbf{k}}\Psi_{n,\mathbf{k}}(\mathbf{r})$$

Show that the function $u_{n,\mathbf{k}}(\mathbf{r})$ satisfies

$$\left[\frac{1}{2m}\left(\mathbf{p}+\hbar\mathbf{k}\right)^2 + V(\mathbf{r})\right]u_{n,\mathbf{k}}(\mathbf{r}) \quad = \quad E_{n,\mathbf{k}}u_{n,\mathbf{k}}(\mathbf{r})$$

The equations are so far exact. It follows that we can use this result to formulate band structure as perturbation problem around $\mathbf{k} = 0$ in the following way. Expanding the result of the problem

$$\left[\frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) + \frac{\hbar}{m}\mathbf{k}.\mathbf{p} + \frac{\hbar^2k^2}{2m}\right]|u_{n,\mathbf{k}}(\mathbf{r})\rangle \quad = \quad E_{n,\mathbf{k}}|u_{n,\mathbf{k}}(\mathbf{r})\rangle$$

We treat the $\mathbf{k}$-independent part as H_0 and the $\mathbf{k}$-dependent part as the perturbation such that

$$H_0 \quad = \quad \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \tag{4.53}$$

$$H_{\mathbf{k}} \quad = \quad \frac{\hbar}{m}\mathbf{k}.\mathbf{p} + \frac{\hbar^2k^2}{2m} \tag{4.54}$$

The eigenfunctions of H_0 are $|u_{n,\mathbf{0}}(\mathbf{r})\rangle$, which must form a complete set, using which we can express the $|u_{n,\mathbf{k}}(\mathbf{r})\rangle$, where n is the band index. This is a very useful fact which at once allows us to write till second order (notice the indices and their meaning carefully, for convenience we now write $E_{n,\mathbf{k}}$ as $E_n(\mathbf{k})$)

$$E_n(\mathbf{k}) = E_n(\mathbf{0}) + \langle u_{n,\mathbf{0}}|H_{\mathbf{k}}|u_{n,\mathbf{0}}\rangle + \sum_{n\neq m}\frac{|\langle u_{n,\mathbf{0}}|H_{\mathbf{k}}|u_{m,\mathbf{0}}\rangle|^2}{E_n(0) - E_m(0)} \tag{4.55}$$

Now

$$\begin{aligned}\langle u_{n,\mathbf{0}}|H_{\mathbf{k}}|u_{n,\mathbf{0}}\rangle \quad &= \quad \hbar\mathbf{k}.\langle u_{n,\mathbf{0}}|\frac{\mathbf{p}}{m}|u_{n,\mathbf{0}}\rangle + \frac{\hbar^2k^2}{2m} \\ &= \quad \hbar\mathbf{k}.\ \nabla_{\mathbf{k}}E_n(\mathbf{k})|_{\mathbf{k}=0} + \frac{\hbar^2k^2}{2m} \\ &= \quad \frac{\hbar^2k^2}{2m}\end{aligned} \tag{4.56}$$

Provided we have an extremum at $\mathbf{k} = 0$, which is usually the case. Now for the second order part , the matrix elements are taken between states such that $n\neq m$, so a number like $\frac{\hbar^2k^2}{2m}$ cannot contribute anything. So we get:

$$\sum_{n\neq m}\frac{|\langle u_{n,\mathbf{0}}|H_{\mathbf{k}}|u_{m,\mathbf{0}}\rangle|^2}{E_n(0) - E_m(0)} \quad = \quad \frac{\hbar^2}{m^2}\sum_{n\neq m}\frac{\langle u_{n,\mathbf{0}}|\mathbf{k}.\mathbf{p}|u_{m,\mathbf{0}}\rangle\langle u_{m,\mathbf{0}}|\mathbf{k}.\mathbf{p}|u_{n,\mathbf{0}}\rangle}{E_n(0) - E_m(0)} \tag{4.57}$$

So we see how the presence of other bands changes the free electron dispersion explicitly by writing:

$$E_n(\mathbf{k}) \quad = \quad E_n(\mathbf{0}) + \frac{\hbar^2 k^2}{2m} + \frac{\hbar^2}{m^2} \sum_{n \neq m} \frac{\langle u_{n,\mathbf{0}} | \mathbf{k}.\mathbf{p} | u_{m,\mathbf{0}} \rangle \langle u_{m,\mathbf{0}} | \mathbf{k}.\mathbf{p} | u_{n,\mathbf{0}} \rangle}{E_n(0) - E_m(0)} \tag{4.58}$$

We can now see how differentiating eqn. 4.58 twice (w.r.t k_i and k_j) that the effective mass matrix is

$$\mathrm{M}_{ij}^{-1} = \frac{1}{\hbar^2} \frac{\partial^2 E_n}{\partial k_i \partial k_j} \quad = \quad \frac{1}{m} \delta_{ij} + \frac{2}{m^2} \sum_{n \neq m} \frac{\langle u_{n,\mathbf{0}} | p_i | u_{m,\mathbf{0}} \rangle \langle u_{m,\mathbf{0}} | p_j | u_{n,\mathbf{0}} \rangle}{E_n(0) - E_m(0)} \tag{4.59}$$

If the bandgaps are large then the effective mass differs very little from the free electron mass. At least qualitatively, you can show why narrow bandgap semiconductors have small effective masses, using eqn. 4.59

Chapter 5

Tight binding or Linear Combination of Atomic Orbitals (LCAO)

In the last chapter we considered the problem of one electron in the periodic potential of the lattice and solved for its energy eigenvalues. It is possible (and useful) to look at the problem from another point of view. We consider that we are building up a solid atom by atom, like building up a molecule.

5.1 Diatomic molecule and Linear chain of atoms

5.1.1 Diatomic molecule

As an initial problem let's consider building up a molecule form two atoms that are not necessarily identical. When they are far apart then the wavefunctions must be same as the wavefunctions of the isolated atoms - we call the atoms a and b. So the Hamiltonian of the system must be

$$H_{ab} = T + V_a + V_b \tag{5.1}$$

Our basis set is going to be the states $|a\rangle$ $|b\rangle$, centered on atom a and atom b respectively, when they are very far apart. So that the basis set satisfies

$$\begin{aligned} (T + V_a)|a\rangle &= E_0^a|a\rangle & (5.2) \\ (T + V_b)|b\rangle &= E_0^b|b\rangle & (5.3) \end{aligned}$$

The Linear Combination of Atomic Orbitals method means looking for solutions of eqn 5.1 of the form

$$|\psi\rangle = \alpha|a\rangle + \beta|b\rangle \tag{5.4}$$

If eqn. 5.4 is a solution then we must have (E is the unknown eigenvalue we want to solve for)

$$\begin{aligned} \langle a|H_{ab}|\psi\rangle &= E\langle a|\psi\rangle \\ \langle b|H_{ab}|\psi\rangle &= E\langle b|\psi\rangle & (5.5) \end{aligned}$$

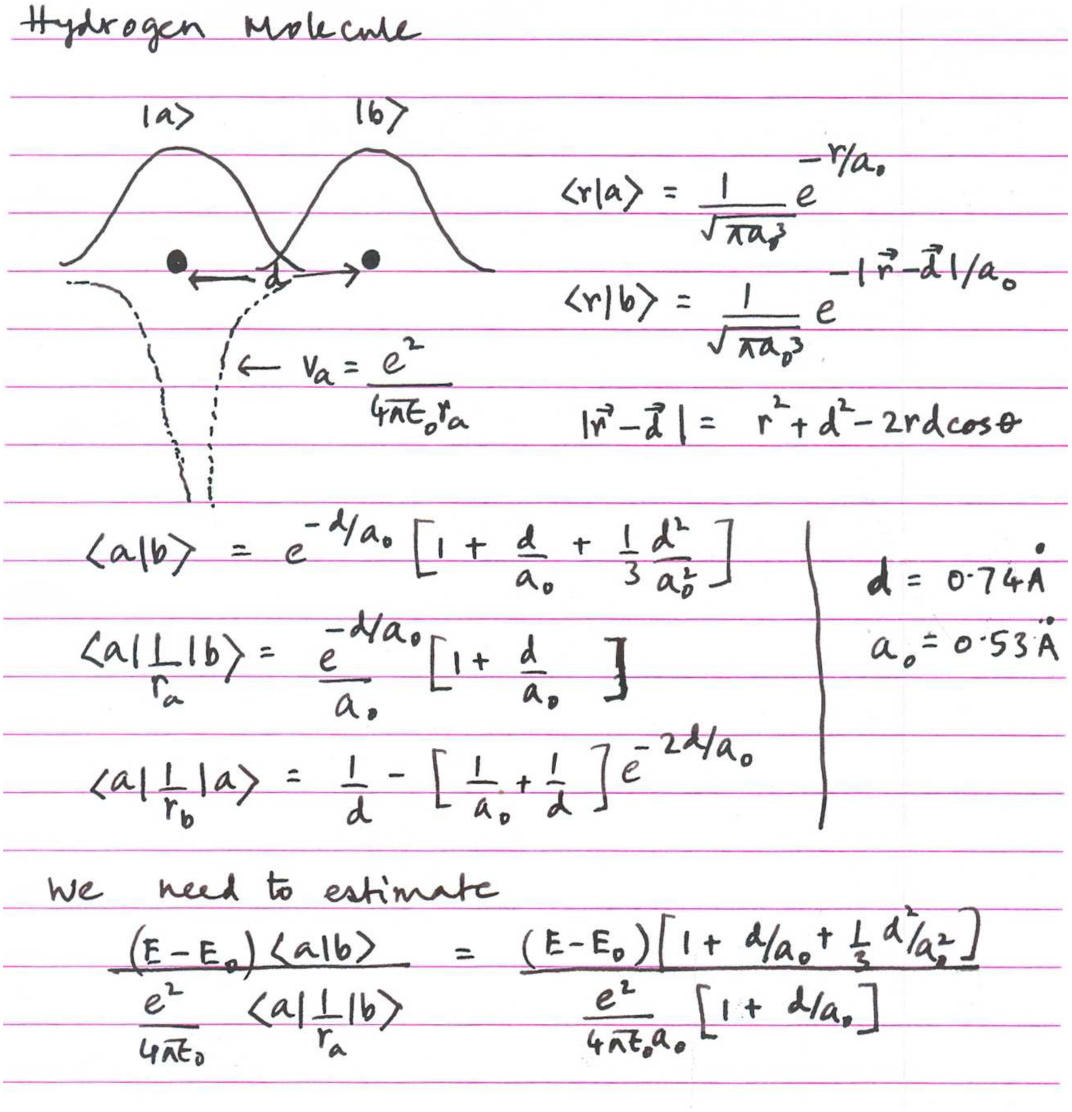
As it stands the set of eqn. 5.5 is exact, but to proceed we need to understand the physical significance of each term and approximate them reasonably.

$$\langle a|b\rangle \approx 0 \tag{5.6}$$

This means that there is negligible overlap between the atomic orbitals.

$$\langle a|T + V_a + V_b|b\rangle \equiv t \tag{5.7}$$

Under the action of the Hamiltonian the state $|a\rangle$ and $|b\rangle$ can mix a little bit. We will come across this type of a term many times in future. A term of this type is called a hopping term. It is important to



Hydrogen Molecule

$\leftarrow V_a = \frac{e^2}{4\pi\epsilon_0 r_a}$

$$\langle r|a\rangle = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}$$

$$\langle r|b\rangle = \frac{1}{\sqrt{\pi a_0^3}} e^{-|\vec{r}-\vec{d}|/a_0}$$

$$|\vec{r}-\vec{d}| = r^2 + d^2 - 2rd\cos\theta$$

$$\langle a|b\rangle = e^{-d/a_0}\left[1 + \frac{d}{a_0} + \frac{1}{3}\frac{d^2}{a_0^2}\right]$$

$$\langle a|\frac{1}{r_a}|b\rangle = \frac{e^{-d/a_0}}{a_0}\left[1 + \frac{d}{a_0}\right]$$

$$\langle a|\frac{1}{r_b}|a\rangle = \frac{1}{d} - \left[\frac{1}{a_0} + \frac{1}{d}\right]e^{-2d/a_0}$$

$d = 0{\cdot}74\,\mathring{A}$

$a_0 \doteq 0{\cdot}53\,\mathring{A}$

We need to estimate

$$\frac{(E-E_0)\langle a|b\rangle}{\frac{e^2}{4\pi\epsilon_0}\langle a|\frac{1}{r_a}|b\rangle} = \frac{(E-E_0)\left[1 + d/a_0 + \frac{1}{3}d^2/a_0^2\right]}{\frac{e^2}{4\pi\epsilon_0 a_0}\left[1 + d/a_0\right]}$$

Figure 5.1: The overlap of the two 1s orbitals in Hydrogen molecule. Some of the intermediate steps are left for you to fill in - you should be able to estimate the term we retained and the term we dropped. Though the wavefunctions are specific to the H_2 molecule, the general conclusion would be true for any two tightly bound states, separated by a not too large an amount. Of course if we keep increasing the separation, then the hopping term would also go to zero.

understand why we claim that that the hopping term (eqn. 5.7) can be larger than the direct overlap term 5.6. See the fig. 5.1 and study it carefully. You should be able to reason out why we could ignore the expression in eq. 5.6 but retain the hopping term.

$$\langle a|H_{ab}|a\rangle = \langle a|T+V_a|a\rangle + \langle a|V_b|a\rangle = E_0^a + \langle a|V_b|a\rangle \equiv \tilde{E}_0^a \tag{5.8}$$

You should now be able to appreciate the physical significance of the approximation (known as the tight binding approximation) show that eqn. 5.5 leads (assuming both the atoms are identical, it does not mean $V_a = V_b$, because they are still centered at different points, though the functional forms will be similar.

$$\begin{pmatrix} \tilde{E}_0 - E & t \\ t^* & \tilde{E}_0 - E \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0 \tag{5.9}$$

The solution is obtained by setting the determinant to zero.

$$E = \tilde{E}_0 \pm |t| \tag{5.10}$$

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|a\rangle \mp |b\rangle\right) \tag{5.11}$$

PROBLEM : Complete the algebra leading to eqn. 5.10 and eqn. 5.11. The lower energy state is called the bonding state (in chemistry) and the higher energy state is called the antibonding state. Which state has the higher electron density at the mid-point between the two atoms?

5.1.2 Linear chain of atoms with nearest neighbour interaction

We now extend the ideas of tight binding with one hopping term to a linear chain of atoms,each spaced by a units. We will always consider a chain that has its ends joined together. Periodic boundary conditions are then obviously easy to apply. This means that the $N+1^{th}$ atom is same as the 1^{st} atom.

The hamiltonian is then

$$H = T + V_1 + V_2 + \ldots + V_N \tag{5.12}$$

Remember that although there are N sites/atoms/potential wells, there is only one particle co-ordinate. We are solving for single particle eigenstates in the potential created by all the atoms.

Single orbital on a site

Since this is a periodic potential, we use Bloch's theorem in combination with the tight binding idea of using wavefunctions ($|\phi_n\rangle$) localized at the n^{th} atomic site as our starting point. The wavefunction is then

$$|\psi_k\rangle = \frac{1}{\sqrt{N}}\sum_n e^{ikna}|\phi_n\rangle \tag{5.13}$$

$$\therefore \langle r|\psi_k\rangle = \frac{1}{\sqrt{N}}\sum_n e^{ikna}\langle r|\phi_n\rangle \tag{5.14}$$

$$\psi_k(r) = \frac{1}{\sqrt{N}}\sum_n e^{ikna}\phi(r-na) \tag{5.15}$$

where $|\phi_n\rangle$ is the wavefunction localized on the n^{th} atom, satisfying

$$(T+V_n)\,|\phi_n\rangle = E_0|\phi_n\rangle \tag{5.16}$$

$$\begin{aligned} \langle\phi_n|H|\phi_n\rangle &= E^0 + \langle\phi_n|V_{n-1}|\phi_n\rangle + \langle\phi_n|V_{n+1}|\phi_n\rangle \\ &= \tilde{E}_0 \end{aligned} \tag{5.17}$$

PROBLEM : Show that the wavefunction 5.13

1. satisfies the Bloch criteria $\psi(r) = e^{ikr}u_k(r)$, where $u_k(r+na) = u_k(r)$
2. is correctly normalized provided a certain assumption is made. What is the assumption?

The periodic boundary condition requires that the values of k be quantized. However you can see that if N is large then the quantization gets more and more finely spaced and k becomes continuous in the large N limit.

$$e^{ik(N+1)a} = e^{ika} \tag{5.18}$$

$$\therefore kNa = 2m\pi \tag{5.19}$$

$$k = \frac{2\pi}{a}\frac{m}{N} \tag{5.20}$$

The problem is now surprisingly straightforward, because there are no unknown parameters in eqn. 5.13, all we need to do is take the expectation value

$$E(k) = \langle\psi_k|H|\psi_k\rangle$$

$$= \langle\psi_k|T + V_1 + V_2 + V_3 + \ldots + V_N|\psi_k\rangle \tag{5.21}$$

$$= \frac{1}{N}\sum_{n,m} e^{-ikna}e^{ikma}\langle\phi_n|H|\phi_m\rangle \tag{5.22}$$

$$= \frac{1}{N}\sum_{n=m}\langle\phi_n|H|\phi_n\rangle + \frac{1}{N}\sum_{n=m\pm1} e^{ik(m-n)a}\langle\phi_n|H|\phi_m\rangle + \xcancel{\frac{1}{N}\sum_{|n-m|>1}} \ldots \tag{5.23}$$

$$\approx \tilde{E}_0 + \left(te^{ika} + te^{-ika}\right) \tag{5.24}$$

$$\therefore E(k) = \tilde{E}_0 + 2t\cos ka \tag{5.25}$$

We have solved the band structure, t has the same significance (nearest neighbour hopping) that we discussed in eqn. 5.7.

Two points to note

1. The bandwidth is proportional to the hopping term.
2. It is surprisingly easy to generalize the result to 2 and 3 dimensions, because *all* we need to do is sum over the nearest neighbours! If they are symmetrically located then the bandwidth would simply be $2zt$ where z is the number of nearest neighbours or the co-ordination number of the lattice.

PROBLEM : Calculate the group velocity of a particle at the bottom of the band and at the corner ($k = \pm\pi/a$). Show that there is a point of inflection (where the second derivative changes sign) somewhere between $k = 0$ and $k = \pm\pi/a$.

How does the Bloch function look?

Here's a plot of how the functions look. The $k = 0$ wavefunction is shown for reference, because that has the maximum resemblance with the "atomic" wavefunctions. Here we assumed that the atomic wavefunction is a gaussian. See Fig. 5.2.

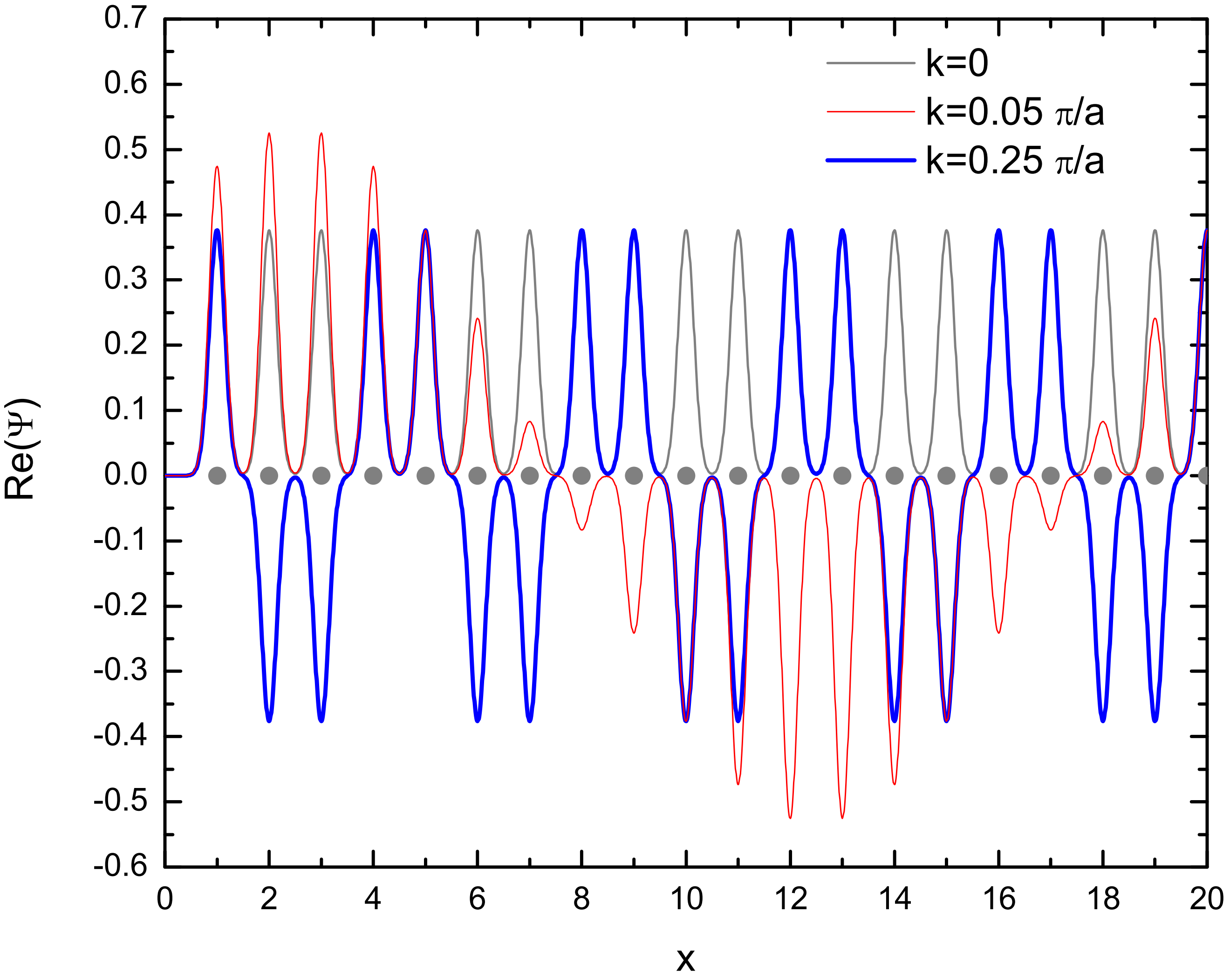


Figure 5.2: The dots are the atomic sites, the $k = 0$ wavefunction shows what the atomic states are like. The other two show what the linear combination of those wavefunctions, as given by Bloch's theorem , would look like. See eqn. 5.13.

Generalising to 2 and 3 dimensions: with 1 oribital per site

The generalisation is easy. To handle 2 and 3d lattices we need to write the wavefunction as

$$\begin{aligned} |\psi_{\mathbf{k}}\rangle &= \frac{1}{\sqrt{N}}\sum_{\mathbf{R}} e^{i\mathbf{k}.\mathbf{R}}|\phi_{\mathbf{R}}\rangle && (5.26)\\ \langle\mathbf{r}|\psi_{\mathbf{k}}\rangle &= \frac{1}{\sqrt{N}}\sum_{\mathbf{R}} e^{i\mathbf{k}.\mathbf{R}}\phi(\mathbf{r}-\mathbf{R}) && (5.27)\end{aligned}$$

Where the sum runs over all direct lattice vectors $\mathbf{R}$ and $|\phi_{\mathbf{R}}\rangle$ is the atomic state centered at $\mathbf{R}$. While taking the expectation value of energy we will group the series of terms into three and ignore the interaction between sites which are not nearest neighbours or next-nearest-neighbours:

$$\begin{aligned} H &= T + V_1 + V_2 + V_3 + \ldots + V_N && (5.28)\\ E(\mathbf{k}) &= \langle\psi_{\mathbf{k}}|H|\psi_{\mathbf{k}}\rangle && (5.29)\\ &= \frac{1}{N}\sum_{\mathbf{R}=\mathbf{R}'}\langle\phi_{\mathbf{R}'}|H|\phi_{\mathbf{R}}\rangle + \frac{1}{N}\sum_{\substack{\mathbf{R},\mathbf{R}'\\ \text{nearest}\\ \text{neigh-}\\ \text{bours}}} e^{i\mathbf{k}.(\mathbf{R}-\mathbf{R}')}\langle\phi_{\mathbf{R}'}|H|\phi_{\mathbf{R}}\rangle + \xcancel{\frac{1}{N}\sum_{\substack{\mathbf{R},\mathbf{R}'\\ \text{further}\\ \text{than}\\ \text{nearest}\\ \text{neigh-}\\ \text{bours}}}\ldots} && (5.30)\\ &\approx E_0 + \sum_{\substack{\text{nearest}\\ \text{neigh-}\\ \text{bours}}} e^{i\mathbf{k}.\mathbf{R}}t_{\mathbf{R}} && (5.31)\end{aligned}$$

Since all sites are identical, it is sufficient to sum over the nearest neighbours of the site at $\mathbf{R} = 0$.

PROBLEM : Consider a 2-d rectangular lattice with sides a and b.

1. Show that following eqn. 5.31 the bandstructure would be of the form

$$E(k_x, k_y) = E_0 - 2t_1\cos(ak_x) - 2t_2\cos(bk_y) \tag{5.32}$$

2. What is the reciprocal lattice? Draw the first Brillouin zone.
3. Plot the constant energy contours, assuming $t_1 > t_2 > 0$ and $a < b$. Why is this physically reasonable?
4. Plot some constant energy contours. How do the contours look for small k? How do the shapes change at slightly larger k? Do all constant energy contours close within the first Brillouin zone?

Similarly for 3d lattices like BCC with 8 nearest neighbours and FCC with 12 neighbours can be summed up.

PROBLEM : Tight-binding bandstructure with a single orbital per site gives on BCC and FCC

1. For Body Centered Cubic lattice write down the co-ordinates of the nearest neighbours of $(0, 0, 0)$
2. Then show, with 8 nearest neighbour hopping terms and a as the side of the cube

$$E(k_x, k_y, k_z) = E_0 + 8t\cos\frac{k_x a}{2}\cos\frac{k_y a}{2}\cos\frac{k_x a}{2} \tag{5.33}$$

3. For Face Centered Cubic lattice write down the co-ordinates of the nearest neighbours of $(0, 0, 0)$

4. Then show, with 12 nearest neighbour hopping terms and a as the side of the cube:

$$E(k_x, k_y, k_z) = E_0 + 4t\left[\cos\frac{k_x a}{2}\cos\frac{k_y a}{2} + \cos\frac{k_y a}{2}\cos\frac{k_z a}{2} + \cos\frac{k_z a}{2}\cos\frac{k_x a}{2}\right] \tag{5.34}$$

Counting the number of states in k space

Geometrically, periodic boundary condition in 1d means, putting all the lattice points on a ring. In 2d it means putting them on a "torus", in 3d it would be some hypersurface that we can only define mathematically.

Now think of a lattice with lattice vectors $\mathbf{a_1}$,$\mathbf{a_2}$, $\mathbf{a_3}$ and corresponding reciprocal lattice vectors $\mathbf{b_1}$, $\mathbf{b_2}$,$\mathbf{b_3}$. Let's take the $N = N_1 N_2 N_3$ as the number of unit cells in the crystal. Note that it doesn't imply that we are taking a cubic/rectangular volume, only. In general (algebraically, in any dimension) periodic (or Born von Karman) boundary conditions means that we require

$$\begin{aligned}
\psi(\mathbf{r} + N_i \mathbf{a_i}) &= \psi(\mathbf{r}) && (5.35)\\
\therefore e^{iN_i \mathbf{k}.\mathbf{a_i}} &= 1 \qquad \text{(Bloch}'\text{s theorem)} && (5.36)\\
\therefore \mathbf{k} &= \sum_i \frac{m_i}{N_i}\mathbf{b_i} \qquad \text{for integer } m && (5.37)\\
& && (5.38)
\end{aligned}$$

The volume of allowed $\mathbf{k}$-space per point is then:

$$\begin{aligned}
\Delta\mathbf{k} &= \frac{\mathbf{b_1}}{N_1}.\frac{\mathbf{b_2}}{N_2} \times \frac{\mathbf{b_3}}{N_1} && (5.39)\\
&= \frac{1}{N}\frac{(2\pi)^3}{v_{unit\ cell}} && (5.40)\\
&= \frac{(2\pi)^3}{V_{crystal}} && (5.41)
\end{aligned}$$

This means:

1. Whenever we need to sum over all states we can thus interchange discrete summation and continuous integration by the following rule

$$\sum_{\mathbf{k}} (\ldots) \rightarrow \int \frac{V}{(2\pi)^3} \mathrm{d}^3\mathbf{k}\,(\ldots) \tag{5.42}$$

2. Since the $\mathbf{k}$-space density is uniform, we can start from here and use the $E(\mathbf{k})$ relation to convert this into density of states in energy $D(E)$.

3. We need to multiply this by 2 for spin 1/2 particles like electrons. In general by $(2s + 1)$ if the particle has spin s, because each k-state can accommodate one particle with spin $-1/2$, one with spin 1/2 etc.

The density of states in energy

We want to write an expression for the number of states between between E to $E + \delta E$. We will call this function $D(E)$.

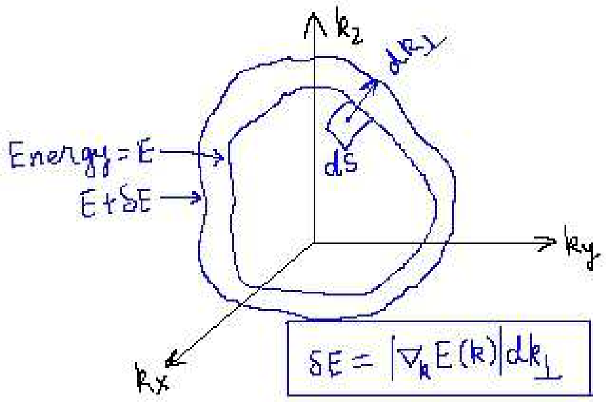


Figure 5.3: Calculation of the k-space area/volume between two equal energy contours differing slightly in energy.

In 1d it is trivial. We choose an interval k to $k+\text{delta}k$ and the corresponding interval in energy $E(k)$ to $E(k+\delta k) = E+\delta E$. The number of states in these two intervals must be equal, so:

$$
\begin{aligned}
D(E)\delta E &= D(k)\delta k && (5.43)\\
&= \frac{L}{2\pi}\frac{\delta k}{\delta E} && (5.44)\\
&= \frac{L}{2\pi}\frac{1}{\mathrm{d}E/\mathrm{d}k} && (5.45)
\end{aligned}
$$

Because there is only one component of k the derivative is simple.

Now, in 2d and 3d we proceed as follows, follow the logic carefully, this would reappear many times in different places.

1. We need to count all the states that lie between the two constant energy contours E and $E+\mathrm{d}E$. These can have quite complex shapes depending on the $E(k_x, k_y, k_z)$ relation.
2. To do this we calculate the k-space volume enclosed by the two contours and multiply the number with $V/(2\pi)^3$. Our volume element here is $\mathrm{d}S\mathrm{d}k_\perp$.
3. The normal to an "equipotential" is given by the gradient. Hence the normal to $E(k_x, k_y, k_z) =$ constant will be given by $\nabla_{\mathbf{k}}E(k_x, k_y, k_z)$ So:

$$
\begin{aligned}
\delta E &= |\nabla_{\mathbf{k}}E(k_x, k_y, k_z)|\mathrm{d}k_\perp && (5.46)\\
\therefore D(E)\delta E &= \frac{V}{(2\pi)^3}\int\limits_S \mathrm{d}S\frac{\delta E}{|\nabla_{\mathbf{k}}E(k_x, k_y, k_z)|} && (5.47)\\
\therefore D(E) &= (2s+1)\frac{V}{(2\pi)^3}\int\limits_S \frac{\mathrm{d}S}{|\nabla_{\mathbf{k}}E(k_x, k_y, k_z)|} && (5.48)
\end{aligned}
$$

In the last step we have included the spin-degeneracy. V is the sample volume. If we want density of states per unit volume, obviously this will be dropped.

4. Points where the group velocity vanishes can give rise to singularities in $D(E)$, but these will be integrable. If the gradient vanishes then we should be able to expand E around this point as

$$E(k_x, k_y, k_z) = E_0 + \frac{\hbar^2}{2m_x}k_x^2 + \frac{\hbar^2}{2m_y}k_y^2 + \frac{\hbar^2}{2m_z}k_z^2 \tag{5.49}$$

If all the coefficients are positive, then this is a band minima, if all are negative it is a band maxima and if they are mixed it is a saddle point. Points where the density of states or its derivative is singular are called *van Hove singularities*.

5.1.3 More than 1 orbital per site

At the beginning of the chapter we calculated the wavefunction/energy levels of a diatomic molecule. We had to solve for the eigenvalues and then get the coefficients of the atomic states which made up the molecular wavefunctions. But no such procedure was needed when we solved the linear chain. Why? The reason is that the symmetry (Bloch's theorem) told us what the coefficients would be. We now ask, what if there are two atoms a and b per lattice site (the basis can of course have more) or two orbitals on the same atom (like a $2s$ and $2p$ orbital or some s and d orbitals..). In these cases we still begin with the atomic wavefunctions, but Bloch's theorem cannot tell us how much of the wavefunction of site a and site b to take. We must solve for those.

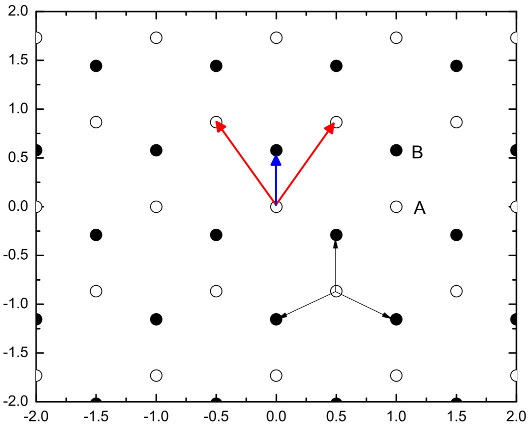


Figure 5.4: Lattice vectors of Graphene. All the atoms are Carbon, but the two types of sites mean that the Bloch functions cannot be written down straightway.

Consider the example of a single sheet from graphite (graphene). The triangular lattice has a two atom basis. We take the following as lattice and vectors:

$$\begin{aligned}\mathbf{a_1} &= \frac{a}{2}(1, \sqrt{3})\\ \mathbf{a_2} &= \frac{a}{2}(-1, \sqrt{3})\end{aligned} \tag{5.50}$$

The two point basis is composed of :

$$\begin{aligned}\text{Type A atoms} &: (0,0)\\ \text{Type B atoms} &: \mathbf{d} = a(0, \frac{1}{\sqrt{3}})\end{aligned} \tag{5.51}$$

Notice that the nearest neighbours of A are B type atoms. Thus the largest hopping terms would occur between A-B overlaps.

If $|\phi_A\rangle$ and $|\phi_B\rangle$ are states centered at A and B. So we form a set of two Bloch functions and make a linear combination with unknown (to be solved for) coefficients α and β

$$|\psi_{\mathbf{k},A}\rangle = \frac{1}{\sqrt{N}}\sum_{\mathbf{R}} e^{i\mathbf{k}.\mathbf{R}}|\phi_A\rangle \tag{5.52}$$

$$|\psi_{\mathbf{k},B}\rangle = \frac{1}{\sqrt{N}}\sum_{\mathbf{R}} e^{i\mathbf{k}.\mathbf{R}}|\phi_B\rangle \tag{5.53}$$

$$|\psi_{\mathbf{k}}\rangle = \alpha|\psi_{\mathbf{k},A}\rangle + \beta|\psi_{\mathbf{k},B}\rangle \tag{5.54}$$

Now we have:

$$\begin{aligned} H|\psi_{\mathbf{k}}\rangle &= E|\psi_{\mathbf{k}}\rangle \\ \langle\psi_{\mathbf{k}A}|H|\psi_{\mathbf{k}}\rangle &= E\langle\psi_{\mathbf{k}A}|\psi_{\mathbf{k}}\rangle \\ \langle\psi_{\mathbf{k}B}|H|\psi_{\mathbf{k}}\rangle &= E\langle\psi_{\mathbf{k}B}|\psi_{\mathbf{k}}\rangle \end{aligned} \tag{5.55}$$

The idea can be extended to more complex basis sets. For this we need to the following (the calculation/justification is left as an exercise).

$$\begin{aligned} \langle\psi_{\mathbf{k}A}|\psi_{\mathbf{k}A}\rangle &= 1 \\ \langle\psi_{\mathbf{k}A}|\psi_{\mathbf{k}B}\rangle &\approx 0 && (5.56) \\ \langle\psi_{\mathbf{k}A}|H|\psi_{\mathbf{k}A}\rangle &= \tilde{E}_0 \\ \langle\psi_{\mathbf{k}A}|H|\psi_{\mathbf{k}B}\rangle &= \sum_{\substack{\text{nearest}\\\text{neighbours}}} e^{i\mathbf{k}.\mathbf{R}}\langle\phi_A|H|\phi_B\rangle && (5.57) \\ &= \left(e^{i\mathbf{k}.\mathbf{0}} + e^{-i\mathbf{k}.\mathbf{a_1}} + e^{-i\mathbf{k}.\mathbf{a_2}}\right)\langle\phi(\mathbf{r})|H|\phi(\mathbf{r}-\mathbf{d})\rangle && (5.58) \\ &= \left(1 + 2\cos\left(\frac{k_x a}{2}\right)e^{-i\frac{\sqrt{3}}{2}k_y a}\right)t && (5.59) \\ &= F(k_x, k_y)t && (5.60) \end{aligned}$$

where t is again the nearest neighbour hopping amplitude. The hopping term occurs between the an atom and its 3 neighbours, shown by black arrows in Fig. 5.4. If we considered next nearest neighbours as well, these would have come from the terms in $\langle\psi_{\mathbf{k}A}|H|\psi_{\mathbf{k}A}\rangle$. (Justify this as an exercise).

With these in place the set of eqns. 5.55 gives the eigenvalue equation:

$$\begin{pmatrix} \tilde{E}_0 & tF(k_x, k_y) \\ tF^*(k_x, k_y) & \tilde{E}_0 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{5.61}$$

PROBLEM :
Show that:

1. the eigenvalues of matrix 5.61 are given by

$$\begin{aligned} E(k_x, k_y) &= \tilde{E}_0 \pm t|F(k_x, k_y)| && (5.62) \\ \text{where} \qquad |F|^2 &= 1 + 4\cos^2\frac{k_x a}{2} + 4\cos\frac{k_x a}{2}\cos\frac{\sqrt{3}}{2}k_y a && (5.63) \end{aligned}$$

2. The reciprocal lattice vectors of the graphene lattice are given by:

$$\begin{aligned} \mathbf{b_1} &= \frac{2\pi}{a}\left(1, \frac{1}{\sqrt{3}}\right) \\ \mathbf{b_2} &= \frac{2\pi}{a}\left(-1, \frac{1}{\sqrt{3}}\right) \end{aligned} \tag{5.64}$$

Two branches have now appeared, this is a common feature in problems where the lattice has a two basis points, pretty much similar situations occur for electron energy bands as well as phonon bands (we will see later).

Now we identify the first Brillouin zone of the triangular lattice and plot the energy eignenvalues in that zone.

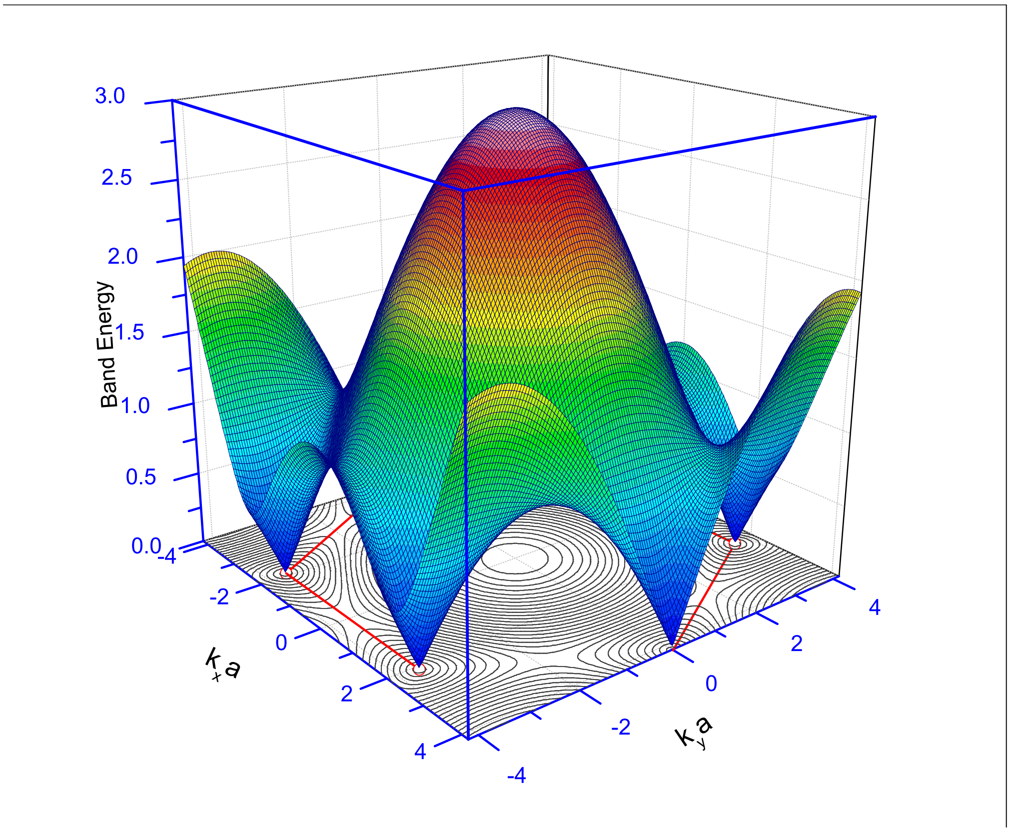


Figure 5.5: Notice that at the six corner points the upper and lower bands touch. The dispersion relation near those points is linear and hence the electrons near those points behave very differently from electrons in most other common substances which have a parabolic dispersion. Also the "touch" implies that we have a zero bandgap material.

5.2 Measuring the effective mass: Cyclotron resonance

Since the effective mass is the curvature of the $E(\boldsymbol{k})$ relation, matching this to experimental data is an important part of refining band structure calculations. How do we measure this? We use the fact that in magnetic field the electron orbits have a frequency of rotation. One way is to measure the resonant frequencies in a magnetic field. We will see that this frequency is related to effective mass in a particular direction. $\boldsymbol{M}$ is 3×3 matrix or the effective mass tensor. It is symmetric. The equation of motion of an electron in a band is:

$$\boldsymbol{M}\frac{d}{dt}\boldsymbol{v} = -e\boldsymbol{v} \times \boldsymbol{B} \tag{5.65}$$

If we look for oscillatory solutions, we must have

$$\boldsymbol{v} = \tilde{\boldsymbol{v}}_0 e^{i\omega t} \tag{5.66}$$

Let's direct the magnetic field along z axis, so that $\boldsymbol{B} = B_0\hat{\boldsymbol{z}}$.

PROBLEM : Show that the last two equations imply

$$\begin{vmatrix} M_{xx} & M_{xy} + i\frac{eB_0}{\omega} & M_{xz} \\ M_{xy} - i\frac{eB_0}{\omega} & M_{yy} & M_{yz} \\ M_{zx} & M_{zy} & M_{zz} \end{vmatrix} = 0 \tag{5.67}$$

And expanding the determinant gives

$$\frac{\det M}{M_{zz}} = \frac{e^2B_0^2}{\omega^2} \tag{5.68}$$

Notice that if we rotate the direction of the magnetic field, we can bring another effective mass into focus.

5.3 Orthogonalised Plane Waves (OPW) and Pseudopotential

First try the following problem:

PROBLEM : Consider a monatomic cubic lattice (again) with lattice constant $a = 5$Å, in which the ion cores sitting in the lattice sites have a Bohr radius of $r = 0.1$Å. Bohr radius of an atom is the radius of the $1s$ state. This state behaves like the tightly bound states (ground state of the simple quantum wells) in the Kronig Penny model that we studied. Estimate by simple arguments, the approximate number of the reciprocal lattice vectors you will have to retain in the plane wave secular determinant (that we derived in the context of Bloch's theorem), to get a good description of this material. The numbers given to you are typical, but not specific to any element. Obviously this will no longer be a "nearly free electron" problem.

The nearly free electron and tight binding approximations represent two opposite limits. The complete solution of the lattice potential should yield the plane wave like delocalised states as well as the localised or tightly bound "core" states. It would require a very large number (upto a million!) of plane wave states to reproduce a deep core state. Also these "plane wave" states and "core" states must be mutually orthogonal, since they belong to different eigenvalues of the same hamiltonian. However if we could somehow write the plane wave states in a way so that they are *by construction* orthogonal to the core states, then the problem would be simplified. Here "simplified" means that a lesser number of components (not a million, but in practice may be about 100) would be needed. This is a crucial point. Let's see how. Some of the intermediate algebra is left for you to complete:

First write down the tight binding state formed out of a core orbital (sum over core states are denoted by c and sum over $\boldsymbol{R}$ denotes sum over all lattice site. Here for simplicity we stick to a monatomic lattice,

H denotes the full lattice hamiltonian that contains the potentials created by all the atoms. $\boldsymbol{G}$ denotes reciprocal lattice vectors and $\boldsymbol{k}$ is within BZ1.

$$\langle \boldsymbol{r}|f_{c\boldsymbol{k}}\rangle = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}.\boldsymbol{R}}\phi_c(\boldsymbol{r}-\boldsymbol{R}) \tag{5.69}$$

First we construct the orthogonalised plane wave (OPW) basis states

$$\begin{aligned} |\chi_{\boldsymbol{k}}\rangle &= |\boldsymbol{k}\rangle - \sum_c \langle f_{c\boldsymbol{k}}|\boldsymbol{k}\rangle|f_{c\boldsymbol{k}}\rangle \\ |\chi_{\boldsymbol{k}-\boldsymbol{G}}\rangle &= |\boldsymbol{k}-\boldsymbol{G}\rangle - \sum_c \langle f_{c\boldsymbol{k}}|\boldsymbol{k}-\boldsymbol{G}\rangle|f_{c\boldsymbol{k}}\rangle \end{aligned} \tag{5.70}$$

We will use these rather than $|\boldsymbol{k}-\boldsymbol{G}\rangle$, to form the eigenvalue equation. The full OPW state is a linear combination

$$|\Psi_{\boldsymbol{k}}\rangle = \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}|\chi_{\boldsymbol{k}-\boldsymbol{G}}\rangle \tag{5.71}$$

We will assume that the core state has no dispersion and the energy of the state $\langle \boldsymbol{r}|f_{c\boldsymbol{k}}\rangle$ can be replaced by its atomic value E_c independent of $\boldsymbol{k}$
The main characteristic of the OPW state (eq 5.70) is that it is like a plane wave far from the ion core, but oscillates very fast close to the cores. This allows it to remain orthogonal to the Bloch states (with very little bandwidth) formed out of the core states (prove this).

$$\begin{aligned} H|\Psi_{\boldsymbol{k}}\rangle &= \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}\left[H|\boldsymbol{k}-\boldsymbol{G}\rangle - \sum_c \langle f_{c\boldsymbol{k}}|\boldsymbol{k}-\boldsymbol{G}\rangle E_c|f_{c\boldsymbol{k}}\rangle\right] \\ &= \lambda|\Psi_{\boldsymbol{k}}\rangle \end{aligned} \tag{5.72}$$

Now left multiply with $\langle \boldsymbol{k}-\boldsymbol{G'}|$ and show that

$$\begin{aligned} \langle \boldsymbol{k}-\boldsymbol{G'}|H|\Psi_{\boldsymbol{k}}\rangle &= \sum_{\boldsymbol{G}} C_{\boldsymbol{G}}\left[\frac{\hbar^2}{2m}(\boldsymbol{k}-\boldsymbol{G})^2\delta_{\boldsymbol{G}\boldsymbol{G'}} + \langle \boldsymbol{G'}|V|\boldsymbol{G}\rangle + \sum_c E_c\langle f_{c\boldsymbol{k}}|\boldsymbol{k}-\boldsymbol{G}\rangle\langle \boldsymbol{k}-\boldsymbol{G'}|f_{c\boldsymbol{k}}\rangle\right] \\ &= \lambda\sum_{\boldsymbol{G}} C_{\boldsymbol{G}}\left[\delta_{\boldsymbol{G}\boldsymbol{G'}} - \sum_c \langle f_{c\boldsymbol{k}}|\boldsymbol{k}-\boldsymbol{G}\rangle\langle \boldsymbol{k}-\boldsymbol{G'}|f_{c\boldsymbol{k}}\rangle\right] \end{aligned} \tag{5.73}$$

This means

$$\sum_{\boldsymbol{G}} C_{\boldsymbol{G}}\left[\left(\frac{\hbar^2}{2m}(\boldsymbol{k}-\boldsymbol{G})^2-\lambda\right)\delta_{\boldsymbol{G}\boldsymbol{G'}} + \langle \boldsymbol{G'}|V|\boldsymbol{G}\rangle + \sum_c(\lambda-E_c)\langle f_{c\boldsymbol{k}}|\boldsymbol{k}-\boldsymbol{G}\rangle\langle \boldsymbol{k}-\boldsymbol{G'}|f_{c\boldsymbol{k}}\rangle\right] = 0 \tag{5.74}$$

- The values of λ that will arise as the roots of the determinental equation, are the desired band energies.
- Notice how the core states have contributed an additional term to the potential. The original potential term $\langle \boldsymbol{G'}|V|\boldsymbol{G}\rangle$ is negative in magnitude because it is the attractive Coulomb potential of the atomic nucleii. But the new term is positive because $\lambda > E_c$.
- Numerical computation show that the cancellation is very good, leaving often only 5% of $\langle \boldsymbol{G'}|V|\boldsymbol{G}\rangle$. This effective potential is called the psedopotential.
- This also tells us why the band structure of real materials still has considerable similarity with the nearly free electron result.

5.3.1 The pseudopotential and the pseudo-wavefunction

We can now ask, whether a simplified equation can be found that the state $\Phi_{\boldsymbol{k}} = \sum_{\boldsymbol{G}} C_{\boldsymbol{G}} |\boldsymbol{k} - \boldsymbol{G}\rangle$ satisfy. This is the smoothly varying part of the OPW wavefunction $\Psi_{\boldsymbol{k}}$. We have

$$\begin{aligned} |\Psi_{\boldsymbol{k}}\rangle &= \sum_{\boldsymbol{G}} C_{\boldsymbol{G}} \left[|\boldsymbol{k} - \boldsymbol{G}\rangle - \sum_c \langle f_{c\boldsymbol{k}} | \boldsymbol{k} - \boldsymbol{G}\rangle |f_{c\boldsymbol{k}}\rangle \right] && (5.75) \\ &= \sum_{\boldsymbol{G}} C_{\boldsymbol{G}} |\boldsymbol{k} - \boldsymbol{G}\rangle - \sum_{\boldsymbol{G}} C_{\boldsymbol{G}} \sum_c \langle f_{c\boldsymbol{k}} | \boldsymbol{k} - \boldsymbol{G}\rangle |f_{c\boldsymbol{k}}\rangle && (5.76) \\ &= |\Phi_{\boldsymbol{k}}\rangle - \sum_c \langle f_{c\boldsymbol{k}} | \Phi_{\boldsymbol{k}}\rangle |f_{c\boldsymbol{k}}\rangle && (5.77) \end{aligned}$$

Therefore

$$\begin{aligned} H|\Psi_{\boldsymbol{k}}\rangle &= H|\Phi_{\boldsymbol{k}}\rangle - \sum_c \langle f_{c\boldsymbol{k}} | \Phi_{\boldsymbol{k}}\rangle E_c |f_{c\boldsymbol{k}}\rangle && (5.78) \\ &= E|\Phi_{\boldsymbol{k}}\rangle - E\sum_c \langle f_{c\boldsymbol{k}} | \Phi_{\boldsymbol{k}}\rangle |f_{c\boldsymbol{k}}\rangle && (5.79) \\ & && (5.80) \end{aligned}$$

Therefore

$$\begin{aligned} H|\Phi_{\boldsymbol{k}}\rangle - \sum_c \langle f_{c\boldsymbol{k}} | \Phi_{\boldsymbol{k}}\rangle E_c |f_{c\boldsymbol{k}}\rangle &= E|\Phi_{\boldsymbol{k}}\rangle - E\sum_c \langle f_{c\boldsymbol{k}} | \Phi_{\boldsymbol{k}}\rangle |f_{c\boldsymbol{k}}\rangle && (5.81) \\ T|\Phi_{\boldsymbol{k}}\rangle + V|\Phi_{\boldsymbol{k}}\rangle + \sum_c (E - E_c) \langle f_{c\boldsymbol{k}} | \Phi_{\boldsymbol{k}}\rangle |f_{c\boldsymbol{k}}\rangle &= E|\Phi_{\boldsymbol{k}}\rangle && (5.82) \\ T|\Phi_{\boldsymbol{k}}\rangle + \underbrace{\left[V + \sum_c (E - E_c) |f_{c\boldsymbol{k}}\rangle\langle f_{c\boldsymbol{k}}| \right]}_{\text{U}} |\Phi_{\boldsymbol{k}}\rangle &= E|\Phi_{\boldsymbol{k}}\rangle && (5.83) \end{aligned}$$

Convince yourself that the operator U is like an integral operator:

$$\begin{aligned} U|\Phi_{\boldsymbol{k}}\rangle &= V|\Phi_{\boldsymbol{k}}\rangle + \sum_c (E - E_c) |f_{c\boldsymbol{k}}\rangle\langle f_{c\boldsymbol{k}} | \Phi_{\boldsymbol{k}}\rangle && (5.84) \\ \langle \boldsymbol{r}|U|\Phi_{\boldsymbol{k}}\rangle &= \langle \boldsymbol{r}|V|\Phi_{\boldsymbol{k}}\rangle + \sum_c (E - E_c) \langle \boldsymbol{r}|f_{c\boldsymbol{k}}\rangle\langle f_{c\boldsymbol{k}} | \boldsymbol{r'}\rangle\langle \boldsymbol{r'}|\Phi_{\boldsymbol{k}}\rangle && (5.85) \\ \therefore U(\boldsymbol{r})\Phi(\boldsymbol{r}) &= V(\boldsymbol{r})\Phi(\boldsymbol{r}) + \sum_c (E - E_c) \int d\boldsymbol{r'}\, \underbrace{f_{c,\boldsymbol{k}}(\boldsymbol{r}) f^*_{c,\boldsymbol{k}}(\boldsymbol{r'})}_{K(\boldsymbol{r},\boldsymbol{r'})} \Phi(\boldsymbol{r'}) && (5.86) \end{aligned}$$

The pseudo-potential acts on the pseudo-wavefunction and produces the correct eigenvalues! The deep coulomb potential near the nuclei and the sharp rise and fall (nodes) of the wavefunction have been taken away. This OPW+Pseudopotenial method can be used to calculate real band structures of many elements.

5.3.2 What have we still left out?

There are two things we have not considered at all. These are beyond the scope of this course.

1. We have ignored the spin of the electrons other than a simple degeneracy factor of 2. The spin enters the band structure (even at zero magnetic field) through an effect called the "spin-orbit" coupling. This is a small perturbation, but can only be correctly deduced by starting with the Dirac equation for an electron.

2. At the end of the day, no band structure problem is a single electron problem. The coulomb repulsion between the electrons is strong, but it is a remarkable fact that they do not fundamentally alter the band picture that we have calculated so far. Why it can happen was systematically explained by Landau.

Chapter 6

Lattice vibrations

References:

1. Ashcroft and Mermin, Chapters 21,22,23,24
2. Kittel

It does not take much to motivate the idea that the lattice cannot be completely static. Let us recall, what we know from the kinetic theory of gases. At any finite temperature the gas molecules move about with kinetic energies proportional to the absolute temperature. In case of a solid such motion must be somewhat restricted, otherwise the solid will just fly apart. So we can quite correctly surmise that the atoms that sit on lattice points must execute some kind of motion about their equilibrium position. At higher temperatures, the motion will be more energetic. It is interesting to point out an old, simple criteria (due to Lindemann) that when the amplitude of vibration of the atoms about their equilibrium position reaches about 10% of the lattice constant, the solid melts.

However, there are also a good number of simple experimental observations that suggest that the lattice must act as a reservoir of energy.

- Consider a frozen inert gas lattice, it has no free electron gas moving around. But even these have a specific heat, that varies in a particular way with temperature. In fact it is in general true for solids that the specific heat (C_v)of a solid approaches zero as $T{\to}0$ and a near constant value at higher temperatures.
- The scattering of X-rays, visible light, neutrons etc from crystals show distinct features that cannot be explained if the lattice was always static.

We approach this problem exactly the way we treat small oscillations (and normal modes) in classical mechanics.

- We first identify the sources of energy of the system.
- Then we write the Hamiltonian taking into account all the particles.
- The kinetic energy is simply $T = \sum_{\text{all}} p_i^2/2m$, as always, where i runs over the index of the particles.
- We then approximate (Taylor expand) the pairwise sum of potential energies retaining the quadratic terms only. The pairwise sum may be restricted to a few nearest neighbours in some (*but not all*) cases.
- Recall that if our expansion is centered about an equilibrium point, then the first order term must be zero.
- We then find the normal mode frequencies.

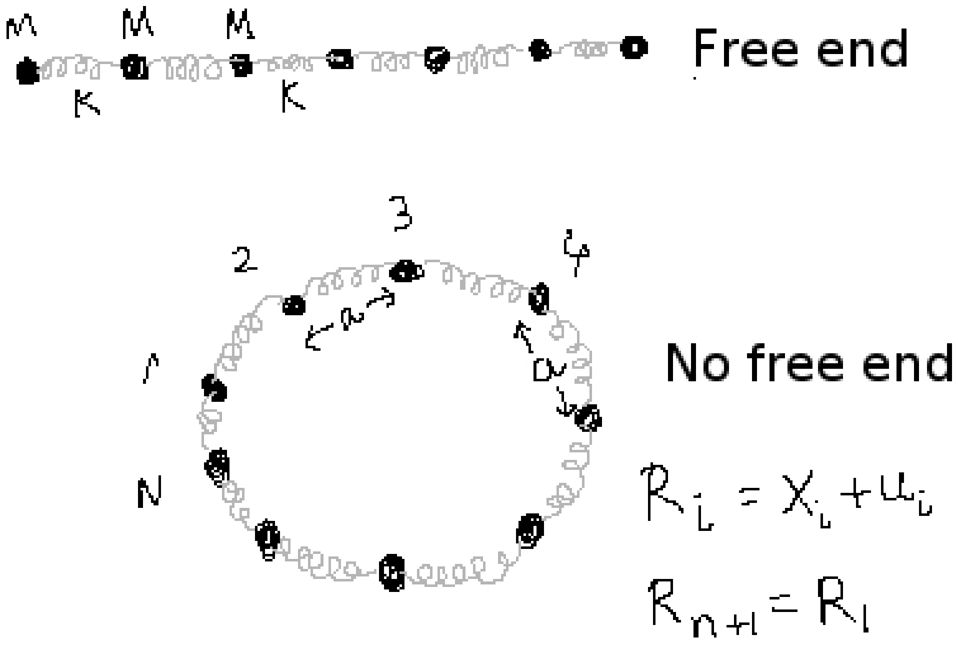


Figure 6.1: The periodic boundary condition, in 1d means using a ring rather than a line. The compression of the spring between i^{th} and $i{+}1^{\text{th}}$ site is given by $u_i - u_{i+1}$. Note that the correct sign of the displacements is needed. Here we will consider displacements to be positive if it is to the right - following the usual convention that the x-axis is positive in the right direction. The displacements are longitudinal.

- Once we have expressed the energy as a sum of harmonic oscillators, we can think about "quantizing" the system.

We first take a look at a (regulation) toy model of an 1-dimensional chain. Even though it is very simplified, it allows us to draw some very general conclusions, that is why it is really important to understand this!

6.1 1-dimensional mass & spring chain

We will deal with the equilibrium positions of many particles and small deviations from these positions. We need a systematic way to label them. We will use the convention that uppercase (X_i, Y_i, $\mathbf{R}_i$) will denote equilibrium positions and lowercase letters (x_i, y_i, u_i) will denote deviations from the equilibrium positions.

Consider N identical particles (mass m, lattice spacing a) arranged on a ring as shown. We consider a ring, to avoid the complexity associated with handling free ends of a chain. We *can* do the problem for a long chain with free ends, but the answer turns out to be same for large N.
So, we write

$$H = \sum_n \frac{p_i^2}{2m} + \frac{K}{2} \sum_n (u_{n+1} - u_n)^2 \tag{6.1}$$

with the understanding that

$$u_N = u_0 \tag{6.2}$$

Note that u_n occurs twice in the potential energy part, once with u_{n+1} and once with u_{n-1}. The equations of motion can be written following the standard procedure:

$$\begin{aligned} m\ddot{u}_n &= \dot{p_n} \\ &= -\frac{\partial H}{\partial u_n} \qquad (6.3) \\ &= K(u_{n-1} - 2u_n + u_{n+1}) \qquad (6.4) \end{aligned}$$

We try a wavelike solution

$$u_n = \epsilon e^{i(kna - \omega t)} \qquad (6.5)$$

where the allowed values of k must satisfy

$$e^{ikNa} = 1 \qquad (6.6)$$

such that when we return to the $n = 0$ position, we don't produce two different values of the displacement. Using the solution 6.5 in eqn 6.4 gives the relation between ω and k.

$$\omega^2 = 2\frac{K}{m} \sin^2 \frac{ka}{2} \qquad (6.7)$$

Comparison with a general "small oscillation" problem

When we deal with small oscillations in an arbitrary system, we cannot anticipate solutions for the mode co-ordinates. In those cases (as in a molecule for example) we must diagonalize the hamiltonian first and *find* the mode co-ordinates. In this case, thanks to the periodicity of the system we could write/guess the mode co-ordinates *before* diagonalizing the hamiltonian.

If we wrote out the set of linear equations 6.4 in a matrix form, with

$$u_n = A_n e^{-i\omega t} \qquad (6.8)$$

then the most general form of the matrix we can get is given in eqn 6.9. Notice that the most generic form is not restricted to nearest neighbour couplings, but all columns will necessarily be rotated versions of the first column. Such a "circulant" matrix has only n independent elements.

We have an $n{\times}n$ matrix of the circulant form:

$$\begin{pmatrix} c_0 & c_{n-1} & c_{n-2} & .. & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & .. & c_3 & c_2 \\ .. & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & .. \\ c_{n-2} & c_{n-3} & .. & .. & .. & c_{n-1} \\ c_{n-1} & c_{n-2} & c_{n-3} & .. & .. & c_0 \end{pmatrix} \begin{pmatrix} A_0 \\ A_1 \\ .. \\ .. \\ A_{n-2} \\ A_{n-1} \end{pmatrix} = \lambda \begin{pmatrix} A_0 \\ A_1 \\ .. \\ .. \\ A_{n-2} \\ A_{n-1} \end{pmatrix} \qquad (6.9)$$

The eigenvectors will be of the form

$$A = \left(1, z, z^2, .., .., z^{n-1}\right)^T \qquad (6.10)$$

where $z^n = 1$. Prove this result by direct substitution. You should then be able to justify what we did in eqn 6.6. Notice that utilising the specific form of the eigenvectors reduced a $n{\times}n$ matrix problem to just one equation. Convince yourself that the n possible values of z will correspond to the different k vectors.

6.1.1 How many solutions to expect

Recall that a system with N degrees of freedom will have N mode frequencies. What are those? We have not yet finished solving the problem. Using eqn 6.6 we see that the k values must be

$$k = 0, \frac{2\pi}{a}\frac{1}{N}, \frac{2\pi}{a}\frac{2}{N} .... \frac{2\pi}{a}\frac{N-1}{N} \qquad (6.11)$$

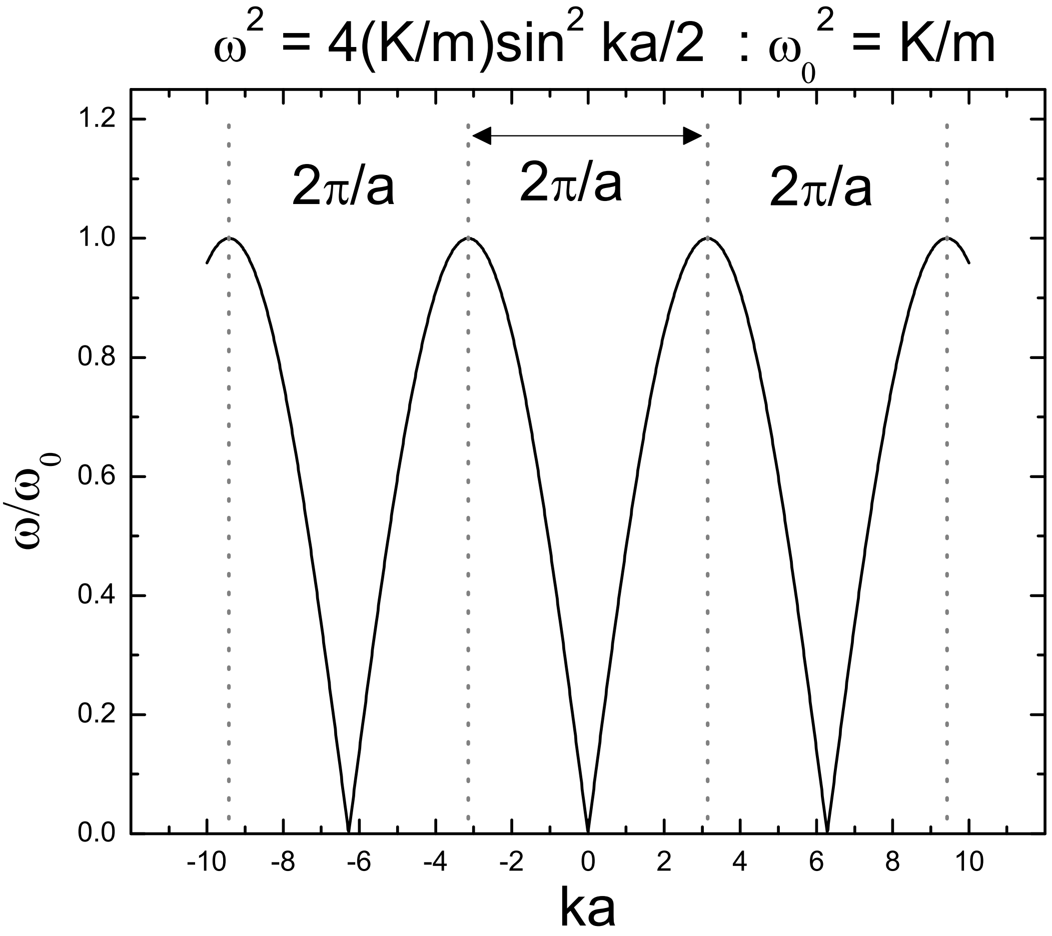


Figure 6.2: Dispersion relation given in equation 6.7. Notice how the values of ω repeats itself.

But there is something more to it. The displacements u_n only make sense at discrete points $R_n = na$. This means that if we make a change

$$k \to k + \frac{2\pi l}{a} \tag{6.12}$$

where l is any integer, the displacements will be indistinguishable. So the "useful" values of k can only span a length $2\pi/a$ in $k-$space. This is also clearly seen by looking at the dispersion relation (Fig. 6.4), which repeats with a similar interval. Usually we choose the region between $[-\pi/a, \pi/a]$ with the zero at the center. This choice also highlights the symmetrical behaviour of $+k$ and $-k$ directions.

PROBLEM: Given that the vibrational frequencies of solids are often in the range of $\sim 10^{12}$Hz, consider a chain of atoms with 50amu. What kind of spring constant (K) does it suggest between atoms? How does it compare with a spring that extends by 1cm if 1kg weight is suspended from it?

This multiple valued character of the possible wavevectors appears in other situations where a displacement is sampled at discrete points in space, or a signal is sampled at discrete intervals in time (say every 1 millisec, for example). Can you see why? Recall the relation between sampling interval and aliasing of high frequencies while doing a discrete Fourier transform.

Small k behaviour of $\omega(k)$

For small k we have

$$\omega = 2\omega_0 \frac{ka}{2} \tag{6.13}$$

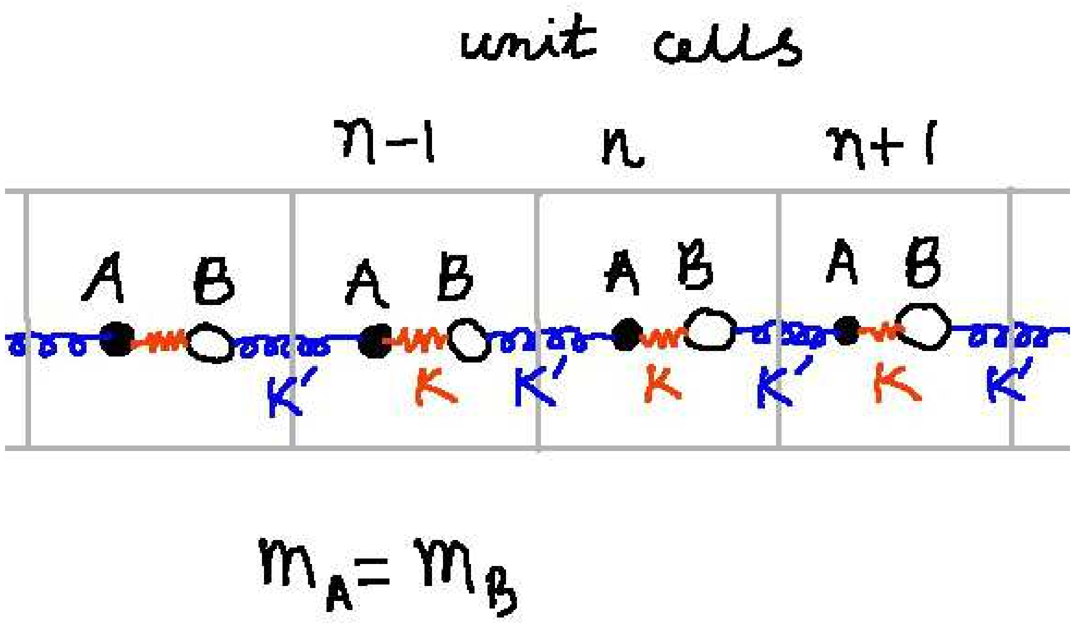


Figure 6.3: The diatomic chain. The lattice has a two-atom basis (labelled A and B), whose instantaneous displacements are $u_{A,n}$ and $u_{B,n}$ respectively. The two masses are still considered equal and the spring constants are K and K'.

The phase velocity is clearly constant, as a result all frequencies travel at the same speed. Such modes are called accoustic modes in obvious analogy with sound waves. It turns out that as long as the lattice is monoatomic, all the vibration modes are of accoustic type. This is true not only for our ”toy model” but also for real 3-dimensional solids.

6.1.2 1 dimensional chain with a two-atom basis

Now consider the system shown in Fig. 6.3, we can write down the sources of energy by adding kinetic and potential parts as before. We write down the energy of one unit cell and sum over all the cells:

$$H = \sum_n \left(\frac{p_{A,n}^2}{2m} + \frac{p_{B,n}^2}{2m} \right) + \frac{K}{2} \sum_n (u_{B,n} - u_{A,n})^2 + \frac{K'}{2} \sum_n (u_{A,n+1} - u_{B,n})^2 \tag{6.14}$$

This leads to the coupled equations of motion:

$$\begin{aligned} m\ddot{u}_{A,n} &= \dot{p}_{A,n} \\ &= -\frac{\partial H}{\partial u_{A,n}} \\ &= K(u_{B,n} - u_{A,n}) - K'(u_{A,n} - u_{B,n-1}) \qquad (6.15) \\ m\ddot{u}_{B,n} &= \dot{p}_{B,n} \\ &= -\frac{\partial H}{\partial u_{B,n}} \\ &= -K(u_{B,n} - u_{A,n}) + K'(u_{A,n+1} - u_{B,n}) \qquad (6.16) \end{aligned}$$

Then we do the trial solutions as :

$$\begin{aligned} u_{A,n} &= \epsilon_A e^{i(kna - \omega t)} \\ u_{B,n} &= \epsilon_B e^{i(kna - \omega t)} \end{aligned} \tag{6.17}$$

which leads to

$$\begin{pmatrix} K + K' - m\omega^2 & -K - K'e^{-ika} \\ -K - K'e^{ika} & K + K' - m\omega^2 \end{pmatrix} \begin{pmatrix} \epsilon_A \\ \epsilon_B \end{pmatrix} = 0 \tag{6.18}$$

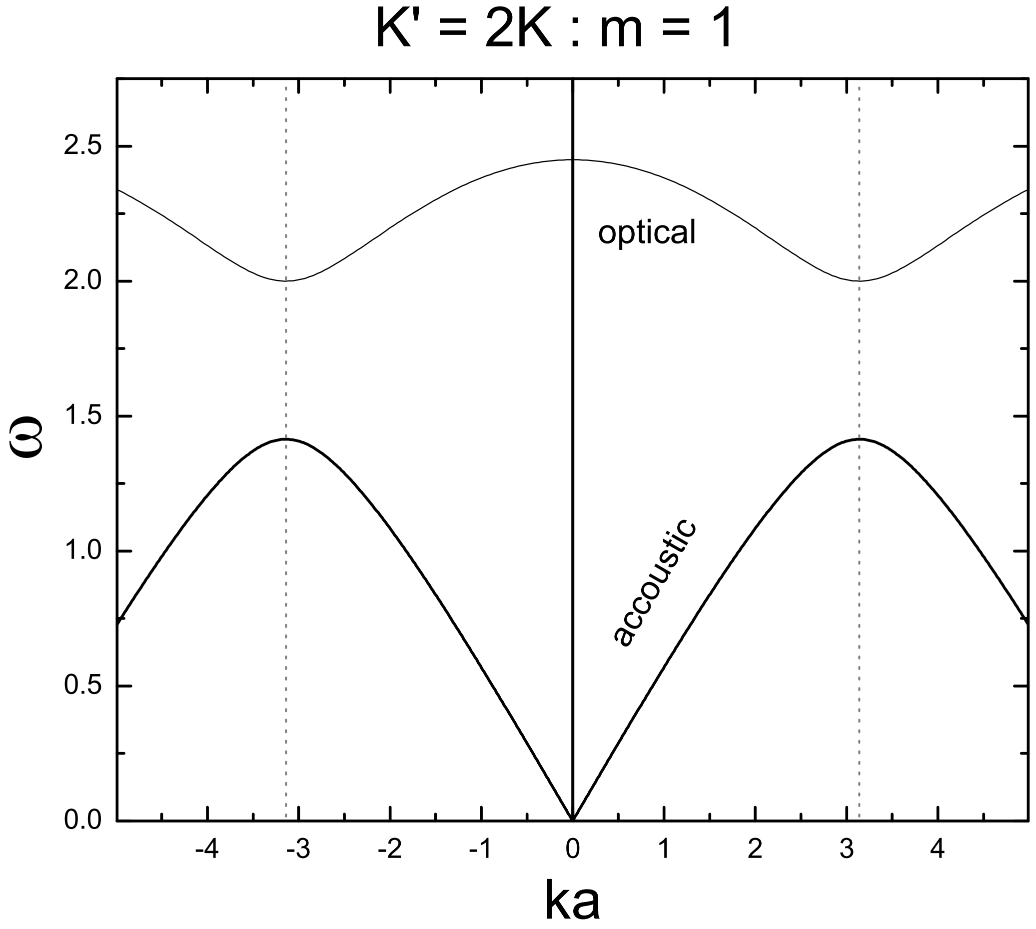


Figure 6.4: Dispersion relation given in equation 6.19. Notice how the values of ω repeats itself as before.

We need the eigenvalues of the matrix to get the mode frequencies. There should be two branches given by :

$$m\omega^2 = (K + K') \pm \sqrt{K^2 + K'^2 + 2KK'\cos ka} \tag{6.19}$$

The ratio of the two displacements, is given by :

$$\frac{\epsilon_A}{\epsilon_B} = \mp \frac{K + K'e^{-ika}}{|K + K'e^{ika}|} \tag{6.20}$$

The first thing to notice is that one of the branches (the +ve solution) goes to zero as $k \rightarrow 0$, but the other one tends to a finite value. For $ka << 1$ we can expand the solutions as:

$$\omega_+ = \sqrt{\frac{2(K + K')}{m}} + \mathcal{O}(ka)^2 \tag{6.21}$$

$$\omega_- = \sqrt{\frac{KK'}{2(K + K')}}\, ka \tag{6.22}$$

The two branches are shown in the figure, plotted for $K' = 2K$ and $m = 1$ in Fig.6.4

Why are optical modes so called

Notice the ω_+ solution in equation 6.22, and compare the ratio of the displacements of the two types of atoms. The ratio is close to -1. This means that they are vibrating out of phase with each other. That can happen, for example if visible light ($\lambda \sim 500$nm) excites a crystal of salts (like NaCl). Because the

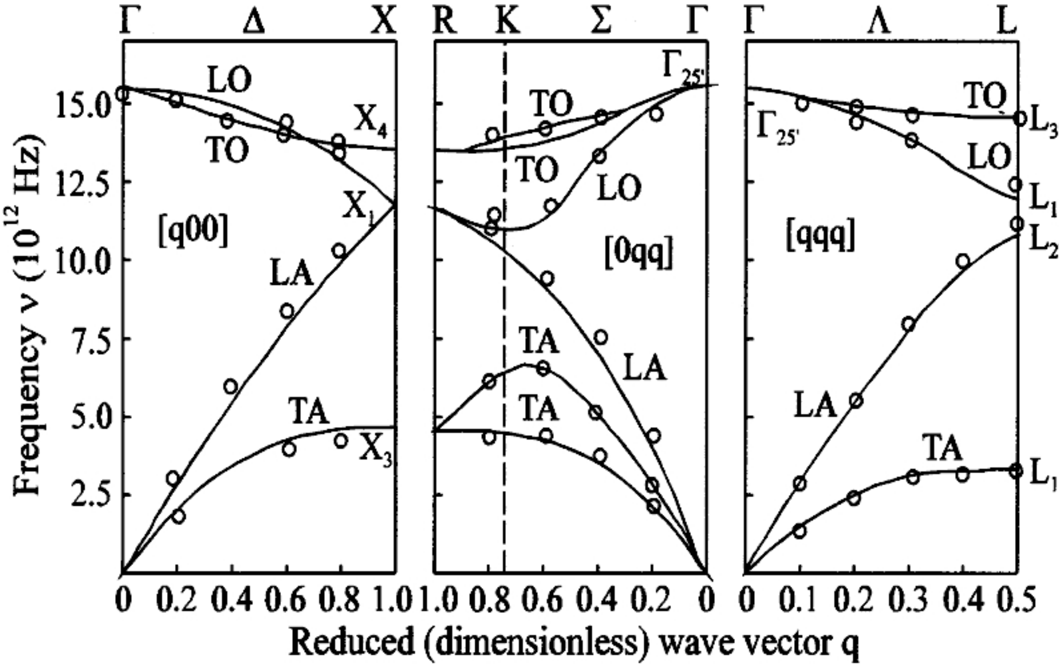


Figure 6.5: Phonon dispersion in Silicon. Notice how the simple model of 2atoms per unit cell, in figure 6.4 already captures the key features. LO stands for Longitudinal Optic, TO for Transverse Optic, and similarly LA, TA denote the acoustic modes. For details refer to R Tubino, L. Piseri, and G. Zerbi, Journal of Chemical Physics **56**, 3 (1972) 1022-1039 and G. Dolling, Proc. Symp. Inelastic Scattering Neutrons in Solids and Liquids, Chalk River, IAEA, Vienna, vol. 2, 1963, p. 37. Data like this is available at <http://www.ioffe.ru/SVA/NSM/Semicond/Si/mechanic.html>

A and B atoms (separated by $\sim$ 1nm) are oppositely charged, the long wavelength electromagnetic field would push them in opposite directions. In the accoustic mode, the nearby atoms would vibrate in-phase. See fig 6.5 for a real phonon spectrum and notice the acoustic and optical modes.

Phonon dispersion and the Fourier transform of the interaction strength

So far we have supposed that the interaction (spring like coupling) is only between the the nearest neighbours. Now if we relax this what happens? We formulate the problem in terms of a Hamiltonian as follows, still in one dimension, and using a ring with N mass points as in fig 6.1 V_{nm} denotes the interaction strength between the sites m and n.

$$\begin{aligned} H &= \frac{1}{2M}\sum_n p_n^2 + \frac{1}{2}\sum_{nm} V_{nm} u_n u_m && (6.23) \\ \therefore \dot{p_n} &= -\frac{\partial H}{\partial x_n} && (6.24) \\ &= -\sum_m V_{nm} u_m && (6.25) \\ & && (6.26) \end{aligned}$$

Then we try an oscillatory solution of the form

$$u_m = Ae^{i(kma-\omega t)} \qquad (6.27)$$

substituting this in eqn 6.25

$$\begin{aligned} -M\omega^2 e^{i(kma-\omega t)} &= -\sum_n V_{mn} e^{i(kna-\omega t)} \\ \therefore M\omega^2(k) &= \sum_n V_{mn} e^{ik(n-m)a} \\ &= \sum_n V(|ma-na|) e^{ik(n-m)a} \\ &= V(k) \end{aligned} \tag{6.28}$$

You should be able to justify the last couple of steps for yourself. Also convince yourself that the earlier result ($\omega \propto \sin ka/2$) is a special case of this result.

6.2 Writing the lattice vibrations in 3D

The Hamiltonian for the 3D Bravais crystal is conceptually not hard to write, but there are a lot more indices to be careful about. The energy of the system U depends on the instantaneous co-ordinates of all the particles. We can write the total potential energy as a pairwise sum as before. But we need to keep in mind the vector nature of all the displacements, since we are in 3D. Upto 2nd order we can write :

$$U(\mathbf{R}_1+\mathbf{u}_1, \mathbf{R}_2+\mathbf{u}_2+\ \ldots.) = U(\mathbf{R}_1, \mathbf{R}_2,\ \ldots.) + \frac{1}{2}\sum_{\mathbf{R}_i\mathbf{R}_j}\sum_{\mu,\nu}\frac{\partial^2 U}{\partial X_{i,\mu}\partial X_{j,\nu}} x_{i,\mu}x_{j,\nu} \tag{6.29}$$

$$= \frac{1}{2}\sum_{\mathbf{R}_i\mathbf{R}_j}\sum_{\mu\nu} D_{\mu\nu}(\mathbf{R}_i,\mathbf{R}_j) x_{i,\mu}x_{j,\nu} \tag{6.30}$$

The indices i, j run over the lattice points, the indices (μ, ν) run over the x, y, z components of each position vector or displacements. The first order term vanishes because the derivatives are evaluated at the equilibrium positions.

This leads to :

$$H = \sum_{\mathbf{R}_i}\frac{\mathbf{P}_i^2}{2m} + \frac{1}{2}\sum_{\mathbf{R}_i\mathbf{R}_j}\sum_{\mu\nu} D_{\mu\nu}(\mathbf{R}_i,\mathbf{R}_j) x_{i,\mu}x_{j,\nu} \tag{6.31}$$

and the equation of motion (there will be $3N$ such equations)

$$m\ddot{u}_{i,\mu} = \sum_{\mathbf{R}_j}\sum_{\nu} D_{\mu\nu}(\mathbf{R}_i,\mathbf{R}_j) x_{j,\nu} \tag{6.32}$$

The form of the equation 6.29 tells us that the elements of the 3×3 matrix $D_{\mu\nu}$ will have certain symmetries/properties:

$$D_{\mu\nu}(\mathbf{R}_i,\mathbf{R}_j) = D_{\mu\nu}(\mathbf{R}_i-\mathbf{R}_j) \tag{6.33}$$

$$D_{\mu\nu}(\mathbf{R}_j-\mathbf{R}_i) = D_{\mu\nu}(\mathbf{R}_i-\mathbf{R}_j) \tag{6.34}$$

$$D_{\mu\nu}(\mathbf{R}_i,\mathbf{R}_j) = D_{\nu\mu}(\mathbf{R}_i,\mathbf{R}_j) \tag{6.35}$$

$$\sum_{\mathbf{R}_i} D(\mathbf{R}_i) = 0 \tag{6.36}$$

We are now working in 3D, so the displacement of the particles is not necessarily along the direction of propagation as we had in 1D. Keeping this in mind, we write the displacements ($\mathbf{u}$ has components x_μ) are :

$$\mathbf{u}_i = \boldsymbol{\epsilon}(\mathbf{k}) e^{i(\mathbf{k}.\mathbf{R}_i-\omega t)} \tag{6.37}$$

Using this form in equation 6.32, we get :

$$\begin{aligned} -m\omega^2(\mathbf{k})\epsilon_\mu(\mathbf{k}) &= \sum_{\mathbf{R}} D_{\mu\nu}(\mathbf{R}) e^{i\mathbf{k}.\mathbf{R}} \epsilon_\nu(\mathbf{k}) & (6.38)\\ &= D_{\mu\nu}(\mathbf{k})\epsilon_\nu(\mathbf{k}) & (6.39) \end{aligned}$$

$D_{\mu\nu}(\mathbf{k})$ is also a 3×3 matrix and can be seen as a Fourier transform. We now have N clearly defined eigenvalue equations, each producing 3 eigenvalues. The boundary conditions, in this case, is also periodic. We set it such that the structure repeats after N_1,N_2 and N_3 units in x,y and z direction respectively, with $N = N_1N_2N_3$. The $\mathbf{k}$ vectors will be chosen such that

$$\mathbf{k} = \frac{n_1}{N_1}\mathbf{b}_1 + \frac{n_2}{N_2}\mathbf{b}_2 + \frac{n_3}{N_3}\mathbf{b}_3 \tag{6.40}$$

This guarantees

$$\mathbf{u}(\mathbf{R} + N_j\mathbf{a}_j) = \mathbf{u}(\mathbf{R}) \tag{6.41}$$

where $\mathbf{a}$ and $\mathbf{b}$ are direct and reciprocal lattice vectors respectively. As we did for the 1-d chain, the convention is to take the k-vector to lie in the first Brillouin zone.

The solutions are our $3N$ mode frequencies. All of these are accoustic modes, but can differ in their slopes near zero. The three eigenvectors corresponding to the eigenvalues give the direction of polarisation $\boldsymbol{\epsilon}(\mathbf{k})$ of the mode. As before the solution will give us the dispersion $\omega(\mathbf{k})$. Although we have not committed to any specific form of $D(\mathbf{R})$ it is clear that as we move away from the lattice site, *i.e* $\mathbf{R}$ becomes large, it must decrease rapidly. If we restrict ourselves to the nearest neighbours only, then (mathematically) we would have something very similar to a "tight binding" model, that we will use quite a bit when we study electronic band structure.

6.3 The lattice as a collection of harmonic oscillators

You can skip this section for a first reading and go to section 6.4. But this is worth knowing how to express a lattice as a collection of harmonic oscillators rigorously...

We have been using a vector $\mathbf{k}$, to label the vibrational (plane wave like) states. The harmonic oscillator like hamiltonian clearly suggests that a state with frequency $\omega(\mathbf{k})$ should have an energy $\hbar\omega$. This is indeed correct but we haven't quite proved it. Secondly if the "quantized" waves are labelled by $\mathbf{k}$ and ω, it is legitimate to ask what the meaning of $\hbar\mathbf{k}$ would be? Clearly it cannot be physical momentum, because we have already seen that there is some freedom is choosing the values - for the 1-d chain we could shift the wavevectors by $2\pi/a$, if $\hbar\mathbf{k}$ was physical momentum (like that of a photon or a moving particle) we could not have done it. To understand what the lattice vibration state labels really mean, we need to develop a little bit of (very useful!) formalism. Without this bit, the explanation of scattering from a crystal would require a lot of handwaving arguments...that we want to avoid.

6.3.1 The harmonic oscillator revisited

Let us start by writing a harmonic oscillator hamiltonian with the bare minimum content. We will set $m = 1$, $\omega = 1$, so that we get

$$H = \frac{\hat{p}^2}{2} + \frac{\hat{x}^2}{2} \tag{6.42}$$

Of course we know how the problem is solved quantum mechanically, but this time we want to think in a slightly different way. Eqn 6.42 reminds us of an expression like $x^2 + y^2$, which we could have *factorized*, (using complex numbers) if these were just numbers. The ˆ reminds us that they are operators, not pure numbers. But how close can we get to factorizing it? Let's see

$$\frac{(\hat{x}-i\hat{p})}{\sqrt{2}}\frac{(\hat{x}+i\hat{p})}{\sqrt{2}} = \frac{\hat{p}^2}{2}+\frac{\hat{x}^2}{2}+\frac{i}{2}(\hat{x}\hat{p}-\hat{p}\hat{x}) \tag{6.43}$$

$$= H - \frac{i}{2}(i\hbar) \tag{6.44}$$

$$\therefore H = \frac{(\hat{x}-i\hat{p})}{\sqrt{2}}\frac{(\hat{x}+i\hat{p})}{\sqrt{2}}+\frac{\hbar}{2} \tag{6.45}$$

Recall that $\hat{x}$ and $\hat{p}$ are hermitian operators, but the presence of the i in the expression means, that if we define (dropping the ˆ from now on)

$$a = x + ip \tag{6.46}$$

then

$$a^\dagger = x - ip \tag{6.47}$$

The definitions immediately give

$$[a, a^\dagger] = \hbar \tag{6.48}$$

$$[a, a^\dagger a] = \hbar a \tag{6.49}$$

$$[a^\dagger, a^\dagger a] = -\hbar a^\dagger \tag{6.50}$$

What is the significance of these operators. Well suppose (with no assumptions about the eigenvalues of a harmonic oscillator being equally spaced etc.)

$$H|\psi\rangle = \lambda|\psi\rangle \tag{6.51}$$

Consider :

$$Ha|\psi\rangle = (aH - \hbar a)|\psi\rangle \tag{6.52}$$

$$= a\lambda|\psi\rangle - \hbar a|\psi\rangle \tag{6.53}$$

$$= (\lambda - \hbar)a|\psi\rangle \tag{6.54}$$

Notice that the action of a on an eigenstate of H takes it to a state with the eigenvalue lowered by $\hbar$. But this his cannot go on indefinitely, because then the lowest eignevalue would be $-\infty$! The only way for this not to happen is for the series to terminate - thus the action of a on the ground state must produce a *null* state.

In a similar way (using the commutation relation 6.50) one can prove that the action of $a^\dagger$ on an eigenstate of H would be to take it to the next higher energy state. The relation between $\hat{x}$, $\hat{p}$ and a, $a^\dagger$ are linear and it is easy to go back and forth between these two. If you compare the algebra required to compute the matrix element of operators using Hermite polynomial type states (in the co-ordinate representation) and the raising/lowering operator method, then you will immediately see the advantages. Any operator that can be expressed as a polynomial in $\hat{x}$, $\hat{p}$ can be easily turned into an equivalent expression consisting of the ladder operators. We will see later that the operators will also be useful in a number of contexts. for obvious reasons these operators are also called creation (raising the eigenvalue) and annihilation (lowering the eigenvalue) operators.

This is a standard topic in any quantum mechanics text. Also at this point , we should note that we had simplified the hamiltonian (eqn. 6.42) a bit too much. We need to put back the factors of m, ω etc. The conventional definition also absorbs the factor of $\hbar$ in the definition of the operators, so that the commutator $[a, a^\dagger] = 1$. The usual definitions are :

$$a = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} + i\sqrt{\frac{1}{2\hbar m\omega}}\hat{p} \tag{6.55}$$

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\hat{x} - i\sqrt{\frac{1}{2\hbar m\omega}}\hat{p} \tag{6.56}$$

which implies that :

$$\begin{aligned} H &= \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) && (6.57)\\ [a, a^\dagger] &= 1 && (6.58)\\ & && (6.59)\end{aligned}$$

The correctly normalised excited states can be written as

$$|n\rangle = \frac{1}{\sqrt{n!}}\left(a^\dagger\right)^n|0\rangle \tag{6.60}$$

PROBLEM:
1. Invert the relations 6.55,6.56

2. Prove the statement made in eqn. 6.60. $|0\rangle$ denotes the ground state.

6.3.2 Using the ladder operators

There is some lengthy and messy algebra here, but it is worth knowing how it is done and how the steps work out. We first note that the 3D hamiltonian of eqn. 6.31 is already in a quadratic form. But in 3D each mode frequency has eigenvectors polarised along different directions. We denote the frequency and polarisation vector by $\omega(\mathbf{k})$ and $\boldsymbol{\epsilon}(\mathbf{k})$ Then we need to move over from spatial co-ordinate to wavevectors. This is done by defining a "fourier transform" of the operator relations as follows (N is the number of sites, $s = 1, 2, 3$ runs over the polarisation. $\mathbf{R}_i$ runs over the lattice sites). Notice that the part inside the bracket is like the operators we discussed in the subsection 6.3.1. One such operator is there for each site and each mode. The total number of such operators must equal the number of degrees of freedom we started with - which is 3N.

$$\begin{aligned} a_{\mathbf{k}s} &= \frac{1}{\sqrt{N}}\sum_{\mathbf{R}_i} e^{-i\mathbf{k}.\mathbf{R}_i}\boldsymbol{\epsilon}_s(\mathbf{k}).\left(\sqrt{\frac{m\omega_s(\mathbf{k})}{2\hbar}}\mathbf{u}(\mathbf{R}_i) + i\sqrt{\frac{1}{2\hbar m\omega_s(\mathbf{k})}}\mathbf{p}(\mathbf{R}_i)\right) && (6.61)\\ a^\dagger_{\mathbf{k}s} &= \frac{1}{\sqrt{N}}\sum_{\mathbf{R}_i} e^{i\mathbf{k}.\mathbf{R}_i}\boldsymbol{\epsilon}_s(\mathbf{k}).\left(\sqrt{\frac{m\omega_s(\mathbf{k})}{2\hbar}}\mathbf{u}(\mathbf{R}_i) - i\sqrt{\frac{1}{2\hbar m\omega_s(\mathbf{k})}}\mathbf{p}(\mathbf{R}_i)\right) && (6.62)\end{aligned}$$

Notice all the indices carefully, sometimes we omit bits of it, but you should at this point be clear about the meaning of each index.
The set of equations need to be inverted. The procedure is outlined here, but some of the details are left as an *extremely important* exercise.

- Add the expressions for $a_{\mathbf{k}s}$ and $a_{-\mathbf{k}s}$ using the eqns 6.61, 6.62.

- Now look at the polarisation index or the dot product. Multiply both sides by $\boldsymbol{\epsilon}_s(\mathbf{k})$ and sum over s. recall that these three vectors are eigenvectors of a hermitian matrix. So these are complete, orthonormal. This should take care of the $\boldsymbol{\epsilon}(\mathbf{k})$ from one side.

- Sum both sides over $\mathbf{k}$ after simplifying. This should produce a delta function because

$$\sum_{\mathbf{k}} e^{i\mathbf{k}.(\mathbf{R}-\mathbf{R}_i)} = N\delta_{\mathbf{R}\mathbf{R}_i} \tag{6.63}$$

You should get the results

$$\mathbf{u}(\mathbf{R}_i) = \frac{1}{\sqrt{N}}\sum_{\mathbf{k}s}\sqrt{\frac{\hbar}{2m\omega_s(\mathbf{k})}}\left(a_{\mathbf{k}s}+a^{\dagger}_{-\mathbf{k}s}\right)\boldsymbol{\epsilon}_s(\mathbf{k})e^{i\mathbf{k}.\mathbf{R}_i} \tag{6.64}$$

$$\mathbf{p}(\mathbf{R}_i) = \frac{-i}{\sqrt{N}}\sum_{\mathbf{k}s}\sqrt{\frac{\hbar m\omega_s(\mathbf{k})}{2}}\left(a_{\mathbf{k}s}-a^{\dagger}_{-\mathbf{k}s}\right)\boldsymbol{\epsilon}_s(\mathbf{k})e^{i\mathbf{k}.\mathbf{R}_i} \tag{6.65}$$

PROBLEM: Deduce eqns. 6.64,6.65 from 6.61,6.62, following the steps outlined in the preceding paragraph.

Now we need to be able to simplify the 3D small oscillation hamiltonian (eqn. 6.31 using the operators, the final result is very elegant and useful, but again there are some non-trivial steps needed to get there.

We need to evaluate the following starting with eqns 6.64,6.65

$$\begin{aligned}\sum_{\mathbf{R}_i}\frac{\mathbf{p}(\mathbf{R}_i).\mathbf{p}(\mathbf{R}_i)}{2m} &= -\frac{1}{N}\sum_{\mathbf{R}_i}\sum_{\mathbf{k}s}\sum_{\mathbf{k}'s'}\sqrt{\frac{\hbar m\omega_s(\mathbf{k})}{2}}\sqrt{\frac{\hbar m\omega_{s'}(\mathbf{k}')}{2}}\left(a_{\mathbf{k}s}-a^{\dagger}_{-\mathbf{k}s}\right)\left(a_{\mathbf{k}'s'}-a^{\dagger}_{-\mathbf{k}'s'}\right)\\ &\quad\times\boldsymbol{\epsilon}_s(\mathbf{k}).\boldsymbol{\epsilon}_{s'}(\mathbf{k}')e^{i(\mathbf{k}+\mathbf{k}').\mathbf{R}_i}\\ &= \frac{1}{4}\sum_{\mathbf{k}s}\hbar\omega_s(\mathbf{k})\left(a_{\mathbf{k}s}-a^{\dagger}_{-\mathbf{k}s}\right)\left(a_{\mathbf{k}'s'}-a^{\dagger}_{-\mathbf{k}'s'}\right)\end{aligned} \tag{6.66}$$

How does the complicated looking first step simplify to eqn. 6.66?

- We first do the sum over $\mathbf{R}_i$ and get a delta function $\delta_{\mathbf{k}+\mathbf{k}'}$ This delta function "clicks" N times and gets rid of the $1/N$ factor at the beginning.
- This still leaves us with the sum over polarisation indices ss'. This is easy because

  $$\boldsymbol{\epsilon}_s(\mathbf{k}).\boldsymbol{\epsilon}_{s'}(-\mathbf{k}) = \delta_{ss'}$$

  We also need to use the fact that $\omega_s(\mathbf{k}) = \omega_s(-\mathbf{k})$
- This leaves us with a sum on only one set of indices.

Now comes the other sum, using eqn. 6.31 as our starting point:

$$\begin{aligned}\frac{1}{2}\sum_{\mathbf{R}_i\mathbf{R}_j}\sum_{\mu\nu}D_{\mu\nu}(\mathbf{R}_i,\mathbf{R}_j)x_{i,\mu}x_{j,\nu} &= \frac{1}{N}\sum_{\mathbf{R}_i,\mathbf{R}_j}\sum_{\mu\nu}\sum_{\mathbf{k}s}\sum_{\mathbf{k}'s'}\sqrt{\frac{\hbar}{2m\omega_s(\mathbf{k})}}\sqrt{\frac{\hbar}{2m\omega_{s'}(\mathbf{k}')}}\\ &\quad\times\left(a_{\mathbf{k}s}+a^{\dagger}_{-\mathbf{k}s}\right)\left(a_{\mathbf{k}'s'}+a^{\dagger}_{-\mathbf{k}'s'}\right)\times\epsilon_{s\mu}(\mathbf{k})\epsilon_{s'\nu}(\mathbf{k}')e^{i\mathbf{k}.\mathbf{R}_i}e^{i\mathbf{k}'.\mathbf{R}_j}\\ &= \frac{1}{4}\sum_{\mathbf{k}s}\hbar\omega_s(\mathbf{k})\left(a_{\mathbf{k}s}+a^{\dagger}_{-\mathbf{k}s}\right)\left(a_{-\mathbf{k}s}+a^{\dagger}_{\mathbf{k}s}\right)\end{aligned} \tag{6.67}$$

How does the algebra work out to get to eqn 6.67 from the first step?

- First do the sum on $\mathbf{R}_i$ (could have done the j sum as well). $D_{\mu\nu}(\mathbf{R}_i-\mathbf{R}_j)$ can be fourier transformed by introducing an extra factor of $e^{-i\mathbf{k}.\mathbf{R}_j}$ and changing the term $e^{i\mathbf{k}'.\mathbf{R}_j}$ to $e^{i\mathbf{k}.\mathbf{R}_j}e^{i\mathbf{k}'.\mathbf{R}_j}$.
- The fourier transform should yield $D_{\mu\nu}(\mathbf{k})$ as in eqn. 6.39.

- But the $\epsilon_{s\mu}(\mathbf{k})$ are components of the eigenvector of $D_{\mu\nu}(\mathbf{k})$ and the eigenvalues are $\omega_s^2(\mathbf{k})$. So we can write :

$$D_{\nu\mu}(\mathbf{k})\epsilon_{s\mu} \quad = \quad -\omega_s^2(\mathbf{k})\epsilon_{s\nu}$$

- Then we do the sum on $\mathbf{R}_j$ and get a delta function as before, allowing us to reduce the expression to a single $\mathbf{k}$ sum.
- The sum on the polarisation indices follow.
- The $\omega_s^2(\mathbf{k})$ in the numerator cancels one factor of $\omega_s(\mathbf{k})$ in the denominator and the result follows...

The final result now (very simply!) follows by adding eqns. 6.66 and 6.67.

$$H = \sum_{\mathbf{k}s} \hbar\omega_s(\mathbf{k}) \left(a^\dagger_{\mathbf{k}s} a_{\mathbf{k}s} + \frac{1}{2} \right) \tag{6.68}$$

Notice that the entire system behaves like $3N$ decoupled harmonic oscillators as we might have expected. This now allows us to bring in a bit of thermodynamics into it and calculate the energy content of the system and its specific heat etc.

6.4 Internal energy and specific heat of the lattice

The mean energy of a system that has allowed energy states E_i at a certain temperature is given (thermodynamically) by :

$$\begin{aligned} \langle E \rangle &= \frac{\sum_i E_i e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} \\ &= -\frac{\partial \ln Z}{\partial \beta} \end{aligned} \tag{6.69}$$

where we have used the standard expression $Z = \sum_i e^{-\beta E_i}$ for the partition function and $\beta = 1/kT$.

Since the allowed energy states for an oscillator with frequency ω are equally spaced in energy and the resulting geometric series for the partition function can be easily summed up.

$$\begin{aligned} E_i &= \left(i + \frac{1}{2} \right) \hbar\omega \\ \therefore\ Z &= \sum_i e^{-\beta\left(i+\frac{1}{2}\right)\hbar\omega} \\ &= \frac{e^{-\frac{\beta\hbar\omega}{2}}}{1 - e^{-\beta\hbar\omega}} \qquad (6.70) \\ \therefore\ -\frac{\partial \ln Z}{\partial \beta} &= \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \qquad (6.71) \end{aligned}$$

If the number of modes per unit volume in an interval ω to $\omega + d\omega$ is $D(\omega)d\omega$, the total (per unit volume) internal energy is obtained by summing over all of them. Then we obtain the specific heat (per unit volume at constant volume, c_v) by differentiating the total energy w.r.t temperature.

$$\begin{aligned} U/V &= \int d\omega D(\omega)\hbar\omega \left(\frac{1}{2} + \frac{1}{e^{\beta\hbar\omega} - 1} \right) \qquad (6.72) \\ \therefore c_v &= \int d\omega D(\omega)\hbar\omega \frac{\partial}{\partial T} \frac{1}{e^{\beta\hbar\omega} - 1} \qquad (6.73) \end{aligned}$$

For our 3D system we must have

$$D(\omega)d\omega = \sum_s \frac{d^3\mathbf{k}}{(2\pi)^3} \tag{6.74}$$

where s runs over the three polarisation indices. We assume that the solid is isotropic and $\omega_s(\mathbf{k}) = v|\mathbf{k}|$ where v is the sound velocity. The second assumption is justified at low temperatures, when only low energy acoustic modes are occupied. The total number of modes must be 3N/V, so we must have a cut-off for the integration such that for the acoustic (sound wave like) modes Eqn. 6.74 leads to

$$D(\omega) = \frac{3\omega^2}{2\pi^2 v^3} \tag{6.75}$$

upto a cutoff frequency till 3N modes are accounted for:

$$\int_0^{\omega_D} d\omega D(\omega) = \frac{3N}{V} \tag{6.76}$$

$$\therefore\ {\omega_D}^3 = 6\pi^2 v^3 \frac{N}{V} \tag{6.77}$$

This cutoff frequency is called the Debye frequency and if $\hbar\omega_D$ is expressed in temperature units, then the corresponding temperature is called the Debye temperature Θ_D. The integral in eqn. 6.73 can now be evaluated with a simple substitution :

$$\hbar\omega_D = k_B \Theta_D \tag{6.78}$$

$$x = \beta\hbar\omega \tag{6.79}$$

$$c_v = 9\frac{N}{V} k_B \left(\frac{T}{\Theta_D}\right)^3 \int_0^{\Theta_D/T} dx \frac{x^4 e^x}{(e^x - 1)^2} \tag{6.80}$$

6.4.1 Low and high temperature limits of the Debye formula

Low temperature limit

At low temperatures $\Theta_D/T \to \infty$, it is left as an exercise to show :

PROBLEM: Show that for $T << \Theta_D$

$$\int_0^{\Theta_D/T} dx \frac{x^4 e^x}{(e^x - 1)^2} = \frac{4\pi^4}{15} \tag{6.81}$$

The lattice specific heat should then vary as $\sim T^3$ at low temperatures ($T << \Theta_D$). This is well verified experimentally. Although we assumed a monatomic lattice, the result is far more general. The reason is that at low temperatures the other (optical) modes do not contribute to the specific heat as they are much higher in energy. So except for the value of the coefficient of T^3, our conclusions remain correct, because we have captured the essential part - the sound wave like modes.

High temperature limit : Dulong Petit value

If all the oscillators behaved classically, then they would have a mean thermal energy of $kT/2$ per degree of freedom. An oscillator in 3D has three kinetic and 3 potential (energy) degrees of freedom and hence should have an energy density of $3nkT$ and hence a specific heat of $3nk$, notice the two regimes shown in Fig. 6.6.

The Einstein formula

To be written...when is it useful etc.

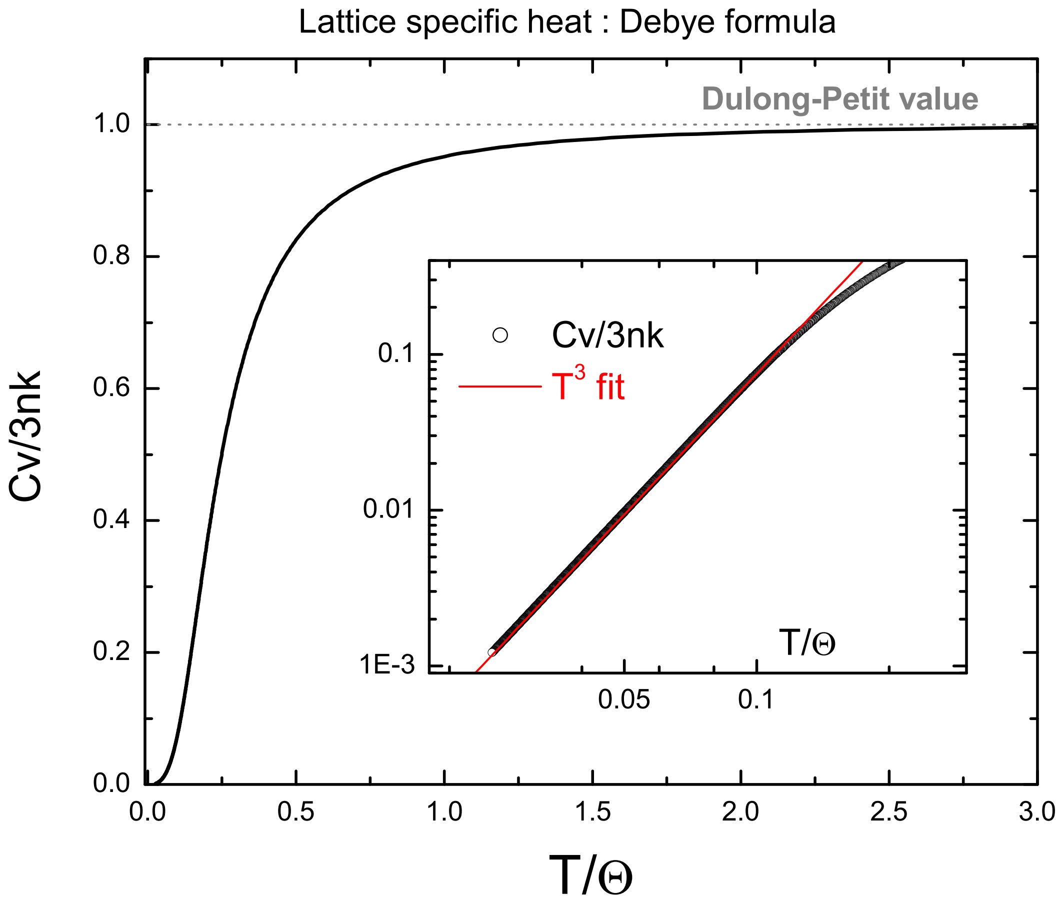


Figure 6.6: The variation of the $c_v(T)$ as predicted by the Debye formula, eqn 6.80. Notice the inset, where the low temperature part is plotted in a log-log scale to bring out the power law nature. The T^3 beahviour extends till about $\sim 0.1\Theta_D$. At high temperatures the formula correctly approaches the Dulon-Petit value.

6.5 Electrodynamics in the polar crystal with no free electrons

When we treated the metal as a free electron gas we calculated the electrodynamic and thermodynamic properties of this gas. We will do a bit of the same.

6.5.1 Forced vibrations of a polar lattice due to an incident electromagnetic wave

we calculated the phonon modes of the lattice before - these were free vibrations of the lattice, now if an electromagnetic wave is present, we need to treat it as follows. Let us consider that atom of type A has a positive charge and atom of type B has a negative charge (salts like NaCl). So we have the equations of motions for a mass-spring system with an additional electric field. The electric field is likely to come from electromagnetic radiation, so we neglect its spatial variation (*i.e.* $q = 0$), as typical interatomic spacing (1 nm) is much smaller than the wavelength of visible light (500 nm). The equations of motion are :

$$\begin{aligned} m_A \ddot{u_{A,i}} &= -K(u_{A,i} - u_{B,i}) - K(u_{A,i} - u_{B,i-1}) + eE \\ m_A \ddot{u_{B,i}} &= -K(u_{B,i} - u_{A,i+1}) - K(u_{B,i} - u_{A,i}) - eE \end{aligned} \tag{6.82}$$

The driving field is $E = \tilde{E}_0 e^{-i\omega t}$ We neglect spatial variation and hence the solutions of this forced vibration problem are expected to be of the form:

$$\begin{aligned} u_A &= \tilde{u}_{A0} e^{-i\omega t} \\ u_B &= \tilde{u}_{B0} e^{-i\omega t} \end{aligned} \tag{6.83}$$

This leads to the set of equations

$$\begin{pmatrix} 2K - m_A\omega^2 & -2K \\ -2K & 2K - m_B\omega^2 \end{pmatrix} \begin{pmatrix} \tilde{u}_{A0} \\ \tilde{u}_{B0} \end{pmatrix} = \begin{pmatrix} e\tilde{E}_0 \\ -e\tilde{E}_0 \end{pmatrix} \tag{6.84}$$

The solution (the algebra is left as problem) bears some resemblance to the plasma oscillation problem for free electrons.

$$\begin{aligned} \tilde{u}_{A0} &= \frac{e\tilde{E}_0}{m_A\left[2K\left(\frac{1}{m_A} + \frac{1}{m_B}\right) - \omega^2\right]} = \frac{e\tilde{E}_0}{m_A\left[\omega_T^2 - \omega^2\right]} \\ \tilde{u}_{B0} &= \frac{-e\tilde{E}_0}{m_B\left[2K\left(\frac{1}{m_A} + \frac{1}{m_B}\right) - \omega^2\right]} = \frac{-e\tilde{E}_0}{m_B\left[\omega_T^2 - \omega^2\right]} \end{aligned} \tag{6.85}$$

ω_T is the optical phonon frequency that we have calculated before. The subscript T means that it is a transverse mode - something that we will justify a little later.

Polarisability of the lattice

Our target is to convert this expression into an expression for polarisability. The dipole moment per unit volume (polarisation) of the lattice must be ;

$$P = Ne(\tilde{u}_{A0} - \tilde{u}_{B0}) = \frac{2Ne}{\omega_T^2 \mu} \frac{1}{1 - \omega^2/\omega_T^2} \tilde{E}_0 \tag{6.86}$$

where we have used the standard expression for the reduced mass (μ). The total polarisability would be due to the movement of the atoms and the electronic polrisability of the core-states of the atoms. (There are no free electrons here.) Using

$$D = \epsilon(\omega)E = \epsilon_0 E + P \tag{6.87}$$

We get

$$\epsilon(\omega) = 1 + \frac{P_{electron}}{\epsilon_0 E} + \frac{2Ne}{\omega_T^2 \mu} \frac{1}{1 - \omega^2/\omega_T^2} \tag{6.88}$$

Clearly this diverges at $\omega = \omega_T$. We can hide the frequency dependence of the second term by defining the dielectric constants at very low and very high frequencies. Clearly

$$\epsilon(0) = 1 + \frac{P_{electron}}{\epsilon_0 E} + \frac{2Ne}{\omega_T^2 \mu} \tag{6.89}$$

$$\epsilon(\infty) = 1 + \frac{P_{electron}}{\epsilon_0 E} \tag{6.90}$$

Hence eqn. 6.88 can be written in the more convenient form:

$$\epsilon(\omega) = \epsilon(\infty) + \frac{\epsilon(0) - \epsilon(\infty)}{1 - \omega^2/\omega_T^2} \tag{6.91}$$

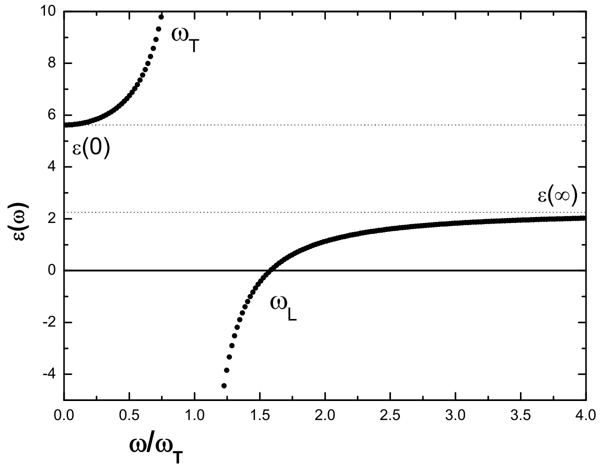


Figure 6.7: Plot of $\epsilon(\omega) = \epsilon(\infty) + \frac{\epsilon(0)-\epsilon(\infty)}{1-\omega^2/\omega_T^2}$. Notice the divergence at ω_T and the zero crossing at ω_L. The two points are related very simply.

The divergence happens at the frequency of the transverse optical mode, the zero crossing is the longitudinal optical mode. The logic that the zero crossing happens at the frequency of longitudinal oscillation is as follows.

- In the insulating medium $\nabla.\mathbf{D} = 0$, because there are no free charges like free electrons. So we must have $\nabla.\mathbf{E} = -\nabla.\mathbf{P}/\epsilon_0$.
- For a transverse mode $\mathbf{P} = \tilde{\boldsymbol{P}} e^{i(\boldsymbol{k}.\boldsymbol{r}-\omega t)}$, we must have $\boldsymbol{k}.\tilde{\boldsymbol{P}} = 0$ in this case. But for a longitudinal mode $\nabla.\boldsymbol{P} \neq 0$, hence $\nabla.\boldsymbol{E} \neq 0$. So to ensure $\nabla.\boldsymbol{D} = 0$, we need $\epsilon(\omega) = 0$.

Now it is easy to see how the divergence and zero-crossing are related

$$\omega_L^2 = \frac{\epsilon(0)}{\epsilon(\infty)} \omega_T^2 \tag{6.92}$$

This relation is called the Lydan-Sachs-Teller (LST) relation. The generic shape of the dielectric response curve is is seen in many polar crystals like alkali halides. If the crystal is non-polar the frequencies merge. Between ω_T and ω_L, we see that $\epsilon(\omega) < 0$. This means that even though there are no free electrons, the electromagnetic wave is attenuated in the medium due to lattice vibrations. The reflectivity of the crystal would be almost 1, in this region.

Light and lattice vibrations together

We just saw that that passage of light of a certain frequencies, polarises the lattice. That means such light waves must transfer some energy to the lattice as well and can get some energy from the lattice vibrations. In any case it means that the $\omega(k)$ relation of the light wave must change. *The lattice vibrations and light waves must mix in some way.* It can be anticipated that the mixing will be strongest in the region where the two $\omega(k)$ relations actually come very close to each other. This indeed happens - see the Fig. 6.8. See how the red and blue lines - the acctual dispersion, deviates from the dotted lines, the "non-interacting" dispersion.

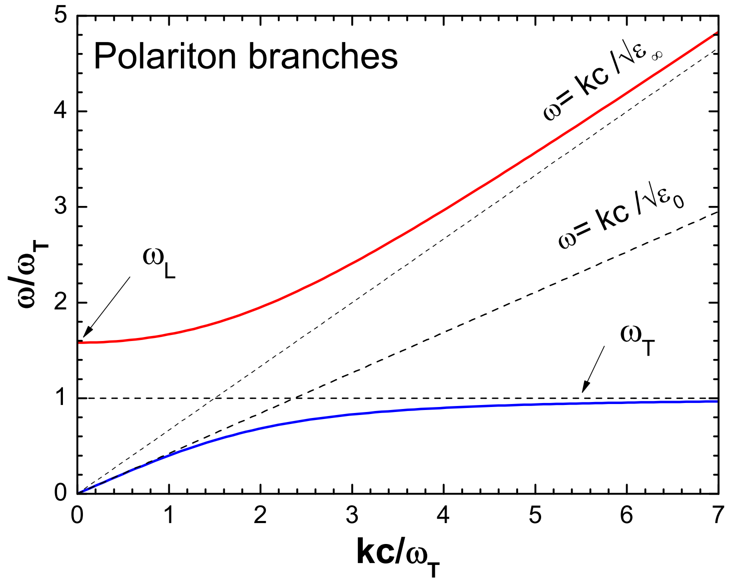


Figure 6.8: Plot of the upper and lower branches of the phonon-photon mixed modes. The two independent modes would have crossed at a certain point if there were no energy exchange possible between the two modes. The maximum distortion happens near the point where they would have crossed. This kind of a situation is often encountered in solid state physics- where two modes mix and then repel each other creating an "anti-crossing".

Since we have already got the relation eqn. 6.90 we can use it to write

$$\begin{aligned} \omega^2 &= k^2\frac{c^2}{n^2} \\ &= \frac{k^2c^2}{\epsilon(\omega)} \\ \therefore kc &= \omega\sqrt{\epsilon_\infty + \frac{\epsilon_0-\epsilon_\infty}{1-(\omega/\omega_T)^2}} \end{aligned} \tag{6.93}$$

For every k we can then have two solutions for ω, these solutions are called the *polariton modes* of the system. They are plotted in Fig. See how the photon like and phonon like modes (the straight lines have got mixed up).

PROBLEM: Invert the eqn. 6.93 and show that

$$\frac{\omega(k)}{\omega_T} = \sqrt{\frac{(\epsilon_0-q^2)\pm\sqrt{(\epsilon_0-q^2)^2-4\epsilon_\infty q^2}}{2\epsilon_\infty}} \tag{6.94}$$

where $q = kc/\omega_T$. Then calculate the small k and large k limits of the equation. Verify that the asymptotes drawn in the Fig. 6.8 are correct.

6.6 Electrons and phonons together: How an electron "sees" a phonon?

Any metal/conducting material would have mobile electrons in them. How do phonons interact with electrons and effect electric current? The problem is formulated in the following way:

- An electron is in an eignestate of the *static* lattice.
- But as the lattice vibrates the lattice potential deviates a little bit from what it was when all the ions were sitting still in their respective equilibrium positions.
- We calculate the change in potential due to a small change of the lattice ions from their equilibrium position.
- This extra part in the potential is the perturbation.
- We can use perturbation theory to calculate the transition rate from an eigenstate, with a certain crystal momentum, to all other eigenstates of the static lattice.
- This transition rate is now equated to the relaxation rate. This relaxation rate is used in the simple formula for resistivity ($1/\rho = Ne^2\tau/m$), to get an expression for the measured resistivity, via the Boltzmann transport equation.
- Some parts of this argument require very careful thinking. While the problem can be formulated in very general terms, obtaining any closed form result requires several approximations that have their regimes of validity. These must be carefully understood!
- The treatment of the problem that we give here, is due to Bloch and Mott.

The electron-phonon perturbation term

The presence of lattice vibrations means that instantaneous positions of the lattice sites are slightly different from the static (equilibrium) positions. This implies that there would be a slight change in the potential for which the Bloch states are eigenstates. However the potential seen by the electrons would not be just this extra (deviation) potential. This potential would be screened by the presence of conduction electrons. While the whole problem can be solved in this way, the entire calculation is quite involved. We will show the form this interaction potential takes and point out an important conservation law resulting from the form of this potential.
The total potential due to the ions in the lattice (assume that this is monatomic) is the sum over all ionic potentials originating from the lattice points $\boldsymbol{R}_n$

$$V(\boldsymbol{r}) = \sum_n v(\boldsymbol{r} - \boldsymbol{R}_n) \tag{6.95}$$

The presence of the lattice vibration implies that these ions are disturbed from their equilibrium positions by an amount $\delta\boldsymbol{x}_n$, so the perturbed potential is:

$$\begin{aligned} V(\boldsymbol{r}) &= \sum_n v(\boldsymbol{r} - \boldsymbol{R}_n - \delta\boldsymbol{x}_n) \\ \therefore \delta V(\boldsymbol{r}) &= \sum_n \delta\boldsymbol{x}_n . \nabla_{\boldsymbol{R}_n} v(\boldsymbol{r} - \boldsymbol{R}_n) \\ \therefore \delta V(\boldsymbol{r}) &= -\sum_n \delta\boldsymbol{x}_n . \nabla_{\boldsymbol{r}} v(\boldsymbol{r} - \boldsymbol{R}_n) \end{aligned} \tag{6.96}$$

For a phonon mode at wavevector $\boldsymbol{q}$ and a certain polarisation $\boldsymbol{\varepsilon_q}$

$$\delta\boldsymbol{x}_n = A_{\boldsymbol{q}}\boldsymbol{\varepsilon_q} e^{i\boldsymbol{q}.\boldsymbol{R_n}} + \text{c.c} \tag{6.97}$$

Hence we can write

$$\delta V(\boldsymbol{r}) = -A_{\boldsymbol{q}}\boldsymbol{\varepsilon_q}.\nabla_{\boldsymbol{r}} \sum_n \left[e^{i\boldsymbol{q}.\boldsymbol{R_n}} v(\boldsymbol{r} - \boldsymbol{R}_n)\right] \tag{6.98}$$

The part within the box brackets looks exactly like a tight binding function and hence should have an expansion of the form

$$\sum_n \left[e^{i\boldsymbol{q}.\boldsymbol{R_n}} v(\boldsymbol{r}-\boldsymbol{R}_n)\right] = \sum_m c_{\boldsymbol{G}_m} e^{i(\boldsymbol{q}+\boldsymbol{G}_m).\boldsymbol{r}} \tag{6.99}$$

where the reciprocal lattice vectors $\boldsymbol{G}_m$ satisfy $e^{i\boldsymbol{R}_n.\boldsymbol{G}_m} = 1$ for all n, m. So we can rewrite eqn 6.98 and sum over all possible $\boldsymbol{q}$ as:

$$\delta V(\boldsymbol{r}) = -i \sum_{\boldsymbol{q}} A_{\boldsymbol{q}} \boldsymbol{\varepsilon}_{\boldsymbol{q}}. \sum_m (\boldsymbol{q}+\boldsymbol{G}_m) c_{\boldsymbol{G}_m} e^{i(\boldsymbol{q}+\boldsymbol{G}_m).\boldsymbol{r}} \tag{6.100}$$

This potential must be screened by the dielectric function of the free electron gas, but we neglect this, since we are not going to calculate something quantitatively here. What kind of scattering between free electron states ($e^{i\boldsymbol{k}_i.\boldsymbol{r}}$) can this potential cause? We need to look at the matrix element

$$M(\boldsymbol{k}_i, \boldsymbol{k}_f) = \langle \boldsymbol{k}_f | \delta V(\boldsymbol{r}) | \boldsymbol{k}_i \rangle \tag{6.101}$$

It is clear that the $M(\boldsymbol{k}_i, \boldsymbol{k}_f)$ can be non zero, only if

$$\boldsymbol{k}_f - \boldsymbol{k}_i = \boldsymbol{q} \tag{6.102}$$

$$\boldsymbol{k}_f - \boldsymbol{k}_i = \boldsymbol{q} + \boldsymbol{G}_m \tag{6.103}$$

Even though phonons carry no physical momentum, the conservation law looks very much like momentum conservation. The change in *crystal momentum* of the electron has been taken up by the crystal momentum phonons. Notice also eqn. 6.103. It is possible to have a situation where the change in *crystal momentum* of the electron takes it to another Brillouin zone, due to the extra translation provided by $\boldsymbol{G}_m$. Such processes have been historically called *umklapp* processes.

PROBLEM: *This problem is meant to illustrate the role/necessity of umklapp processes for large angle scattering of electrons.* Imagine a monoatomic, monovalent solid that freezes in an isotropic cubic lattice. Denote the density of electrons (and also the atoms) by $n = N/V$. First show that the Fermi wavevector (k_F) and the maximum wavevector of the phonons within the Debye model (k_D) are simply related. Then show that if the deflection of an electron by larger than $\sim 79^o$ occurs then the process needs to be an umklapp scattering. The number ($\sim 79^o$) is specific to the Debye model.

Relating the matrix element to resistivity

We need to relate transition probability integrated over all possible final states to calculate the relaxation (decay) rate of a certain state. Recall the standard perturbative treatment of a constant perturbation (switched on at time $t = 0$). This means that the electron would be able to make transitions to states of equal energy only.

Consider the hamiltonian

$$i\hbar \frac{d}{dt}|\Psi\rangle = H_0|\Psi\rangle + V|\Psi\rangle \tag{6.104}$$

We assume that we know the eigenfunctions of H_0 and they are normalized over a volume Ω such that

$$H_0|u_{\boldsymbol{k}}\rangle = E_{\boldsymbol{k}}|u_{\boldsymbol{k}}\rangle \tag{6.105}$$

We need to trace the evolution of Ψ over time such that

$$|\Psi(t)\rangle = \frac{1}{\sqrt{\Omega}} \sum_{\boldsymbol{k}} a_{\boldsymbol{k}}(t)|u_{\boldsymbol{k}}\rangle e^{-iE_{\boldsymbol{k}}t/\hbar} \tag{6.106}$$

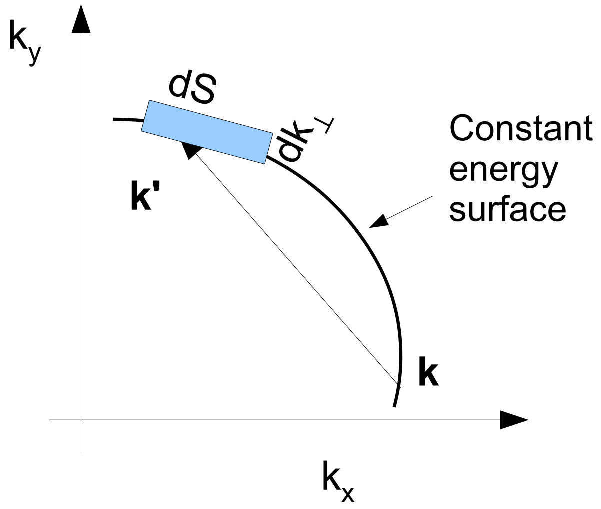


Figure 6.9: The transitions over which we need to integrate to get the total transition rate out of the state $|u_{\boldsymbol{k}}\rangle$. Note that dS is a surface area and $dk_\perp$ has the same direction as the gradient $\nabla_{\boldsymbol{k}} E(\boldsymbol{k})$, hence the group velocity.

The initial conditions are

$$\begin{aligned} a_{\boldsymbol{k}}(t=0) &= 1 \\ a_{\boldsymbol{k'}}(0) &= 0 \ (\boldsymbol{k} \neq \boldsymbol{k'}) \end{aligned} \tag{6.107}$$

Substitute eqn. 6.106 in eqn. 6.104 and then left multiply by $\langle u_{\boldsymbol{k'}}|e^{iE_{\boldsymbol{k'}}t}$, and integrate over Ω, neglect terms which are second order like $Va_{\boldsymbol{k'}}$, to get the time evolution of the coefficients

$$i\hbar\frac{d}{dt}a_{\boldsymbol{k'}} = \frac{1}{\Omega}\langle u_{\boldsymbol{k'}}|V|u_{\boldsymbol{k}}\rangle e^{i(E_{\boldsymbol{k'}}-E_{\boldsymbol{k}})t/\hbar} \tag{6.108}$$

This gives

$$|a_{\boldsymbol{k'}}|^2 = \frac{1}{\hbar^2\Omega^2}\,|\langle u_{\boldsymbol{k'}}|V|u_{\boldsymbol{k}}\rangle|^2\,\frac{2(1-\cos xt)}{x^2} \tag{6.109}$$

where $x = (E_{\boldsymbol{k'}} - E_{\boldsymbol{k}})/\hbar$.
The number of state in the volume $dSdk_\perp$ is given by the usual density of states formula for $\boldsymbol{k}$ space. We can then write for the transition probability per unit time

$$P(\boldsymbol{k}\to\boldsymbol{k'})dS \;=\; \frac{dS}{8\pi^3\hbar^2\Omega}\frac{\partial}{\partial t}\int|\langle u_{\boldsymbol{k'}}|V|u_{\boldsymbol{k}}\rangle|^2\frac{2(1-\cos xt)}{x^2}dk_\perp \tag{6.110}$$

For points close to the surface we can write

$$dx = \frac{\delta(E_{\boldsymbol{k'}}-E_{\boldsymbol{k}})}{\hbar} = \frac{1}{\hbar}|\nabla_{\boldsymbol{k}}E(\boldsymbol{k})|dk_\perp = |\boldsymbol{v}_g|dk_\perp \tag{6.111}$$

We then note that the integral 6.110 has a strong maxima at $x = 0$ and get (see standard quantum mechanics books for the details). We need to use the fact that

$$\frac{\sin^2\frac{xt}{2}}{x^2} = \frac{\pi t}{2}\delta(x) \qquad \text{as t} \to \infty \tag{6.112}$$

to complete the integration in equation 6.110. We then get:

$$P(\boldsymbol{k}\to\boldsymbol{k'})dS' = \frac{1}{4\pi^2\hbar^2\Omega}|\langle u_{\boldsymbol{k'}}|V|u_{\boldsymbol{k}}\rangle|^2\frac{dS'}{|\boldsymbol{v}_g|} \tag{6.113}$$

We state without detailed proof that the expression can be used (assuming spherical symmetry and the simplifying assumption that the transition rate depends only on $|\boldsymbol{k}-\boldsymbol{k'}|$) to write the relaxation time as:

$$\frac{1}{\tau} = \frac{1}{\Omega h^2|\boldsymbol{v}_g|}\int|\langle u_{\boldsymbol{k'}}|V|u_{\boldsymbol{k}}\rangle|^2(1-\cos\theta)dS' \tag{6.114}$$

where θ is the angle of deviation in the scattering $\boldsymbol{k} \to \boldsymbol{k'}$ The purpose of writing it in this form is to show that small angle scattering is not very effective in contributing to resistivity due to the emergence of the $1 - \cos\theta$ factor.

6.6.1 Using the Boltzmann Transport equation

The BTE tells us how the distribution function would change as a result of applied electric fields and internal scattering processes. In a steady state the sum of these two must balance each other.

$$\left.\frac{\delta f}{\delta t}\right|_{\text{collision}} = \frac{d\boldsymbol{r}}{dt}.\nabla_r f + \frac{d\boldsymbol{k}}{dt}.\nabla_{\boldsymbol{k}} f + \frac{\partial f}{\partial t} \tag{6.115}$$

Using the semiclassical equation of motion for the $\boldsymbol{k}$ vector in an electric field $\boldsymbol{F}$, for a spatially uniform, steady state distribution we have:

$$\left.\frac{\delta f}{\delta t}\right|_{\text{collision}} = -\frac{|e|\boldsymbol{F}}{\hbar}.\nabla_{\boldsymbol{k}} f \tag{6.116}$$

The simplest approach to equilibrium (with characteristic relaxation time τ) for a distribution (f) slightly pushed out of equilibrium (f^0) is:

$$\left.\frac{\delta f}{\delta t}\right|_{\text{collision}} = -\frac{f - f^0}{\tau} \tag{6.117}$$

Let $\boldsymbol{F} = (F, 0, 0)$, then equations 6.116,6.117 can be combined to give

$$\begin{aligned} f(\boldsymbol{k}) &= f^0\left(k_x + \frac{|e|F\tau}{\hbar}, k_y, k_z\right) \\ &= f^0(k_x, k_y, k_z) + \frac{\partial f^0}{\partial k_x}\frac{|e|F\tau}{\hbar} && (6.118) \\ &= f^0(\boldsymbol{k}) + g(\boldsymbol{k}) && (6.119) \end{aligned}$$

The last equation also serves as the definition of $g(\boldsymbol{k})$, the deviation from equilibrium. If the only variation is along x then the partial derivative can be replaced by a total derivative.By using the chain rule for derivatives we can write the derivative $\frac{df^0}{dk_x}$ as (this is a useful identity, that we will use)

$$\begin{aligned} \frac{df}{dk_x} &= \frac{df}{d|k|}\frac{d|k|}{dk_x} && (6.120) \\ &= \frac{df}{d|k|}\frac{k_x}{|k|} && (6.121) \end{aligned}$$

Change due to collisions in terms of $P(\boldsymbol{kk'})$

A particle entering the volume element around $\boldsymbol{k}$ comes from some $\boldsymbol{k'}$, provided there was an initial filled state at $\boldsymbol{k'}$ and an empty state at $\boldsymbol{k}$. This is because electrons are fermions.

$$N_{in} = 2\frac{d^3\boldsymbol{k}}{(2\pi)^3}\int dS'\ f(\boldsymbol{k'})(1 - f(\boldsymbol{k}))P(\boldsymbol{k'} \to \boldsymbol{k}) \tag{6.122}$$

Similarly the number of electrons leaving the volume element at $\boldsymbol{k}$ can go to any $\boldsymbol{k'}$ where it finds an empty state. So,

$$N_{out} = 2\frac{d^3\boldsymbol{k}}{(2\pi)^3}\int dS'\ f(\boldsymbol{k})(1 - f(\boldsymbol{k'}))P(\boldsymbol{k} \to \boldsymbol{k'}) \tag{6.123}$$

Since $P(\boldsymbol{k} \to \boldsymbol{k'}) = P(\boldsymbol{k'} \to \boldsymbol{k})$, we get

$$\begin{aligned}
\frac{df}{dt} &= N_{in} - N_{out} \\
&= \int dS' \ \left[f(\boldsymbol{k'}) - f(\boldsymbol{k})\right] P(\boldsymbol{kk'}) \\
&= \int dS' \ \left[g(\boldsymbol{k'}) - g(\boldsymbol{k})\right] P(\boldsymbol{kk'}) \\
&= \int dS' \ \frac{|e|F\tau}{\hbar}\frac{df}{d|k|}\left(\frac{k'_x - k_x}{|k|}\right) P(\boldsymbol{kk'}) \qquad (6.124) \\
&= -\frac{|e|F}{\hbar}\frac{df}{d|k|}\frac{k_x}{|k|} \qquad (6.125)
\end{aligned}$$

In the last two steps we have used 6.119,6.121. Equating the last two steps we get the important relation

$$\int dS' \left(1 - \frac{k'_x}{k_x}\right) P(\boldsymbol{kk'}) = \frac{1}{\tau} \qquad (6.126)$$

Notice that the integration is over $\boldsymbol{k}'$, whereas $\boldsymbol{k}$ remains fixed. We have taken the electric field F to be along $\hat{\boldsymbol{x}}$ direction. There is no loss of generality in taking $\boldsymbol{k}$ to be along $\hat{\boldsymbol{x}}$ also.

- Let the angle between $\boldsymbol{k}'$ and $\boldsymbol{k}$ to be θ, such that $\frac{k'_x}{k_x} = \cos\theta$.
- We assume that the scattering does not depend on the azimuthal angle, so $P(\boldsymbol{kk'}) = P(\theta)$.
- $dS' = k^2 \sin\theta \ d\theta d\phi$, where ϕ is the angle that changes if we keep the projection of $\boldsymbol{k'}$ on $\boldsymbol{k}$ fixed and rotate $\boldsymbol{k'}$.
- We then get

$$\begin{aligned}
\frac{1}{\tau} &= \int \ k^2 \sin\theta \ d\theta d\phi (1 - \cos\theta) P(\theta) \\
&= 2\pi k^2 \int \ d\theta \sin\theta \ (1 - \cos\theta) P(\theta) \qquad (6.127)
\end{aligned}$$

- The relaxation time has now been related to a microscopic scattering mechanism. The value of $\frac{1}{\tau}$ can be plugged in the simple formula for resistivity $\rho = \frac{m}{ne^2\tau}$

What if there are many scattering processes?

We can also see if there are many scattering processes then the total rate can be obtained by adding the rates obtained by integrating over the individual processes with scattering probabilities $P_1(\theta)$, $P_1(\theta)$ etc. The result would be

$$\frac{1}{\tau_{total}} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + ... \qquad (6.128)$$

This result is called the Matthiesen's rule, when applied to the contribution of many scattering processes to the total resistivity.

How can we get "resistance" from an elastic scattering calculation?

Resistance of a piece of wire produces the Joule heating. Clearly there is energy loss involved in the process. Yet we used a formulation of elastic scattering throughout. Why did it work?

We could restrict ourselves (the integrations) to the fermi surface alone, because each scattering event was approximately energy conserving. We are dealing with electrons with few electron volts of energy, giving up some energy to the lattice. The quanta of lattice vibrations have energies of a few milli eV. Now each scattering (transition) can cause emission and absorption of one phonon only and hence the fractional change in the electron's energy is very small. That is why the method works. But we need to justify why emission/absorption of 100 phonons does not happen in a single event.

Let us consider a lattice site (actually a mode co-ordinate) as a simple harmonic oscillator initially in a state $|n\rangle$ and an electron in a state $|\boldsymbol{k}\rangle$. Write the expression for the perturbing potential by displacing the atom's co-ordinate by a small amount. We have two sets of co-ordinates in the problem now - one for the electron and the other for the phonon/harmonic oscillator. Show that the only allowed transitions are those in which $|n\rangle \rightarrow |n \pm 1\rangle$. If you have set up the problem correctly, this should follow from preperties of the Hermite polynomials. Since the energy of the electron+phonon system must remain constant and the phonon can only take up a tiny amount of energy in one event, it follows that the electron can only lose a tiny amount of energy in one event. Of course the sum total of all the energy lost by the electron will show up as Joule heat. This is why we are justified in considering the scattering as "elastic". Of course if they were perfectly elastic then there could be no thermalisation or appearance of any Joule heat.

The low and high temperature limits of resistivity: Bloch-Grüneisen formula

Using the result we deduced in eqn 6.127 and some other simplifying assumptions (Due to Bloch, Gruneisen, Sommerfeld and Bethe) an explicit formula for the resistivity due to electron-phonon scattering can be deduced:

$$\rho(T) = \rho(0) + \text{const.} \left(\frac{T}{\Theta_D}\right)^5 G\left(\Theta_D/T\right) \tag{6.129}$$

$$\text{where} \quad G(x) = \int_0^x dz \frac{z^5}{(e^z - 1)(1 - e^{-z})} \tag{6.130}$$

It is left as an exercise to show that at high temperature ($T > \Theta_D$), $\rho \propto T$ and at low temperature $\rho \propto T^5$. Notice also that there is a lot of similarity between how the specific heat varies with temperature and how $\rho(T)/T$ varies with temperature.

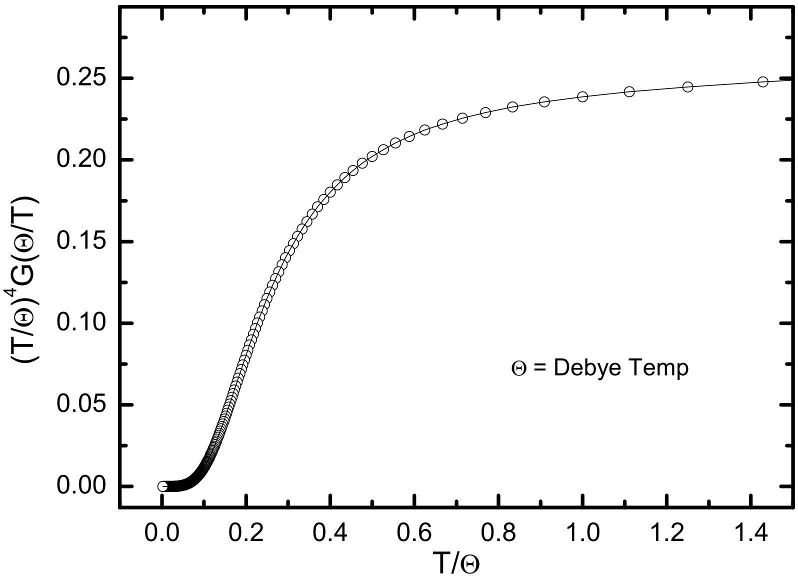


Figure 6.10: The integral in formula 6.130 plotted as a function of $\frac{T}{\Theta_D}$. Notice the similarity between the shape of the curve and the specific heat of a solid.

To evaluate the relations just derived in the last two sections is algebraically quite difficult and there are many details which need to be taken into account. for example

- The real phonon spectrum - not the Debye approximation
- The probability of both the normal and the umklapp processes.
- The fact that the fermi surface of the electrons is not always spherical.
- There are many other effects like impurity scattering, electron-electron scattering which also contribute to the resistivity of metals.
- The variation resistivity at low temperatures can also be affected by small amounts of magnetic impurities.

Two results however qualitatively stand out.

- At high temperatures $R \propto T$ for most metals. Typical values for Platinum is $0.385\Omega/\ 100\Omega$ for every Kelvin near room temperature. This can be qualitatively understood by looking at the amplitude A_q of the phonon mode at temperature T.

PROBLEM: Show using summing over all the harmonic oscillator levels that the mean energy of the oscillator, and hence its amplitude of vibration is given by

$$\langle E\rangle = M\omega_q^2 A_q^2 = \frac{\hbar\omega_q}{e^{\beta\hbar\omega_q} - 1} \tag{6.131}$$

Show that the high temperature limit of this expression is $\propto T$ and hence the modulus of the matrix element 6.101 is also $\propto T$

- At low temperatures the phonon contribution falls off as $R \propto T^5$

Chapter 7

Magnetism

We know that the thermodynamic equilibrium of a system (at constant volume and temperature) is governed by the condition that (Helmholtz) Free energy, F should be minimum. F can be a function of many variables. Magnetisation per unit volume of a sample is defined as

$$M(T,B) = -\frac{1}{V}\frac{\partial F}{\partial B} \tag{7.1}$$

Since $F = U - TS$, if the entropy contribution is small then energy and free energy are often used interchangeably. Magnetic susceptibility is then the second derivative defined as

$$\chi(T) = \frac{\partial M}{\partial B} = -\frac{1}{V}\frac{\partial^2 F}{\partial B^2} \tag{7.2}$$

Units of magnetic quantities

The conventional definition of susceptibility uses $\boldsymbol{H}$ and $\boldsymbol{M}$. Since these two quantities have the same units, $\chi = \frac{\partial M}{\partial H}$ is dimensionless. Recall that $\boldsymbol{B} = \mu_0(\boldsymbol{H} + \boldsymbol{M})$. The correct SI unit for $\boldsymbol{B}$ is Tesla, for $\boldsymbol{H}$ and $\boldsymbol{M}$ the unit is ampere/mt. For some historical reasons the published data on χ is written after dividing the (volume susceptibility) $\chi_v = \frac{\partial M}{\partial H}$ by *density.* It is called the mass susceptibility $\chi_m = \chi_v/\rho$. So it is common to see χ expressed in $\mathrm{m}^3\mathrm{kg}^{-1}$. In these units many common diamagnetic substance (*e.g.* water) has $\chi_m \sim -10^{-8}$, Oxygen (gas) is slightly paramagnetic with $\chi_m \sim 10^{-6}$. It is in general true that paramagnetic susceptibilities are somewhat larger than diamagnetic susceptibilities. We will see the reason behind this.

Analysing magnetic behaviour

We can try to understand magnetic behaviour of materials in a few steps.

- Understand what magnetic field does to individual ions sitting in a lattice, with no free electrons around.
- What magnetic field does to a "free" electron gas in 2d and 3D.
- What happens when free electrons and ions with some magnetic moment come together.

7.1 Quantum mechanical effect of the magnetic field

The Hamiltonian with a magnetic field is written using the vector potential $\boldsymbol{A}$ as:

$$H = \frac{1}{2m}\left(\hat{\mathbf{p}} + |e|\boldsymbol{A}\right)^2 + V + g\left(\frac{e}{2m}\right)\boldsymbol{S.B} \tag{7.3}$$

We will assume that the magnetic field is uniform and given by $B_0\hat{\mathbf{z}}$, and the following conditions

$$\begin{aligned} \boldsymbol{B} &= \nabla \times \boldsymbol{A} && (7.4)\\ \nabla.\boldsymbol{A} &= 0 && (7.5)\\ \boldsymbol{A} &= -\frac{1}{2}\boldsymbol{r} \times \boldsymbol{B} && (7.6) \end{aligned}$$

As far as the spin of the electron is concerned we need to recall the following:

$$\boldsymbol{\mu} = g\left(\frac{e}{2m}\right)\boldsymbol{S} \tag{7.7}$$

The spin operator for spin-$\frac{1}{2}$ particles is

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma} \tag{7.8}$$

where $\boldsymbol{\sigma}$ is a vector formed of the Pauli spin matrices.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

It is left as an exercise to show that this leads to Hamiltonian for n independent particles

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m} + V + \mu_B\left(\frac{\boldsymbol{L} + g_0\boldsymbol{S}}{\hbar}\right).\boldsymbol{B} + \frac{e^2}{8m}B^2\sum_i\left(x_i^2 + y_i^2\right) \tag{7.9}$$

where $\boldsymbol{L}$ and $\boldsymbol{S}$ are the total angular momentum and total spin operator of the system. The Bohr magneton is defined as $|e|\hbar/2m$ and has a value 9.27×10^{-24} Joule/Tesla, would often appear in many calculations Notice that the spin part plays a role in the additional (perturbing) part but is not essential.

PROBLEM: Show using standard perturbation theory, that 7.9, that the shift in the energy levels due to the presence of the magnetic field

$$\delta E_n = \mu_B\langle n|\boldsymbol{B}.\left(\frac{\boldsymbol{L} + g_0\boldsymbol{S}}{\hbar}\right)|n\rangle + \mu_B^2\sum_{n\neq n'}\frac{\left|\langle n|\boldsymbol{B}.\left(\frac{\boldsymbol{L}+g_0\boldsymbol{S}}{\hbar}\right)|n'\rangle\right|^2}{E_n - E_{n'}} + \frac{e^2}{8m}B^2\langle n|\sum_i\left(x_i^2 + y_i^2\right)|n\rangle \tag{7.10}$$

This simple result can be used for situations where the system can be represented as a collection of slightly deformed individual ions.

Order of the perturbing terms

Consider a field $B \sim 1$Tesla. The term proportional to $\boldsymbol{B}$ is of the order of $\mu_B B \approx 10^{-4}$eV. On the other hand the last term in 7.10 will be about 10^4 smaller than that at $B \sim 1$Tesla. Thus unless the the first term vanishes, it is necessarily much less important.

However consider an important case where the first term is zero. The ground state of crystals composed of ions/atoms with filled shells (inert gases and alkali halides) should have zero orbital and and spin angular momentum. The equation 7.10 says that the energy of the system should increase a little with magnetic field - this means that the response of these should be *diamagnetic*. This prediction is quantitatively correct for inert gas and alkali halide crystals. This phenomena is usually called *Larmor diamagnetism*.

When the atomic shells are partially filled the second order term arising from the $\boldsymbol{L}$ and $\boldsymbol{S}$ contributions with a negative sign will come into the picture for the ground state. Whether the (insulating) crystal will be diamagnetic or paramagnetic depends on the balance between these two competing effects. The paramagnetic term is attributed to *Van Vleck*.

7.1.1 Forces on a magnetic dipole

We can now prove an interesting result. A small paramagnet cannot float (stably) in a magnetic field, but a small diamagnet can. Let's analyse the forces on a small dipole and try to see if there can be a stable point. The energy of a dipole is

$$U = -\boldsymbol{m}.\boldsymbol{B} \tag{7.11}$$

We need to see if the energy has a stable minima. So the question is what is the sign of

$$\nabla^2(-\boldsymbol{m}.\boldsymbol{B}) = ? \tag{7.12}$$

For a dipole with a fixed moment (that is not effected by a $\boldsymbol{B}$), this is

$$\nabla^2 U = -\left[m_x \nabla^2 B_x + m_y \nabla^2 B_y + m_z \nabla^2 B_z\right] \tag{7.13}$$

But in free space, with no radiation/ time varying electric field we have

$$\begin{aligned} \nabla^2 \boldsymbol{B} &= \nabla\nabla.\boldsymbol{B} - \nabla \times \nabla \times \boldsymbol{B} \\ &= 0 \end{aligned} \tag{7.14}$$

Hence the laplacian of the individual components of the magnetic field are all zero and there is no point of stable equilibrium. Now if the dipole is itself induced by the magnetic field, then we need to look at a potential energy of the type:

$$\nabla^2 U = -\nabla^2 \chi \boldsymbol{B}.\boldsymbol{B} = -\chi \nabla^2 \left(B_x^2 + B_y^2 + B_z^2\right) \tag{7.15}$$

Now it is left as an exercise to prove the result (using the previous one) that :

$$\nabla^2 \boldsymbol{B}.\boldsymbol{B} = 2\left(|\nabla B_x|^2 + |\nabla B_y|^2 + |\nabla B_z|^2\right) > 0 \tag{7.16}$$

Hence for a diamagnet ($\chi < 0$) we can find stable equilibrium points in a magnetic field, but not for paramagnets. As an interesting aside, since animal bodies are mostly water (which is slightly diamagnetic), one can make a small frog float in strong (about 15-20Tesla) magnetic field!

PROBLEM: Calculate the susceptibility of an ion with angular momentum J as follows. Consider an ion with total angular momentum J- such that it has $2J+1$ states. The partition function would be given by

$$\begin{aligned} Z &= \sum_{j_z=-J}^{J} e^{-\beta g^* \mu_B B j_z} \\ \therefore M &= -\frac{N}{V}\frac{\partial F}{\partial B} = \frac{N}{V} g^* \mu_B J B_J\left(\frac{g*\mu_B J B}{k_B T}\right) \end{aligned} \tag{7.17}$$

where $B_J(x)$ denotes the Brillouin function

$$B_J(x) = \frac{2J+1}{2J}\coth\frac{2J+1}{2J}x - \frac{1}{2J}\coth\frac{1}{2J}x \tag{7.18}$$

This leads to *Curie's law* for high temperature ($k_B T >> g^* \mu_B B$) susceptibility

$$\chi = \frac{N}{V}\frac{(g^*\mu_B)^2}{3}\frac{J(J+1)}{k_B T} \tag{7.19}$$

7.2 Magnetic moment of an isolated atom: Lande g-factor & Hund's rules

An atom can have many electrons. These will have spin moment $\boldsymbol{s}_i$ and orbital moment $\boldsymbol{l}_i$. To get the total moments, we need to add

$$\begin{aligned} \boldsymbol{L} &= \sum \boldsymbol{l}_i \\ \boldsymbol{S} &= \sum \boldsymbol{s}_i \qquad (7.20) \\ \boldsymbol{J} &= \boldsymbol{L} + \boldsymbol{S} \qquad (7.21) \end{aligned}$$

There could be various values of $\boldsymbol{L}$ and $\boldsymbol{S}$ depending on the relative orientation of the individual momenta. Predicting the correct value of $\boldsymbol{L}$ and $\boldsymbol{S}$ is what Hund'd rule does - we will do this in the next section. Now, suppose we have been given the correct $\boldsymbol{J}$, $\boldsymbol{L}$ and $\boldsymbol{S}$ values. We then need to relate it to the total magnetic moment of the atom, so that we can predict its energy level in a weak magnetic field. The problem is slightly non-trivial because the gyromagnetic ratios associated with $\boldsymbol{L}$ and $\boldsymbol{S}$ are not the same, so the magnetic moment will not be, in a simple vector addition sense, point along $\boldsymbol{J}$. We define the Bohr magneton and the other relevant quantities as

$$\begin{aligned} \mu_B &= \frac{|e|\hbar}{2m} && (7.22) \\ \boldsymbol{\mu_S} &= -2\frac{\mu_B}{\hbar}\boldsymbol{S} && (7.23) \\ \boldsymbol{\mu_L} &= -\frac{\mu_B}{\hbar}\boldsymbol{L} && (7.24) \\ & && (7.25) \end{aligned}$$

The total moment is then

$$\begin{aligned} \boldsymbol{\mu} &= -\frac{\mu_B}{\hbar}(2\boldsymbol{S}+\boldsymbol{L}) \\ &= -\frac{\mu_B}{\hbar}(\boldsymbol{J}+\boldsymbol{S}) && (7.26) \end{aligned}$$

7.2.1 Lande g-factor

Now we place it in a magnetic field. We will see how the problem is treated "semi-classically" and then quantum mechanically.

Semiclassical vector model

See fig 7.1. The vector $\boldsymbol{J}$ is assumed to go around the direction of the magnetic field (taken to be $\hat{\boldsymbol{z}}$). We want to take the projection of $\boldsymbol{\mu}$ and $\boldsymbol{B}$ along $\boldsymbol{J}$. We will use the "quantum" aspect that the z-component of angular momenta can only be $J_z = M\hbar$ and that the magnitude of the angular momentum vector is $\sqrt{J(J+1)}\hbar$. But otherwise we treat the vectors classically.

$$\begin{aligned} \Delta E &= -\boldsymbol{\mu}.\boldsymbol{B} \\ &= -\left(\boldsymbol{\mu}.\frac{\boldsymbol{J}}{|\boldsymbol{J}|}\right)\left(\boldsymbol{B}.\frac{\boldsymbol{J}}{|\boldsymbol{J}|}\right) && (7.27) \\ &= \frac{\mu_B}{\hbar}\left(\frac{(\boldsymbol{L}+2\boldsymbol{S}).\boldsymbol{J}}{|\boldsymbol{J}|^2}\right)BJ_z && (7.28) \\ &= \frac{\mu_B}{\hbar}\left(\frac{|\boldsymbol{J}|^2+\boldsymbol{J}.\boldsymbol{S}}{|\boldsymbol{J}|^2}\right)M\hbar B && (7.29) \\ &= \mu_B\left(1+\frac{\boldsymbol{J}.\boldsymbol{S}}{|\boldsymbol{J}|^2}\right)MB && (7.30) \\ &= \mu_B\left(1+\frac{|\boldsymbol{J}|^2-|\boldsymbol{L}|^2+|\boldsymbol{S}|^2}{2|\boldsymbol{J}|^2}\right)MB && (7.31) \\ &= \mu_B\left(1+\frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)}\right)MB && (7.32) \\ &= Mg_J\mu_B B && (7.33) \end{aligned}$$

The definition of the Lande g-factor is then

$$g_J = 1+\frac{J(J+1)-L(L+1)+S(S+1)}{2J(J+1)} \qquad (7.34)$$

It tells us how the spacing between the discreet Zeeman levels depend on $\boldsymbol{J}$,$\boldsymbol{L}$ and $\boldsymbol{S}$.

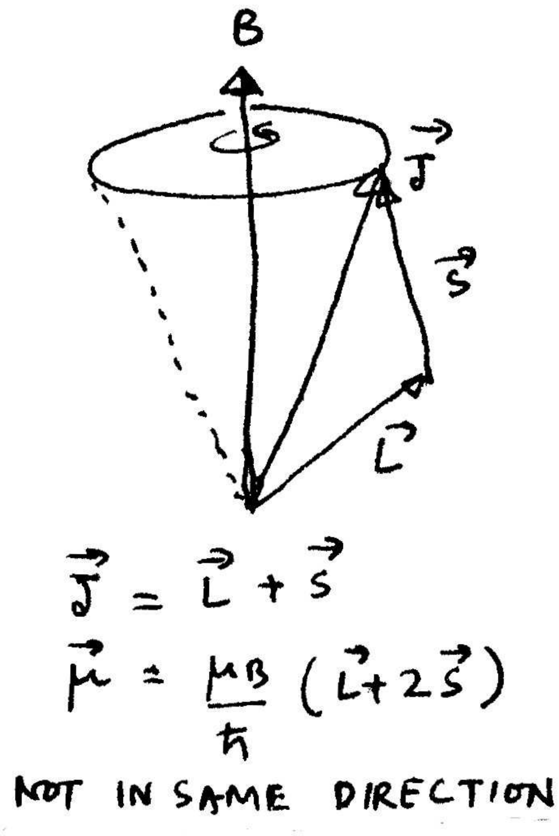


Figure 7.1: Notice that the vectors $\boldsymbol{J}$ and $\boldsymbol{\mu}$ are not in general in the same direction.

The quantum mechanical treatment of the problem

Can this idea of a vector rolling on a cone and then taking its average value be placed on a firmer footing? The obvious hamiltonian with a small magnetic field added as a perturbation is going to be:

$$\begin{aligned} H &= H_0 - \boldsymbol{\mu}.\boldsymbol{B} \\ &= H_0 - B\mu_z \\ &= H_0 + V \end{aligned} \tag{7.35}$$

In absence of B, J and M are good quantum numbers and the electronic states are characterized by state kets like $|\alpha, JM\rangle$ where α is the set of all other quantum numbers - like radial quantum no etc. Obviously:

$$J_z|\alpha, JM\rangle = M\hbar|\alpha, JM\rangle \tag{7.36}$$

$$\boldsymbol{J}^2|\alpha, JM\rangle = J(J+1)\hbar^2|\alpha, JM\rangle \tag{7.37}$$

All m states are obviously degenerate at $B = 0$. The standard problem is now to apply first order perturbation using $V = -B\mu_z$ as the perturbing potential. The operator for the magnetic moment is given by:

$$\begin{aligned} \boldsymbol{\mu} &= -\frac{\mu_B}{\hbar}(2\boldsymbol{S} + \boldsymbol{L}) \\ &= -\frac{\mu_B}{\hbar}(\boldsymbol{J} + \boldsymbol{S}) \end{aligned} \tag{7.38}$$

$$\Delta E = \langle \alpha, JM|V|\alpha, JM\rangle \tag{7.39}$$

$$= -B\langle \alpha, JM|\mu_z|\alpha, JM\rangle \tag{7.40}$$

Now in place of the "averaging" along the the direction of $\boldsymbol{J}$ that we used, we will utilise a theorem that holds for the expectation value of any vector operator between angular momentum eigenstates - the Wigner-Eckart theorem. The derivation of the result is given in many textbooks (*e.g* see Sakurai for "projection theorem" for vector operators). If $\boldsymbol{T}$ is any operator whose expectation value transforms under rotation like that of a vector - then its expectation in angular momentum eigenstates is related to the expectation of the angular momentum operator itself, in a simple way:

$$\langle\alpha, JM|T_\nu|\alpha, JM\rangle = \frac{\langle\alpha, JM|J_\nu|\alpha, JM\rangle}{\hbar J(J+1)}\langle\alpha, JM|\boldsymbol{J}.\boldsymbol{T}|\alpha, JM\rangle \tag{7.41}$$

There is no summation over M or J involved, it is a simple proportionality. Then taking $T_\nu = \mu_z$, we get for eqn 7.40 to be

$$\begin{aligned} \Delta E &= -B\langle\alpha, JM|\mu_z|\alpha, JM\rangle & (7.42)\\ &= -B\frac{M}{\hbar J(J+1)}\langle\alpha, JM|\boldsymbol{J}.\boldsymbol{\mu}|\alpha, JM\rangle & (7.43)\end{aligned}$$

Now using

$$\boldsymbol{J}.\boldsymbol{\mu} = -\frac{\mu_B}{\hbar}(\boldsymbol{J}^2 + \boldsymbol{J}.\boldsymbol{S}) \tag{7.44}$$

we get

$$\begin{aligned} \Delta E &= \mu_B BM\frac{\langle\alpha, JM|\boldsymbol{J}^2 + \boldsymbol{J}.\boldsymbol{S}|\alpha, JM\rangle}{\hbar^2 J(J+1)}\\ &= \mu_B BM\left(1 + \frac{\langle\alpha, JM|\boldsymbol{J}.\boldsymbol{S}|\alpha, JM\rangle}{\hbar^2 J(J+1)}\right) & (7.45)\end{aligned}$$

The quantity within the brackets would lead to exactly the same expression as eqn 7.34, the steps are trivial because we can express $\boldsymbol{J}.\boldsymbol{S}$ in terms of $\boldsymbol{J}^2$, $\boldsymbol{L}^2$ and $\boldsymbol{S}^2$. Notice how the Wigner-Eckart theorem, which is exact, replaces our vector model "assumptions".

7.2.2 Hund's rules

In the calculation of the Lande g-factor, we assumed that the values of L, S and J are given to us. But since two angular momenta can be combined in various ways, we need a principle to tell us which is the favoured combination. This principle (for a many electron atom) is the Hund's rule.

The rule states that "every orbital in a subshell is singly occupied with one electron before any one orbital is doubly occupied, and all electrons in singly occupied orbitals have the same spin."

This needs to be applied to the outermost (unfilled) shell of an atom only, since the fully filled shells have zero net angular momentum. We do it in the following steps (these are partly empirical)

1. The z-component (S_z) of the total spin should have the maximum possible value consistent with Pauli exclusion principle. This means that if we have 6 d-electrons, then we should put 5 of them in up-state and one remaining one in the down state, pairing up with any one up electron as shown in fig. The qualitative justification is that if two spins point in the same direction then they would try to stay farther apart in real space. This would reduce the Coulomb repulsion contribution.

2. The orbital angular momenta combine to give the highest possible total L, consistent with the previous rule. This electron wavefunctions at large L are more spread out and this again reduces the repulsive Coulomb interaction.

3. The angular momenta $\boldsymbol{L}$ and $\boldsymbol{S}$ combine to give $J = |L - S|$ if the shell is less than half filled and $J = L + S$ when the shell is more than half filled.

$^{2S+1}L_J$ notation

To specify the LSJ values, for historic reasons the following notation is used.

- The value of S is specified in terms of its multiplicity $2S+1$ as a superscript.
- The L value is specified by a letter $S=0$, $P=1$, $D=2$, $F=3$, $G=4$ and then H, J, K etc.
- The J value is specified as a subscript - unlike the S value, it is not the multiplicity that is written, but the value itself.

As an example the Fe^{2+} ion has a spin state $^5D_4$. It means it has $S=2$, $L=2$, $J=4$.

How well do these rules work?

The rules are stated for an atom, but magnetic materials are generally solids, so we really have to apply the rule to ions sitting in a lattice. The rule also tells us that atoms with incomplete shells with large capacity for holding electrons (d,f) are likely to have large moments. This is indeed correct - the transition metals of the periodic table are where we find magnetic behaviour.

One would experimentally measure either the magnetic moment or the susceptibility of a certain amount of the material and then try to see how much each moment must be contributing. See eqn 7.19. We can use the Hund's rule and Lande g factor to predict what the value of $g^*\sqrt{J(J+1)}$ should be and test it against experimental data - for obvious reasons this is called the effective Bohr magneton.

See the following two tables for the f and the d block elements. This shows that the predictions are pretty good for most f-elements but not so good for many of the d-elements. Somewhat surprisingly the data seems to suggest that a better match would be obtained if we consider only S and set $L=0$ for the d elements. Why should it be so?

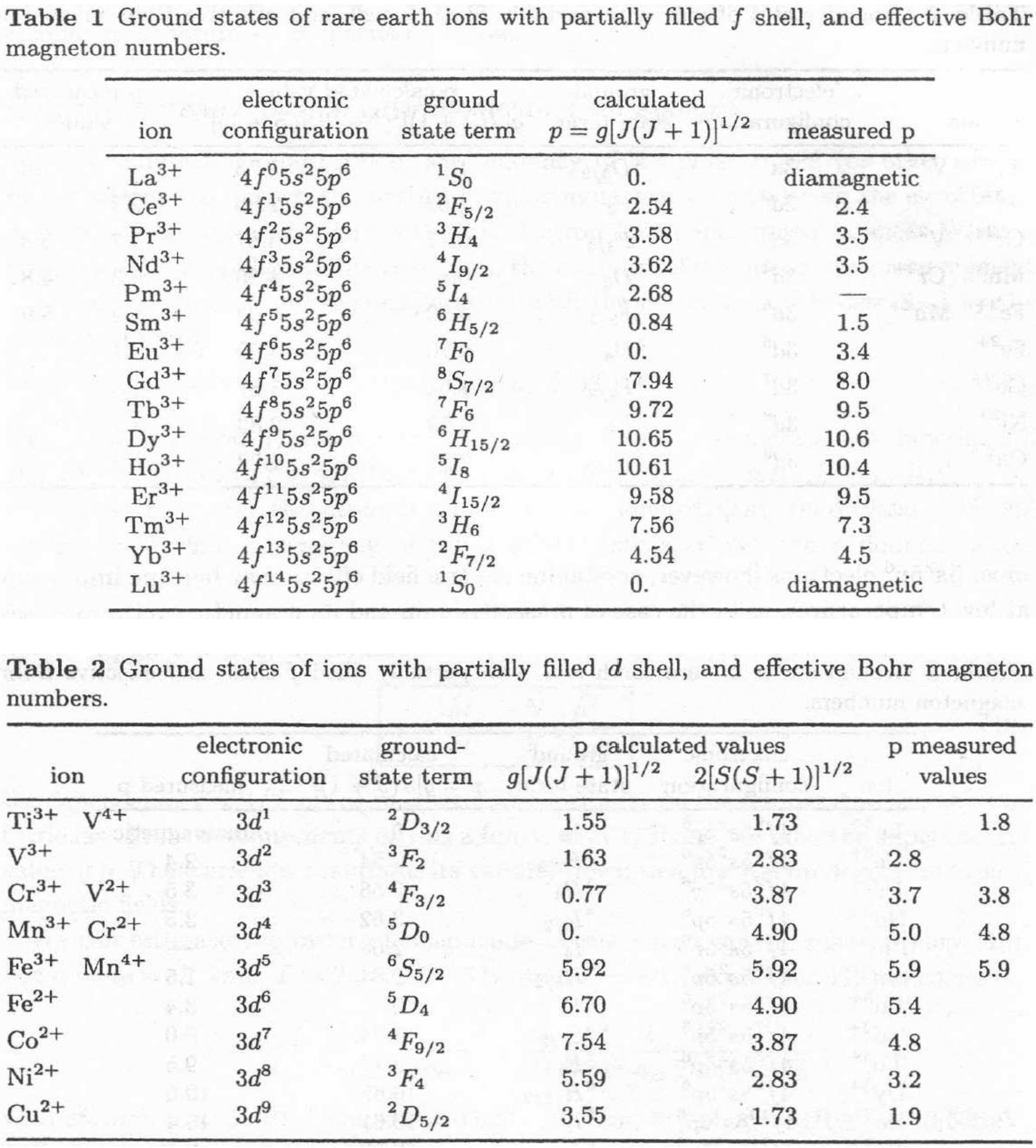


Table 1 Ground states of rare earth ions with partially filled f shell, and effective Bohr magneton numbers.

ion	electronic configuration	ground state term	calculated $p = g[J(J+1)]^{1/2}$	measured p
La^{3+}	$4f^0 5s^2 5p^6$	$^1S_0$	0.	diamagnetic
Ce^{3+}	$4f^1 5s^2 5p^6$	$^2F_{5/2}$	2.54	2.4
Pr^{3+}	$4f^2 5s^2 5p^6$	$^3H_4$	3.58	3.5
Nd^{3+}	$4f^3 5s^2 5p^6$	$^4I_{9/2}$	3.62	3.5
Pm^{3+}	$4f^4 5s^2 5p^6$	$^5I_4$	2.68	—
Sm^{3+}	$4f^5 5s^2 5p^6$	$^6H_{5/2}$	0.84	1.5
Eu^{3+}	$4f^6 5s^2 5p^6$	$^7F_0$	0.	3.4
Gd^{3+}	$4f^7 5s^2 5p^6$	$^8S_{7/2}$	7.94	8.0
Tb^{3+}	$4f^8 5s^2 5p^6$	$^7F_6$	9.72	9.5
Dy^{3+}	$4f^9 5s^2 5p^6$	$^6H_{15/2}$	10.65	10.6
Ho^{3+}	$4f^{10} 5s^2 5p^6$	$^5I_8$	10.61	10.4
Er^{3+}	$4f^{11} 5s^2 5p^6$	$^4I_{15/2}$	9.58	9.5
Tm^{3+}	$4f^{12} 5s^2 5p^6$	$^3H_6$	7.56	7.3
Yb^{3+}	$4f^{13} 5s^2 5p^6$	$^2F_{7/2}$	4.54	4.5
Lu^{3+}	$4f^{14} 5s^2 5p^6$	$^1S_0$	0.	diamagnetic

Table 2 Ground states of ions with partially filled d shell, and effective Bohr magneton numbers.

ion	electronic configuration	ground-state term	p calculated values $g[J(J+1)]^{1/2}$	p calculated values $2[S(S+1)]^{1/2}$	p measured values	
Ti^{3+} V^{4+}	$3d^1$	$^2D_{3/2}$	1.55	1.73		1.8
V^{3+}	$3d^2$	$^3F_2$	1.63	2.83	2.8	
Cr^{3+} V^{2+}	$3d^3$	$^4F_{3/2}$	0.77	3.87	3.7	3.8
Mn^{3+} Cr^{2+}	$3d^4$	$^5D_0$	0.	4.90	5.0	4.8
Fe^{3+} Mn^{4+}	$3d^5$	$^6S_{5/2}$	5.92	5.92	5.9	5.9
Fe^{2+}	$3d^6$	$^5D_4$	6.70	4.90	5.4	
Co^{2+}	$3d^7$	$^4F_{9/2}$	7.54	3.87	4.8	
Ni^{2+}	$3d^8$	$^3F_4$	5.59	2.83	3.2	
Cu^{2+}	$3d^9$	$^2D_{5/2}$	3.55	1.73	1.9	

Figure 7.2: The list is reproduced from a book by Grosso & Parravicini. See the comparison between the experimental values and the predictions based on Hund's rule. Try to calculate the the LSJ values from the electronic configurations by applying the rules, yourself.

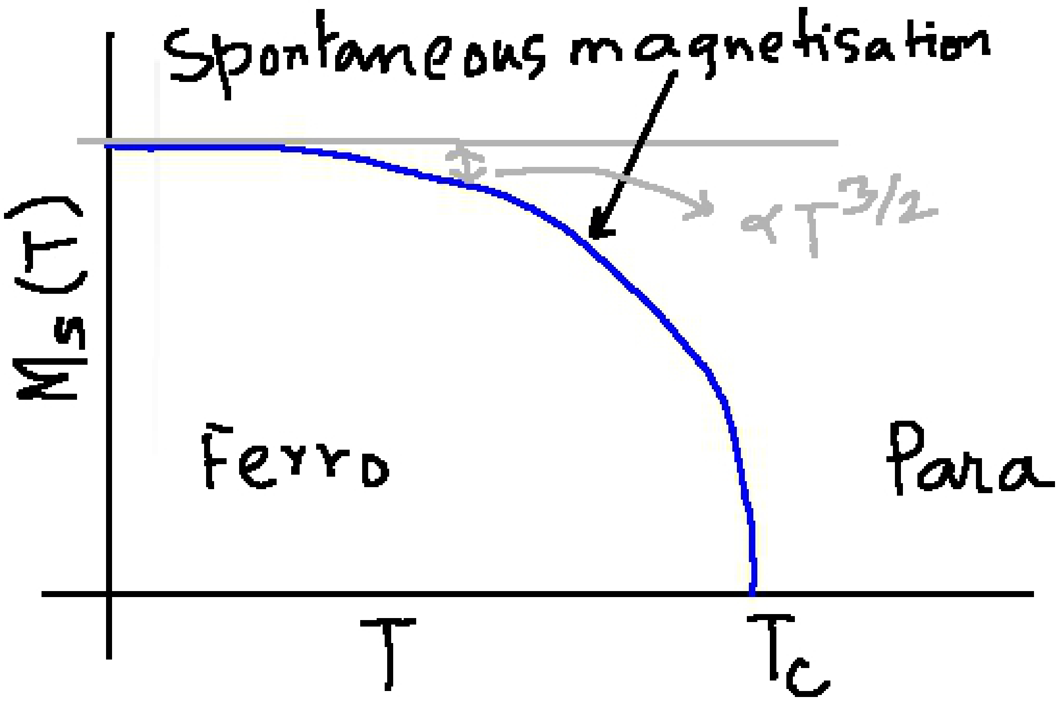


Figure 7.3: The typical behaviour of spontaneous magnetisation of a ferromagnet. This magnetisation is the maximum one can measure for a sample of the ferromagnet with no applied field. In general the measured value would not be this because of the presence of magnetic domains. We will deal with this very important practical aspect of domain formation later.

Adiabetic demagnetisation method of producing very low temperatures

To be written

7.3 The electron gas in a magnetic field: oscillatory phenomena

To be written

7.4 Exchange interaction and ferromagnetism

A few elements in the periodic table - namely *Fe*, *Co*, *Ni* can have very large magnetisation without any applied external magnetic field. These are metals, so they have free electrons as well as some electrons which form a very narrow band - arising from the *d* electrons of the atom. It turns out that to understand the origin of this effect, it is necessary to consider all the electrons at once. In all the problems we have done so far, we treated *one* electron in the field of all the atoms.

Why dipolar interaction is not the cause of ferromagnetism

The interaction energy between two dipoles cannot account for ferromagnetism.

- Dipole-dipole would favour antiparallel alignment
- The energy scale is too small to account for ferromagnetic Curie temperature of ∼ 1000K. Curie temperature is that above which there is no spontaneous magnetisation in the system - see fig 7.3. Above T_c iron would behave like a paramagnet.

  Recall that each dipole is of the order of the Bohr magneton. Their energy of interaction is

$$U_{12} = \frac{\mu_0}{4\pi}\frac{1}{r^3}\left[3(\boldsymbol{m_1}.\hat{\boldsymbol{r}})(\boldsymbol{m_2}.\hat{\boldsymbol{r}}) - \boldsymbol{m_1}.\boldsymbol{m_2}\right] \tag{7.46}$$

Table 7.1: List of some ferromagnetic Curie temperatures (Wikipedia)

Substance	Curie temp (Kelvin)
Iron (Fe)	1043
Cobalt (Co)	1403
Nickel (Ni)	631
Iron Oxide (Fe_2O_3)	895
Gadolinium (Gd)	292
Fe_2O_3	948
$FeOFe_2O_3$	858
$NiOFe_2O_3$	858
$CuOFe_2O_3$	728
$MgOFe_2O_3$	713
$MnBi$	630
$MnSb$	587
$MnOFe_2O_3$	573
$Y_3Fe_5O_{12}$	560
CrO_2	386
$MnAs$	318
Dy	88
EuO	69

Table 7.2: List of some anti-ferromagnetic Neel temperatures (Wikipedia)

substance	Neel temperature (Kelvin)
MnO	116
MnS	160
$MnTe$	307
MnF_2	67
FeF_2	79
$FeCl_2$	24
FeO	198
$CoCl_2$	25
CoO	291
$NiCl_2$	50
NiO	525
Cr	308

We can estimate this expression assuming a separation of $r \sim 1\text{nm}$, $U_{12} \approx \frac{\mu_0}{4\pi}\frac{{\mu_B}^2}{r^3}$ is a very small energy.

- The only source of such a strong interaction is the electronic Coulomb interaction, but how does it come into play in magnetism?

We have the following problem

$$H = \sum_i \frac{\boldsymbol{p_i}^2}{2m} + \sum_f \sum_i V_f(\boldsymbol{r_i}) + \frac{1}{4\pi\epsilon}\frac{1}{2}\sum_i \sum_j \frac{e^2}{|\boldsymbol{r_i} - \boldsymbol{r_j}|} \tag{7.47}$$

- Here i runs over all the electrons and f runs over all the lattice sites.
- $V_f(\boldsymbol{r_i})$ means the potential due to lattice site f on particle i.
- Wavefunctions centred on different sites will be taken to be orthogonal - which is not strictly true, but we have done the same for all our tight-binding type calculations.

The hamiltonian (7.47) can also be written as a sum of single particle and two particle (interaction) terms.

$$H = \sum_i h_i(\boldsymbol{r_i}) + \sum_{i,j} h_{ij}(\boldsymbol{r_i}, \boldsymbol{r_i})$$

where

$$h(\boldsymbol{r_i}) = \frac{\boldsymbol{p_i}^2}{2m} + \sum_f V_f(\boldsymbol{r_i}) \tag{7.48}$$

$$h_{ij}(\boldsymbol{r_i}, \boldsymbol{r_i}) = \frac{1}{4\pi\epsilon}\frac{1}{2}\frac{e^2}{|\boldsymbol{r_i} - \boldsymbol{r_j}|} \tag{7.49}$$

In all previous chapters we ignored the fact that two electrons will repel (interact with) each other - this we can no longer do. We shall see that it is Coulomb interaction in conjunction with antisymmetry requirements which ultimately give rise to magnetism. We need to develop a way of writing wavefunction of many electrons together and then take expectation values of such many particle wavefunctions to identify the correct ground state of the Hamiltonian.

7.4.1 Origin of the exchange interaction between two electrons

The Hydrogen molecule

The H_2 molecule has two electrons which may be thought of as coming from two atoms, marked A and B in fig. We write the wavefunction of the two electron state, using the atomic states as the basis functions. Let ϕ_A denote a $1s$ wavefunction centered at site A. The spin part of the wavefunctions are denoted by α and β. There are possible four combinations, where we do not need to put two electrons at the same site. The state are not normalised at this point. The factor of $\frac{1}{\sqrt{2}}$ would have given us the correct normalisation, if the states were orthogonal.

$$|\phi_{A\uparrow}, \phi_{B\uparrow}\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_A(\boldsymbol{r}_1)\alpha 1 & \phi_A(\boldsymbol{r}_2)\alpha 2 \\ \phi_B(\boldsymbol{r}_1)\alpha 1 & \phi_B(\boldsymbol{r}_2)\alpha 2 \end{vmatrix} \tag{7.50}$$

$$|\phi_{A\uparrow}, \phi_{B\downarrow}\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_A(\boldsymbol{r}_1)\alpha 1 & \phi_A(\boldsymbol{r}_2)\alpha 2 \\ \phi_B(\boldsymbol{r}_1)\beta 1 & \phi_B(\boldsymbol{r}_2)\beta 2 \end{vmatrix} \tag{7.51}$$

$$|\phi_{A\downarrow}, \phi_{B\uparrow}\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_A(\boldsymbol{r}_1)\beta 1 & \phi_A(\boldsymbol{r}_2)\beta 2 \\ \phi_B(\boldsymbol{r}_1)\alpha 1 & \phi_B(\boldsymbol{r}_2)\alpha 2 \end{vmatrix} \tag{7.52}$$

$$|\phi_{A\downarrow}, \phi_{B\downarrow}\rangle = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_A(\boldsymbol{r}_1)\beta 1 & \phi_A(\boldsymbol{r}_2)\beta 2 \\ \phi_B(\boldsymbol{r}_1)\beta 1 & \phi_B(\boldsymbol{r}_2)\beta 2 \end{vmatrix} \tag{7.53}$$

The states ϕ_A and ϕ_B are not orthogonal, because they have a little overlap at the middle. Thus our basis states are not orthogonal, but there is nothing wrong in that.

Now, since we know that isolated atom solutions (call the ground state energy of each atom to be E_0), we know that the energy levels of the system, when the A-B distance is very large must be $2E_0$. The 2-electron hamiltonian with the Coulomb interaction included can be written as (notice the sign of the various interaction terms):

$$H = \frac{\boldsymbol{p}_1^2}{2m} + \frac{\boldsymbol{p}_2^2}{2m} - \frac{1}{4\pi\epsilon_0}\left(\frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_1|} + \frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_2|} + \frac{1}{|\boldsymbol{r}_B - \boldsymbol{r}_1|} + \frac{1}{|\boldsymbol{r}_B - \boldsymbol{r}_2|}\right) + \frac{1}{4\pi\epsilon_0}\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \tag{7.54}$$

Here $\boldsymbol{r}_A$, $\boldsymbol{r}_B$ denote the coordinate of the A and B site. What are the expectation values of this Hamiltonian in the Slate determinant states that we have just written down? Notice that from this point we cannot draw any analogy with classical behaviour, because there is no analogue of symmetric or anti-symmetric wavefunctions in classical physics. Before we go ahead and calculate the expectation values there is a bit more work to do with the wavefunctions. Recall that the two electron wavefunctions can be written either in the $|S_1 S_2 Sz_1 Sz_2\rangle$, or $|S_1 S_2 S\ Sz\rangle$ form. We prefer the second form. It is left as an important exercise to verify that :

The state $S = 1$, $Sz = 1$

$$\begin{aligned} \Psi_{1,1} &= |\phi_{A\uparrow}, \phi_{B\uparrow}\rangle \\ &= \frac{1}{\sqrt{2}}\left[\phi_A(r_1)\phi_B(r_2) - \phi_B(r_1)\phi_A(r_2)\right]\alpha(1)\alpha(2) \end{aligned} \tag{7.55}$$

The state $S = 1$, $Sz = 0$

$$\begin{aligned} \Psi_{1,0} &= \frac{|\phi_{A\uparrow}, \phi_{B\downarrow}\rangle - |\phi_{A\downarrow}, \phi_{B\uparrow}\rangle}{\sqrt{2}} \\ &= \frac{1}{\sqrt{2}}.\frac{1}{\sqrt{2}}\left[\phi_A(r_1)\phi_B(r_2) - \phi_B(r_1)\phi_A(r_2)\right](\alpha(1)\beta(2) + \beta(1)\alpha(2)) \end{aligned} \tag{7.56}$$

The state $S = 1$, $Sz = -1$

$$\begin{aligned} \Psi_{1,-1} &= |\phi_{A\downarrow}, \phi_{B\downarrow}\rangle \\ &= \frac{1}{\sqrt{2}} \left[\phi_A(r_1)\phi_B(r_2) - \phi_B(r_1)\phi_A(r_2)\right] \beta(1)\beta(2) \end{aligned} \tag{7.57}$$

The state $S = 0$, $Sz = 0$

$$\begin{aligned} \Psi_{0,0} &= \frac{|\phi_{A\uparrow}, \phi_{B\downarrow}\rangle + |\phi_{A\downarrow}, \phi_{B\uparrow}\rangle}{\sqrt{2}} \\ &= \frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}} \left[\phi_A(r_1)\phi_B(r_2) + \phi_B(r_1)\phi_A(r_2)\right] (\alpha(1)\beta(2) - \beta(1)\alpha(2)) \end{aligned} \tag{7.58}$$

Normalisation of these states

What is $\langle\Psi_{1,1}|\Psi_{1,1}\rangle$? To compute this we have to remember that the wavefunciton contains both spatial and spin parts.

$$\begin{aligned} \langle\Psi_{1,1}|\Psi_{1,1}\rangle &= \frac{1}{2} \left(\langle\phi_A(r_1)|\phi_A(r_1)\rangle\langle\phi_B(r_2)|\phi_B(r_2)\rangle - \langle\phi_A(r_1)|\phi_B(r_1)\rangle\langle\phi_B(r_2)|\phi_A(r_2)\rangle\right. \\ &+ \left.\langle\phi_B(r_1)|\phi_B(r_1)\rangle\langle\phi_A(r_2)|\phi_A(r_2)\rangle - \langle\phi_B(r_1)|\phi_A(r_1)\rangle\langle\phi_A(r_2)|\phi_B(r_2)\rangle\right) \\ &= 1 - W^2 \end{aligned} \tag{7.59}$$

where

$$W = \langle\phi_A(r)|\phi_B(r)\rangle \tag{7.60}$$

is the overlap integral. But the normalisation is *not* the same for all the states. Following the same process (but the algebra has to be carefully done) we can show the full set of results:

$$\begin{aligned} \langle\Psi_{1,1}|\Psi_{1,1}\rangle &= 1 - W^2 \\ \langle\Psi_{1,0}|\Psi_{1,0}\rangle &= 1 + W^2 \\ \langle\Psi_{1,-1}|\Psi_{1,-1}\rangle &= 1 - W^2 \\ \langle\Psi_{0,0}|\Psi_{0,0}\rangle &= 1 + W^2 \end{aligned} \tag{7.61}$$

It is left as an exercise to work out the full algebra - notice the difference in the sign of W^2. When both spins are up or both are down, then it has a negative sign, otherwise it is positive.
Now we get back to the Hamiltonian and rewrite it in two rather obvious parts. We define the isolated atom part, (which does not depend on how far the atoms are) of it as

$$H_0 = \left(\frac{\boldsymbol{p}_1^2}{2m} - \frac{e^2}{4\pi\epsilon_0}\frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_1|}\right) + \left(\frac{\boldsymbol{p}_2^2}{2m} - \frac{e^2}{4\pi\epsilon_0}\frac{1}{|\boldsymbol{r}_B - \boldsymbol{r}_2|}\right) \tag{7.62}$$

and the interacting part of it as

$$\Delta H = -\frac{e^2}{4\pi\epsilon_0}\left(\frac{1}{|\boldsymbol{r}_B - \boldsymbol{r}_1|} + \frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_2|} - \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} - \frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_B|}\right) \tag{7.63}$$

We have now formulated the problem clearly. We need to evaluate for each $|S_1 S_2 S Sz\rangle$ state

$$\frac{\langle\Psi|\Delta H|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{7.64}$$

For example:

$$\begin{aligned} \langle\Psi_{1,1}|\Delta H|\Psi_{1,1}\rangle &= \frac{1}{2}\left[\langle\phi_A(r_1)\phi_B(r_2)|\Delta H|\phi_A(r_1)\phi_B(r_2)\rangle + \langle\phi_A(r_1)\phi_B(r_2)|\Delta H|\phi_A(r_1)\phi_B(r_2)\rangle\right] \\ &- \frac{1}{2}\left[\langle\phi_A(r_1)\phi_B(r_2)|\Delta H|\phi_B(r_1)\phi_A(r_2)\rangle + \langle\phi_B(r_1)\phi_A(r_2)|\Delta H|\phi_A(r_1)\phi_B(r_2)\rangle\right] \\ &= K - J \end{aligned}$$

where

$$
\begin{aligned}
K &= \langle \phi_A(r_1)\phi_B(r_2)|\Delta H|\phi_A(r_1)\phi_B(r_2)\rangle &(7.65)\\
&= \int\!\!\int d^3\boldsymbol{r_1} d^3\boldsymbol{r_2} |\phi_A(r_1)|^2 |\phi_B(r_2)|^2 \Delta H \\
J &= \langle \phi_B(r_1)\phi_A(r_2)|\Delta H|\phi_A(r_1)\phi_B(r_2)\rangle &(7.66)\\
&= \int\!\!\int d^3\boldsymbol{r_1} d^3\boldsymbol{r_2} \phi_A^*(r_1)\phi_B^*(r_2)\Delta H \phi_B(r_1)\phi_A(r_2)
\end{aligned}
$$

Notice that the K integral is like the perturbative part of the potential averaged over the charge density. But the J integral has no such simple interpretation. Notice how the arguments of the two wavefunctions have been interchanged or exchanged - this is called the exchange integral. Now we can work out the other three integrals. We get

$$
\begin{aligned}
\frac{\langle \Psi_{1,1}|\Delta H|\Psi_{1,1}\rangle}{\langle \Psi_{1,1}|\Psi_{1,1}\rangle} &= \frac{K-J}{1-W^2}\\
\frac{\langle \Psi_{1,0}|\Delta H|\Psi_{1,0}\rangle}{\langle \Psi_{1,0}|\Psi_{1,0}\rangle} &= \frac{K-J}{1+W^2}\\
\frac{\langle \Psi_{1,-1}|\Delta H|\Psi_{1,-1}\rangle}{\langle \Psi_{1,-1}|\Psi_{1,-1}\rangle} &= \frac{K-J}{1+W^2}\\
\frac{\langle \Psi_{0,0}|\Delta H|\Psi_{0,0}\rangle}{\langle \Psi_{0,0}|\Psi_{0,0}\rangle} &= \frac{K+J}{1-W^2}
\end{aligned}
$$

Convince yourself that the signs of the exchange integral is different in the triplet state (= 1) and the singlet state ($S = 0$). Both, the K and J terms are functions of $|\boldsymbol{r}_A - \boldsymbol{r}_B|$. Explicit calculation of the terms show that $J > K$. Then we see:

$$
E_{singlet} - E_{triplet} = \frac{J - KW^2}{1 - W^4} \tag{7.67}
$$

If W is small, J is negative then the singlet lies lower in energy. An antisymmetric spatial part and a symmetric spin part is favoured.

The sign of J

The integral J can be written (from eqn 7.67 and using the form of ΔH)

$$
\begin{aligned}
J &= -\frac{e^2}{4\pi\epsilon_0}\langle \phi_B(r_1)\phi_A(r_2)|\frac{1}{|\boldsymbol{r}_B - \boldsymbol{r}_1|} + \frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_2|} - \frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} - \frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_B|}\phi_A(r_1)\phi_B(r_2)\rangle \\
&= \frac{e^2}{4\pi\epsilon_0}\Bigg[-W\langle \phi_A(\boldsymbol{r}_1)|\frac{1}{|\boldsymbol{r}_B - \boldsymbol{r}_1|}|\phi_B(\boldsymbol{r}_1)\rangle - W\langle \phi_B(\boldsymbol{r}_2)|\frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_2|}|\phi_A(\boldsymbol{r}_1)\rangle &(7.68)\\
&+ W^2\frac{1}{|\boldsymbol{r}_A - \boldsymbol{r}_2|} + \langle \phi_A(\boldsymbol{r}_1)\phi_B(\boldsymbol{r}_2)|\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}|\phi_B(\boldsymbol{r}_1)\phi_A(\boldsymbol{r}_2)\rangle\Bigg]
\end{aligned}
$$

Can you see that if $W = 0$, then $J > 0$. Thus if the basis functions are orthogonal, the exchange term must be positive. This is a very important conclusion.

The excited states of Helium atom

The He atom contains two electrons and in the ground state both these electrons must be in the $1s$ state. Thus the spins must be oppositely directed and there is only one possibility consistent with the exclusion principle - the singlet state ${}^1S_0$ following the LSJ notation. Now consider an excited state where one of the electrons is put in a higher state (may be $2s$, may be $2p$). So the two electron states will have to be built out of one $1s$ and one $2s$ state. Following the convention of naming atomic wavefunctions with nlm, we can write

$$
E = E_{100} + E_{nlm} + \Delta E \tag{7.69}
$$

We can write the full wavefunction by constructing a Slater determinant out of two one electron states Ψ_{100} and Ψ_{nlm} along with the spin functions as before. But now the two wavefunctions are necessarily orthogonal and the exchange integral must be positive - unlike the H_2 molecule. Thus the lower excited state should be a triplet. This striking prediction of quantum mechanics is indeed known to be correct. (See Sakurai's QM book for details)

Formulation of the equivalent spin hamiltonian

So we see that the exchange integral behaves differently in the singlet and triplet states. If we ignore the "small" term contributed by non-orthogonality (W) in some cases, then we can say that the sign of J is the principal contribution that changes. Is there an operator that will take on eigenvalue $+1$ in the triplet state and -1 in the singlet state? Then that operator multiplied by the magnitude of J would reproduce the correct two electron energy levels. Heisenberg and Dirac pointed out that the operator $\boldsymbol{S}_1.\boldsymbol{S}_2$ does just that. Write the total spin for two spin-$\frac{1}{2}$ particles as

$$\begin{aligned} \boldsymbol{S}^2 &= (\boldsymbol{S}_1+\boldsymbol{S}_2)^2 \\ &= \boldsymbol{S}_1^2+\boldsymbol{S}_2^2+2\boldsymbol{S}_1.\boldsymbol{S}_2 \\ \therefore \boldsymbol{S}_1.\boldsymbol{S}_2 &= \frac{1}{2}\left[S(S+1)-2.\frac{1}{2}\left(\frac{1}{2}+1\right)\right] \qquad (7.70) \\ \therefore \frac{1}{2}+2\boldsymbol{S}_1.\boldsymbol{S}_2 &= \begin{cases} -1 & \text{if } S=0 \\ \ \ 1 & \text{if } S=1 \end{cases} \qquad (7.71) \end{aligned}$$

The hamiltonian of the form

$$H = K - J\left(\frac{1}{2}+\boldsymbol{S}_1.\boldsymbol{S}_2\right) \tag{7.72}$$

will thus reproduce the observed energy eigenvalues in the triplet and singlet states. An apparent dipole-dipole like interaction term has arisen out of Coulomb interaction and antisymmetry requirements of the total wavefunction (Slater determinant). Now we need to extend these ideas to many electron systems.

7.5 The ferromagnetic state from Coulomb interaction

We have N electrons and N lattice sites. We label the lattice sites by a,b,c,d.... and the electron spatial co-ordinates by $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, $\boldsymbol{r_3}$ ...$\boldsymbol{r_N}$. The Hamiltonian is

$$H = \underbrace{\sum_i \frac{\boldsymbol{p_i}^2}{2m}}_{\text{K.E.}} + \underbrace{\sum_f \sum_i V_f(\boldsymbol{r_i})}_{\text{lattice potential}} + \underbrace{\frac{1}{2}.\frac{1}{4\pi\epsilon_0}\sum_{\substack{ij\\ i\neq j}} \frac{e^2}{|\boldsymbol{r_i}-\boldsymbol{r_j}|}}_{\text{e-e Coulomb interaction}} \tag{7.73}$$

We need to take the expectation value of this H is fully antisymmetrized Slater determinant states. We will start by taking the expectation in a state where all the sites have spin ↑. The electrons are not completely localized but do not have large nearest neighbour hopping matrix element. In other words these electrons reside in a narrow band.

We denote the single electron eigenstates by

$$\phi_a,\ \phi_b,\ \phi_c....\phi_f$$

implying that the wavefunction is centered around the site a,b,c etc, which appear as the subscript. We will also write the terms in such a way that $\boldsymbol{r_1}$, $\boldsymbol{r_2}$, $\boldsymbol{r_3}$ ...$\boldsymbol{r_N}$ occur in their natural order. So each term of the Slater determinant (except for the sign) will be built up as follows, see fig 7.4

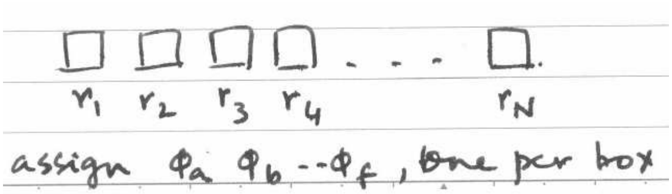


Figure 7.4: If ϕ_a is assigned to box for $\boldsymbol{r_1}$ it gives the function $\phi_a(\boldsymbol{r_1})$

7.6 The state with all spins pointing in the same direction

The Slater determinant with all spin up states is defined as :

$$\Phi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_a(\boldsymbol{r_1}) & \phi_a(\boldsymbol{r_2}) & .. & .. & \phi_a(\boldsymbol{r_N}) \\ \phi_b(\boldsymbol{r_1}) & \phi_b(\boldsymbol{r_2}) & .. & .. & \phi_b(\boldsymbol{r_N}) \\ .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. \\ \phi_N(\boldsymbol{r_1}) & \phi_N(\boldsymbol{r_2}) & .. & .. & \phi_N(\boldsymbol{r_N}) \end{vmatrix} \tag{7.74}$$

When expanded this will have $N!$ terms and if

$$\langle \phi_x(\boldsymbol{r}) | \phi_y(\boldsymbol{r}) \rangle \approx \delta_{xy} \tag{7.75}$$

then $\dfrac{1}{\sqrt{N!}}$ is also the correct normalisation.

If fully expanded an expectation value of the type $\langle \Phi | H | \Phi \rangle$ will have at least $N! \times N!$ terms. Obviously tabulating all of these and finding a trend or pattern among them would be just impossible. We group them in a particular way and the important fact is that only a small fraction of them are different from zero.

We first group the terms in the hamiltonian (eqn 7.73) as follows:

$$H = \sum_i \left(\frac{\boldsymbol{p_i}^2}{2m} + \sum_f V_f(\boldsymbol{r_i}) + \frac{1}{2} . \frac{1}{4\pi\epsilon_0} \sum_{\substack{j \\ i \neq j}} \frac{e^2}{|\boldsymbol{r_i} - \boldsymbol{r_j}|} \right) \tag{7.76}$$

Then we pick any one *ket* state from the $N!$ terms. Say

$$\chi = \phi_e(\boldsymbol{r_1}) \phi_g(\boldsymbol{r_2}) \phi_b(\boldsymbol{r_3}) \phi_d(\boldsymbol{r_4}) \phi_f(\boldsymbol{r_5}) .... \phi_x(\boldsymbol{r_N}) \tag{7.77}$$

Then from the *bra* side we pick a term with exactly the same sequence

$$\chi^* = \phi_e^*(\boldsymbol{r_1}) \phi_g^*(\boldsymbol{r_2}) \phi_b^*(\boldsymbol{r_3}) \phi_d^*(\boldsymbol{r_4}) \phi_f^*(\boldsymbol{r_5}) .... \phi_x^*(\boldsymbol{r_N}) \tag{7.78}$$

Now consider:

$$\begin{aligned} \left\langle \chi \left| \frac{\boldsymbol{p_1}^2}{2m} + \sum_f V_f(\boldsymbol{r_1}) \right| \chi \right\rangle &= \left\langle \phi_e \left| \frac{\boldsymbol{p_1}^2}{2m} + \sum_f V_f(\boldsymbol{r_1}) \right| \phi_e \right\rangle \\ &= E_0 \end{aligned} \tag{7.79}$$

All the other terms just pass through, since no term in the hamiltonian part contains the co-ordinates of $\boldsymbol{r_2}$, $\boldsymbol{r_3}$ ...$\boldsymbol{r_N}$. Because all lattice sites are equivalent, the number E_0 is independent of which lattice site (e) appeared with $\boldsymbol{r_1}$.

To this we need to add the contribution of the coulomb part $\frac{1}{2}.\frac{1}{4\pi\epsilon_0}\sum_{\substack{j\\ i\neq j}}\frac{e^2}{|\boldsymbol{r_i}-\boldsymbol{r_j}|}$ This gives

$$K = \frac{1}{2}\frac{1}{4\pi\epsilon_0}\sum_{\substack{j\\ 1\neq j}}\int d^3\boldsymbol{r_1}..d^3\boldsymbol{r_N}\frac{e^2}{|\boldsymbol{r_1}-\boldsymbol{r_j}|}\chi\chi^* \tag{7.80}$$

$$= \frac{1}{2}\frac{1}{4\pi\epsilon_0}\sum_{\substack{j\\ 1\neq j}}\int d^3\boldsymbol{r_1}d^3\boldsymbol{r_j}\frac{e^2}{|\boldsymbol{r_1}-\boldsymbol{r_j}|}\rho_e(\boldsymbol{r_1})\rho_x(\boldsymbol{r_j}) \tag{7.81}$$

Where $\rho_e = \phi_e\phi_e^*$ and x is the state associated with $\boldsymbol{r_j}$. As we sum over j we will find x taking all the indices $a, b, c, d, f, ...$ except (e). Again the result is independent of our choice of the electron co-ordinate $\boldsymbol{r_1}$

Hence we can write

$$\begin{aligned}&\sum_i\left\langle\chi\left|\frac{\boldsymbol{p_i}^2}{2m}+\sum_f V_f(\boldsymbol{r_i})+\frac{1}{2}.\frac{1}{4\pi\epsilon_0}\sum_{\substack{j\\ i\neq j}}\frac{e^2}{|\boldsymbol{r_i}-\boldsymbol{r_j}|}\right|\chi\right\rangle\\ =&\sum_i(E_0+K)\\ =&N\tilde{E}_0\end{aligned} \tag{7.82}$$

But we started picking an arbitrary permutation *egbdf*... So the same arguments work for each of the $N!$ permutations. The sum over all the $N!$ permutations then cancels exactly with the $\frac{1}{\sqrt{N!}}\frac{1}{\sqrt{N!}}$ arising from the normalisation of Φ.

So the energy expectation from the direct terms, including the coulomb part scales with N as one would expect.

Because the number of site indices and the number of electron co-ordinates are the same there can be no pair of terms (one from the ket side , one from the bra side) that differ only in one position.

But there can be terms which differ in two places leaving the other $N-2$ terms same. To evaluate these terms we will group the hamiltonian differently

$$\begin{aligned}H =& \sum_i\frac{\boldsymbol{p_i}^2}{2m}\\ &+\sum_{\langle ij\rangle\ \text{pair}}\left(\frac{1}{4\pi\epsilon_0}\frac{e^2}{|\boldsymbol{r_i}-\boldsymbol{r_j}|}\right)+\sum_f\sum_i V_f(\boldsymbol{r_i})\end{aligned} \tag{7.83}$$

The expectation of the first part in any cross term is zero. This is because there would be at least one ϕ state in the ket side that will hit a different state in the bra side.

for the second part, the $N-2$ state (sites) other than those which go with $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are identical on both sides. Let's assume we have

- $\phi_m(\boldsymbol{r_1})\phi_n(\boldsymbol{r_2})$ on the ket side
- $\phi_n^*(\boldsymbol{r_1})\phi_m^*(\boldsymbol{r_2})$ on the bra side

There are $N(N-1)$ ways of choosing these states. The value of the term will depend on the (m, n) pair that is chosen.

$$J_{mn} = \left\langle \phi_n(\boldsymbol{r_1})\phi_m(\boldsymbol{r_2}) \left| \frac{1}{4\pi\epsilon_0}\frac{e^2}{|\boldsymbol{r_1}-\boldsymbol{r_2}|} + \sum_f V_f(\boldsymbol{r_1}) + \sum_f V_f(\boldsymbol{r_2}) \right| \phi_m(\boldsymbol{r_1})\phi_n(\boldsymbol{r_2}) \right\rangle \tag{7.84}$$

We retain only the potential arising from the ions at the m and n sites. Hence

$$J_{mn} = \left\langle \phi_n(\boldsymbol{r_1})\phi_m(\boldsymbol{r_2}) \left| \frac{1}{4\pi\epsilon_0}\frac{e^2}{|\boldsymbol{r_1}-\boldsymbol{r_2}|} + V_m(\boldsymbol{r_1}) + V_n(\boldsymbol{r_1}) + V_m(\boldsymbol{r_2}) + V_n(\boldsymbol{r_2}) \right| \phi_m(\boldsymbol{r_1})\phi_n(\boldsymbol{r_2}) \right\rangle \tag{7.85}$$

J_{mn} will depend on the relative locations of m and n. If the electron states are spherically symmetric like s states, then they will depend on the distance alone but for d or f orbitals the relative orientation will also be a factor.

Now it may appear there is some inconsistency in claiming that single particle operators cannot contribute to terms where the bra and ket states differ in two locations and still retaining the the single particle potential in the expression. If the states were exactly orthogonal this would be correct, but as an exercise you can try to figure out how the small non-orthogonality plays a role.

The next figure (fig. 7.5 shows how many ways the exchange term can be chosen. The $N!$ factor will again cancel with the normalisation. Then, because J_{mn} also falls off with distance ($|m-n|$), we will retain the nearest neighbour exchange only. We can retain more terms, but this will not change the conclusions.

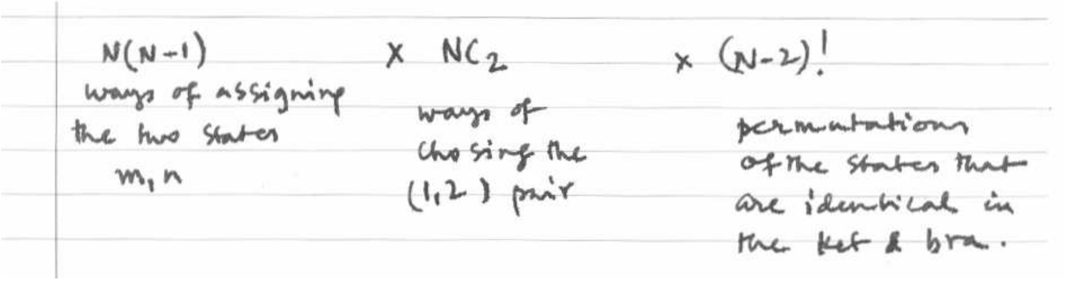


Figure 7.5: No of ways of the exchange term can be constructed.

If we retain only nearest neighbour terms, then the first factor $N(N-1) \to Nz$ where z is the number of nearest neighbours in the lattice.

The numerical value of J can be positive or negative depending on a lot of fine details of the wavefunction and the potential.

Also we can absorb the factor of 1/2 in the definition of J

Now comes a crucial point. These terms will all come with a negative sign in the energy expectation. This is because $\phi_m\phi_n$ appears with their order reversed on two sides. *This means in the expansion of the determinant, they differ by exactly one "exchange" and hence must occur with opposite sign.* So finally we can write for the energy of a state with all spins pointing in the same direction:

$$\langle\Phi|H|\Phi\rangle = N(\tilde{E}_0 - zJ) \tag{7.86}$$

We need not consider terms with three (or more) positions differing in the bra and ket. Because we only have potentials (operators) that involve the co-ordinate of two particles at once, there will be some electron co-ordinate that will "pass through" and find itself associated with a different state (site) on the other side. This will make such terms zero or very small. So we do not look for terms any more complex than the exchange terms.

7.7 A state with one spin flipped

Now suppose we have a state with one site spin (say site labelled f) flipped. How does the energy expectation change? The state is

$$\Phi_f = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{a\uparrow}(\boldsymbol{r_1}) & \phi_{a\uparrow}(\boldsymbol{r_2}) & .. & .. & \phi_{a\uparrow}(\boldsymbol{r_N}) \\ \phi_{b\uparrow}(\boldsymbol{r_1}) & \phi_{b\uparrow}(\boldsymbol{r_2}) & .. & .. & \phi_{b\uparrow}(\boldsymbol{r_N}) \\ .. & .. & .. & .. & .. \\ \phi_{f\downarrow}(\boldsymbol{r_1}) & \phi_{f\downarrow}(\boldsymbol{r_2}) & .. & .. & \phi_{f\downarrow}(\boldsymbol{r_N}) \\ .. & .. & .. & .. & .. \\ \phi_{N\uparrow}(\boldsymbol{r_1}) & \phi_{N\uparrow}(\boldsymbol{r_2}) & .. & .. & \phi_{N\uparrow}(\boldsymbol{r_N}) \end{vmatrix} \tag{7.87}$$

$$\langle \Phi_f | H | \Phi_f \rangle = ?$$

The first set of terms, which have the orbitals appearing in the same order in ket and bra, will remain the same. The downspin state will hit a downspin state.

The exchange integral will vanish if either of the states m or n contains the flipped spin. The hamiltonian contains no spin-flipping term, the spin co-ordinate just passes through. Hence the down spin associated with $\boldsymbol{r_1}$ say will find itself with an upspin state. This orthogonality is exact, unlike the site to site (approximate) orthogonality.

So some of the terms which were being subtracted before are no longer going to be subtracted. Their number is equal to the number of the nearest neighbours of the flipped spin state. We thus have:

$$\langle \Phi_f | H | \Phi_f \rangle = \langle \Phi | H | \Phi \rangle + zJ \tag{7.88}$$

If $J > 0$ then the all spin up state lies lower in energy, even though there is no magnetic interaction at all in the hamiltonian. This will be the key element in the explanation of ferromagnetism.

But this excites state is highly degenerate, since we could have flipped any one of the N spins. So we must make a linear combination of all the possibilites and write:

$$\Psi = \sum_f a_f \Phi_f \tag{7.89}$$

Now, to solve:

$$H\Psi = E\Psi \tag{7.90}$$

$$H \sum_{f'} a_{f'} \Phi_{f'} = E \sum_{f'} a_{f'} \Phi_{f'} \tag{7.91}$$

$$\sum_{f'} a_{f'} \langle \Phi_{f'} | H | \Phi_f \rangle = E \sum_{f'} a_{f'} \underbrace{\langle \Phi_f | \Phi_{f'} \rangle}_{\approx \delta_{ff'}} \tag{7.92}$$

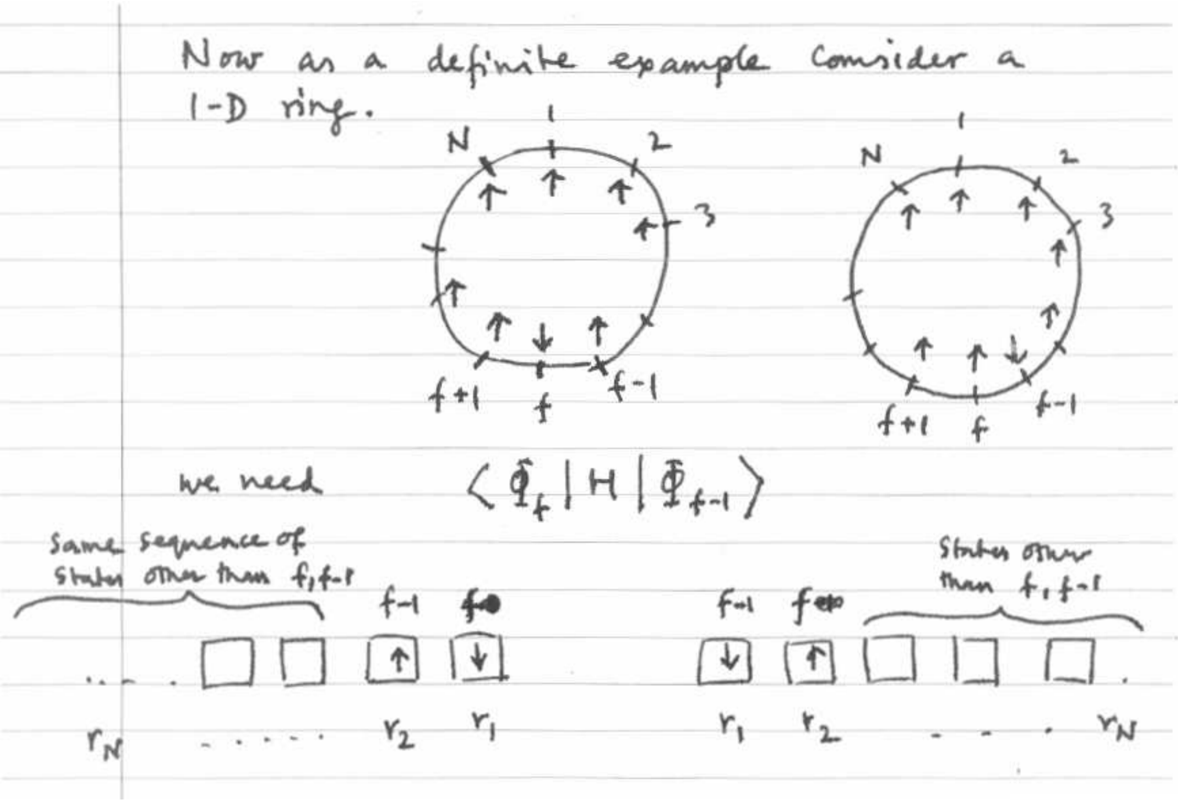
We would then need to solve a linear set of equations, to get the coefficients and the eignevector, but first we need to calculate $\langle \Phi_{f'} | H | \Phi_f \rangle$.

Since the flipped spin is at different sites in the ket and the bra, it implies that the sequence of terms must differ in at leat two places, because unless f hits f' the spin-spin orthogonality will make it zero. But to ensure that we need to make at least one place different (f in place of f'). Since no term with one place different can contribute, we need to look into only those terms (exchange) with two places different.

We have already shown that it the sequences differ at two places, then the $\frac{\boldsymbol{p}^2}{2m}$ type terms cannot contribute.

The exchange terms are not too small only if f and f' are nearest neighbours.

Now consider a 1D ring, so that there are only two nearest neighbours:

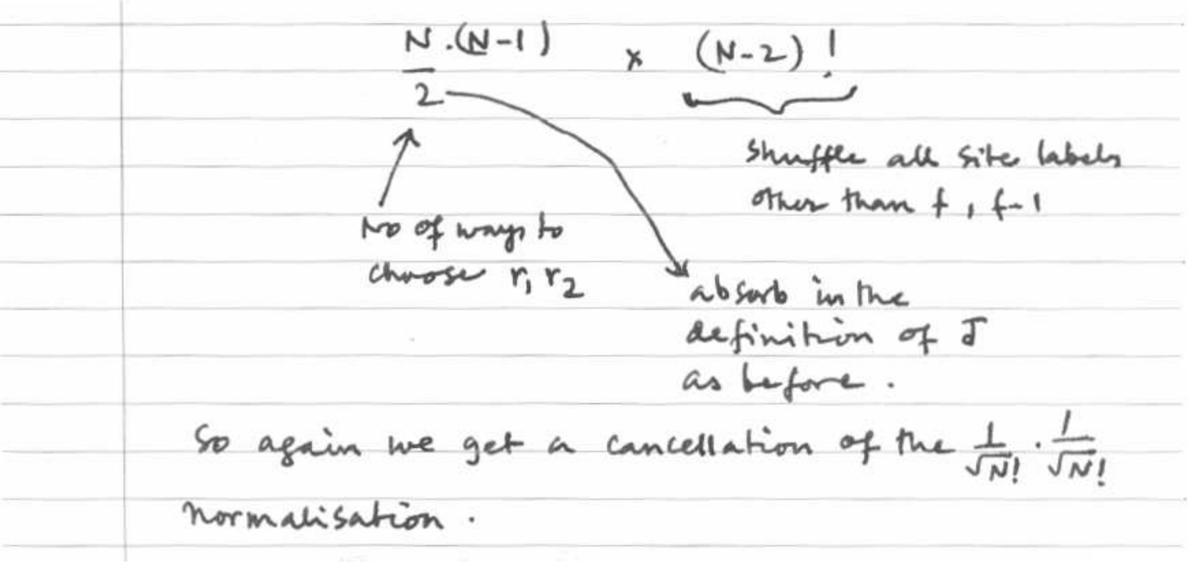


Now as a definite example consider a 1-D ring.

we need $\langle \Phi_f | H | \Phi_{f-1} \rangle$

Figure 7.6: The only exchange terms between $\langle \Phi_{f'} | H | \Phi_f \rangle$ will be non-zero if $f' = f \pm 1$.

Again the choice of coordinates can be made in ${}^N C_2$ ways after f is chosen. The next figure (fig 7.7)shows the cancellations with normalisation etc.



$$\frac{N.(N-1)}{2} \times \underbrace{(N-2)!}$$

So again we get a cancellation of the $\frac{1}{\sqrt{N!}} \cdot \frac{1}{\sqrt{N!}}$ normalisation.

Figure 7.7: The counting of the number of terms....

So finally we get :

$$\begin{aligned} \langle\Phi_{f'}|H|\Psi\rangle &= -a_{f-1}J - a_{f+1}J && (7.93) \\ \langle\Phi_f|H|\Phi_f\rangle = &= \langle\Phi|H|\Phi\rangle + 2J && (7.94) \end{aligned}$$

Thus the set of linear equations again reduce to the "circulant" type matrix that is typical of many "nearest neighbour" coupling situations. The eigenvectors are formed of N^{th} roots of unity.

$$(E - E_0 - 2)a_f + (a_{f-1} + a_{f+1})\,J \tag{7.95}$$

Hence with $k = \dfrac{2\pi}{N}j$,

$$\begin{aligned} a_f &= e^{ikfa} && (7.96) \\ \Psi &= \sum_f e^{ikfa}\Phi_f && (7.97) \\ E - E_0 &= 2J(1 - \cos ka) && (7.98) \end{aligned}$$

These excited states are called spin wave states. The appearance of k makes the origin of the "wave" aspect clear. In the later sections we will see that the "equivalent" Heisenberg spin hamiltonian reproduces the same result, provided the exchange parameter J is used as a phenomenological input. Within the "spin hamiltonian" there is no explanation of the origin of J.

7.8 Magnitude of J : Why is Iron ferromagnetic?

The answer to this apparently very basic question is not very straightforward. Let's recall the expression for J_{mn} in equation 7.85

$$J_{mn} = \left\langle \phi_n(\boldsymbol{r_1})\phi_m(\boldsymbol{r_2}) \left| \frac{1}{4\pi\epsilon_0}\frac{e^2}{|\boldsymbol{r_1} - \boldsymbol{r_2}|} + V_m(\boldsymbol{r_1}) + V_n(\boldsymbol{r_1}) + V_m(\boldsymbol{r_2}) + V_n(\boldsymbol{r_2}) \right| \phi_m(\boldsymbol{r_1})\phi_n(\boldsymbol{r_2}) \right\rangle$$

The interaction between the electrons give positive contribution (the first term) the other attractive terms would give negative results. If the charge is somewhat far from the nucleii and concentrated in the in-between region then we may expect the result to be positive. However what is true for direct "charge" like terms ($\rho = \phi^*\phi$) is an indication of what may happen for the "cross terms" and not a certainty. d and f shells fulfill this criteria better - but only a full numerical calculation can give the answer. Broadly speaking we require:

- Atoms with incomplete shells of high azimuthal quantum number
- The atoms should not be too close - then the negative contribution of the attractive potential would dominate.
- The atoms should not be too far - then there can be very little overlap of the orbitals.
- This fine balance between the attractive and repulsive terms determine the final value of J, it turns out, among the elements, it is best satisfied by Iron, Nickel and Cobalt.

The ferromagnetic transition temperature

When there are lots of excitations above the ground state then the ferromagnetic state will be destroyed. Since the energy of these excitations are related to the exchange integral, the ferromagnetic T_c is ultimately related to the strength of this quantity. Loosely speaking $I \sim k_B T_c$. See table 7.1.

7.8.1 The Heisenberg spin hamiltonian finally

All the discussion can now summarised in a single and remarkably simple expression. Since each spin flip of the electron on the atomic site costs energy we write including an external field in the z-direction.

$$H = \sum_{ij} -J_{ij}\boldsymbol{S^i}.\boldsymbol{S^j} - \mu_B B_0 \sum_i S_z^i \tag{7.99}$$

- In the simplest case i, j are restricted to being nearest neighbours.
- But the interaction can spread beyond that. It is entirely possible for the J_{ij} to change sign as a function of distance.
- In such a case the nearest neighbour interaction is ferromagnetic but the next nearest neighbour interaction can be antiferromagnetic.
- This kind of competing interactions can give rise to a very rich variety of magnetic ground states.

Excited sates of the Heisenberg hamiltonian : Spin waves, magnons

Recall that spin and angular momentum have similar quantum mechanical properties. We do the analogue of the the determinant eqn. 7.89, in terms of the spin raising and spin lowering operator. The combinations

$$\begin{aligned} S_+ &= S_x + iS_y && (7.100)\\ S_- &= S_x - iS_y && (7.101)\end{aligned}$$

have the properties that they change the S_z value of an angular momentum state by one unit ($\hbar$). The minimum and maximum values of s_z are $-S$ and S in steps of one. The *magnitude* of the total spin is $S^2 = S(S+1)$, as usual

$$\begin{aligned} S_+(\boldsymbol{R})|S_z\rangle_{\boldsymbol{R}} &= \sqrt{(S-S_z)(S+1+S_z)}|S_z+1\rangle_{\boldsymbol{R}} && (7.102)\\ S_-(\boldsymbol{R})|S_z\rangle_{\boldsymbol{R}} &= \sqrt{(S+S_z)(S+1-S_z)}|S_z-1\rangle_{\boldsymbol{R}} && (7.103)\end{aligned}$$

The Hamiltonian can be written as

$$\begin{aligned} H &= \frac{1}{2}\sum_{\boldsymbol{R},\boldsymbol{R'}} -J(\boldsymbol{R}-\boldsymbol{R'})S_{\boldsymbol{R}}.\boldsymbol{S}_{\boldsymbol{R'}}\\ &= \frac{1}{2}\sum_{\boldsymbol{R},\boldsymbol{R'}} -J(\boldsymbol{R}-\boldsymbol{R'})\left[\frac{1}{2}S_-(\boldsymbol{R})S_+(\boldsymbol{R'}) + \frac{1}{2}S_+(\boldsymbol{R})S_-(\boldsymbol{R'}) + S_z(\boldsymbol{R})S_z(\boldsymbol{R'})\right]\\ &= \frac{1}{2}\sum_{\boldsymbol{R},\boldsymbol{R'}} -J(\boldsymbol{R}-\boldsymbol{R'})\left[S_+(\boldsymbol{R})S_-(\boldsymbol{R'}) + S_z(\boldsymbol{R})S_z(\boldsymbol{R'})\right] && (7.104)\end{aligned}$$

The ground state is

$$|0\rangle = \prod_{\boldsymbol{R}} |S\rangle_{\boldsymbol{R}} \tag{7.105}$$

where S is the maximum eigenvalue of S_z, hence

$$S_z(\boldsymbol{R}) = S|S\rangle_{\boldsymbol{R}} \tag{7.106}$$

Now construct a normalised low lying state differing from the ground state only at the site $\boldsymbol{R}$ by one unit as:

$$|\boldsymbol{R}\rangle = \frac{1}{\sqrt{2S}}S_-(\boldsymbol{R})|0\rangle \tag{7.107}$$

and the spin wave state, using the expression 7.89 as

$$|\boldsymbol{k}\rangle = \frac{1}{\sqrt{N}}\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}.\boldsymbol{R}}|\boldsymbol{R}\rangle \tag{7.108}$$

PROBLEM: Using eqn 7.103 and 7.102 show that

$$S_{-}(\boldsymbol{R'})S_{+}(\boldsymbol{R})|\boldsymbol{R}\rangle = 2S|\boldsymbol{R'}\rangle \tag{7.109}$$

This simple results can be used to work with the hamiltonian 7.104 which has been reduced to a very similar form.
Show that the ground sate energy of H as in eqn. 7.104 is

$$E_0 = -S^2 \sum_{\boldsymbol{R},\boldsymbol{R'}} J(\boldsymbol{R}-\boldsymbol{R'}) \tag{7.110}$$

The two results eqn 7.109 and 7.110 can now be used together to sum over all sites in the Hamiltonian to show that

$$H|\boldsymbol{R}\rangle = E_0|\boldsymbol{R}\rangle + S\sum_{R'} J(\boldsymbol{R}-\boldsymbol{R'})[|\boldsymbol{R}\rangle - |\boldsymbol{R'}\rangle] \tag{7.111}$$

The problem shows that the action of the Hamiltonian takes a flipped spin from the site $\boldsymbol{R}$ to $\boldsymbol{R'}$. We have encountered this situation a few times before (in writing Bloch states for example). Thus a linear combination of the $|\boldsymbol{R}\rangle$ can be used to write (excited) eignestates of the Hamiltonian. It turns out that the combination is as given in equation 7.108. It is not necessary to assume this form, it can be proved, but we would assume this form and go ahead.
Using the result eqn. 7.111 we can write:

$$\begin{aligned}
H\left(\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}.\boldsymbol{R}}|\boldsymbol{R}\rangle\right) &= E_0\left(\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}.\boldsymbol{R}}|\boldsymbol{R}\rangle\right) + S\sum_{\boldsymbol{R},\boldsymbol{R'}} J(\boldsymbol{R}-\boldsymbol{R'})e^{i\boldsymbol{k}.\boldsymbol{R}}\left(|\boldsymbol{R}\rangle - |\boldsymbol{R'}\rangle\right) && (7.112)\\
&= E_0\left(\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}.\boldsymbol{R}}|\boldsymbol{R}\rangle\right) + \sum_{\boldsymbol{X}=\boldsymbol{R}-\boldsymbol{R'}}\left(1-e^{i\boldsymbol{k}.\boldsymbol{X}}J(\boldsymbol{X})\right)\left(\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}.\boldsymbol{R}}|\boldsymbol{R}\rangle\right) && (7.113)\\
\therefore H|\boldsymbol{k}\rangle &= \sum_{\boldsymbol{X}}\left(1-e^{i\boldsymbol{k}.\boldsymbol{X}}J(\boldsymbol{X})\right)|\boldsymbol{k}\rangle && (7.114)\\
\therefore E(\boldsymbol{k}) &= \sum_{\boldsymbol{X}} J(\boldsymbol{X})\sin^2\frac{\boldsymbol{k}.\boldsymbol{X}}{2} && (7.115)
\end{aligned}$$

The intermediate steps are left as as exercise. The energy of the spin wave state can thus be reduced to a sum over the exchange integrals. The summation can be done once the lattice structure is known. Often we can restrict the sum the just over nearest neighbours.

7.8.2 Decrease of magnetisation of the excited states

Notice the similarity of this energy dispersion relation with phonons. The similarity is bit more than mathematical- both these states (lattice waves 7 spin waves) describe bosonic excitations about a ground state. We call the magnetic excitation a "magnon". The reduction of magnetisation from the saturation value (at $T=0$, $M(0)=NS$) can be understood as the total number of magnons present in the system. This is given by summing over the magnon population following a Bose-Einstein distribution

$$\begin{aligned}
\frac{M(T)}{M(0)} &= 1-\frac{1}{NS}\sum_{\boldsymbol{k}} n(\boldsymbol{k}) && (7.116)\\
&= 1-\frac{V}{NS}\int_0^\infty \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\frac{1}{e^{E(\boldsymbol{k})/k_BT}-1} && (7.117)\\
&= 1-\frac{V}{NS}(k_BT)^{3/2}\frac{1}{4\pi^2\alpha^{3/2}}\int_0^\infty \frac{\sqrt{x}}{e^x-1}\mathrm{d}x && (7.118)
\end{aligned}$$

The final expression assumes that the temperature is small. We have written $E = \alpha k^2$ for small $|\boldsymbol{k}|$ by aproximating the dispersion relation. This $T^{3/2}$ reduction of saturation magnetisation is well verified experimentally and is often called Bloch's $T^{3/2}$ law.

Antiferromagnetic and ferromagnetic magnons

To be written

Further simplifications of the Heisenberg spin hamiltonian: the Ising model

To be written

Dimensionality, magnetic order and phase transitions

To be written

Magnetic domain: how does it occur

To be written

Hysteresis in ferromagnets

To be written

Appendix A

Fermi Surface, electrons and holes

We have seen in earlier chapters that most of the response of an electron gas in a substance comes from the region around the Fermi energy. This means that the shape and structure of the Fermi surface is something very crucial. The problem is also instrinsically linked to the band-structure problem, since the filling of the levels (upto the Fermi energy) requires a knowledge of the density of states. That is why we are dealing with the fermi surface shape after doing a few things with calculating the band structure.

We would again start by drawing the free electron fermi surfaces - but in cases where it spans more than one Brillouin zone. What is the utility of plotting free electron surfaces? The answer is that in many cases the band structure of metals differs significantly from the free electron result only near the zone boundaries. Thus, if one can draw the surfaces for the free electron case, then with relatively minor tweaks around the zone boundaries, one can make a reasonable sketch of how it will look for a real metal. That is why we do it.

A.1 The free electron "squarium"

Consider a hypothetical 2D "metal" with one, two or three electrons per atom. The real space lattice is a square. What would its fermi circle look like? Let us assume that the lattice constant is a. Then its reciprocal lattice is also a square with lattice constant $2\pi/a$. Each $\boldsymbol{k}$-state can accommodate two electrons - one of each spin. If we have x electrons per atom then the radius of the Fermi circle is (show this as a simple problem)

$$k_F = \frac{1}{a}\sqrt{2\pi x} \tag{A.1}$$

Depending on the value of x it can span multiple Brillouin zones. The interpretation of the next figure A.1 is left to you as a simple exercise. There is no reason to believe that x must be an integer. Using "dopants" it is possible to create a situation where there are 1.5 electrons per unit cell or something like that. However, if we restrict ourselves to simple monatomic lattice, then we can have only integer values of x.

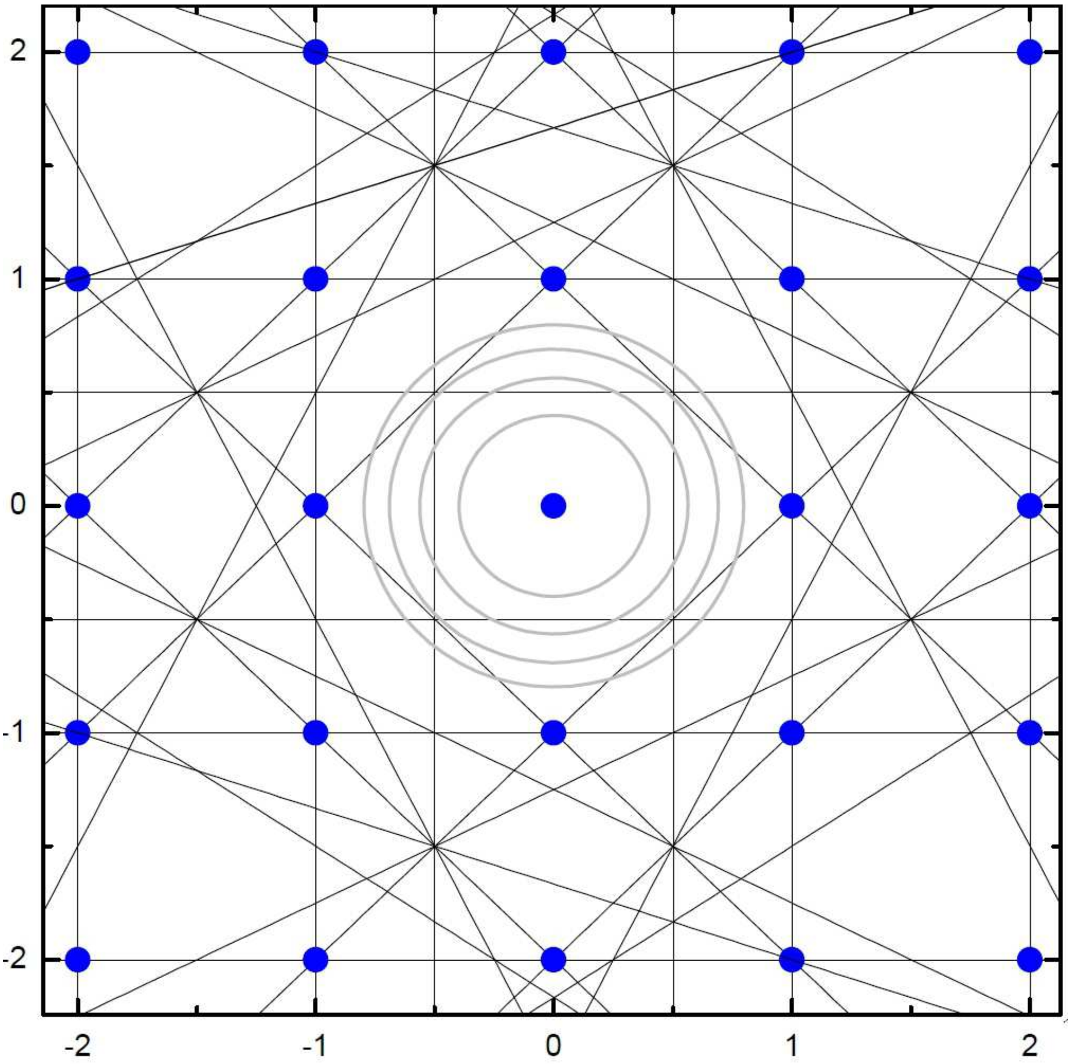


Figure A.1: The Fermi circles for 1,2,3,4 electrons per site. Notice how the circle spans into several Brillouin zones. You should be able to interpret the meaning of the various points and lines.

A.1.1 1 electron per atom

In this case the Fermi circle is fully contained in the first zone. Fig A.2 shows this.

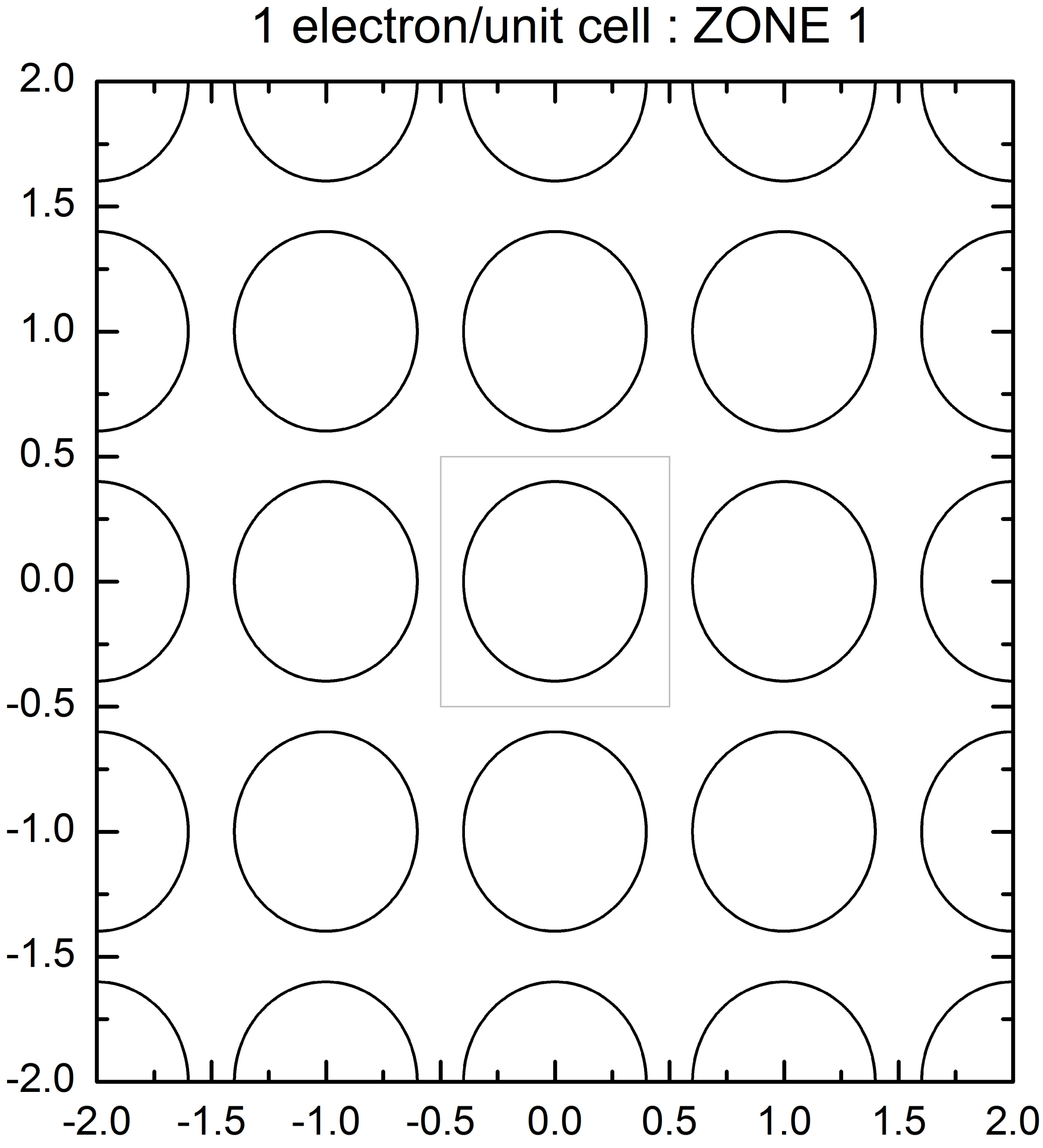


Figure A.2: The Fermi circles for 1 electrons per site. All occupied states are in the first zone. The k_x and k_y axes are in units of $2\pi/a$.

A.1.2 2 electrons per atom

In this case the Fermi circle spread over first and second zone. Fig. A.3 and Fig. A.4 show this. Notice how the branches can be translated back into the first zone if we want. Using the figure, you can deduce what reciprocal lattice vector translation would do the job for various segments.

A point to notice is that the branch in zone 2 starts filling from the periphery towards the centre. We will see later that such an arrangement of the states, where the occupied states enclose the unoccupied higher energy states lead to a strikingly different behaviour of the electrons. By comparing the figures A.4, A.5 and A.6 you can see how the BZ2 branch grows inwards as the number of electrons increases.

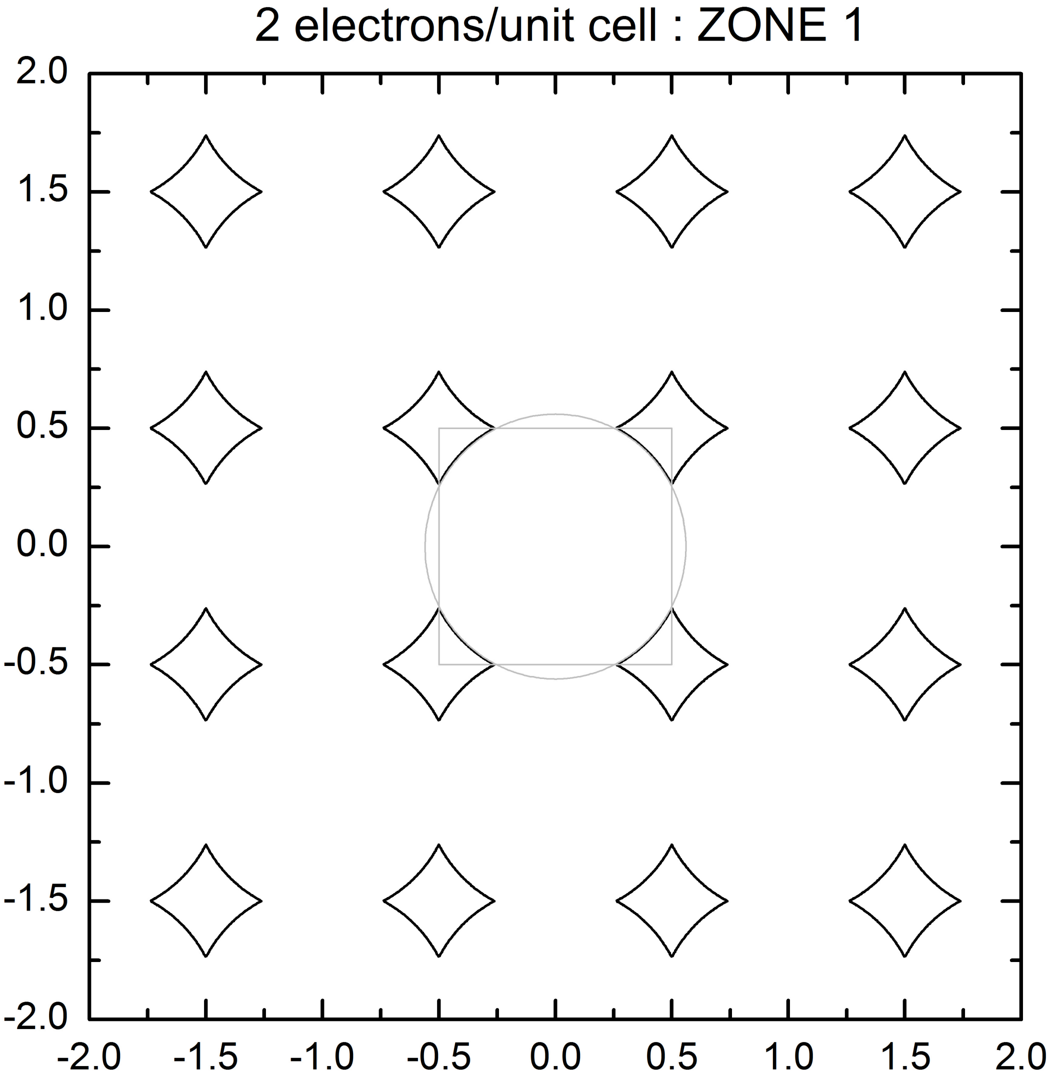


Figure A.3: The Fermi circles for 2 electrons per site. Occupied states are in the first and second zone. The k_x and k_y axes are in units of $2\pi/a$.

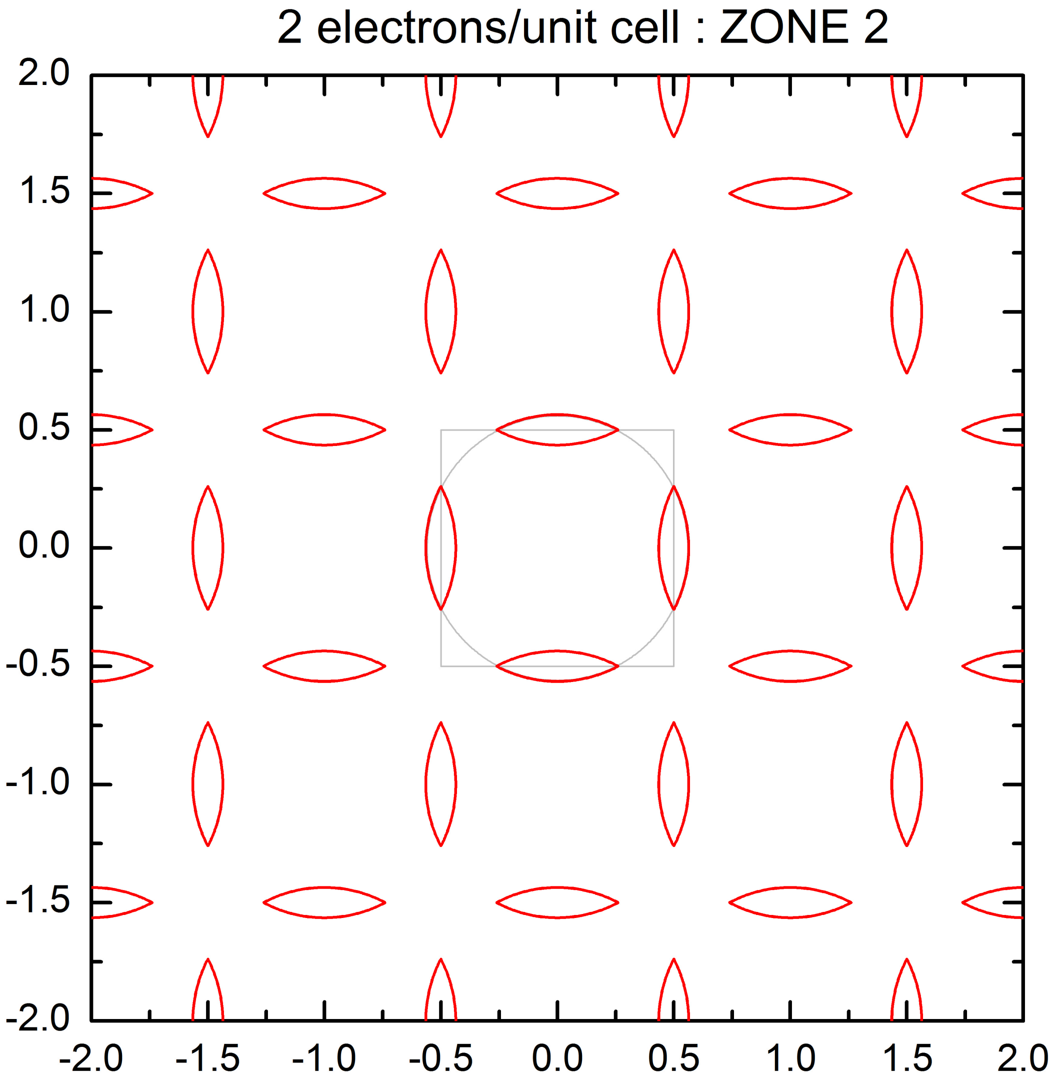


Figure A.4: The Fermi circles for 2 electrons per site. The parts in the second zone. The k_x and k_y axes are in units of $2\pi/a$.

A.1.3 3 electrons per atom

In this case the Fermi circle fills the first zone completely and spreads over the second zone. Fig. A.5 and Fig. A.5 show this. Notice how the branches can be translated back into the first zone if we want. Using the figure, you can deduce what reciprocal lattice vector translation would do the job for various segments.

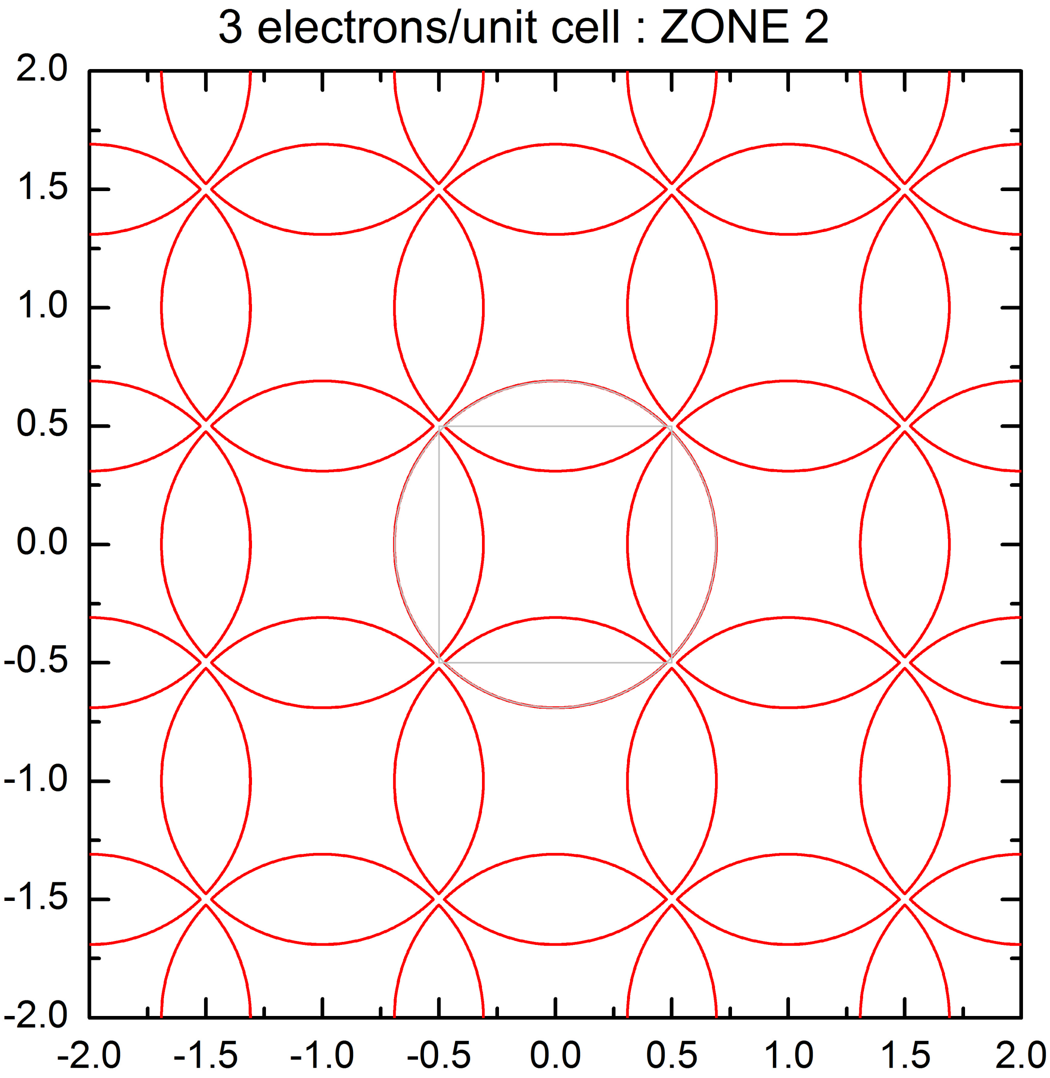


Figure A.5: The Fermi circles for 3 electrons per site. The parts in the second zone. The k_x and k_y axes are in units of $2\pi/a$.

A.2 4 electrons per atom

The fermi circle now covers BZ1 fully. BZ2, BZ3 and BZ4 are partially filled as the figures A.6, A.7 and A.8 show.

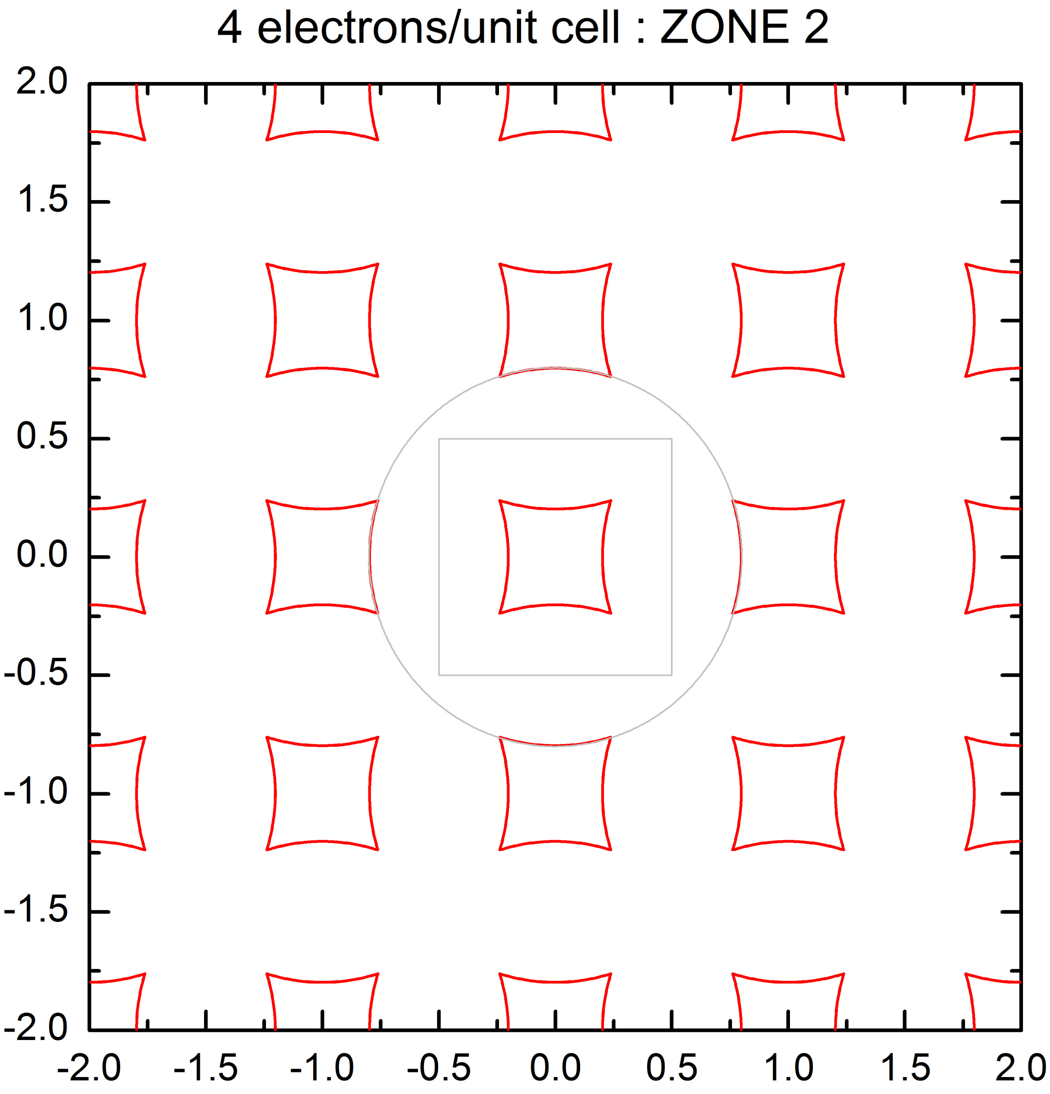


Figure A.6: The Fermi circles for 4 electrons per site. The parts in the second zone. The k_x and k_y axes are in units of $2\pi/a$.

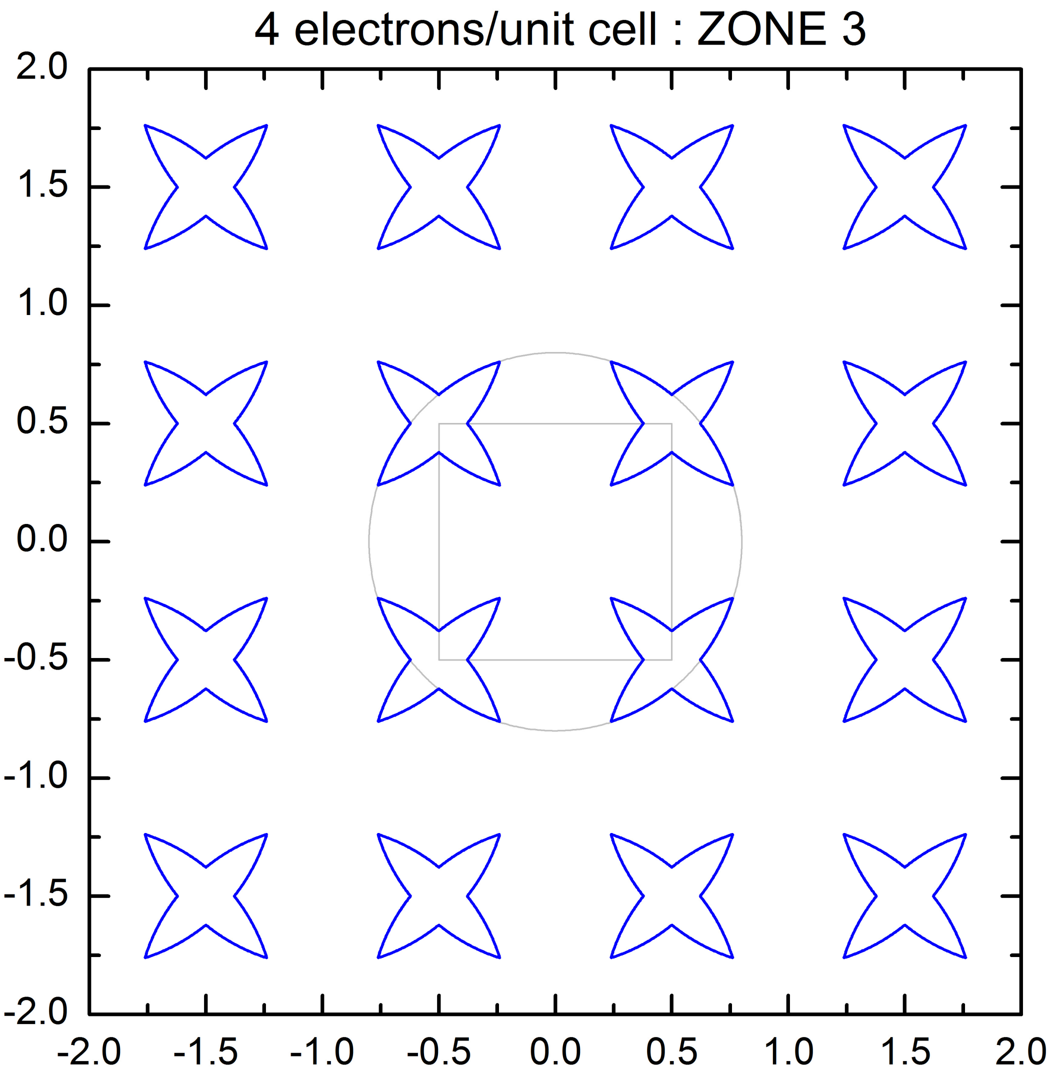


Figure A.7: The Fermi circles for 4 electrons per site. The parts in the third zone. The k_x and k_y axes are in units of $2\pi/a$.

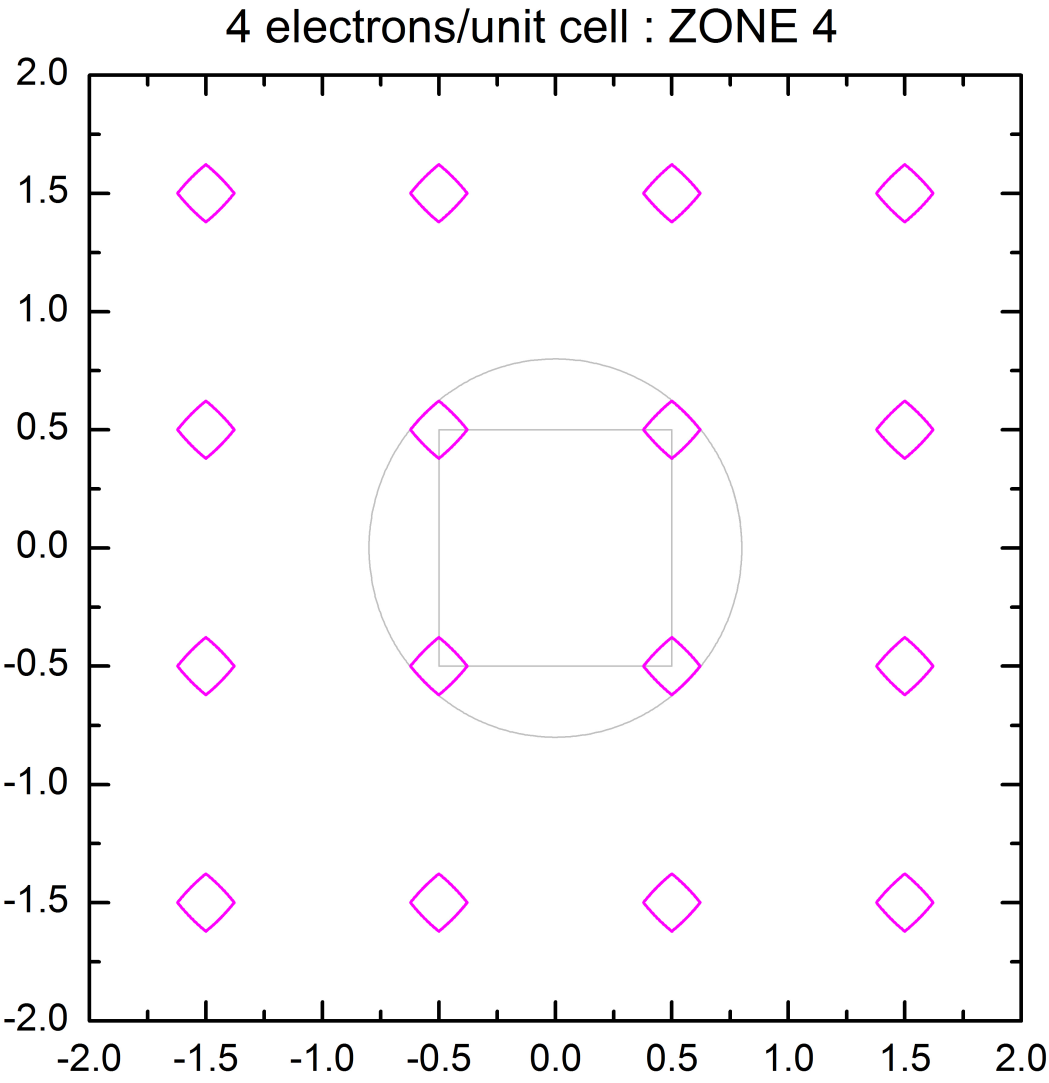


Figure A.8: The Fermi circles for 4 electrons per site. The parts in the fourth zone. The k_x and k_y axes are in units of $2\pi/a$.

A.3 Doing the same in 3D

We can now work out how things will be in 3D. We will look at a FCC (real space) lattice with 1, 2 and then 3 electrons per atom. If the conventional cubic lattice constant is a, then the sides of the conventional cube in reciprocal space would be $4\pi/a$.

A.3.1 1 electron per atom

In this case the fermi sphere is within the first zone.

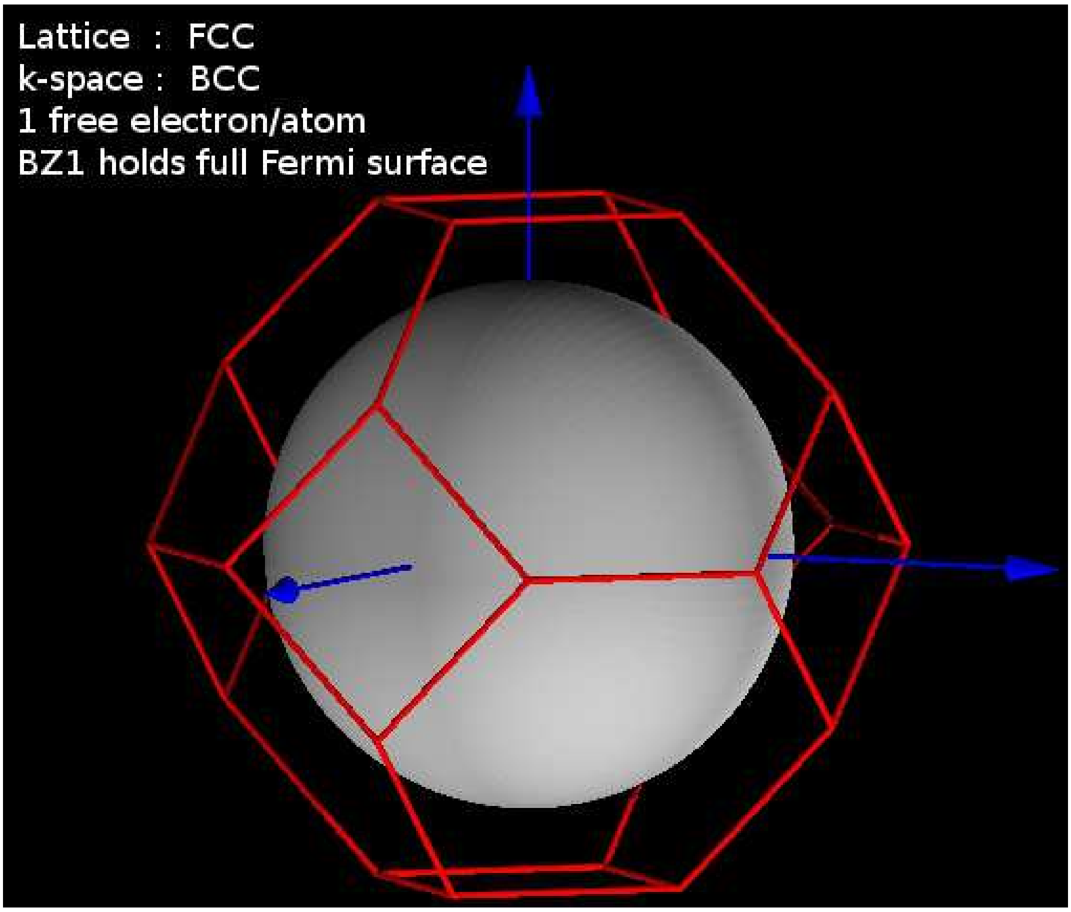


Figure A.9: The Fermi sphere for 1 electron per site.

A.3.2 2 electrons per atom

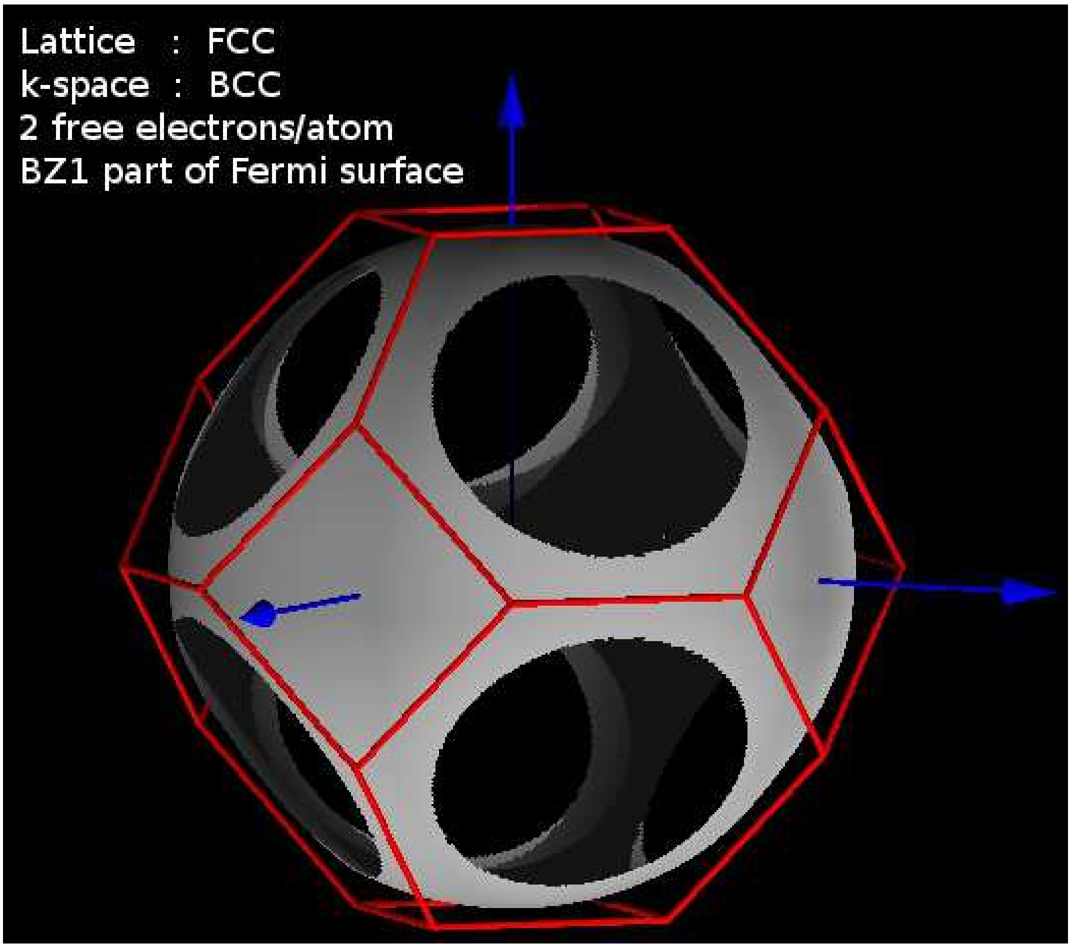


Figure A.10: The Fermi sphere for 2 electrons per site - the BZ1 part.

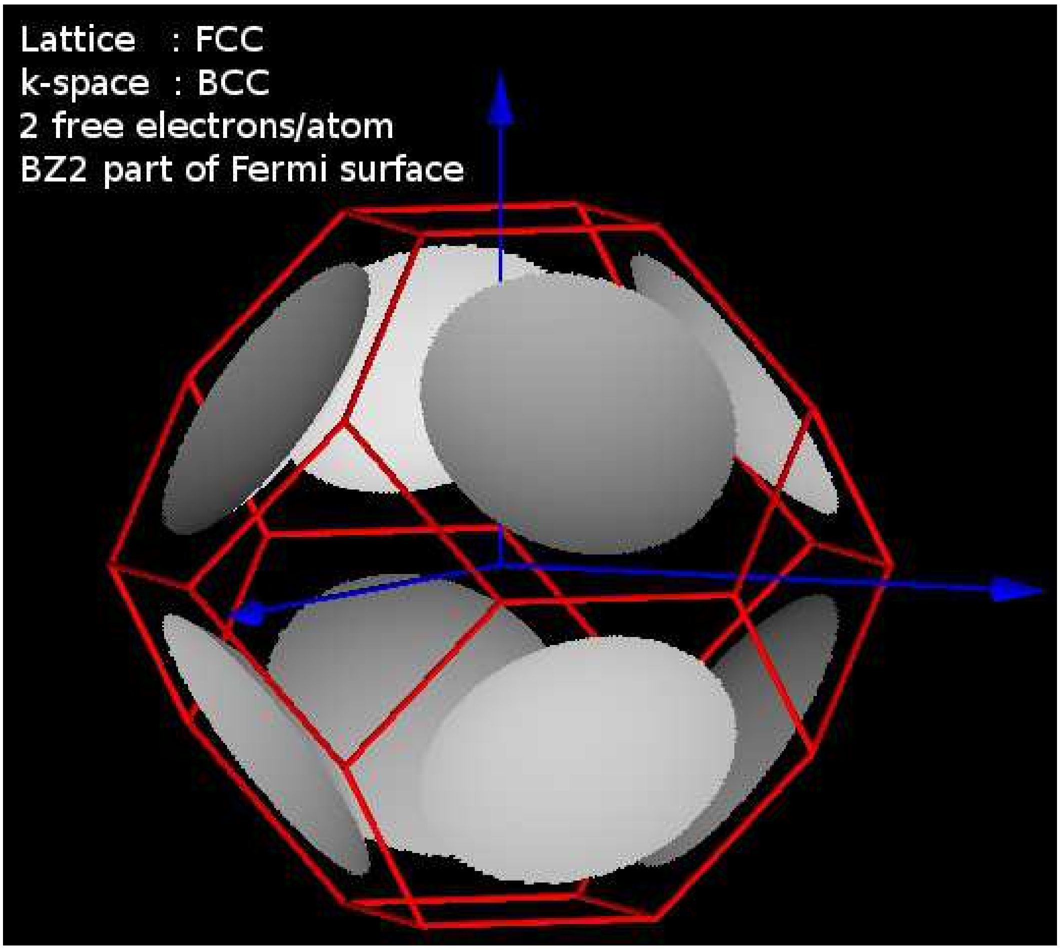


Figure A.11: The Fermi sphere for 2 electrons per site - the BZ2 part, translated back to BZ1.

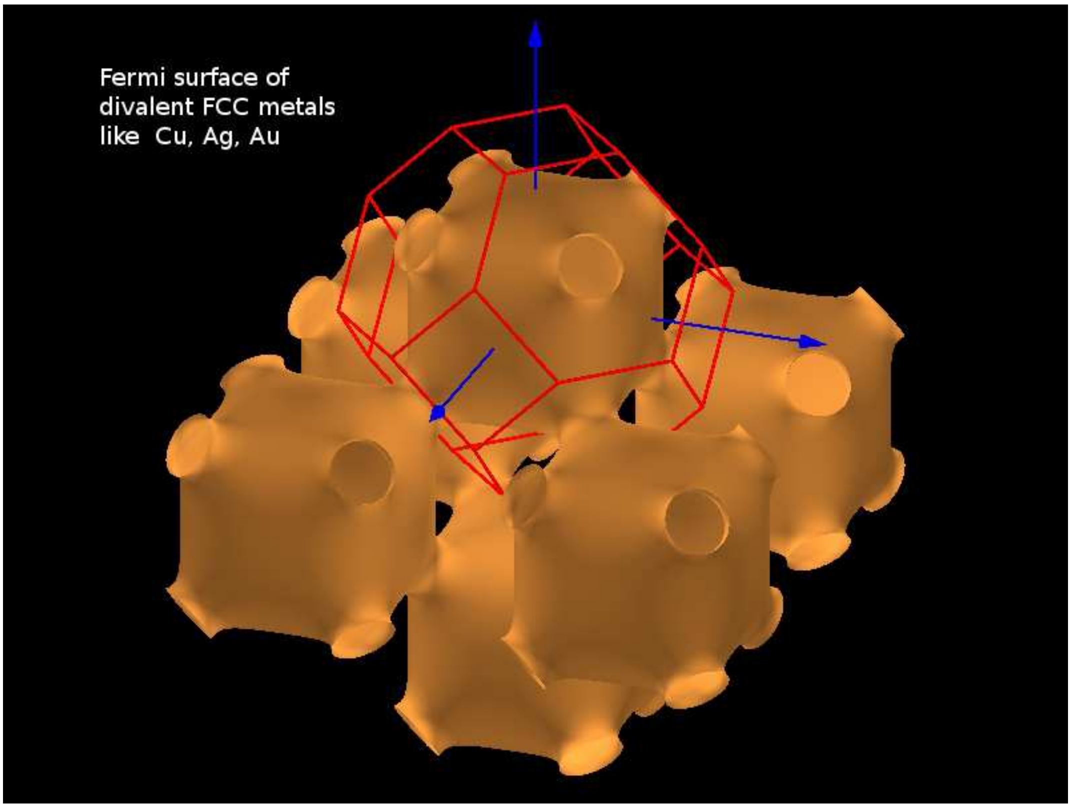


Figure A.12: The more realistic Fermi sphere for a divalent metal like Cu, Ag, Au. The most important feature is the "necking" near the hexagonal faces of the zone boundary.

In reality the divalent metals have a necking of the fermi sphere near the hexagonal faces of the zone boundary rather than a clean cut of the fermi sphere. See the figure A.12.

A.3.3 3 electrons per atom

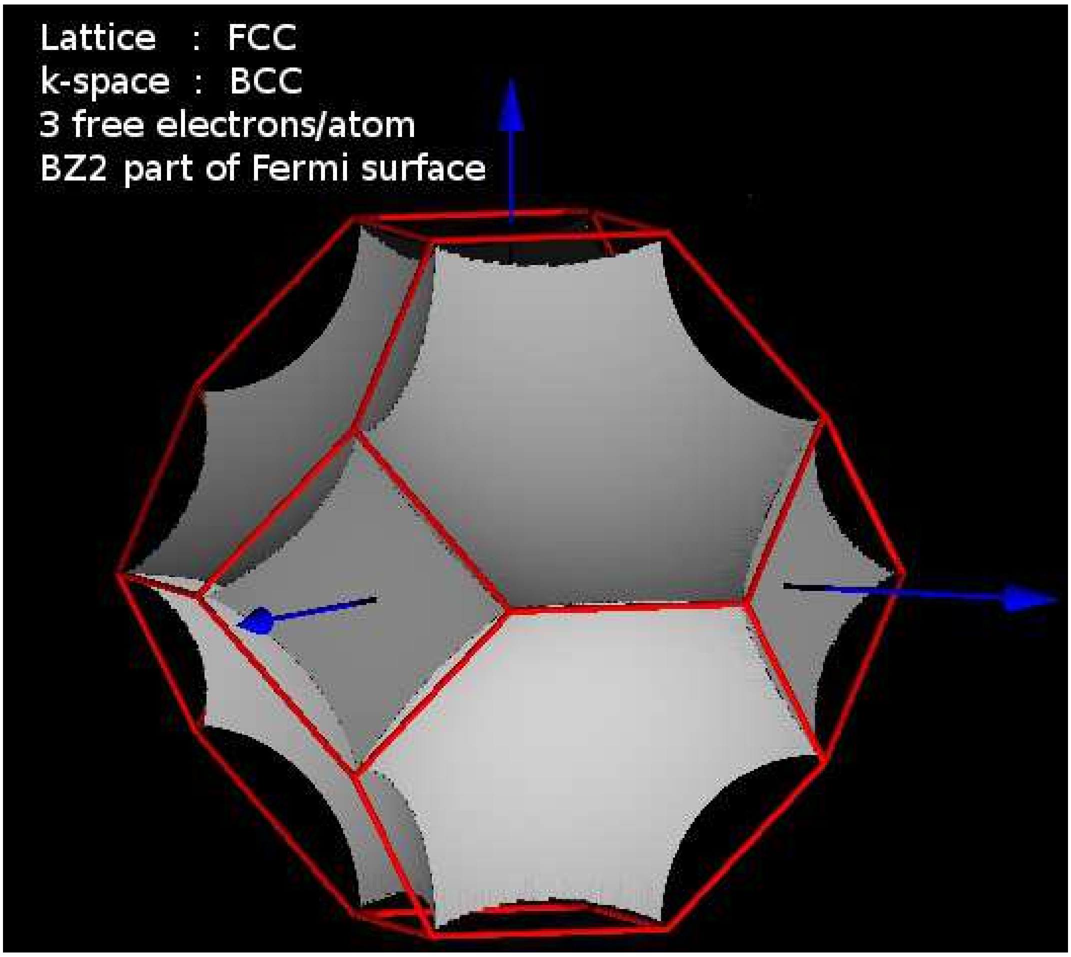


Figure A.13: The Fermi sphere for 3 electrons per site - the BZ2 part, translated back to BZ1.

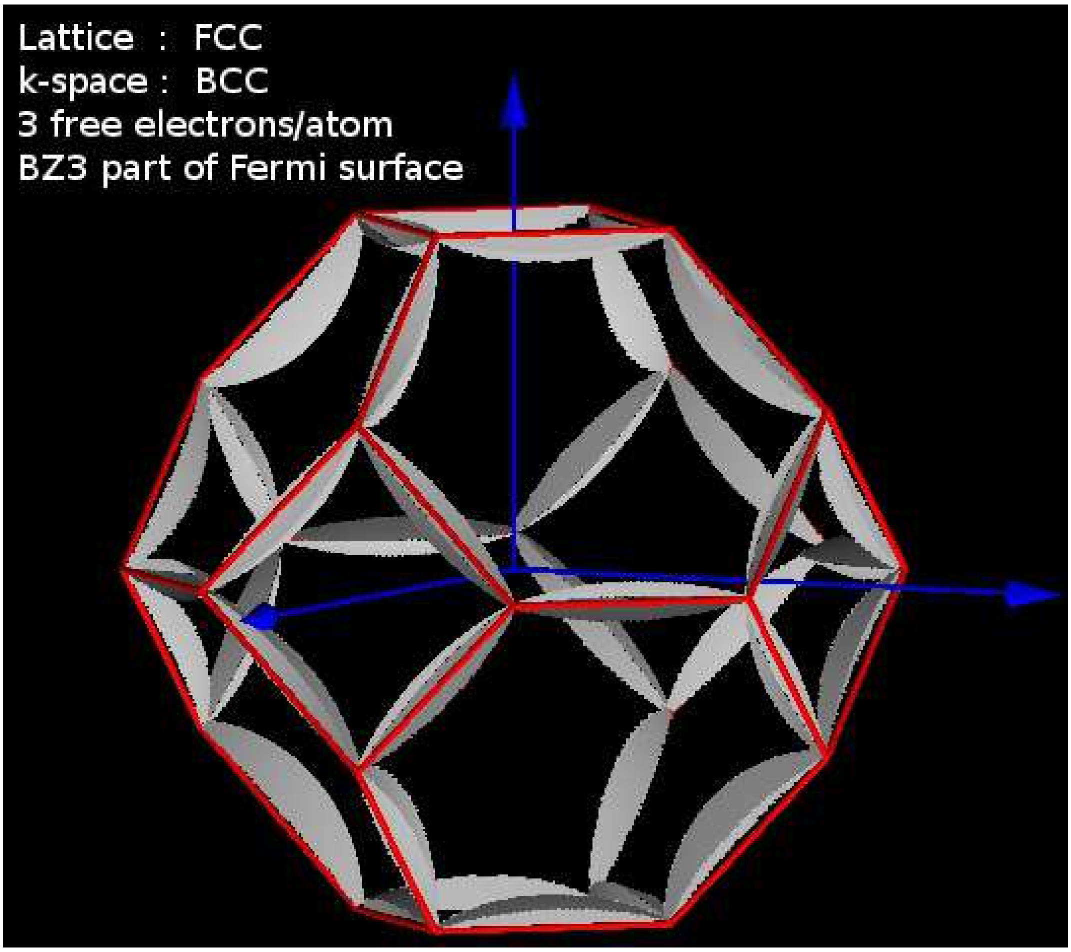


Figure A.14: The Fermi sphere for 3 electrons per site - the BZ3 part, translated back to BZ1.

A.4 Electrons and holes

The word "hole" comes up in more than one context. It is important to know the exact meaning of this. It is *not* always merely an absence of an electron.

A.4.1 The equation of motion of a particle in a band

The components of the equation of motion of a particle in a certain $\boldsymbol{k}$ state in a band $\mathcal{E}(\boldsymbol{k})$ can be written as :

$$\frac{dv_i}{dt} = \boldsymbol{M}_{ij}^{-\boldsymbol{1}}.F_j \tag{A.2}$$

where $\boldsymbol{F}$ is the "force" on the particle, given by

$$\boldsymbol{F} = -|e|\left(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B}\right) \tag{A.3}$$

The elements of the inverse effective mass matrix are given by

$$\boldsymbol{M}_{ij}^{-\boldsymbol{1}} = \frac{1}{\hbar^2}\frac{\partial^2\mathcal{E}}{\partial k_i \partial k_j} \tag{A.4}$$

The velocity is the group velocity given by

$$v_j = \frac{1}{\hbar}\frac{\partial\mathcal{E}}{\partial k_j} \tag{A.5}$$

A.4.2 Sign of the effective mass

Consider a simple one-dimensional band structure shown in the figure. We can see for the lowermost band near $k = 0$ all the derivatives are positive. In 1d all the derivatives become total derivatives and the things are simple. Since all the derivatives are simple the equation of motion is appropriate to that of a particle with negative charge and positive mass.

Now consider the same band near the Brillouin zone. the first derivative is positive but the second derivative is negative. Thus the equation of motion is appropriate to that of a "particle" with negative charge and negative mass. The important quantity driving the motion of a charge is the e/m ratio. Because of the essential change in the sign of the second derivative the ratio has changed sign. We can think of the motion, as if it is that of a positive charge and positive mass.

Notice that the equation of motion does not care whether there actually is a particle in that state or not. Also important is that somewhere in between there must be a point (of inflection) where the second derivative changes sign.

Now consider the band just above (marked 2). Due to the shape of the band, we must have $d\mathcal{E}/dk < 0$ always. You can see that at the lower energies the sign of the second derivative is positive this time.

Comparing the two situations you also see that in the first case the higher energy states surround the lower energy states. In the second case it is just the opposite.

A.4.3 Direction of the gradient

Consider a 2D case for simplicity, where the filling of a band has just started from the edge. Here the higher energy states are inside. Notice how the sign of the derivatives are different in the two cases, shown in the figure.

Now consider a state whose evolution is given by

$$\begin{aligned} \frac{d\boldsymbol{k}}{dt} &= -\frac{|e|}{\hbar}\left(\boldsymbol{E} + \boldsymbol{v}\times\boldsymbol{B}\right) \\ \boldsymbol{v} &= \frac{1}{\hbar}\nabla_{\boldsymbol{k}}\mathcal{E} \end{aligned} \tag{A.6}$$

You can show by taking the dot product of first equation with $\boldsymbol{B}$, that the component of $\boldsymbol{k}$ in the direction of $\boldsymbol{B}$ does not change. Similarly, by taking the dot product and $\boldsymbol{v}$, we can show that there is no component of $d\boldsymbol{k}/dt$ in the direction of the gradient to a constant energy surface. The electron thus stays on a constant energy surface. So geometrically it executes an orbit on a constant energy surface.

The sense of rotation would be different if $\nabla_{\boldsymbol{k}}\mathcal{E}$ points inwards or outwards. Here again you can convince yourself that an orbit that encloses higher energy empty states, experiences an opposite push. This is also "hole like" behaviour. The effective mass doesn't come into consideration here.

A.4.4 Current carried by a band

The current carried by a band can be written as a sum over filled states or as a sum over unfilled states.

$$\begin{aligned} \boldsymbol{j} &= -|e| \int_{filled} \boldsymbol{v} \frac{d^3\boldsymbol{k}}{(2\pi)^3} \\ &= -|e| \int_{all} \boldsymbol{v} \frac{d^3\boldsymbol{k}}{(2\pi)^3} + |e| \int_{unfilled} \boldsymbol{v} \frac{d^3\boldsymbol{k}}{(2\pi)^3} \\ &= 0 + |e| \int_{unfilled} \boldsymbol{v} \frac{d^3\boldsymbol{k}}{(2\pi)^3} \end{aligned} \tag{A.7}$$

The integral of group velocity over a full band will vanish. This is in general true that the integral of a derivative of a periodic function over a full period vanishes. It is left as an exercise to supply the argument.

We cannot mix the "electron" and "hole" type description within a same band. But if two partially filled bands exist, we can treat one of them as an electron band and the other as a hole band. In semiconductors, this is often done and we write when n electrons (in conduction band) and p holes (in valence band) are driven by an electric field:

$$j = (n|e|\mu_n + p|e|\mu_p)\, E \tag{A.8}$$

A.4.5 The semiclassical evolution of a $\boldsymbol{k}$-state

The evolution of a $\boldsymbol{k}$-state is governed by

$$\dot{\boldsymbol{k}} = -\frac{|e|}{\hbar}\left(\boldsymbol{E} + \frac{1}{\hbar}\nabla_{\boldsymbol{k}}\mathcal{E} \times \boldsymbol{B}\right) \tag{A.9}$$

In a case where $\boldsymbol{B}$ and $\boldsymbol{E}$ are perpendicular, it is useful if we define an auxiliary function

$$\boldsymbol{w} = \frac{\boldsymbol{E} \times \boldsymbol{B}}{B^2} \tag{A.10}$$

It can be shown that the "Lorentz force" like equation can be written in a compact form

$$\dot{\boldsymbol{k}} = -\frac{|e|}{\hbar}\left(\nabla_{\boldsymbol{k}}[\mathcal{E} - \hbar\boldsymbol{k}.\boldsymbol{w}] \times \boldsymbol{B}\right) \tag{A.11}$$

The algebra is left as an exercise. The useful physical picture is that the k-space trajectories are given by intersections of planes perpendicular to the magnetic field with surfaces on which the function $\mathcal{E} - \hbar\boldsymbol{k}.\boldsymbol{w}$ is constant.

Some of these trajectories are closed and some of them are open. In a band it is not possible for all the trajectories to be open or all of them to be closed. Though certainly not obvious, it turns out that in the high magnetic field limit it is only the closed orbits that contribute to the Hall current (voltage). Thus if the unoccupied levels are closed orbits, then the band will behave like a hole band, if the occupied levels are closed then the band will behave as an electron band. The fact that some metals show "hole like" Hall voltage, arises from this.

Hall voltage in a two band case

We consider the most common geometry where the material is in the form of a thin rectangular slab, the electric field driving the current is along the length of the slab. The magnetic field is perpendicular. Let's consider the case of FCC Aluminium, which has three electrons per atom. We have seen in such case BZ1 is completely filled. BZ2 and BZ3 are partially filled. A fully filled band will carry no current. We can add the current carried by the two active bands in matrix form by writing

$$\boldsymbol{j} = (\boldsymbol{\sigma}_2 + \boldsymbol{\sigma}_3)\,.\boldsymbol{E} \tag{A.12}$$

where

$$\begin{aligned} \boldsymbol{\sigma_2} &= \frac{1}{{\rho_2}^2 + {R_2}^2 B^2} \begin{pmatrix} \rho_2 & R_2 B \\ -R_2 B & \rho_2 \end{pmatrix} \\ \boldsymbol{\sigma_3} &= \frac{1}{{\rho_3}^2 + {R_3}^2 B^2} \begin{pmatrix} \rho_2 & R_3 B \\ -R_3 B & \rho_2 \end{pmatrix} \end{aligned} \tag{A.13}$$

We have used the shorthand $R = 1/ne$ for the Hall coefficient. We now need to invert $\boldsymbol{\sigma}_2 + \boldsymbol{\sigma}_3$ to get the composite resistivity matrix. The algebra is certainly long drawn, but it gives the final result

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} \rho & RB \\ -RB & \rho \end{pmatrix} \begin{pmatrix} j_x \\ j_y \end{pmatrix} \tag{A.14}$$

where

$$R = \frac{R_2{\rho_3}^2 + R_3{\rho_2}^2 + R_2R_3(R_2+R_3)B^2}{(\rho_2+\rho_3)^2 + (R_2+R_3)^2B^2} \tag{A.15}$$

$$\rho = \frac{\rho_2\rho_3(\rho_2+\rho_3) + (\rho_2{R_3}^2 + \rho_3{R_2}^2)B^2}{(\rho_2+\rho_3)^2 + (R_2+R_3)^2B^2} \tag{A.16}$$

What is the high field limit $|RB| >> \rho$ of R in this case? Clearly we need to retain only the B^2 terms. We then get

$$R = \frac{1}{(n_2+n_3)|e|} \tag{A.17}$$

But if one of the bands act like a hole band the the plus sign would be replaced by a minus. This is the key to understanding why some "free electron metals" like Aluminium give the *opposite* sign of the Hall voltage.

Hall voltage of trivalent Aluminium

In the chapter on free electrons (table :), we saw that the carrier density of Al, calculated from the Hall voltage at high magnetic field, turns out to be about one-third of the expected value and of opposite sign. So the charge carriers appear to be holes. We can understand this in terms of the Fermi surfaces of a 3e/atom FCC solid that we have seen.

- Let's say there are N unit cells, hence $3N$ electrons.
- BZ1 is full and that holds $2N$ electrons. This full band is inert.
- The remaining N electrons are now partly in BZ2 and partly in BZ3.
- BZ2 is hole like (closed orbits are unoccupied.) So we have

$$n_{e,BZ2} + n_{h,BZ2} = 2N \tag{A.18}$$

$$n_{e,BZ2} + n_{e,BZ3} = N \tag{A.19}$$

$$\therefore n_{e,BZ3} - n_{h,BZ2} = -N \tag{A.20}$$

- The apparent carrier density is thus $-\frac{1}{3}$ of the total number of free electrons.