PHYS431 Notes

Hersh Kumar

Contents

1 Crystal Structures

This course deals with ordinary matter in its solid, condensed state. The key thing is that in the solid phase, atoms or ions which make up the solid, have more or less fixed positions with respect to each other. The definition of crystal structure is an arrangement of atoms or ions in space. The key questions that we will be asking for the first third of the course are:

- How are the atoms arranged in solids? (Chapter 1)
- How is the structure of a solid found experimentally? (Chapter 2)
- What holds a solid together? (Chapter 3)

Solids have three kinds of structures:

1. Crystalline - The atoms or ions have a highly ordered and repetitive/periodic arrangement. Examples include NaCl, quartz, $BaTiO_3$, diamonds, etc.
2. Amorphous/Disordered - There is no relation between the positions of atoms which are sufficiently far apart. Examples include amorphous Si, amorphous phase change memory (e.g. GeSbTe)
3. Quasi-crystalline - Atoms have an ordered but non-repetitive/non-periodic arrangement. Examples include $Mg_{32}\,(AlZn)_{49}$, $Al-Pd-Mn$.

When we talk about periodic arrangements in crystalline structures, we mean that we have translational invariance. Another key term in the crystalline definition is the word “ordered”. Most of the time we will deal with crystalline materials, because amorphous materials are more complicated and mysterious. Until a few decades ago, crystalline and amorphous materials were the only types thought to exist, until the discovery of quasi-crystalline solids, which were discovered in 1984. For example, the first example of a quasi-crystalline solid that is listed in the examples has pentagonal structure, which cannot infinitely repeat, they do not tile a plane infinitely without breaking at some point. Thus the structure is ordered, but non-periodic.

In this course, we will only cover crystalline solids, because the periodicity of crystalline solids is the basis for our mathematical description of the solid. Let us do some definitions. A **Lattice** is a repetitive/periodic arrangement of points in space. A **Bravais Lattice** is a lattice which “looks the same” from any lattice point. The first example of this is shown in Figure 1.

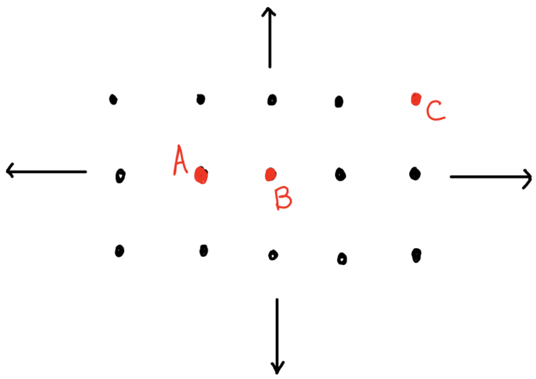

Figure 1: An example of a 2D Bravais Lattice, the lattice looks identical when at any point.

We can see that if we are at different points on this lattice, the lattice looks exactly the same at any point we pick. A second example of a lattice is shown in Figure 2.

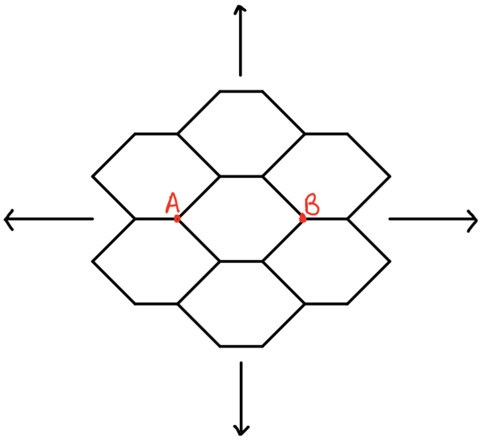

Figure 2: An example of a 2D hexagonal lattice, which tiles infinitely.

This lattice is not a Bravais Lattice, as we can see that points A and B are not equivalent.

Primitive Translation Vectors (PTVs) are a set of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ such that all the points on the Bravais lattice can be written as

$$\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$$

Where n_1, n_2, n_3 are any and all integers. For a given Bravais lattice, PTVs are not unique, there is more than 1 way to traverse to every point. This is demonstrated in Figure 3.

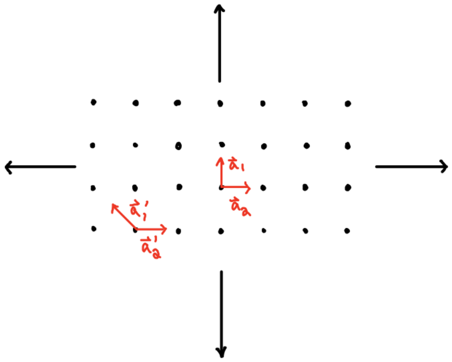

Figure 3: Two sets of primitive translation vectors on a 2D lattice.

The **Primitive Unit Cell** is the volume enclosed by a parallelpiped (usually) formed by a set of PTVs, $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$, as displayed in Figure 4. This is the smallest volume cell which can be repeated using PTVs to exactly fill the space. Shapewise, a primitive unit cell is not unique, but all of them have the same volume:

$$V = |\boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)|$$

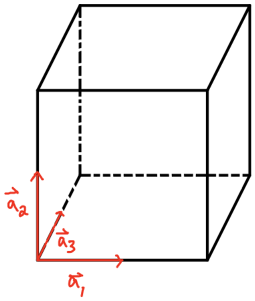

Figure 4: An example of the parallelpiped generated by a set of primitive translation vectors.

One important type of primitive unit cell is the **Wigner-Seitz Cell**, which is important for understanding electronic structure of solids. This is a primitive cell volume formed by regions of space closest to a single lattice point. The recipe for building this cell first begin with a lattice. We then pick a lattice point, and draw lines to all nearest neighbors. We then draw perpendicular bisectors, and then connect them to enclose the cell. A 2D example is shown in Figure 5:

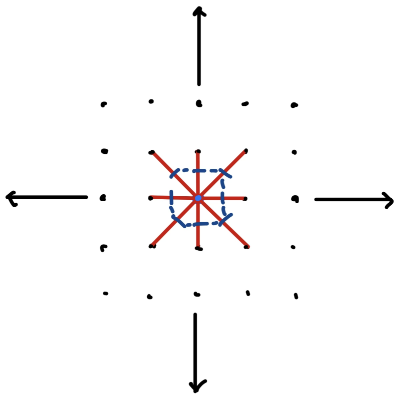

Figure 5: The construction of a Wigner-Seitz cell in a simple 2D lattice.

Now let us think about attaching atoms to points on the lattice. One important point is that most real materials/crystals are not Bravais lattices. Most solids are composed of more than 1 type of atom, and thus we have sites with different atoms, and therefore they are no longer equivalent. To fully specify a crystal structure, we do need a Bravais lattice, but we also need a **basis**. In other words, the recipe for creating a crystal is a Bravais lattice and a basis.

The **basis** is a set of vectors which specify positions of atoms or ions in a primitive cell. An example of a simple crystal structure is shown in Figure 6.

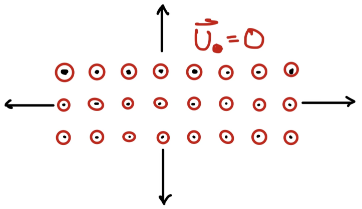

Figure 6: An example of a simple crystal structure using a basic lattice and basis vectors.

We see that when $\boldsymbol{u}_{\text{filled}} = 0$, we just place an atom on every lattice point. We also note that for a particular structure, a set of lattice and basis vectors is not unique. We could have just as easily done $\boldsymbol{u}_{\text{filled}} = \frac{1}{2}a\hat{x} + \frac{1}{2}a\hat{y}$, where a is the lattice spacing. The result of this choice of basis is shown in Figure 7.

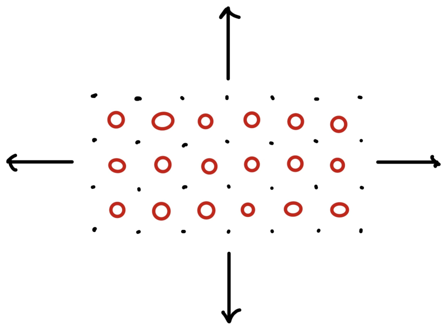

Figure 7: An example of the same crystal structure shown in Figure 6 with a different basis.

We can look at another example, where we are delineating the distances between lattice points, marked by xs. We set the basis vectors to have an unfilled atom to the left of every lattice point, and a filled atom to the right. The result is shown in Figure 8.

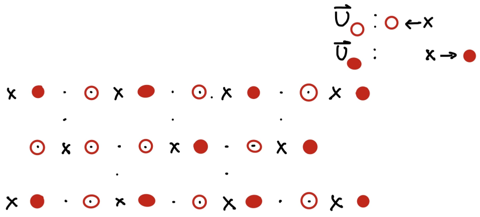

Figure 8: An example of a crystal structure with 2 basis vectors, each corresponding to a different type of atom.

There are infinitely many possible crystal structures, but only 14 types of Bravais lattices. The different types were established in 1845. These are known as the irreducible representations.

1. **Cubic** (3 types):
   a) Simple Cubic (SC): 3 sides have the same length, and 3 angles are right angles ($\alpha, \beta, \gamma = 90$ degrees).
   b) Body Centered Cubic (BCC): Atoms at corners and 1 at center.
   c) Face Centered Cubic (FCC): Atoms at corners and center of each face.
2. **Tetragonal** (2 types): Sides are $a = a \neq b$, $\alpha = \beta = \gamma = 90$ degrees
   a) Simple
   b) Body Centered (BCT)
3. **Orthorhombic** (4 types): $a \neq b \neq c$, $\alpha = \beta = \gamma = 90$ degrees
   a) Simple
   b) Face Centered
   c) Base Centered (Atoms only on center of 2 of the 6 faces)

   d) Body Centered

4. **Hexagonal** (1 type): angles a_1 and a_2 are 120 degrees, and $a_1 = a_2 \neq a_3$.
5. **Monoclinic** (2 types): $a \neq b \neq c$, $\alpha = \beta = 90$ degrees, $\gamma \neq 90$ degrees.
   a) Simple
   b) Base Centered
6. **Trigonal/Rhombohedral** (1 type): $a = b = c$, $\alpha = \beta = \gamma \neq 90$ degrees.
7. **Triclinic** (1 type): $a \neq b \neq c$, $\alpha \neq \beta \neq \gamma$

We see that the triclinic crystals have the "lowest" symmetry. We will see that as we increase the temperature of a solid, it will work its way up the list, undergoing structural transitions.

Let us look at examples of structures of real solids. The first structure we will look at is CsCl, which is a simple cubic lattice with a 2 atom basis, as shown in Figure 9.

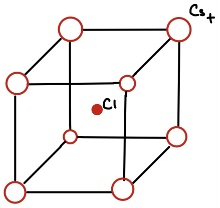

Figure 9: The crystal structure of CsCl, which is built on a simple cubic lattice with a 2 atom basis.

This is an example of a perfectly ordered crystal. Note that this is not a BCC lattice, since the center atom is not the same as the corner atom. If we had a disordered crystal, we don't have the same pattern. In that case, we say that each atom could be a "gray" atom, an average between the two atoms. If we do this, then suddenly every atom is the same, and we can use the BCC model.

Let us look at another example of a real structure. This is known as Hexagonal close-packed structure (HCP). Examples include Be, Cs, and Cd. Close-packed means that space is filled to maximal extent. An example is shown in Figure 10:

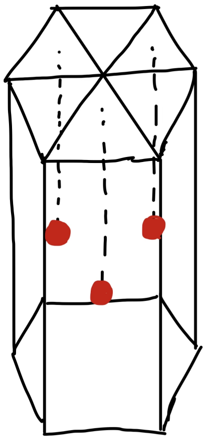

Figure 10: The crystal structure of a Hexagonal Close-Packed (HCP) structure.

HCP is a hexagonal lattics and a 2 atom basis, where the atoms are identical. We need two basis vectors, even though both of the two types of atoms must be the same. There are two ways to get maximum packing, the first being HCP. The other method is FCC. Imagine packing spheres with minimal wasted space. The diagram for maximal packing in two dimensions is shown in Figure 11.

We can easily pack the first layer, and the second layer is constructed by placing the centers of the second layer spheres on top of the open space between 3 spheres. For the third layer, we have two choices. The first choice is to place the third layer where the first layer was placed, while the second choice is to place the third layer above the space between three second layer spheres. We expect both of these methods to eventually repeat, since we have a crystal. If we consider the first layer locations as A, and the second layer choice as B, the choice of repeating the first layer choice obtains $ABABABA\ldots$. This generates the HCP structure. The other option is $ABCABCABCA\ldots$. This generates the FCC structure. This is why we say that HCP and FCC are "cousins". The packing fraction for these two methods is 0.74.

A single crystal has a perfect lattice that goes on for a long distance (think of calcium carbonate, if it remains ordered, we see that the macroscopic crystal looks like the crystal structure.) If instead the lattice is broken, we have powder, which is just a collection of "small" single crystals (for calcium carbonate that gives us chalk). When we look at single crystals and we do things like diffraction, we are essentially looking at one plane that intersects the crystal, whereas when we have a powder, we are essentially shooting our beams at a bunch of small crystals that have arbitrary orientations, and thus we are essentially getting the information about all possible orientations of the single crystal. In single crystal materials, we have "perfect" ordering, up to macroscopic scales (calcium carbonate crystals for example), and at such large scales, properties and shapes of materials reflect those at the microscopic level.

We can now move onto describing planes of crystals. We use the convention of the Miller index, as shown in Figure 12.

The plane intersects our 3 axes at 3 points, some distance along each axis, $x\boldsymbol{a_1}$, $y\boldsymbol{a_2}$, and $z\boldsymbol{a_3}$. The Miller index is given by $\left(\frac{1}{x}, \frac{1}{y}, \frac{1}{z}\right)$, and we multiply by a number such that we have the smallest set of integers. In the case shown in Figure 12, our intersections are given by

$$\left(\frac{1}{4}, \frac{1}{1}, \frac{1}{5}\right)$$

We can then multiply by 20, and we have the smallest set of integers, which is the Miller index for this plane:

$$\left(\frac{1}{4}, \frac{1}{1}, \frac{1}{5}\right) \times 20 = (5, 20, 4)$$

Typically we look at planes such as the one that has Miller index $(1, 1, 1)$, as well as the one that intersects the $\boldsymbol{a_1}$ axis, and is parallel to the $\boldsymbol{a_2}$ $\boldsymbol{a_3}$ plane. This has Miller index $(1, 0, 0)$ (where we use the fact that $1/\infty = 0$). When we have Miller indices that are negative, we instead place a bar on top of the negative, so we can have indices like $(1, \bar{1}, 0)$.

Now we note that $(1, 0, 0)$ and $(\bar{1}, 0, 0)$ are equivalent planes, since they are just translations of each other. To denote a family of planes, we use curly brackets, so the family of planes that are equivalent to $(1, 0, 0)$ would be written as $\{1, 0, 0\}$. In the demonstration of the Si crystal wafer, we used cleaving to expose the $(1, 0, 0)$ plane.

2 Crystal Diffraction and the Reciprocal Lattice

The key question for Chapter 2 is how the crystal structure/atomic arrangement determined experimentally. When solid state physics first started, 60-70 years ago, the only answer to this was diffraction. Nowadays there are other techniques, such as electron microscopy and scanning tunnelling microscopy. In this class we will only cover diffraction. This can be done with x-rays, electrons, and neutrons, and this method uses reciprocal space techniques. Nowadays there are real space techniques, which can directly resolve the structure of the atomic arrangement.

Diffraction can be used to determine the microscopic arrangement of any crystalline materials. For this reason, diffraction is used in biology, medicine, etc. Diffraction relies on both the wave and particle nature of light, and it turns out that we can also do this with electrons.

Recall that c is the speed of light, $c = 10^8$ meters per second. We also have the relationship between the speed, wavelength, and frequency:

$$c = \nu\lambda$$

For any kind of light, we have that the energy is given by $E = h\nu$, where $h = 6.626 \times 10^{-34}$ Joule seconds. Thus, for light, we can write the wavelength in terms of the energy:

$$\lambda = \frac{hc}{E}$$

For particles, like in the case of electrons and neutrons, we write the energy in terms of the momentum:

$$E = \frac{p^2}{2m}$$

We then use the idea of DeBroglie wavelengths, where any momentum has an associated DeBroglie wavelength:

$$\lambda = \frac{h}{p}$$

From this, we can write the energy in terms of the wavelength:

$$E = \frac{h^2}{\lambda^2}\frac{1}{2m}$$

Which then allows us to control the wavelength by changing the energy:

$$\lambda = \frac{h}{\sqrt{2mE}}$$

We then introduce the idea of a wavenumber, k, which is defined as $k = \frac{2\pi}{\lambda}$. With this, we can write the momentum in terms of the wavenumber:

$$p = \frac{h}{2\pi}k$$

For x-rays, we have that $\lambda = \frac{12.4\text{Å}}{E}$, where E is measured in kiloelectronvolts.

We will see that we want λ to be on the order of, or slightly smaller than the lattice spacing a, in order to see strong diffraction effects.

2.1 Review of Wave Mechanics

For a wave in 1 dimension, we have $\psi = A\cos(kx)$. We have that the wavenumber k is related to the wavelength:

$$k = \frac{2\pi}{\lambda}$$

The wave propagation speed is given by $c = \nu\lambda = \frac{\omega}{k}$, where ω is the angular frequency, $\omega = 2\pi\nu$. The relationship between ω and k is known as the dispersion relation.

For travelling waves, we use Euler notation (complex exponentials):

$$A\cos(kx) \to Ae^{ikx}$$

A travelling wave in 1 dimension is described as

$$\psi = Ae^{i(kx-\omega t)}$$

In particular, that is a wave moving in the positive x direction, and if we have an exponential of the form $Ae^{i(kx+\omega t)}$, we have a wave moving in the negative x direction.

In 3 dimensions, we use wave vectors instead of wavenumbers, and we can write a plane wave:

$$\psi = e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$$

Where $\boldsymbol{x} = (x, y, z)$. $\boldsymbol{k}$ points perpendicularly to the wavefront. The "intensity" of the wave is given by the square of the amplitude of the wave.

The basis of diffraction is the idea of constructive and destructive interference. Waves add constructively if they are in phase, but will interfere with each other if they are out of phase.

2.2 Bragg Diffraction

Let us begin with the experimental arrangement for powder diffraction. The x-ray source is sent through a monochromater, and collimated into a beam, and then shot at the sample, which is slowly rotating. The film records the location of the diffracted beams. The diagram for the arrangement is shown in Figure 13.

Afterwards, we take off the film strip. We then have that $\alpha = 0$ is at the center. The reason we rotate the sample is because we want to obtain all possible planes that will give us a diffraction peak. The peaks occur due to the x-rays constructively interfering, while most rays destructively interfere, giving nothing on the film. An important point is that each material or structure gives a unique pattern of lines. The intensity of a peak depends on the Z of atoms, the number of electrons. The real space structure can be determined from the pattern.

Another type of setup is a θ to 2θ scan. In this case, the sample is held fixed, and the angle of incidence of the x-rays and the angle of reflection are identical. A diagram of this setup is shown in Figure 14.

This is also known as a locked scan, since the angles between the incidence and reflected rays are locked to be equal to each other. In a scan, the θ increases continuously, and the resulting data looks like the plot in Figure 15.

The idea behind Bragg diffraction (1912) is based on the planes that make up the crystal. We can zoom in on the surface and see the different planes of the crystal, with some spacing d. We can model light rays striking the crystal, parallel rays that each strike a different layer of the crystal, and bounce out of the crystal. We are assuming that the crystal is aligned correctly, and the incident angle is the same as the reflected angle, $\theta_i = \theta_r$. If we look at rays that are on the plane below the top plane, we have that the ray travels a slightly longer distance than the ray above it, and likewise for all rays below. The second layer ray has to travel an extra $2d\sin\theta$ overall, to keep up with the top level ray. The third layer ray has to travel an extra $4d\sin\theta$. Thus we have that our condition for the rays meeting the detector at the same time, we have that the extra distance is an integer multiple of the wavelength. The intensity of the reflected rays together will add constructively only if we have that

$$\begin{gathered}2d\sin\theta = c_0\lambda \\ 4d\sin\theta = c_1\lambda \\ 8d\sin\theta = c_2\lambda \\ \vdots\end{gathered}$$

where the cs are integers. Thus the simplest solution to this is that $2d\sin\theta = n\lambda$, and this is the only way that we have a “hit”.

We can rewrite this, and we have that

$$\sin\theta = \frac{n\lambda}{2d}$$

The left side is bounded in magnitude by 1, and thus we have that

$$\frac{n\lambda}{2d} \leq 1$$

Thus, in order for a diffracted intensity to happen, $\lambda < 2d$. This is why we need the wavelength to be of the order of the crystal lattice spacing. Depending on d and λ, we can have multiple allowed values of n that will get us intensity peaks.

Now we have to explain how we get different intensity peaks, and we also need to explain how x-rays reflect off of atomic planes. To answer these questions, we need to go beyond the simple Bragg picture. To do this, we will need to introduce the concept of the reciprocal lattice.

2.3 Reciprocal Lattice

The real lattice is in Cartesian space, and the basic idea is that the reciprocal lattice is in the $\boldsymbol{k}$-space, which can also be thought of as the momentum space, since $p = \hbar k$.

Definition 2.1. *Given a Bravais lattice with primitive translation vectors* $\boldsymbol{a}_1$,$\boldsymbol{a}_2$, *and* $\boldsymbol{a}_3$, *the reciprocal lattice is a set of vectors* $\boldsymbol{G}$, *such that*

$$\boldsymbol{G} = l_1\boldsymbol{b}_1 + l_2\boldsymbol{b}_2 + l_3\boldsymbol{b}_3$$

where l*s are* $1, \pm 1, \pm 2, \pm 3, \dots,$ *and the* $\boldsymbol{b}$*s are defined as*

$$\begin{aligned}\boldsymbol{b}_1 &= \frac{2\pi}{V}\boldsymbol{a}_2 \times \boldsymbol{a}_3\\ \boldsymbol{b}_2 &= \frac{2\pi}{V}\boldsymbol{a}_3 \times \boldsymbol{a}_1\\ \boldsymbol{b}_3 &= \frac{2\pi}{V}\boldsymbol{a}_1 \times \boldsymbol{a}_2\end{aligned}$$

Where $V = \boldsymbol{a}_1 \cdot (\boldsymbol{a}_2 \times \boldsymbol{a}_3)$, *the volume of the primitive cell.*

We denote the real position on the lattice as $\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3$. $\boldsymbol{G}$'s have one important property:

$$\boldsymbol{b}_i \cdot \boldsymbol{a}_j = 2\pi\delta_{ij}$$

Which is easy to prove, since we define the $\boldsymbol{b}$ in terms of cross products of two of the three $\boldsymbol{a}$ vectors. We can also prove that

$$e^{i\boldsymbol{G}\cdot\boldsymbol{R}} = 1$$

for any $\boldsymbol{G}$ and $\boldsymbol{R}$ for a given crystal.

We can write this out:

$$\begin{aligned}e^{i\boldsymbol{G}\cdot\boldsymbol{R}} &= e^{i[n_1 l_1 \boldsymbol{a}_1\cdot\boldsymbol{b}_1 + n_2 l_2 \boldsymbol{a}_2\cdot\boldsymbol{b}_2 + n_3 l_3 \boldsymbol{a}_3\cdot\boldsymbol{b}_3]}\\ &= e^{2\pi i[n_1 l_1 + n_2 l_2 + n_3 l_3]}\\ &= 1\end{aligned}$$

Let us look at some examples of a reciprocal lattice. For a simple cubic:

$$\begin{aligned}\boldsymbol{a}_1 &= a\hat{x}\\ \boldsymbol{a}_2 &= a\hat{y}\end{aligned}$$

$$\boldsymbol{a}_3 = a\hat{z}$$

The volume of the unit cell in this case is just a^3. We can then compute the reciprocal lattice vectors:

$$\begin{aligned}\boldsymbol{b}_1 &= \frac{2\pi}{a^3}a^2(\hat{y}\times\hat{z})\\ &= \frac{2\pi}{a}\hat{x}\\ \boldsymbol{b}_2 &= \frac{2\pi}{a}\hat{y}\\ \boldsymbol{b}_3 &= \frac{2\pi}{a}\hat{z}\end{aligned}$$

In this case, outside of the scaling factor, the reciprocal space of the simple cubic seems to also be a simple cubic lattice.

Let us now compute the reciprocal space vectors for the fcc case. We have that

$$\begin{aligned}\boldsymbol{a}_1 &= \frac{a}{2}(\hat{y}+\hat{z})\\ \boldsymbol{a}_2 &= \frac{a}{2}(\hat{x}+\hat{z})\\ \boldsymbol{a}_3 &= \frac{a}{2}(\hat{x}+\hat{y})\end{aligned}$$

We can then compute the volume, and we find that $V = \boldsymbol{a}_1\cdot(\boldsymbol{a}_2\times\boldsymbol{a}_3) = \frac{a^3}{4}$. We can then use this to compute the reciprocal space vectors:

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

It turns out that there are scaled versions of the BCC primitive translation vectors, and the inverse relation holds, the reciprocal space vectors of BCC gets the FCC PTVs. This also shows that if we take the reciprocal of the reciprocal lattice vectors, we get the original real space vectors.

2.4 The Von Laue Diffraction Condition

Previously, we did the simple Bragg diffraction picture. Now let us derive diffraction using the reciprocal picture. Let us zoom further into the picture of the Bragg diagram. Suppose an x-ray scatters off of 2 atoms, each on their own plane. We have the vector $\boldsymbol{R}$ between the two atoms, and we have an incoming wave with wavevector $\boldsymbol{k}$. the incident angle is θ_i, and the reflected angle is θ_r, with some reflected wavevector $\boldsymbol{k}'$. We assume elastic scattering, so $|\boldsymbol{k}| = |\boldsymbol{k}'|$. The intensity of the x-ray scattered is given by

$$\begin{aligned}I(\boldsymbol{k}\to\boldsymbol{k}') &= |F|^2\\ &= |F_1+F_2|^2\end{aligned}$$

Where F_1 is the amplitude of the x-ray scattering off of the first atom, and F_2 is the amplitude of the x-ray scattering off of the second atom. Recall that x-rays are electromagnetic waves, which means that we have a travelling electric field and travelling magnetic field:

$$\boldsymbol{E} = E_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$$
$$\boldsymbol{B} = B_0 e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$$

We can rewrite the amplitude/intensity relationship:

$$F_1 = \sqrt{I_0} e^{ik(S_i+S_r)}$$
$$F_2 = \sqrt{I_0} e^{ik(S_i+S_r+\Delta S)}$$

Where I_0 is the initial intensity, S_i is the distance from the x-ray source to atom 1, S_r is the distance from atom 1 to the detector, and $\Delta S = \Delta S_i + \Delta S_r$ (the extra distance the wave travels for hitting the second atom). We can compute the values of ΔS_i and ΔS_r:

$$\Delta S_i = \boldsymbol{R} \cdot \frac{\boldsymbol{k}}{|\boldsymbol{k}|}$$
$$\Delta S_r = -\boldsymbol{R} \cdot \frac{\boldsymbol{k'}}{|\boldsymbol{k'}|}$$

We can then write out the expression for the scattered intensity:

$$\begin{aligned}|F_1 + F_2|^2 &= |F_1|^2 \left| 1 + \frac{F_2}{F_1} \right|^2 \\ &= I_0 \left| 1 + \frac{e^{ik(S_i+S_r+\Delta S)}}{e^{ik(S_i+S_r)}} \right|^2 \\ &= I_0 \left| 1 + e^{i\boldsymbol{R}(\boldsymbol{k}-\boldsymbol{k'})} \right|^2\end{aligned}$$

Notice that the intensity can range from 0 to $4I_0$. The maximum intensity occurs when $e^{i\boldsymbol{R}(\boldsymbol{k}-\boldsymbol{k'})} = 1$. Now recall that the reciprocal lattice has the property that $e^{i\boldsymbol{R}\cdot\boldsymbol{G}} = 1$, which matches this condition. This means that we have a maximum whenever we have $\boldsymbol{k} - \boldsymbol{k'} = \boldsymbol{G}$.

We can generalize this picture to N atoms, and we have some F_i for each atom i. Let us first look more into the maximal condition. We have that $\boldsymbol{k} \cdot \boldsymbol{k'} = \boldsymbol{G}$, and thus $\boldsymbol{k'} = \boldsymbol{k} - \boldsymbol{G}$. We can then find that $|\boldsymbol{k'}|^2 = |\boldsymbol{k}|^2 + |\boldsymbol{G}|^2 - 2\boldsymbol{k} \cdot \boldsymbol{G}$, which then lets us write that $|\boldsymbol{G}|^2 = 2\boldsymbol{k} \cdot \boldsymbol{G}$. Any $\boldsymbol{k}$ that satisfies this condition for any $\boldsymbol{G}$ will result in a diffraction point. This leads us to Brillouin zones. This condition has a very useful geometrical interpretation, it lets you find all $\boldsymbol{k}$s that satisfy the condition. Brillouin zones are Wigner-Seitz cells in the reciprocal lattice. We first draw the lattice (reciprocal), and then extend dotted lines from a point to all nearest neighbors. We then draw the perpendicular bisectors and connect them. The central region is the first Brillouin zone, and the next closest region is the second Brillouin zone, and every Brillouin zone covers the same amount of volume/area in the reciprocal lattice.

Brillouin zone boundaries (BZB) represent all k values that satisfy the diffraction condition. We can prove this for any $\boldsymbol{k}$ that lands on the BZB. we have that the perpendicular distance from the center point to the BZB is $\frac{|G|}{2}$, and thus we have that

$$\frac{\boldsymbol{G}}{2} \cdot \boldsymbol{k} = \frac{G}{2} k \cos\theta$$

and we know that $\cos\theta = \frac{Gk}{2}$, and thus we can rewrite our statement, and we find that

$$2\boldsymbol{G} \cdot \boldsymbol{k} = G^2$$

Thus, Brillouin zones are a graphical representation of all $\boldsymbol{k}$s that diffract off of a crystal and produce peaks. In fact, in the Von Laue picture, the reciprocal picture of the Brillouin zones has the same geometry as the real space picture, with the incident wavevector $\boldsymbol{k}$ and the reflected wavevector $\boldsymbol{k}'$ being orthogonal to each other.

How is this information useful for figuring out the real-space crystal structure?

If we can get the full reciprocal lattice, then we can just use the reciprocal lattice definition and find the real space lattice vectors. We can get the full reciprocal lattice by mapping the BZBs in the lattice. How do we do this? We first fix the wavelength $\lambda = \frac{2\pi}{k}$, and we sweep the $\boldsymbol{k}$ direction, (can be thought of as sweeping θ). We get “hits” whenever the swept $\boldsymbol{k}$ lands on a BZB.

We can do a variation of this, instead of fixing λ, we fix θ, and we modify the magnitude of $\boldsymbol{k}$. This is basically pointing in a certain direction in reciprocal space, and marking all the intersections with the BZBs. With these experiments, we can map out where the boundaries of the Brillouin zones are, and from those we can find the $\boldsymbol{G}$ vectors, and then use the definition of the reciprocal lattice to find the real space lattice vectors.

2.5 N Atom Diffraction

We looked at the case of 2 atom diffraction, and we now want to extend this to N atoms. This process will help us figure out the differences in intensity between different peaks. It turns out that it is really electrons and electron clouds that are doing the scattering, and they are spread out. When we look at the structure of a crystal in the real world, we have electron clouds, and the electrons make the crystal “blurry”, they have probabilistic positions. Therefore, we consider the concentration of electrons at a particular position in the lattice. To do this, we have to of course take into account electrons being contributed from atoms that are nearby, but in fact we should take into account the contribution of all atoms. We will integrate across the entire crystal.

We define $n(\boldsymbol{r})$ to be the number of electrons per unit volume at position $\boldsymbol{r}$. $\boldsymbol{r}$ is measured from some global origin in the crystal. We index each atom via the center/nucleus of the atom. In fact, we know the electron concentration of particular atoms, relative to their centers, for example, we know $n_{\text{Li}}(\boldsymbol{r})'$, where the prime denotes that it is relative to the center of the atom, not the global origin for the crystal.

Recall that for 2 atoms, $I(\boldsymbol{k} \to \boldsymbol{k}') = I_0|1 + e^{i\boldsymbol{R}\cdot(\boldsymbol{k}-\boldsymbol{k}')}|^2$. $\boldsymbol{R}$ determined the relative position between the two atoms. We can say that we got 2 terms from two atoms, located at $\boldsymbol{R}_1 = 0$ and $\boldsymbol{r}_2 = \boldsymbol{R}$. We can then generalize this to N atoms:

$$\begin{aligned} I &= I_0|1 + e^{i\boldsymbol{R}_2\cdot(\boldsymbol{k}-\boldsymbol{k}')} + e^{i\boldsymbol{R}_3\cdot(\boldsymbol{k}-\boldsymbol{k}')} + \dots|^2 \\ &= I_0|\sum_{\text{atoms}} e^{i\boldsymbol{R}_i\cdot(\boldsymbol{k}-\boldsymbol{k}')}|^2 \end{aligned}$$

But really we know that it is the electrons that scatter the x-rays, so we should factor that in:

$$\begin{aligned} I &= I_0 \,\Big|\sum_{\text{electrons}} e^{i\boldsymbol{r}\cdot(\boldsymbol{k}-\boldsymbol{k}')}\Big|^2 \\ &= I_0 \,\Big|\int n(\boldsymbol{r}) e^{i\boldsymbol{r}\cdot(\boldsymbol{k}-\boldsymbol{k}')}\, d^3\boldsymbol{r}\Big|^2 \end{aligned}$$

We can then rewrite this for the amplitude:

$$F = \sqrt{I_0}\int n(\boldsymbol{r}) e^{i\boldsymbol{r}\cdot(\boldsymbol{k}-\boldsymbol{k}')} d^3\boldsymbol{r}$$

Let us now think about $n(\boldsymbol{r})$. Electron concentrations are associated with atoms, and they are quantified relative to the origin of the atoms.

For every point in the space, we need to think about the contribution from all atoms:

$$n(\boldsymbol{r}) = \sum_{\text{all atoms}} n_i(\boldsymbol{r} - \boldsymbol{r}_i)$$

Where $\boldsymbol{r}_i$ is the origin or nuclear center of atom i, and $n_i(\boldsymbol{r}')$ is the electron concentration for atom i located at $\boldsymbol{r}' = 0$.

Remember that a crystal is a Bravais lattice and a basis. We can then break down the summation into one that references the real space lattice vectors. The position of atom i can be written as

$$\boldsymbol{r}_i = \boldsymbol{R}_{m_1m_2m_3} + \boldsymbol{u}_j$$

Where $\boldsymbol{R}_{m_1m_2m_3} = m_1\boldsymbol{a}_1 + m_2\boldsymbol{a}_2 + m_3\boldsymbol{a}_3$, and $\boldsymbol{u}_j$ is the basis vector. The basis determines the type of atom, and each atom of the same type will have the same basis vector. We then define $n_j(\boldsymbol{r}')$ as the electron concentration at $\boldsymbol{r}'$ due to atom type j whose center is at $\boldsymbol{r}' = 0$. Using this, we can write out the total electron concentration:

$$n(\boldsymbol{r}) = \sum_{m_1m_2m_3}\sum_j n_j(\boldsymbol{r} - \boldsymbol{R}_{m_1m_2m_3} - \boldsymbol{u}_j)$$

We can then write out the amplitude:

$$\begin{aligned} F &= \sqrt{I_0}\int n(\boldsymbol{r}) e^{i\boldsymbol{r}\cdot(\boldsymbol{k}-\boldsymbol{k}')}\, d^3\boldsymbol{r} \\ &= \sqrt{I_0}\sum_{m_1m_2m_3,\, j}\int n_j(\boldsymbol{r} - \boldsymbol{R}_{m_1m_2m_3} - u_j(\boldsymbol{r})) e^{i\boldsymbol{r}\cdot(\boldsymbol{k}-\boldsymbol{k}')}\, d^3\boldsymbol{r} \end{aligned}$$

We then do a change of variables, $\boldsymbol{r}' = \boldsymbol{r} - \boldsymbol{R}_{m_1m_2m_3} - \boldsymbol{u}_j$, and we can rewrite the expression as

$$\begin{aligned} F &= \sqrt{I_0}\sum_{m_1m_2m_3,\, j}\int e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot(\boldsymbol{r}+\boldsymbol{R}_{m_1m_2m_3}+\boldsymbol{u}_j)} n_j(\boldsymbol{r}')\, d^3\boldsymbol{r}' \\ &= \sqrt{I_0}\sum_{m_1m_2m_3\, j} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot(\boldsymbol{R}_{m_1m_2m_3}+\boldsymbol{u}_j)} \in e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} n_j(\boldsymbol{r}')\, d^3\boldsymbol{r}' \end{aligned}$$

The integral is related only to the inner structure of a single atom. This is known as the atomic form factor for that particular atom. We know (by experiment) what the electronic concentration is, so we essentially separate out the integrated part:

$$\begin{aligned}\int e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{r}'} n_j(\boldsymbol{r}')\, d^3\boldsymbol{r}' &= f_j(\boldsymbol{k}-\boldsymbol{k}') \\ &= f_j\end{aligned}$$

This is the atomic form factor, and is the fourier transform of $n_j(\boldsymbol{r}')$. Now let us go to the total intensity:

$$\begin{aligned}I &= |F|^2 \\ &= I_0 \left| \sum_{m_1 m_2 m_3\, j} e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{R}_{m_1 m_2 m_3}} f_j(\boldsymbol{k}-\boldsymbol{k}') e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{u}_j} \right|^2\end{aligned}$$

We are doing all of this to understand where relative intensity peaks come from, and therefore we are interested in peaks, which is when $\boldsymbol{k}-\boldsymbol{k}' = \boldsymbol{G}$. Now let us look at the sum over ms:

$$\begin{aligned}&\sum_{m_1 m_2 m_3} e^{i\boldsymbol{G}\cdot\boldsymbol{R}_{m_1 m_2 m_3}} \\ &\qquad = \sum_{m_1 m_2 m_3} 1 \\ &\qquad = N\end{aligned}$$

This is the total number of lattice points. Thus the intensity will be written as

$$I = N^2 I_0 \left| \sum_j f_j(\boldsymbol{G}) e^{i\boldsymbol{u}_j\cdot\boldsymbol{G}} \right|^2$$

We now define $S(\boldsymbol{G})$, called the structure factor:

$$S = \sum_j f_j(\boldsymbol{G}) e^{i\boldsymbol{u}_j\cdot\boldsymbol{G}}$$

This is information about a unit cell in reciprocal space. Thus we have that the intensity is given by

$$I = N^2 I_0 \left| S(\boldsymbol{G}) \right|^2$$

This defines the intensity of a dffraction peak specified by $\boldsymbol{G}$. Let's do an example. We will look at CsCl structure, which is quite common. This is a simple cubic structure, with a 2 atom basis. We place the Chlorine at the lattice points, and a Caesium in the center of the unit cell. Thus we have that

$$\begin{aligned}\boldsymbol{u}_{Cl} &= 0 \\ \boldsymbol{u}_{Cs} &= \frac{a}{2}(\hat{x}+\hat{y}+\hat{z})\end{aligned}$$

The lattice vectors are given by

$$\boldsymbol{R} = a(m_1\hat{x} + m_2\hat{y} + m_3\hat{z})$$

$$\boldsymbol{G} = \frac{2\pi}{a}(l_1\hat{x} + l_2\hat{y} + l_3\hat{z})$$

We can now compute $S(\boldsymbol{G})$:

$$\begin{aligned} S(\boldsymbol{G}) &= \sum_j f_j(\boldsymbol{G})e^{i\boldsymbol{u}_j\cdot\boldsymbol{G}} \\ &= f_{Cl}(\boldsymbol{G})e^{i\boldsymbol{0}\cdot\boldsymbol{G}} + f_{Cs}e^{i\frac{2\pi}{a}(l_1\hat{x}+l_2\hat{y}+l_3\hat{z})\cdot\frac{a}{2}(\hat{x}+\hat{y}+\hat{z})} \\ &= f_{Cl} + f_{Cs}e^{i\pi(l_1+l_2+l_3)} \end{aligned}$$

Now we note that if $l_1 + l_2 + l_3$ is even, then we get $S(\boldsymbol{G}) = f_{Cl} + f_{Cs}$, and if $l_1 + l_2 + l_3$ is odd, then we get $f_{Cl} - f_{Cs}$. These are the conditions that cause the relative intensities of the diffraction peaks.

We see the CsCl structure in many other binary compounds, suppose an arbitrary one called AB. If $f_A \sim f_B$, then $S(\boldsymbol{G})$ will be $2f_A$ for $l_1 + l_2 + l_3$ even, and we will get 0 if $l_1 + l_2 + l_3$ is odd. This is known as extinction.

Last time we found the equivalency between $\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G}$ and the Bragg picture, where we noted that whenever there is a set of planes, there is a separation associated with the planes, d. There is some θ where diffraction takes place, and we have that $2d\sin\theta = n\lambda$.

We know that every $\boldsymbol{G}$ gives a BZB to satisfy $\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G}$. We also know that a particular plane separation d is connected to G, $d = \frac{2\pi}{G}$. We also know that each peak is connected to a particular $\boldsymbol{G}$, based on the miller index (l_1, l_2, l_3).

If we decrease d, we see that we will increase G for each peak. Therefore an increasing order of 2θ, the peaks (planes) go as (100), (110), (111), (200), (210), etc, in the order of increasing $\sqrt{l_1^2 + l_2^2 + l_3^2}$. Note that this only applies for powder diffraction, since that is the only time we will see all possible peaks.

Experimentally, we can find $f_j(\boldsymbol{G})$ via diffraction. Once we have a complete mapping of $f_j(\boldsymbol{G})$, we can do an inverse Fourier transform to get the $n_j(\boldsymbol{r})$, which then gets the real lattice structure.

For a monoatomic lattice, with 1 atom per unit cell, then $S(\boldsymbol{G}) = f(\boldsymbol{G})$, and then the intensity directly varies with the atomic form factor. We can also do a glancing angle diffraction, where the angle of incidence and reflection is very shallow. In this case, $\boldsymbol{G} \sim 0$, and then we find that $f(G) = Z$, the number of electrons.

3 Forces in Crystals

So far, we have created arrangements of atoms in crystals, and we have figured out how to experimentally determine the structure. Now let us ask some different questions. What holds a solid together? Why does a crystal have one structure and not another? For instance, monoatomic Fe crystals (α-Fe), is BCC. Why is it not FCC or HCP?

Let us first write down some properties of solids that we know. The first is that solids don't flow or spread out, so the atoms inside the solids are in fixed positions with respect to each other. It is also difficult to pull a solid apart, so there is something holding it together. We also know that it is difficult to compress a solid, so there is something keeping it apart. Thus we have that there are repulsive and attractive forces involved in solids, so the crystals assume structures where attractive and repulsive forces on each atom are balanced. It turns out that this arrangement is usually the lowest energy arrangement. Crystals can exist in meta-stable structures, local minima where forces are balanced, but there is a global minimum of energy which is the most stable.

For example, graphite is a more stable form of carbon than diamond. We can think of diamond and graphite as two wells in the plot of the energy in configuration space, and the graphite well is deeper than the diamond well.

3.1 Pauli Exclusion Principle

The Pauli Exclusion Principle states that no 2 particles can be in the same state at the same time. Each ψ is a state of electron, and each state is an orbital, denoted ψ_s, ψ_p, ψ_d, etc. Now when we go from a single particle wavefunction to a multiple particle wavefunction, $\psi(r_1, r_2)$ is the wavefunction that describes the probability amplitude of finding the first particle at r_1 and the second at r_2, at the same time. Electrons have two important properties in quantum mechanics. The first is that they are all identical to each other, there is nothing that distinguishes one electron from any other electron. In the classical picture, particles are billiard balls, and we can keep track of them. For example, in a classical collision, we can keep track of which particle is which. However, in QM, when we have a collision we cannot keep track of which particle is which. For this reason, our multi-particle wavefunctions are invariant under particle interchange, $\psi(r_1, r_2) = \pm\psi(r_2, r_1)$.

The second important property is that electrons are fermions, and because of this, they need to take the anti-symmetric case, so $\psi(r_1, r_2) = -\psi(r_2, r_1)$.

The Pauli Exclusion Principle arises when we attempt to have two fermions in the same state, the wavefunction will be 0 (since we assume that the wavefunction is written as $\psi(r_1, r_2) = \psi_a(r_1)\psi_b(r_2) - \psi_a(r_2)\psi_b(r_1)$). Another important thing to remember in physics is that high energy states are generally not favored, things settle to a lower energy state. Positive or increasing potential energy is associated with a repulsive force, and inversely, negative or decreasing energy is associated with an attractive force.

3.2 Repulsive Forces between Atoms/Ions

There are several difference sources of repulsion between atoms/ions. They can arise from

1. Electrostatic Force - Ions having the same sign charges will repel against each other.
2. Pauli Exclusion Principle- Prevents putting more than one electron in a given state.
3. Heisenberg Uncertainty Principle - $\Delta x \Delta p \geq \frac{\hbar}{2}$, and the energy is proportional to p^2, so squeezing any particles into a small volume costs energy, which generates repulsion.

The PEP and the Uncertainty principle prevent a crystal from collapsing in on itself, and why condensed matters are incompressible.

3.3 Attractive Forces between Atoms/Ions

These are essentially electrostatic in origin, the magnetic effects are generally quite small, and thus result in only small corrections to the electrostatic forces. We give different names for the types of interactions depending on their details.

Bond Type	Example	Strength	Binding Energy per Atom
Van der Waals	Ar	Weak	~ 0.1 eV/atom
Ionic	NaCl	Strong	~ 5 eV/atom
Covalent	Si	Strong	~ 5 eV/atom
Metallic	Na	Moderate	~ 1 eV/atom

3.3.1 Van der Waals Solids

These solids are not very strong, and only exist at low temperatures. Generally, they are uncharged atoms with closed shells. Some examples include noble gases. The melting point, (which is related to the strength of the bond), of the noble gases are shown below:

Element	Melting Point
Ne	24.5 K
Ar	83.8K
Kr	116 K
Xe	161 K

The repulsive force generated by the Van der Waals force in these solids is based on the fact that noble gases like Ne are closed shell:

$$\text{Ne} : 1S^2 2S^2 2P^6$$

When 2 Ne atoms are pushed together (when the electron clouds begin to overlap), they get become part of the same state, which means that the PEP begins to apply. To avoid this, one of the electrons will increase its energy (the shells are completely full so it must go to a higher energy), which is something that the system does not want. This is because it is a big energy expense for one electron to be in a higher shell, so this results in a repulsive force.

The repulsive force takes the form of

$$U(r) = \frac{A}{r^{12}}$$

Or $U(r) = U_0 e^{-\lambda r}$, where r is the distance between the atoms, and A, U_0, λ are constants. Here, for Van der Waals, we use the $\frac{A}{r^{12}}$ model. We note that it is positive and increases rapidly as r is decreased.

Now let us look at the attractive force. Where does it come from? The shells are full, and the atoms are neutral so its not electrostatic. The semi-classical explanation of the attractive force in Van der Waals solids is that we can treat atoms as dipoles, where the positive charge is the nucleus and the negative charge is the center of mass of the negative charge. The electron clouds are fluctuating, so at any given time, the center of mass of the cloud is not coincident with the nucleus position, so we can treat it like a dipole, without any real issues.

We can consider the dipole moment, p, which has magnitude $l(t)q$, where $l(t)$ is the time dependent distance between the two charges, and q is the charge.

We have the Lennard-Jones potential, which is an empirical formula:

$$J(R) = 4\epsilon\left[\left(\frac{\sigma}{R}\right)^{12} - \left(\frac{\sigma}{R}\right)^{6}\right]$$

Where the first term is repulsive and the second term is attractive, based on their signs. This is the potential between any two atoms. The total potential is given via a summation:

$$U_{\text{total}} = \frac{1}{2}4\epsilon N \sum_j \left[\left(\frac{\sigma}{R_{ij}}\right)^{12} - \left(\frac{\sigma}{R_{ij}}\right)^6\right]$$

Where we are summing over all atoms, and there are N of them, and the $\frac{1}{2}$ is there to stop the doublecounting.

If we plot the energy as a function of R, the atomic distance, we have a repulsive term that goes as R^{-12}, and also an attractive part that goes as R^{-6}. If we add these together, we have a bound state at the minimum of the potential, where we have a shallow potential well. To find this minimum, all we have to do is take the derivative and set it equal to 0, then solve for R, and we have R_0, the equilibrium distance. From there, we can find the equilibrium energy, $U_{\text{total}}(R_0)$. This is the cohesive energy, the amount of energy it takes to separate an atom from the crystal. If we compare the results of this model to the experimentally derived value of $\frac{R_0}{\sigma}$, where R_0 is found via x-ray diffraction and σ is found via gas properties, we find that the values are in good agreement.

Why is it that it forms an FCC solid at this point, and not something else like BCC? If we look at the energy expression U_{total}:

$$U_{\text{total}} = 2\epsilon N \left[\sum_j \left(\frac{1}{p_{ij}}\right)^{12} \left(\frac{\sigma}{R}\right)^{12} + \left(\frac{1}{p_{ij}}\right)^6 \left(\frac{\sigma}{R}\right)^6\right]$$

We see that the values of p_{ij} are purely geometrical, and for bcc we have different values than fcc. If we go through and compute the minimum energy distance, we find that R_0 for bcc is smaller than the R_0 for fcc, and therefore the energy for bcc is larger than the energy for fcc. This tells us that fcc structure has lower energy. This is why fcc is chosen over bcc.

Let us now discuss the effect of pressure on the Van der Waals crystals. As we squeeze the crystal, R will decrease. If we apply enough pressure, crystals will undergo a phase transition to a new structure. This makes sense if we draw a plot of U against R, where we see that as we decrease R, we will be pushed up to the left of the potential curve. Because of this, at some point, fcc will no longer be lower energy than bcc, and thus if we apply enough pressure, the crystal will change from fcc to bcc, since that is now lower energy than the fcc structure.

3.3.2 Ionic Crystals

Many crystals consist of ions, held together by electrostatic forces, such as the Coulombic attraction. These materials are often electrical insulators. Ionic crystals are commonly formed by an alkali metal (Na, K, Cr), and a halogen (Fl, Cl, I, Br, etc). Where do ions come from? Let us look at NaCl as an example. It is energetically favorable to transfer the electron from sodium to chlorine. We add some amount of energy to the sodium (the ionization energy) and we are left with Na^+ and an electron. On the other hand, if we give neutral chlorine an electron, it releases energy and we are left with Cl^-, and the amount is given by the electron affinity. Once the ions are formed, they will electrostatically attract each other, and repel like ions. The cohesive energy is once again the energy needed to separate the ions infinitely far from each other.

Ions have closed shells, so we have a strong repulsive interaction at close distances, due to the Pauli Exclusion Principle. We model the potential due to the PEP as

$$U_{\text{Pauli}} = \lambda e^{-R/\rho}$$

Where R is the distance between ions and λ and ρ are constants. The ions will form so that they have local neutrality. We have to look at the nearest neighbors, where every other neighbor is repulsive rather than attractive. We have a sort of checkerboard pattern of pluses and minuses. There will be a positive energy for the like-sign ions, and a negative energy for the differing sign ions, given by the Coulombic energy:

$$U_{\text{Coulombic}} = \frac{q_1 q_2}{4\pi\epsilon_0 R_{ij}}$$

Then we have the Pauli Exclusion Principle, which we only think about the nearest neighbors being repulsed. This is essentially because we're thinking of the ions as bouncy balls, and the only ones that we need to avoid overlaps with are the nearest neighbors:

$$U_{\text{Pauli}} = z\lambda e^{-R/\rho}$$

Where z is the number of nearest neighbors. We can then write out the total energy:

$$U_{\text{total}} = \frac{N}{2}\left(\sum_j -\frac{\pm e^2}{4\pi\epsilon R_{ij}} + z\lambda e^{-R/\rho}\right)$$

Where the ± alternates.

If we now let $R_{ij} = Rp_{ij}$, where R is the distance to the nearest neighbor:

$$U_{\text{total}} = \frac{N}{2}\left[\frac{e^2}{4\pi\epsilon_0 R}\sum_j (-1)\frac{\pm 1}{p_{ij}} + z\lambda e^{-R/\rho}\right]$$

The sum $\sum \frac{\pm 1}{p_{ij}}$ is known as the Madelung constant, and is denoted α, and only depends on the geometry of the structure.

3.3.3 Covalent Crystals

For covalently bonded crystals, they are more difficult to work with, since we need to deal with quantum chemistry. This is because we are no longer dealing with closed shells like in the Van der Waals and ionic cases. The filled shells aren't perturbed by adjacent atoms as much, so we can have simple models, but when we have partially filled shells, things get a lot more complicated, and there are no simple model potentials, we need quantum mechanics. In covalently bonded crystals, electrons rearrange to occupy available states. We must use QM to find ψ, and then we can calculate the energy of the state. The rough picture is electrons pair up to occupy space in between bonded atoms, and the electrons in a pair have opposite spins. The pairing takes place in such a way that each atom "sees" filled shells, this is the octet rule. We can use qualitative QM to understand what happens.

Suppose we have H_2, which has two protons and 2 electrons. First we need to find the correct Hamiltonian, which we can use to solve for the stationary states. In this case, the Hamiltonian will be of the form

$$\hat{H} = H_{\text{H}_1} + H_{\text{H}_2} + H'$$

Where the first two terms represent the Hamiltonians of the two atoms, and the third term is the interaction term. From this, we can obtain the solution for the two particle wavefunction. Two possible solutions come out:

$$\psi_1 = \chi_{\text{spin antiparallel}}\psi_{\text{spatial symmetric}}$$

$$\psi_2 = \chi_{\text{spin parallel}}\psi_{\text{spatial antisymmetric}}$$

We can see that these both obey the particle-interchange rule, since one term is anti-symmetric and another is symmetric in each of the wavefunctions. We can plot the symmetric and anti-symmetric wavefunctions, and we find that the symmetric case implies that there is a high probability that electrons exist in between nuclei, while in the anti-symmetric case, the probability amplitude goes to 0 in between the nuclei. The symmetric case provides a bound state, while the anti-symmetric case does not give a bound state. Thus we have that the crystals are formed when the wavefunction is the spatially symmetric wavefunction, which means that the spins are anti-parallel, as we stated before. The key is that by forming a pair, the electrons exist in a space between the atoms, rather than individually surrounding atoms, like in the closed shells. Since the electrons are in the space between, we have increases Δx. The Uncertainty Principle tells us that the energy will be on the order of $\frac{\hbar^2}{8m(\Delta x)^2}$, and we see that increasing Δx increases the energy. Thus the covalent bonds decrease the total energy, and thus the state is energetically favorable. Let us discuss properties of covalent bonds. They are strong, relative to the Van der Waals forces, they are “stiff”, they don’t bend, and are “directional”, they repel each other and spread out in space.

Let’s look at CH_4, methane as an example. C has 4 unfilled states, and thus will form 4 bonds. CH_4 has a tetrahedral arrangement, all the bonds are maximally separated from each other. This is known as the diamond structure, since it is the structure of the carbon in diamond, and it turns out that silicon also follows this structure. Essentially, covalent bonds are a mixture of existing states to produce density in between atoms. This is once again due to the uncertainty principle, in metallic bonding, all the electrons are shared by all the atoms, so Δx is very large. Because of this, the energy, which is $\frac{p^2}{2m} \geq \frac{(\Delta p)^2}{2m}$, and we can use the uncertainty principle to derive that the energy decreases as we increase Δx.

4 Phonons

Previously, we plotted potential wells, and we had states at the bottom of the well, jiggling about. We can equivalently describe that as a harmonic oscillator. The harmonic oscillator is obtained via $F = -kx$, and the associated potential energy is $\frac{1}{2}kx^2$. We can write out Newton’s second law:

$$m\frac{d^2x}{dt^2} + kx = 0$$

Which has solution $x = e^{i\omega t}$, where $\omega = \sqrt{\frac{k}{m}}$. The approximation that we make is that at every local minimum of a well, we can approximate it as a harmonic oscillator potential well, so $U(R)$ near $R = R_0$ is given via a Taylor expansion:

$$U(R) = U(R)|_{R_0} + \frac{\partial U}{\partial R}|_{R_0}(R - R_0) + \frac{\partial^2 U}{\partial R^2}|_{R_0}\frac{(R - R_0)^2}{2} + \dots$$

We note that the first term is just $U(R_0)$, the second term is 0, and the third term is of the form of the harmonic oscillator:

$$U(R) = U_{\min} + \alpha x^2$$

Where $x = R - R_0$. The higher order terms make this an anharmonic oscillator, and we will see in chapter 5 that those terms cause things like thermal expansion. Thus we have that a good starting model for an atom is a mass on a spring. Thus when we generalize to all atoms in the crystal, the model is masses connected to each other by springs. The motions of the lattice are called elastic waves, sound, or vibrations. We also introduce a particle, a lattice excitation called a phonon. Phonons are quantum mechanical descriptions of crystal vibrations. They are a manifestation of elastic waves, just like photons are electromagnetic waves. The phonons are responsible for heat transmission in insulators, thermal expansion, resistance in metals, and superconductivity in metallic superconductors.

We begin with a 1D monoatomic crystal. Let us make some assumptions. We assume that each atom couples only to its nearest neighbors, and we only have small oscillations/vibrations. Let us look at our chain of atoms. The sth atom has some displacement u_s from displacement R_s. We also have some lattice spacing a. The potential energy of the sth atom is given by $E_s = \frac{1}{2}c(u_s - u_{s+1})^2 + \frac{1}{2}c(u_s - u_{s-1})^2$, where we have taken into account the spring to the right and the spring to the left. The force on the atom is given by $F = -\frac{\partial E_s}{\partial u_s}$. From this, we have that

$$F_s = cu_{s+1} + cu_{s-1} - 2cu_s$$

We have N coupled equations, since we have a force on every single atom, and each will take into account the motion of two other atoms. We guess that the solution should have the form of oscillations in time and space:

$$u \propto e^{-i\omega t}$$
$$u \propto e^{ikx}$$

Except in this case we don't have x, we use sa instead. However, we note that x is usually a continuous variable, but sa denotes a discrete set of positions. Together, our guess is now

$$u_s(t) = ue^{-i(\omega t - ksa)}$$

Where $k = \frac{2\pi}{\lambda}$, and is the wavenumber, $\omega = 2\pi f$, the angular frequency, and the $\omega t - hsa$ implies that we are moving to the right.

We can insert this solution into the expression for the force on atom s, to check if the solution really does work:

$$\begin{aligned}
m\frac{d^2u_s}{dt^2} &= cu_{s+1} + cu_{s-1} - 2cu_s \\
m\left(-i\omega\right)^2 ue^{-i(\omega t - hsa)} &= cue^{-i(\omega t - h(s+1)a)} + cue^{-i(\omega t - h(s-1)a)} - 2cue^{-i(\omega t - hsa)} \\
-m\omega^2 &= ce^{ika} + ce^{-ika} - 2c \\
-m\omega^2 &= 2c\cos ka - 2c \\
-m\omega^2 &= 2c(\cos ka - 1)
\end{aligned}$$

From this, we have that

$$\omega^2 = \frac{2c}{m}(1 - \cos ka)$$

We see that we are nearing an expression that relates ω to k, the dispersion relation. We can use a trig identity:

$$\frac{1}{2} - \frac{1}{2}\cos\theta = \sin^2\frac{\theta}{2}$$

$$\omega^2 = \frac{4c}{m}\sin^2\frac{ka}{2}$$

Which leaves us with

$$\omega = \sqrt{\frac{4c}{m}}\left|\sin^2\frac{ka}{2}\right|$$

What does this mean? Provided that we have this relationship between ω and k, we have a valid wave-like solution. Recall that for waves, the phase velocity is given by $\frac{\omega}{k}$, and the group velocity is given by $\frac{d\omega}{dk}$. The group velocity can be thought of as the speed at which information/energy propagates. In this case, the group velocity is the speed of lattice vibrations, or the speed of phonons. Note that $\frac{d\omega}{dk}$ is not constant, it depends on k. We note that at the boundaries of Brillouin zones ($k = \pm\frac{\pi}{a}$), the phonons do not propagate, they cannot exist. This isn't a physical place, this is in reciprocal space.

In the case of sound, sound is low frequency, and long wavelength. In this limit, $\frac{2\pi}{\lambda} = k \to 0$, and thus $\sin\frac{ka}{2} \to \frac{ka}{2}$. Thus we have that $\omega \approx a\sqrt{\frac{c}{m}}k$, and the group velocity approaches a constant, $\frac{d\omega}{dk} \approx a\sqrt{\frac{c}{m}}$. This is also called the continuum limit, you get the same result if $\lambda \gg a$:

$$\frac{2\pi a}{\lambda} \ll 1 \to \frac{ka}{2} \ll 1$$

We also note that if k gives us a solution, then so does $k + \frac{2n\pi}{a}$, for any integer n. Thus by convention, we specify the range $-\frac{\pi}{a} \leq k \leq \frac{\pi}{a}$, which covers the first Brillouin Zone.

Let's look at $k = \frac{\pi}{a}$, which is where $\frac{d\omega}{dk} = 0$. We can then write out the equation for the wave:

$$\begin{aligned} u_s &= ue^{-i\omega t}e^{i\pi s} \\ &= (-1)^s ue^{-i\omega t} \end{aligned}$$

This tells us that every other atom is flipped with respect to the neighbors. In time, atoms are oscillating in the normal way, but spatially we have alternating positions. We can think of the atoms as being out of phase with their neighbors in k space. The solution at $k = -\frac{\pi}{a}$ is identical. This is "similar" to what happens to x-rays when $k = \pm\frac{\pi}{a}$, we are in phase and we have diffraction peaks. This is sort of like a resonance.

If we have that $k_{max} = \frac{\pi}{a}$, which is related to the wavelength:

$$\begin{aligned} k_{max} &= \frac{\pi}{a} \\ &= \frac{2\pi}{\lambda} \end{aligned}$$

We have that $\lambda_{min} = 2a$. We can also think about the frequency:

$$f_{max} = \frac{\text{speed of vibration}}{\lambda_{min}}$$

And this is generally around 6×10^{12} Hertz.

4.1 Longitudinal and Transverse Waves

Suppose now our 1D crystal is in 3D space. There are 3 possible types of waves:

1. Longitudinal Waves: Oscillating along the length of the crystal, along the direction of wave propagation.
2. Transverse y Waves: Oscillate perpendicular to the propagation direction in the y direction ("Horizontally"). Can also be referred to as being polarized in the y direction.
3. Transverse z Waves: Oscillate perpendicular to the propagation direction, in the z direction ("Vertically"). Can also be referred to as being polarized in the z direction.

Atom displacement can now be represented as a vector:

$$\boldsymbol{u}_s(t) = \hat{e}_i A e^{-i(\omega t - ksa)}$$

Where $\hat{e}_i$ can be either $\hat{x}$, $\hat{y}$, or $\hat{z}$. These are known as acoustic phonons, because they are associated with sound waves. It turns out that when we go from 1D crystals in 3D space to 3D crystal in a 3D space, you still get Longitudinal Acoustic branch, and 2 Transverse Acoustic branches on the plot of ω vs k. Instead of a line of atoms oscillating, we have a plane of atoms oscillating together.

4.2 Optical Modes

Suppose we have two different atoms per unit cell. So far, we have only looked at monoatomic crystals. Let's suppose as before that we have nearest neighbor interactions, and atoms with masses m and M. Suppose we denote x_s to be the displacement of the s'th small atom, and v_s to be the displacement of the s'th large atom. The potential energy of the two s'th atoms is given by

$$\begin{aligned} E_{x_s} &= \frac{1}{2}C(x_s - v_{s-1})^2 + \frac{1}{2}C(x_s - v_s)^2 \\ E_{v_s} &= \frac{1}{2}C(v_s - x_s)^2 + \frac{1}{2}C(v_s - x_{s+1})^2 \end{aligned}$$

We can then write out the equations of motion:

$$\begin{aligned} F_{x_s} &= m\frac{d^2x_s}{dt^2} \\ &= -\frac{\partial E_{x_s}}{\partial x_s} \\ &= -2cx_s + cv_{s-1} + cv_s \\ F_{v_s} &= M\frac{d^2v_s}{dt^2} \\ &= -\frac{\partial E_{v_s}}{\partial v_s} \\ &= -2cv_s + cx_s + cx_{s+1} \end{aligned}$$

We see that we have two coupled equations, and we guess a wavelike solution, in time and space:

$$\begin{aligned} x_s(t) &= xe^{i(\omega t - ksa)} \\ v_s(t) &= ve^{i(\omega t - ksa)} \end{aligned}$$

We can plug these in, then drop the exponentials, since those pull out of the diffeq, and we are left with

$$\begin{aligned}-m\omega^2 x &= -2Cx + Cv + Cve^{ika}\\ -M\omega^2 v &= -2Cv + Cx + Cxe^{-ika}\end{aligned}$$

Now solving these for x:

$$x = \frac{v(Ce^{ika} + C)}{-m\omega^2 + 2C}$$

And we can then generate a quadratic in terms of ω^2:

$$4 + \frac{mM\omega^4}{C^2} - \frac{2m\omega^2}{C} - \frac{2M\omega^2}{C} = 2(1 + \cos ka)$$

Which gives us that

$$\omega^2 = \left(\frac{C}{M} + \frac{C}{m}\right) \pm \sqrt{\left(\frac{C}{M} + \frac{C}{m}\right)^2 - \frac{2C^2(1-\cos ka)}{mM}}$$

This gives us the full dispersion relation for 2 atoms in a unit cell. To see basic solutions, we look at limits, such as $k \to 0$, where we get that

$$\omega^2(k=0) = \frac{C}{M} + \frac{C}{m} \pm \left(\frac{C}{M} + \frac{C}{m}\right)$$

The $-$ solution gives us $\omega = 0$, which is the same acoustic solution as before. The $+$ solution gives a new type of solutions, the longitudinal optical mode.

If we look at the limit where $k \to 0$, we find that

$$\omega = \sqrt{2\left(\frac{C}{M} + \frac{C}{m}\right)}$$

And from this we find that

$$x = -\frac{M}{m}v$$

The smaller atoms and the larger atoms are moving in opposite directions in the optical mode, and the small mass moves $\frac{M}{m}$ times more. These are longitudinal waves, but there are transverse modes. Why is this called the optical mode? Suppose m has a positive charge and M has a negative charge. This is the setup for an ionic crystal. If this crystal oscillates in the TO mode, every half wavelength the charges will clump together, and we have a dipole wave, which gives us an electric field. This wave fits the profiles of a travelling EM wave, and thus is conducive to being set up by light. There are always 1 LA mode and 2 TA modes, but optical modes are different, they arise from having different atoms in a unit cell. Thus if we have n atoms in our unit cell, we will have $n-1$ LO modes, and $2(n-1)$ TO modes. Adding them together, the total number of modes/branches is $3(n-1)$ optical modes, and then the three acoustic modes.

4.3 Quantization of Lattice Vibrations

Everything that we have done so far was based on classical mechanics, we had the beads on a string model. We need to start using quantum mechanics to understand collective behavior, such as thermal properties. We will also need QM to deal with interaction with other excitations/particles, namely photons and neutrons. Classically, we used collections of harmonic oscillators coupled together, but phonons are quantum mechanical entities. Formal treatment of phonons requires the quantum harmonic oscillator Hamiltonian. The results of the Schrodinger equation tells us that the energy of a vibration/elastic wave mode with a particular ω (with a corresponding k) is given by $\varepsilon = \left(n + \frac{1}{2}\right)\hbar\omega$. The energy depends on n, which is the number of phonons in the state.

When we have n phonons, each carrying energy $\hbar\omega$, when n is large, we say that this state is excited, and the modes occupied by n phonons. For each mode, there is a "zero point" energy $\frac{1}{2}\hbar\omega$. This arises from fluctuations of wavefunctions, together with the uncertainty principle. The total energy of the lattice, taking into account all modes, is given by

$$U = \sum_k \left(n + \frac{1}{2}\right)\hbar\omega(k)$$

Another important concept is the crystal momentum:

$$\boldsymbol{p} = \hbar\boldsymbol{k}$$

This is associated with the phonon, and is used to understand scattering phenomena. This is decidedly different from the real, physical momentum associated with atoms vibrating.

4.4 Neutron Scattering

When a neutron strikes the crystal, we have the emission of a phonon. We have some incident energy ε_n, and wavevector $\boldsymbol{k}_n$. The neutron is reflected, so we have a reflected energy, ε_n' and wavevector $\boldsymbol{k}_n'$. The induced phonon has some energy $\varepsilon_{ph} = \hbar\omega$, and some wavevector $\boldsymbol{k}_{ph}$, which is related to the crystal momentum $\boldsymbol{p} = \hbar\boldsymbol{k}_{ph}$. To analyze scattering, we do the usual conservation laws (energy and momentum), and we find ω and $\boldsymbol{k}_{ph}$:

$$\varepsilon_n = \varepsilon_n' + \hbar\omega$$

However momentum is not conserved in the usual way, which is an indication that the crystal momentum is not the usual momentum, we have to add a new vector $\boldsymbol{G}$:

$$\boldsymbol{k}_n = \boldsymbol{k}_n' + \boldsymbol{k}_{ph} + \boldsymbol{G}$$

We "pick" $\boldsymbol{G}$ so that all $\boldsymbol{k}$ s remain inside the first BZ. Similar experiments can be conducted using light/laser, which is known as Raman scattering. Many important physical properties (functional properties) are directly tied to phonons. One example of this is phonon mediation giving rise to superconductivity, which is known as the BCS mechanism.

4.5 Ferroelectricity

Another example, which we will learn more about, is ferroelectricity. This is associated with the lowest frequency TO mode. This mode is called the soft mode. The onset of ferroelectricity, with the lowering of temperature, is accompanied by the softening of the TO mode, the lowering of the frequency. A classic example of this is $BaTiO_3$, which is not ferroelectric above 130 degrees Celsius, it is a paraelectric. Above this temperature, we can drive the TO mode with a laser/light. As we lower the temperature, the mode softens more and more, and at 130 degrees, the mode dies, freezing the crystal in its current position. The temperature 130 degrees is known as the Curie temperature for the crystal. Below this temperature, the material is tetragonal.

When the structure freezes, the crystal structure generates a dipole and thus the crystal becomes one big dipole, we have a sheet of positive charge on the top, and a sheet of negative charge on the bottom, and inside, the individual unit cells cancel each other out, and we are neutral. This is static and non-volatile, the structure stays even as time passes.

By applying an electric field pulse, we can flip the charge once. These two states can be used as memory states (non-volatile memory). In order to explore new ferroelectric materials, phonon behavior acts as an important indicator, softening modes may indicate a new ferroelectric transition.

4.6 Phonons and Thermal Properties

We will do statistical mechanics (classical and quantum), in order to discuss the thermal properties of a material induced by phonons. Phonons will help us explain heat capacity, thermal conductivity, and thermal expansion. Thermal energy is the energy associated with temperature, and we relate this to the energy of phonons. In electrical insulators, phonons account for all/most of the thermal properties, but in electrical conductors, they contribute to thermal properties.

The **Heat Capacity** is the amount of energy required to raise the temperature of a material by 1 Kelvin per mole. This is a real important property which tells us fundamental properties, such as phases, phase transitions, etc. This is difficult to measure experimentally, but is very important for characterizing materials. Understanding of heat capacity was lacking until about 100 years ago. Classical physics could not explain it, and the problem was solved via the introduction of quantum mechanics, first by Einstein, and then by Debye. The problem of heat capacity was one of the first successful applications of QM. We will begin with the classical treatment, then move to the Einstein method, and then finally Debye's method and the density of states.

4.6.1 Classical Method

The classical model of phonons is vibration of atoms connected to each other with springs. The total energy is given by the kinetic energy and the potential energy:

$$E = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + \frac{C}{2}(x^2 + y^2 + z^2)$$

Where p denotes the momentum in a direction, and x, y, and z represent displacement from equilibrium positions.

From classical statistical physics, the probability of finding a system in a state with energy E is proportional to $e^{-\frac{E}{k_B T}}$, which is known as the Boltzmann factor.

Let there be a parameter A for the system, which takes on different values A_r based on which state it is in. If the system is in state r, then the associated value of A is A_r. The average value of the parameter will be

$$\langle A\rangle = \frac{\sum_r A_r e^{-\frac{E_r}{k_BT}}}{\sum_r e^{-\frac{E_r}{k_BT}}}$$

In statistical physics, momenta are also degrees of freedom, so we have 6 degrees of freedom in our classical system. The average energy will be given via

$$\langle E\rangle = \sum_f \langle E_f\rangle$$

Where f is a degree of freedom. We can then write this as

$$\langle E\rangle = \langle\frac{1}{2m}p_x^2\rangle + \langle\frac{1}{2m}p_y^2\rangle + \langle\frac{1}{2m}p_z^2\rangle + \langle\frac{c}{2}x^2\rangle + \langle\frac{c}{2}y^2\rangle + \langle\frac{c}{2}z^2\rangle$$

Now using the equipartition theorem, which states that each degree of freedom whose energy contribution is quadratic, has average energy $\frac{1}{2}k_BT$:

$$\langle U\rangle_{\text{classical, per atom}} = 3k_BT$$

Thus for the entire crystal, which has $6N$ degrees of freedom:

$$U_{\text{Thermal}} = 3Nk_BT$$

We then relate this to the heat capacity, C_V (where we hold volume constant):

$$\begin{aligned}C_V &= \frac{\partial U_{\text{thermal}}}{\partial T}\\ &= 3Nk_B\end{aligned}$$

Classical physics says that all crystals have the same heat capacity at any temperature, there is no temperature dependence. This is clearly not true, in real life, different crystals have different heat capacities, and there are temperatures dependencies. It turns out that experimentally defined curves match the classical answer, but only at high temperatures. As $T \to 0$, we expect that $C_V \to 0$.

4.6.2 Einstein's Model

Einstein made some assumptions when using quantum mechanics. The first was that we were using the quantum harmonic oscillator, rather than the classical oscillator. The second was that each atom vibrates at a fixed frequency ω_0, in all 3 directions. Next is that ω_0 is the same for all atoms in a crystal.

The thermal energy according to Einstein will be the number of atoms, N, times the average energy of a QHO in 1 dimension, times the number of dimensions, in this case 3.

Recall that for a QHO:

$$E = \hbar\omega_0\left(n + \frac{1}{2}\right)$$

We now need to compute the average of this energy:

$$\langle E\rangle = \hbar\omega_0\left(\langle n\rangle + \frac{1}{2}\right)$$

We need to compute the average excitation of the oscillator:

$$\langle n\rangle = \frac{\sum_n ne^{-\hbar\omega_0(n+\frac{1}{2})/k_BT}}{\sum_n e^{-\hbar\omega_0(n+\frac{1}{2})/k_BT}}$$

Where the denominator is known as the partition function. We can rewrite this as

$$\begin{aligned}\langle n\rangle &= \frac{\sum_n ne^{-\hbar\omega_0 n/k_BT}}{\sum_n e^{-\hbar\omega_0 n/k_BT}}\\ &= -K_BT\frac{\frac{\partial}{\partial\hbar\omega_0}\sum e^{-\hbar\omega_0 n/k_BT}}{\sum e^{-\hbar\omega_0 n/k_BT}}\\ &= -k_BT\frac{\partial}{\partial\hbar\omega_0}\ln\sum_n e^{-\hbar\omega_0 n/k_BT}\end{aligned}$$

Where we have used the fact that

$$\frac{\partial}{\partial\hbar\omega_0}\left[e^{-\hbar\omega_0 n/k_BT}\right] = -\frac{n}{k_BT}e^{-\hbar\omega_0 n/k_BT}$$

Now looking at $\sum_n e^{-\hbar\omega_0 n/k_BT}$, this is the same as

$$\sum_n e^{-\hbar\omega_0 n/k_BT} = 1 + e^{-\hbar\omega_0/k_BT} + e^{-2\hbar\omega_0/k_BT} + \dots$$

This is an infinite geometric series, and converges:

$$\begin{aligned}\langle n\rangle &= -k_BT\frac{\partial}{\partial\hbar\omega_0}\ln\left[\frac{1}{1-e^{-\hbar\omega_0/k_BT}}\right]\\ \sum_n e^{-\hbar\omega_0 n/k_BT} &= k_BT\frac{\partial}{\partial\hbar\omega_0}\ln\left[1-e^{-\hbar\omega_0/k_BT}\right]\\ &= k_BT\frac{\frac{1}{k_BT}e^{-\hbar\omega_0/k_BT}}{1-e^{-\hbar\omega_0/k_BT}}\\ &= \frac{1}{e^{\hbar\omega_0/k_BT}-1}\end{aligned}$$

This is the Planck distribution. Now going back to Einstein's thermal energy:

$$\begin{aligned}U_{\text{Einstein}} &= 3N\langle E\rangle\\ &= \hbar\omega_0 3N\left[\frac{1}{e^{\hbar\omega_0/k_BT}-1} + \frac{1}{2}\right]\end{aligned}$$

Thus the heat capacity will be the derivative of this:

$$C_V = \frac{\partial U}{\partial T}$$

$$
\begin{aligned}
&= \frac{3N\hbar\omega_0 e^{\hbar\omega_0/k_BT}}{\left[e^{\hbar\omega/k_BT}-1\right]^2}\frac{\hbar\omega_0}{k_BT^2} \\
&= 3Nk_B\left[\left(\frac{\hbar\omega_0}{k_BT}\right)^2\frac{e^{\hbar\omega_0/k_BT}}{\left[e^{\hbar\omega_0/k_BT}-1\right]^2}\right]
\end{aligned}
$$

This is a much better approximation. Looking at the low temperature limit, as $T \to 0$, $\frac{\hbar\omega_0}{k_BT} \gg 1$, so we have that

$$C_V \approx 3Nk_B\left(\frac{\hbar\omega_0}{k_BT}\right)^2 e^{-\hbar\omega_0/k_BT}$$

For the high temperature limit, as $T \to \infty$, $\frac{\hbar\omega_0}{k_BT} \ll 1$, and we can look at $\langle n\rangle$:

$$\langle n\rangle \approx \frac{k_BT}{\hbar\omega_0}$$

If we then compute U:

$$
\begin{aligned}
U &\approx 3N\hbar\omega_0\left[\frac{k_BT}{\hbar\omega_0}+\frac{1}{2}\right] \\
&\approx 3Nk_BT
\end{aligned}
$$

We see that in the high temperature limit, the Einstein model matches the classical model. Einstein's model is qualitatively correct, but quantitatively, it is not quite correct.

Experiments find that $C_V \propto T^3$, as $T \to 0$, not the exponential dependence that Einstein's model derives.

4.6.3 Density of States and the Debye Model

Debye's argument was the solids/phonons don't oscillate at just one frequency ω_0. We know, for the acoustic branch, ω runs from 0 to $\sqrt{\frac{2C}{M}}$, we need to consider a range of ω, and we also need to obey the distribution of allowed values of ω.

We define $D(\omega)d\omega$ to be the number of states or modes in a frequency interval given by $[\omega, \omega + d\omega]$. $D(\omega)$ is the density of states in frequency space. We can once again write out the thermal energy:

$$
\begin{aligned}
U_{\text{thermal}} &= \sum_k\left(n_k+\frac{1}{2}\right)\hbar\omega_k \\
&= \sum_\omega\left(n_\omega+\frac{1}{2}\right)\hbar\omega \\
&\approx \sum_\omega n_\omega\hbar\omega \\
&= \sum_\omega\langle n_\omega\rangle\hbar\omega
\end{aligned}
$$

If the sum hits a value of ω that is not allowed, it will not contribute to the sum, which is governed by $D(\omega)$. We will use this when we convert this sum to an integral, using the Debye approximation:

$$U_{\text{thermal}} = \int_0^\infty \langle n_\omega\rangle\hbar\omega D(\omega)\,d\omega$$

We know what $\langle n_\omega \rangle$ is, its just the Planck distribution, so all that is left to do is finding $D(\omega)$. This comes from the boundary condition, we have a finite number of allowed states/modes, because of boundary conditions in the system.

Let us do an example, a simple natural boundary condition for our 1D monatomic crystal. The boundary condition is that the ends of the crystal are fixed, they do not oscillate. We have N atoms, and the total length is $L = Na$. What are the allowed values of k? To do this, we look for values of k that fit the system. We will intuitively have standing waves. The first is one with wavelength $2L$, which gives a value of $k_{min} = \frac{2\pi}{\lambda} = \frac{\pi}{Na}$. The second mode will be the one that fits a full wavelength into the crystal, so $\lambda = Na$, and thus $k_2 = \frac{2\pi}{Na}$. k_3 will be $\frac{3\pi}{Na}$. Looking at the trend, we have that the allowed values of k increment by $\frac{\pi}{Na}$.

Using the Debye approximation, we say that $\frac{d\omega}{dk}$ is a constant, v. Thus we have that $\omega = vk$. From this, we have that

$$\omega_n = \frac{nv\pi}{Na}$$

The density of states, $D(\omega)$, can be written as

$$\begin{aligned} D(\omega) &= D(\omega)\frac{d\omega}{d\omega} \\ &= \frac{1}{\Delta\omega} \\ &= \frac{1}{\frac{v\pi}{Na}} \\ &= \frac{Na}{v\pi} \end{aligned}$$

Where we have used the fact that $\Delta\omega$ is the distance between allowed modes, so the reciprocal will be the density. The lesson here is that we have applied a boundary condition, and we obtained discretized k and ω.

Let's do another example. This is the standard boundary condition, the periodic boundary condition. In this case, when we reach an edge, we loop around, and connect it to the other end. In the 1D case, this means that the displacement of the 0th atom must be the same as the displacement of the Nth atom:

$$x_0(t) = x_N(t)$$

Or more generally:

$$x_S(t) = x_{(S+N)}(t)$$

Applying that to the displacement solution $x_S(t) = Ae^{i(kSa-\omega t)}$:

$$Ae^{i(kSa-\omega t)} = Ae^{i(k(S+N)a-\omega t)}$$

Which leads us to $kNa = 2\pi n$, where $n = 0, \pm 1, \pm 2, \cdots \pm \frac{N}{2}$. These limits come about because k is bounded from $-\frac{\pi}{a}$ to $\frac{\pi}{a}$. We can then write the density of states in k space, since we have that $\Delta k = \frac{2\pi}{L}$:

$$D(k) = \frac{D(k)dk}{dk}$$

$$\begin{aligned}&= \frac{1}{\Delta k}\\ &= \frac{L}{2\pi}\end{aligned}$$

Now we need to convert this to the density of states in terms of ω. To do this, we look at the relationship between k and ω. $D(\omega)\Delta\omega$ is the number of allowed states between ω and $\omega + \Delta\omega$. This is the same as the number of states between k and $k + \Delta k$, but we also have to account for the number of states between $-k$ and $-k - \Delta k$, which is $D(k)\Delta k + D(-k)\Delta k$. We can compute this to be $\frac{L}{\pi}\Delta k$.

Thus we can write out $D(\omega)$:

$$\begin{aligned}D(\omega) &= \frac{L}{\pi}\frac{\Delta k}{\Delta\omega}\\ &= \frac{L}{\pi}\frac{dk}{d\omega}\\ &= \frac{L}{\pi}\frac{1}{v}\end{aligned}$$

Let us now move to the 3D case. We can use the same argument. In 1D, we have $D(k) = \frac{L}{2\pi}$, because there is 1 allowed state in every $\frac{2\pi}{L}$. In 3D, this becomes 1 state per $\left(\frac{2\pi}{L}\right)^3$ region of volume in k space:

$$D_{3D}(k) = \left(\frac{L}{2\pi}\right)^3$$

We now once again want to convert this to be in terms of ω:

$$D(\omega)d\omega = D(k)d^3\boldsymbol{k}$$

Since the total number of states is spherically symmetric, we decide to switch to spherical coordinates. Working through the Jacobian transformation, we convert $d^3\boldsymbol{k}$ into $4\pi k^2 dk$. Another way of thinking about it is that we have a spherical shell, with surface area $4\pi k^2$. Thus we have that $d^3\boldsymbol{k} = 4\pi k^2 dk$, and we need $\frac{d^3\boldsymbol{k}}{d\omega}$, which is then $\frac{4\pi k^2 dk}{d\omega}$. Now using the Debye approximation, we have that $\frac{d^3\boldsymbol{k}}{d\omega} = \frac{4\pi\omega^2}{v^3}$. Putting this all together, we have that

$$D_{3D}(\omega) = \left(\frac{L}{2\pi}\right)^3 \frac{4\pi\omega^2}{v^2}$$

for $\omega \in (0, \omega_{max})$.

We can then compute the thermal energy in 3 dimensions:

$$\begin{aligned}U_{\text{thermal}} &= 3\int_0^\infty \langle n_\omega\rangle \hbar\omega D(\omega)d\omega\\ &= 3\left(\frac{L}{2\pi}\right)^3 \frac{4\pi}{v^3}\int_0^{\omega_D} \omega^2\hbar\omega \langle n_\omega\rangle \, d\omega\end{aligned}$$

Where the 3 comes from the 3 different branches of modes. Now inserting the Planck distribution in for the average, and then doing some work to rewrite the integral, we find that

$$U_{\text{thermal}} = \frac{12\pi}{v^3}\left(\frac{L}{2\pi}\right)^3 \hbar\left(\frac{k_B T}{\hbar}\right)^4 \int_0^{\Theta_D} \frac{x^3\,dx}{e^x - 1}$$

Where $\Theta_D = \frac{\hbar\omega_D}{k_B}$. Now looking at limits, as $T \to \infty$, $x \to 0$, so $e^x \approx 1 + x$. In this limit, we find that $U_{\text{thermal}} \propto T$, and thus C_V is constant. This is the classical limit. For the opposite limit, $T \to 0$ means that $x \to \infty$, and the integral converges to a constant. In this case, we have that $U_{\text{thermal}} \propto T^4$, and $C_V \propto T^3$, which is the relation that we are looking for, this matches the experimental temperature dependence.

4.7 Superconductivity

BCS superconductivity is phonon mediated pairing of electrons. Electrons are negatively charged, so normally they wouldn't pair up, but experimentally, in superconductors, we see them pairing up. When paired, they are no longer fermions, and they can fall into the same state. How do we get them to pair? Suppose we have positively charged ions at the lattice sites, and we have an electron flying through. The ions become slightly attracted to the electron as it moves through, and that leaves a wake. This is a distortion of the lattice, which is a phonon. This wake generates a concentration of positive charges, which in turn attracts another electron. This then overcomes the Coulombic interaction, and gives us our electron pair. Phonons go as $\omega = \sqrt{\frac{c}{m}}$, and thus the highest energy is $\hbar\omega_D$, which is proportional to $\sqrt{\frac{c}{m}}$. This tells us that the superconducting transition temperature T_C should scale inversely with m. This is why the first thing people do when a new superconductor is discovered is looking at isotopes.

4.8 Anharmonic Phonon Effects

So far, we have been working with the harmonic oscillator model, the potential is just a quadratic term. In this case, it is only valid for small oscillations, we assume that the atom doesn't oscillate far out of the well, and we assume the well looks like a quadratic well. Recall that we obtained this via a Taylor expansion:

$$U(r) = U(r_0) + \frac{dU}{dr}\Big|_{r_0}(r - r_0) + \frac{1}{2}\frac{d^2U}{dr^2}\Big|_{r_0}(r - r_0)^2 + \dots$$

Where the second term is 0, and thus we have a linear term and a quadratic term. Suppose we now look at the next term:

$$U(r) = U(r_0) + \frac{1}{2}\frac{d^2U}{dr^2}\Big|_{r_0}(r - r_0)^2 + \frac{1}{3!}\frac{d^3U}{dr^3}\Big|_{r_0}(r - r_0)^3 + \dots$$

Thus we have that

$$U(r) = \text{Constant} + \frac{1}{2}c(r - r_0)^2 + \frac{1}{6}g(r - r_0)^3$$

Where g is a constant, and is negative (based on the shape of the potential).

4.8.1 Thermal Expansion

Let us now look at what the anharmonic term gives us, and we will see that it provides the process behind thermal expansion. For the pure quadratic oscillator, at low temperatures, the atoms oscillate within the confines of the well, and at higher temperatures, the atoms oscillate over longer and longer ranges. Since we have a quadratic curve, the range of the oscillations is always symmetric, and we see that the midpoint of the oscillations will always remain the same, we see no thermal expansion effects as temperature increases. Let us now look at the model with the anharmonic term. If we increase the temperature, we increase the range of oscillations, and we see that, since the potential is not symmetric, the median position shifts as the temperature changes, it shifts to the right as temperature increases. This provides a visual explanation of why we obtain thermal expansion from the introduction of the anharmonic term, the average displacement between atoms increases.

Now let us do a more mathematical explanation. The average displacement at a particular temperature is given by a statistical average:

$$\langle r\rangle = \frac{\int rP(r)\,dr}{\int P(r)\,dr}$$

Where $P(r)$ is the probability of finding the displacement to be r. Skipping the computation and displaying the result, we have that

$$\frac{\langle r\rangle}{r_0} = 1 + \alpha\Delta T$$

Where ΔT is the change in temperature, and α is the thermal expansion coefficient :

$$\alpha = \frac{1}{r_0}\frac{\partial\langle r\rangle}{\partial T}$$

We can also express α in terms of the constants c and g from the definition of $U(r)$:

$$\alpha = \frac{A|g|}{c^2}$$

Where A is another constant. Based on experimental values, we see that smaller α (and therefore larger c) means that a material is harder (Diamond has a small value of α, while aluminum has a large value of α).

4.8.2 Thermal Conductivity

Let us now look at what the anharmonic term gives us in terms of explaining thermal conductivity (for electrical insulators). Thermal conductivity is a measure of how well a material transmits heat. For example, suppose we have a rod of material, and one end is being heated up via a flame, and the other end is not being heated. The thermal conductivity provides a measure for how the temperature on the unheated end will change, heat begins to spread to the cold end.

Let us look at a rod of material where the left end is at temperature T_1, and the right end is at temperature T_2, where $T_1 > T_2$ and the temperatures are fixed. We see that heat will always transfer from the high temperature end to the low temperature end.

We can describe the flux of the thermal energy per unit area, per unit time:

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

Where κ is the thermal conductivity coefficient, and $\frac{dT}{dx}$ is the temperature gradient across the material. κ can be fully described using kinetic theory. For an insulator:

$$\kappa = \frac{1}{3}cvl$$

Where c is the heat capacity, v is the phonon velocity, and l is the mean free path, the distance between phonon scattering events (can be thought of as the distance a phonon travels between striking another phonon, where we consider the phonon to be a particle).

So what is the relationship between the anharmonic terms and thermal conductivity? Using QM, we can relate the anharmonicity to the mean free path. The anharmonic term is a perturbation to the harmonic oscillator Hamiltonian. The perturbation causes transitions between the original energy eigenstates, and these transitions are where the phonon scattering takes place. The key process is known as the 3 phonon process, where we have 2 phonons that interact, and we see 1 phonon as output. There are 2 different types of 3 phonon processes, which have slightly different momentum conservation. The first kind of 3 phonon process is known as the normal process:

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_3$$

Where $\boldsymbol{k}_i$ is the momentum of the ith phonon. The second type of process is called the Umklapp process:

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_3 + \boldsymbol{G}$$

In the normal process, the phonon wave vector is conserved, and in the Umklapp process, we have to take into account the $\boldsymbol{G}$ vector that we saw before when discussing phonon momentum. This Umklapp process affects the mean free path, because $l \propto \frac{1}{\text{Number of Umklapp Phonons}}$.

Where does the $\boldsymbol{G}$ come from? In the normal process (often denoted N process), the phonon momentum vectors of the first two phonons sum to exactly the resultant phonon's momentum, we have standard collision momentum and energy conservation, no loss of momentum or energy. In the Umklapp process (often denoted U process), the resultant phonon momentum is different than the sum of the momenta of the first two phonons, by a factor of $\boldsymbol{G}$. What is happening is that the sum of the phonon momenta $\boldsymbol{k}_1 + \boldsymbol{k}_2$ is going outside of the boundaries of the first Brillouin zone, and so $\boldsymbol{G}$ works to keep the phonon inside the zone. If $\boldsymbol{k}_1 + \boldsymbol{k}_2$ remains inside the BZ, then $\boldsymbol{G}$ will be 0. Our final conclusion is that the Umklapp phonon reduces the mean free path, since it restricts phonons to the first Brillouin zone. Thus it reduces κ, and therefore it reduces the thermal conductivity.

5 Free Electron Model of Metals

To begin, we will first look at the classical model of metals, and then we will proceed to a quantum mechanical treatment of metals. The behaviour of electrons in solids is the basis for electrical devices. Electronics didn't exist 100 years ago, QM is essential for understanding electrical properties.

The Drude Model/Theory was established around 1900, and is the basic classical model for solids. In 1897, JJ Thomson discovered electrons, and the basic picture was that electrons together acted like a classical gas of particles, bouncing around. Thomson correctly figured out that the positive ions were heavy, and were fixed in position. This is a very unusual gas, the density is very high, 10^{22} to 10^{23} electrons per cubic centimeter, rather than typical real gases, which have densities around 10^{19} particles per cubic centimeter. This density is high enough that QM becomes necessary. Based on this simple idea, combined with classical kinetic theory and statistical physics, many predictions can be made, some of which are correct (Ohm's Law is predicted correctly), and some of which are incorrect. Ohm's Law in this gas model is written as

$$\boldsymbol{J} = \sigma \boldsymbol{E}$$

Where $\boldsymbol{J}$ is the current density, $\boldsymbol{E}$ is the electric field, and $\sigma = \frac{ne^2\tau}{m}$ is the conductivity, where τ is the collision time between scattering of electrons and n is the carrier density. One prediction that failes is the heat capacity due to electrons:

$$U_{\text{thermal, classical}}^{\text{element}} = \frac{3}{2}Nk_BT \to C_V = \frac{3}{2}Nk_B$$

Where we have the same problem as before, this says that all materials have the same heat capacity, which we know to be physically untrue. In real materials, there is a linear relationship with temperature:

$$C_V^{\text{electron}} = \gamma T$$

Where γ is a constant.

In order to resolve the discrepancies, QM was applied, where metals are treated as free electron Fermi gases. We make several assumptions:

1. Electrons must be treated as QM entities.
2. Neglect interactions between electrons.
3. Ignore presence of ions (for now).
4. There is an average collision time τ.

The key questions are, how do we find electronic properties using quantum mechanics? The way we proceed is to note that some things are the “same” as what we did for phonons, but many things are vastly different, for example, phonons are bosons while electrons are fermions. To begin, we determine the allowed quantum states for electrons. We then will determine the allowed energy levels, and then the density of states of electrons, and then the occupancy of the states.

The Hamiltonian for this system is of a particle in a box with side length L. If we disregard interaction between electrons, we have that

$$H = \sum_i^N H_i$$

Where we have no potential acting on the particles, so $H_i = \frac{p_i^2}{2m}$. Thus we have that

$$H = \sum_i^N \frac{p_i^2}{2m}$$

Every electron will have that same term. Thus, all we have to do is to solve the Schrodinger equation for a single particle:

$$\begin{aligned}\hat{H}_i &= \frac{p_i^2}{2m} \\ &= -\frac{\hbar^2}{2m}\nabla^2\end{aligned}$$

As usual, we guess the solution for the single-particle wavefunction:

$$\begin{aligned} \psi &= Ae^{i\boldsymbol{k}\cdot\boldsymbol{r}} \\ &= \psi_{\boldsymbol{k}}(\boldsymbol{r}) \end{aligned}$$

In this case, $\boldsymbol{k}$ is the electron wavevector, and $|\boldsymbol{k}| = k = \frac{2\pi}{\lambda}$. $\hbar\boldsymbol{k}$ is the electron momentum, and A is a normalization constant such that

$$\int_{\text{Volume}} |\psi|^2 \, d^3\boldsymbol{r} = 1$$

This integral just tells us that the electron must be somewhere in the box, and thus

$$A = \frac{1}{\sqrt{L^3}}$$

Now we need to find the allowed values of $\boldsymbol{k}$. Let us assume periodic boundary conditions. If we do this, we end up with

$$e^{ik_xL} = e^{ik_yL} = e^{ik_zL} = 1$$

And thus we have that

$$\begin{aligned} k_x &= \frac{2\pi}{L}n_x \\ k_y &= \frac{2\pi}{L}n_y \\ k_z &= \frac{2\pi}{L}n_z \end{aligned}$$

Where the ns can take on values $0, \pm1, \pm2, \ldots$. Note that there is no bound on the values of the ns, since we are not confined to the first Brillouin zone. In general, we can express this as a vector:

$$\boldsymbol{k} = \frac{2\pi}{L}\boldsymbol{n}$$

We can plot allowed values in k space (as a 3 dimensional plot), and we see that we have an allowed value at every point where all 3 values are integers, and the distance between allowed values is $\frac{2\pi}{L}$ in each direction. We have 1 allowed value of $\boldsymbol{k}$ for every $\left(\frac{2\pi}{L}\right)^3$ chunk of volume in k space. For each of these points, we can fit 2 electrons, one with spin up and one with spin down. Thus the density of states in 3 dimensions is given by

$$D(\boldsymbol{k}) = \frac{2}{\left(\frac{2\pi}{L}\right)^3}$$

Now let us compute the allowed energies, which we can do by going back to the Schrodinger equation:

$$\begin{aligned} \hat{H}\psi_{\boldsymbol{k}} &= E_{\boldsymbol{k}}\psi_{\boldsymbol{k}} \\ -\frac{\hbar^2}{2m}\nabla^2\psi_{\boldsymbol{k}} &= E_{\boldsymbol{k}}\psi_{\boldsymbol{k}} \\ -\frac{\hbar^2}{2m}\nabla^2\left[Ae^{i\boldsymbol{k}\cdot\boldsymbol{r}}\right] &= E_{\boldsymbol{k}}Ae^{i\boldsymbol{k}\cdot\boldsymbol{r}} \end{aligned}$$

We can then apply the Laplacian to the exponential:

$$\begin{aligned}\frac{\hbar^2 k^2}{2m}\frac{1}{\sqrt{L^3}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} &= E_{\boldsymbol{k}}\frac{1}{\sqrt{L^3}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \\ E_{\boldsymbol{k}} &= \frac{\hbar^2|\boldsymbol{k}|^2}{2m} - \frac{\hbar^2}{2m}\left(\frac{2\pi}{L}\right)^2\left(n_x^2+n_y^2+n_z^2\right)\end{aligned}$$

Up next, we need to find the density of states in energy space:

$$D(E)\,dE = D(E)\Delta E$$

This is the number of allowed states with energy between E and $E+\Delta E$. This includes all allowed k values which fit in this range, for spin. Previously, we did the conversion from $D(\boldsymbol{k})$ to $D(\omega)$, and now we are doing the conversion from $D(\boldsymbol{k})$ to $D(E)$:

$$\begin{aligned}D(E)\Delta(E) &= D(k) \\ &= D(k)4\pi k^2\Delta k \\ &= \frac{2}{\left(\frac{2\pi}{L}\right)^3}4\pi k^2\Delta k\end{aligned}$$

Thus we have that

$$D(E) = 2\left(\frac{L}{2\pi}\right)^3 4\pi k^2\left(\frac{\Delta k}{\Delta E}\right)$$

Now noting that $E = \frac{\hbar^2 k^2}{2m}$, we can compute $\frac{dk}{dE}$:

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

Inserting this, we have that

$$D(E) = \frac{L^3 m}{\pi^2\hbar^3}\sqrt{2mE}$$

Now looking at E versus k, we see that we have a parabolic dispersion. Starting at $T=0$, the electrons fill the lowest energy states. We define the Fermi energy E_F to be the energy of the highest energy electron. The electrons fill up the k space from the origin up to $k=k_F$, the Fermi wavevector. We can compute the number of electrons:

$$\begin{aligned}N_e &= \int_0^{E_f} D(E)\,dE \\ &= \int_0^{E_F}\frac{L^3}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}E^{1/2}\,dE \\ &= \frac{L^3}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}E_F^{3/2}\end{aligned}$$

Rewriting this to solve for the Fermi energy:

$$E_F = \frac{\hbar^2}{2m}\left[3\pi^2\frac{N_3}{L^3}\right]^{2/3}$$

This is sometimes also written in terms of the electron concentation, n_e, where $n_e = \frac{N_e}{L^3}$. We can also write out the Fermi wavevector:

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

We also define the Fermi velocity:

$$v_F = \frac{\hbar k_F}{m}$$

And the Fermi temperature:

$$T_F = \frac{E_F}{k_B}$$

The average number of electrons in a state is known as the occupancy of the state, and is going to be anywhere from 0 to 1. This is governed by the Fermi-Dirac distribution, and depends on the energy of the state, as well as the temperature. Recall that for phonons, the Planck distribution is

$$\langle n_\omega \rangle = \frac{1}{e^{\hbar\omega/k_B T} - 1}$$

For electrons/Fermions, we have the Fermi-Dirac distribution:

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

where we have defined the chemical potential, μ. $f(E)$ tells us the average number of electrons in the state with energy E. At $T = 0$, if $E > \mu$, then this distribution states that the average number of electrons is 0. On the other hand, at $T = 0$, if $E < \mu$, the distribution states that we will have 1 electron on average. If we plot this, this looks like a step function. Interestingly, if we increase the temperature, we "smear" the jump from filled to empty, it is less a step function, and more of a transition. It's also interesting to note that the width of the "smeared" region is $k_B T$.

Now what is the chemical potential? For all intents and purposes, $\mu = E_F$, since for most technologically relevant temperatures, this is true. Suppose we are at $T = 0$, and all states up to μ are filled. If we add another electron, it will require at least energy μ to be added. There will be a charge (increase) in energy in this case:

$$dU = \mu dN$$

The change in energy is the chemical potential times the number of added electrons. From this, we have that

$$\frac{dU}{dN} = \mu$$

Which is what we see in statistical mechanics.

We can also find the number of electrons in another way:

$$N_e = \int_0^\infty D(E)f(E)\,dE$$

At $T = 0$, we can rewrite this:

$$N_e = \int_0^\mu D(E)\,dE$$

since $f(E) = 1$ up until $E = \mu$, and is 0 after $E = \mu$. Doing this integral out, we find that

$$N_e = \frac{L^3}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\frac{2}{3}\mu^{3/2}$$

Which can be rewritten:

$$\mu = \frac{\hbar^2}{2m}\left(3\pi^2 n_e\right)^{2/3}$$

What we mean by the first definition of μ is that, for most metals/materials, $\mu \approx E_F$ up to really high temperatures.

Now let us consider the heat capacity of the free electron gas. We can write out the total energy of the free electron gas:

$$U = \int_0^\infty ED(E)f(E)\,dE$$

We can graphically determine what the change in energy will be:

$$\Delta U \propto D(E_F)(k_BT)^2$$

Taking the derivative to obtain the heat capacity:

$$\begin{aligned} C_V &= 2D(E_F)k_B^2T \\ &= \gamma T \end{aligned}$$

This is the correct dependence on temperature.

5.1 Electrical Conductivity

We have that the current density is related to the electron density n nand the velocity, v:

$$\boldsymbol{J} = ne\boldsymbol{v}$$

We can rewrite this using the definition of momentum:

$$\boldsymbol{J} = -ne\frac{\hbar\boldsymbol{k}}{m}$$

In the presence of an electric field, $\boldsymbol{F} = -e\boldsymbol{E}$, and thus by Newton's Second Law, $\frac{d\boldsymbol{p}}{dt} = \hbar\frac{d\boldsymbol{k}}{dt}$. Thus we have that

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

Integrating this, we have that

$$\boldsymbol{k}(t) = \boldsymbol{k}(0) - \frac{e\boldsymbol{E}t}{\hbar}$$

Now putting this together with $\boldsymbol{J}$:

$$\begin{aligned}\boldsymbol{J} &= -\frac{ne^2\tau}{m}\boldsymbol{E} \\ &= \sigma\boldsymbol{E}\end{aligned}$$

Where $\sigma = \frac{ne^2\tau}{m}$ is the conductivity.

Suppose we look at the Fermi surface, and we apply an electric field. We have a shift in the wavevectpr $\Delta\boldsymbol{k} = -\frac{e\boldsymbol{E}\tau}{\hbar}$. If the electric field is the in the $+x$ direction, then the Fermi surface starts to shift in the $-k_x$ direction. There are also collisions/scattering of electrons, which happen every τ period. This scattering has the effect of putting every electron back to equilibrium, where they originally were. We have two things at work here, the electric field, which tries to shift the Fermi surface to the side, and the collisions, which try to knock the electrons back to their original positions.

But what causes the scattering? The main mechanisms of scattering are phonons and defects. These each have a collision time, τ_L for lattice phonons, and τ_i for defects or impurities. As a result of these 2 differrent collision times, there is a total collision time:

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

We also define the resistivity:

$$\rho = \frac{1}{\sigma}$$

Which can be written as

$$\begin{aligned}\rho &= \frac{m}{ne^2\tau} \\ &= \frac{m}{ne^2}\frac{1}{\tau_L} + \frac{m}{ne^2}\frac{1}{\tau_i} \\ &= \rho_L + \rho_i\end{aligned}$$

This is Mathiessen's Rule, the total resistivity is the sum of resistivity components, which arise from different scattering times.

Note that ρ_L has strong temperature dependence, and ρ_i is temperature independent.

The free electron model of a metal can explain many things, but there are discrepancies with experimental observations. We had to introduce scattering, and the free electron model doesn't explain why some materials are insulators and some are conductors. The model also doesn't explain superconductivity and ferromagnetism. The model also doesn't explain any optical properties, such as why metals are colored the way they are (why is copper red, why is gold gold, why is silver silver?). The problem with the free electron picture is that we largely ignore the presence of positive ions.

6 Band Theory

Now let us look at what happens when electrons interact with the lattice. The key finding is that there will be energy ranges where electrons will not be allowed to exist. This is what we call a band gap. First let us look at a simple picture of the formation of bands and band gaps, as atoms come together to form a solid.

Consider 1 atom, a Li atom. The electronic configuration of this atom is $1\,\mathrm{S}^2\,2\,\mathrm{S}^1$. Suppose we bring N of these atoms together, where N is a very large number. When we bring these together, we form the Lithium metal. What happens is that the energy levels get split into N levels, so that the electrons don't violate the Pauli Exclusion Principle. As the number of electrons approaches a very large number like N, we obtain a continuous spectrum, or band.

This continuous band forms because N is very large, and only half of the band is filled, from the lowest energy part. Because of this, we know it is a metal. We started with discrete quantized energy levels of individual atoms, and in the process of bringing atoms together, each level has now split and broadened to be a band of energy regions where electrons can occupy states. This happens to all of the energy levels, and the regions where there are no bands are known as band gaps. This explanation rises from the quantum mechanical nature of electrons, but this isn't really a solid explanation.

6.1 Second Explanation

The second explanation of the band gap structure is based on Bragg diffraction of electrons, we will apply the same condition as X-ray diffraction to the electrons. The simple qualitative quantum mechanical explanation of band gaps starts with a 1D lattice. This is already a strong departure from the free electron model. The lattice structure produces a periodic variation in the potential energy, as the electrons like to be near the positive ions that are at the lattice points. This potential can be treated as a perturbation of the free electron Hamiltonian:

$$\begin{aligned}\hat{H} &= H_0 + H' \\ &= \frac{p^2}{2m} + U(x)\end{aligned}$$

Recall that in the phonon case, adding the anharmonic perturbation caused scattering of phonons, and something similar will happen here, where we have the scattering of electrons, terms and new states form as a combination of old states. In general in QM, under the influence of a perturbation, the original eigenstates are no longer the new eigenstates. However, in this case, it turns out that everywhere other than the BZB, the eigenstates are about the same, and they only really vary from the original states near the Brillouin Zone Boundary.

If k satisfies the Bragg condition, the new state is different from the old state, otherwise the new state is approximately the old state. Now let us recall what the Bragg condition is when applied to electrons. Previously, the Bragg condition was

$$\boldsymbol{k} - \boldsymbol{k}' = \boldsymbol{G}$$

In the case of X-rays, $\boldsymbol{k}$ was the incoming wavevector, and $\boldsymbol{k}'$ was the outgoing wavevector, but in this case, we have something different. In the case of electrons, $\boldsymbol{k}$ and $\boldsymbol{k}'$ represent the wavevectors of two electrons, and if the condition is met, the energy levels mix to form a new state. Let us choose $\boldsymbol{k}$ and $\boldsymbol{k'}$ that satisfy this condition, $k = \frac{\pi}{a}$ and $k' = -\frac{\pi}{a}$. The mixing produces a linear combination of the two states. Note that the two states are degenerate in energy:

$$E_k = \frac{\hbar^2 k^2}{2m} = E_{k'}$$

From these 2 states, we get 2 new mixed states:

$$\begin{aligned}\psi_+ &= \frac{1}{\sqrt{2}}\left[\psi_{\frac{\pi}{a}} + \psi_{-\frac{\pi}{a}}\right] \\ &= \frac{1}{\sqrt{2L}}[e^{i\frac{\pi}{a}x} + e^{-i\frac{\pi}{a}x}] \\ &= \sqrt{\frac{2}{L}}\cos\left(\frac{\pi}{a}x\right)\end{aligned}$$

$$\begin{aligned}\psi_- &= \frac{1}{\sqrt{2}}\left[\psi_{\frac{\pi}{a}} - \psi_{-\frac{\pi}{a}}\right] \\ &= \frac{1}{\sqrt{2L}}[e^{i\frac{\pi}{a}x} - e^{-i\frac{\pi}{a}x}] \\ &= \sqrt{\frac{2}{L}}i\sin\left(\frac{\pi}{a}x\right)\end{aligned}$$

Note that both of these states have equal amounts of rightward-travelling wave and leftward-travelling waves. Therefore, these wavefunctions are then standing waves. Note that this is just like how the phonons were unable to propagate past the BZB. In the phonon case, at the BZB, adjacent atoms were oscillating against each other. Standing waves happen because waves happen to "fit" the lattice and get stuck, much like resonance. Looking at the standing waves, they look very similar to each other, and if we compare the two against the potential energy that we started with, which had troughs right where the ions were, we see that the cosine solution fits into the lower energy regions of the potential.

The sine solution places the electrons where the repulsive energy is maximized, which is high energy, and is therefore a "bad" solution, unlike the cosine solution. We see that we started with two states degenerate in energy, and due to the perturbation causing mixing, we ended up with two states that are very different in energy. There are no other states which consist of these original ks, which creates a gap between the two energies at $k = \pm\frac{\pi}{a}$ on the energy plot. The gap that is caused by the splitting of the energies of the two states has no allowed states in between.

6.2 Bloch Theorem

Let us assume that we have a 1D system (the 3D case is very similar). Consider the potential energy of an electron in a crystal, $U(x)$. By the definition of crystal structure, $U(x)$ must be periodic with periodicity a (lattice spacing). In other words, the potential is invariant under translation:

$$U(x) = U(x + na)$$

for all integer n.

The Fourier Theorem states that any periodic function can be decomposed into sinusoidal functions with wavevectors that are multiples of $\frac{2\pi}{a}$. In this case, multiples of $\frac{2\pi}{a}$ are just multiples of $\boldsymbol{G}$, the reciprocal vectors, and thus we have that

$$U(x) = \sum_G U_G e^{i\boldsymbol{G}x}$$

Where $\boldsymbol{G} = \frac{2\pi}{a}n$, for $n = 0, \pm 1, \pm 2, \ldots$. Now shifting the potential over by a lattice spacing:

$$\begin{aligned} U(x+a) &= \sum_G U_G e^{iG(x+a)} \\ &= \sum_G U_G e^{iGx} e^{iGa} \end{aligned}$$

Now due to translational invariance, we have that $e^{iGa} = 1$, which checks out (since $Ga = 2\pi n$). Now writing out the Hamiltonian, we have that

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

Which we can insert into the Schrodinger equation:

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

We could guess e^{ikx}, but this is too simple, it is too close to the original solution (solution for the equation without the potential term). We have introduced a perturbation, and as a result, we expect to have some mixing of solutions to happen. As such, a good solution will be some combination of the original solution:

$$\psi(x) = \sum_k c_k e^{ikx}$$

Where $k = \frac{2\pi}{L}n$, for all integer n. All we have to do now is find what c_k is. Inserting our solution into the Schrodinger equation:

$$\sum_k c_k \left[\frac{\hbar^2}{2m}k^2 e^{ikx} + \sum_G U_G e^{iGx} e^{ikx}\right] = E\sum_k c_k e^{ikx}$$

From here, we pick off the terms that have e^{ikx}, and match them. However, we also have an x dependence in the e^{iGx}. Let us look closer at this term:

$$\begin{aligned} &\to \sum_k \sum_G c_k U_G e^{iGx} e^{ikx} \\ &= \sum_G \sum_k c_k U_G e^{i(G+k)x} \end{aligned}$$

Now we note that $k = \frac{2\pi}{L}n$, and $G = \frac{2\pi}{a}m$, where $n, m \in \mathbb{Z}$. The thing we note is that $L = Na$, and thus is very large, which tells us that there are many more ks than there are Gs, and in fact, the Gs are actually a subset of the values of k (which we intuitively already knew, the Gs are special cases of the wavevectors). In fact, $G = \frac{2\pi}{a}m = \frac{2\pi Nm}{Na} = \frac{2\pi}{L}(Nm) = k$, since N is just some integer. Now using this as a substitution:

$$\begin{aligned} k' &= k + G \\ &= \frac{2\pi}{L}n + \frac{2\pi}{a}m \\ &= \frac{2\pi}{L}n + \frac{2\pi}{L}(Nm) \\ &= \frac{2\pi}{L}(n + Nm) \\ &= \frac{2\pi}{L}n' \end{aligned}$$

Where we have defined $n' = n + Nm$. thus we have that

$$\begin{aligned} &\to \sum_G \sum_k c_k U_G e^{i(G+k)x} \\ &= \sum_G \sum_{k'} c_{k'-G} U_G e^{ik'x} \end{aligned}$$

The next move is a little bit of a magic trick, we can actually get rid of the prime notations, since the sum over k and a sum over k' will account for the same things. Thus, we have that

$$\begin{aligned} &= \sum_G \sum_k c_{k-G} U_G e^{ikx} \\ &= \sum_k \sum_G c_{k-G} U_G e^{ikx} \end{aligned}$$

Now we're home free, we can look at the Schrodinger equation and sort things out term by term:

$$\sum_k c_k \left[\frac{\hbar^2}{2m}k^2 e^{ikx}\right] + \sum_k e^{ikx}\left[\sum_G c_{k-G} U_G\right] = E\sum_k c_k e^{ikx}$$

Now we have a summation over k with an e^{ikx} term for each of the terms in our equation. Equating them term by term, for the same e^{ikx}, we can drop the sums over k, as well as the e^{ikx}:

$$c_k \frac{\hbar^2}{2m}k^2 + \sum_G c_{k-G} U_G = E c_k$$

Which holds true for each k. This is known as the central equation. Note that this tells us how the mixing occurs, the state represented by k will be mixed with the state given by $k - G$. In other words, if we had two states that are mixing, k_1 and k_2, then the difference must be G:

$$k_1 - k_2 = G$$

Which is just the Bragg Law.

Now we can state the Bloch Theorem:

Theorem 6.1. *Bloch Theorem. Given a periodic potential* $U(x) = \sum_G U_G e^{iGx}$*, with* $G = \frac{2\pi}{a}n$*, the wavefunction which solves Schrodinger's equation is of the form*

$$\psi_k = \sum_G c_{k-G} e^{i(k-G)x}$$

Which is a sum of waves with wavevectors differing by G*s.*

Let's now apply the central equation in some specific cases.

For the free electron, $U = 0$, and thus $U_G = 0$ for all G. The central equation then states that

$$c_k \frac{\hbar^2 k^2}{2m} = Ec_k$$
$$E = \frac{\hbar^2 k^2}{2m}$$

Which is what we expect.

In the case of a periodic potential of the form $U(x) = -U_0 \cos\left(\frac{2\pi}{a}x\right)$, we find that $u_G = -\frac{U_0}{2}$ for $G = \pm\frac{2\pi}{a}$, and 0 for all other Gs. Now looking at the boundaries for $k = \pm\frac{\pi}{a}$, we have that

$$c_k \frac{\hbar^2}{2m} k^2 + c_{k-G_1} U_{G_1} + c_{k-G_2} U_{G_2} = Ec_k$$

Where $G_1 = \frac{2\pi}{a}$ and $G_2 = -\frac{2\pi}{a}$. This will be two equations, one for $k = \frac{\pi}{a}$, and another for $-\frac{\pi}{a}$:

$$c_{\frac{\pi}{a}} \frac{\hbar^2}{2m}\left(\frac{\pi}{a}\right)^2 + c_{-\frac{\pi}{a}} U_{\frac{2\pi}{a}} + c_{\frac{3\pi}{a}} U_{-\frac{2\pi}{a}} = Ec_{\frac{\pi}{a}}$$
$$c_{-\frac{\pi}{a}} \frac{\hbar^2}{2m}\left(-\frac{\pi}{a}\right)^2 + c_{\frac{3\pi}{a}} U_{\frac{2\pi}{a}} + c_{\frac{\pi}{a}} U_{-\frac{2\pi}{a}} = Ec_{-\frac{\pi}{a}}$$

We see that we now need to compute $c_{\frac{3\pi}{a}}$ and $c_{-\frac{3\pi}{a}}$, but it turns out that, due to the energy difference between these states, we can actually get rid of those mixings.

In the extended zone scheme, we have that the band gaps separate bands, and each Brillouin Zone contains one band. Each state is uniquely defined by k. In the reduced zone scheme, which is the most common, we "fold" in the bands outside of the first Brillouin zone, to make a more compact representation. Each state is now defined by k and the band number. This is convenient because all bands are now in the first Brillouin Zone.

Now lets think about the 3 dimensional case. For the free electron case, the energy is given by

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

Where $k^2 = k_x^2 + k_y^2 + k_z^2$. An equation for an equi-energy contour line (surface of constant energy) will be spherical. This means that in k space, all of the states on the same curve are degenerate in energy. Each curve is a sphere with radius $k = \sqrt{\frac{2mE}{\hbar^2}}$.

Now let us include interaction between electrons and the lattice. Previously (in the 1D case), we saw that the depth of the potential is the size of the band gap. As a result, the energies of the electrons got modified, for k near BZBs. We saw that in the 1D case, when we had a gap, near the gap, to the left, the energy got pushed down, and to the right of the boundary, the energy got pushed up. In other words, for $k < k_{\text{BZB}}$, $E(k)$ is reduced, and for $k > k_{\text{BZB}}$, $E(k)$ is increased, where we are discussing values of k that are close to k_{BZB}.

We can think of the BZB in 2D, which is in the form of a square, and at the boundaries of the square, the states have decreased energies, and directly outside the boundary of the square, the states have their energies increased.

The states that are near the center of the zone are not affected, so we still have spherical/circular contours. However, the equi-energy surfaces that are closer to the boundaries have segments that are closer to the boundary, and segments that are further. The segments that are further (the ones that are near the corners) will remain unaffected, while those that are closest to the boundary (sides), will get closer to the boundary (compensating for the decrease in energy). Essentially, that part of the curve has to "go further" in order to maintain the energy of the curve. Note that there are now discontinuities in some of these curves, since there existed shells that would intersect the BZB. This is the emergence of the band gaps.

In the free electron case, the location of the energy contours was given by $k = \sqrt{\frac{2mE}{\hbar^2}}$, but now the distance is different for different directions, we no longer have spherical symmetry.

Let us look at how electrons fill up states. Consider a 3D crystal with Np primitive cells. Suppose there are m atoms per primitive cell, and there are z_v valence electrons, the electrons that are sittings in the outermost shell, and the only electrons that can get stripped, and participate in conduction. Based on this setup, there are $z_v mNp$ valence electrons. Let us now consider the number of states.

In 1D, the acceptable values for k are between $-\frac{\pi}{a}$ and $\frac{\pi}{a}$ for the first zone, and recall that the ks that satisfy the periodic boundary conditions are $\frac{2\pi}{L}n$, where $n \in \mathbb{Z}$. We have that $\frac{2\pi}{L}n = \frac{\pi}{a}$, but $L = Na$, so

$$\frac{2\pi}{L}n = \frac{n\pi}{L}$$

Thus for one side, we can go from $n = 0$ to $n = \frac{N}{2}$. However, this is one side, so we take into account symmetry and double it, and then we have to take into account spin, so we double it again. Each BZB therefore gives us $2N$ states. Generalizing to 3D crystals, there are $2Np$ states in each Brillouin zone. Note that the number of states is not affected by z_v or m, it is solely reliant on Np.

Thus we have that the number of states is proportional to just the volume of the Brillouin zone, since the density of states is constant. Recall that each Brillouin zone has the same volume, so thus they have the same number of states.

6.3 Metals, Insulators, and Semiconductors

Now we can finally look at the question of why some objects are metals and why some are insulators, as well as why some are semiconductors. At $T = 0$, all electron states up to $E = E_F$ are filled. The order of filling of states is determined by the order of the energy of the states.

- A material is a metal if it has a partially filled band
- A material is an insulator if it has no partially filled bands

In insulators, all bands are either completely filled or completely empty. Metals are materials with "exposed" Fermi surface, the Fermi surface exists inside a band, so that electrons can pick up "small" energies and see action (participate in electrical current). Semiconductors are insulators where the band gap is small, so if we have enough energy, we can jump to the next band.

We can think of the bands as tanks of liquid, if a tank is full or empty, we can't do anything to it, but if its partially full, we see behavior at the boundary of the exposed liquid. When there is a partially filled band, electrons at the Fermi surface can easily pick up small energy and participate in transport/current.

How do we know if a material will be a metal or an insulator?

Each band has $2Np$ states. A material has $z_v mNp$ valence electrons. From this, we see that if $z_v m$ is odd, we have a metal, whereas if $z_v m$ is even, we don't necessarily have an insulator, because energy levels can fill up out of order, so a new band can start to be filled before the previous one is filled completely.

For example, Si is basically an insulator, but Mg, Ca, and Ba are metallic. If we look at Na, which has electronic structure $1\,\mathrm{S}^2\,2\,\mathrm{S}^2\,2\,\mathrm{P}^6\,3\,\mathrm{S}^1$. We see that we have $z_v = 1$. For a monatomic crystal, we have that $z_v m = 1$, there are Np electrons and $2Np$ states. Here we have only enough electrons to fill half of a band, so this is a metal. For Be, it turns out that states in the second Brillouin Zone have lower energy than some states in the first Brillouin zone, so those fill up first, and so we have partially filled bands, meaning that we have a metal. Thus the lesson is that only detailed investigation of the Fermi surface can tell us whether a divalent material is an insulator or not. There are different methods for mapping out the Fermi surface, such as angle-resolved photo emission.

We often take the free electron energy and write it in terms of the mass:

$$m = \frac{\hbar^2}{\frac{\partial^2 E}{\partial k^2}}$$

And in the non-free electron case, where we add in interactions, we no longer have the same energy, but we can still consider the effect mass, m^*:

$$m^* = \frac{\hbar^2}{\frac{\partial^2 E}{\partial k^2}}$$

Note that this second derivative, $\frac{\partial^2 E}{\partial k^2}$ is the curvature of $E(k)$, and thus when the equi-energy curves are very curved, the electrons have more interaction, and when they are less curved, the electrons have lighter interaction. Also note that we can have negative effective mass, $m^* < 0$, and these particles move in opposite directions as normal electrons, these are holes. In fact, this is consistent with the idea of electrons "jumping" to the next band, because this is equivalent to having a negative curvature.

6.4 Silicon Doping

Let us now look at Si, the most important electronic material. Silicon has $1\,\mathrm{S}^2\,2\,\mathrm{S}^2\,2\,\mathrm{P}^2\,3\,\mathrm{S}^2\,3\,\mathrm{P}^2$, where the last two hybridize to form $4\,\mathrm{SP}^3$ states, like Carbon. Silicon likes to form 4 bonds (according to the octet rule), which are equivalent to tetrahedral coordinates and bonding, forming a diamond structure, which is like FCC with four extra atoms on the inside. Si has a valence of 4+, so lets do doping to make it conducting. There is donor doping (giving an electron to it), so we use

As, which has a valence of 5+, it has 5 electrons. that can be stripped. We substitute the central Si for As, which has an extra “leg”, which can be broken easily. The extra electron can be stripped with 49 meV, so the addition of the Arsenic has created a new state in the band gap, if we add 49 meV, we can jump to the next band.

There is also acceptor doping with B, which has valence 3, one less than Si. In this case, when we substitute it in, it is missing a leg, and with a little bit of energy, we can move an electron from elsewhere to fill that leg, or rather we say that the hole moves from the B to elsewhere. Instead of a new state being created at the bottom of the band above, we create a state above the lower band.