Lecture 14

Thursday, February 23, 2012

Nearly Free Electron Model

Let us look back at the sub-matrix of the Hamiltonian we diagonalized at the BZ boundary: with the two states as our basis, $|1\rangle = \left|k = -\frac{\pi}{a}\right\rangle$ and $|2\rangle = \left|k = \frac{\pi}{a}\right\rangle$

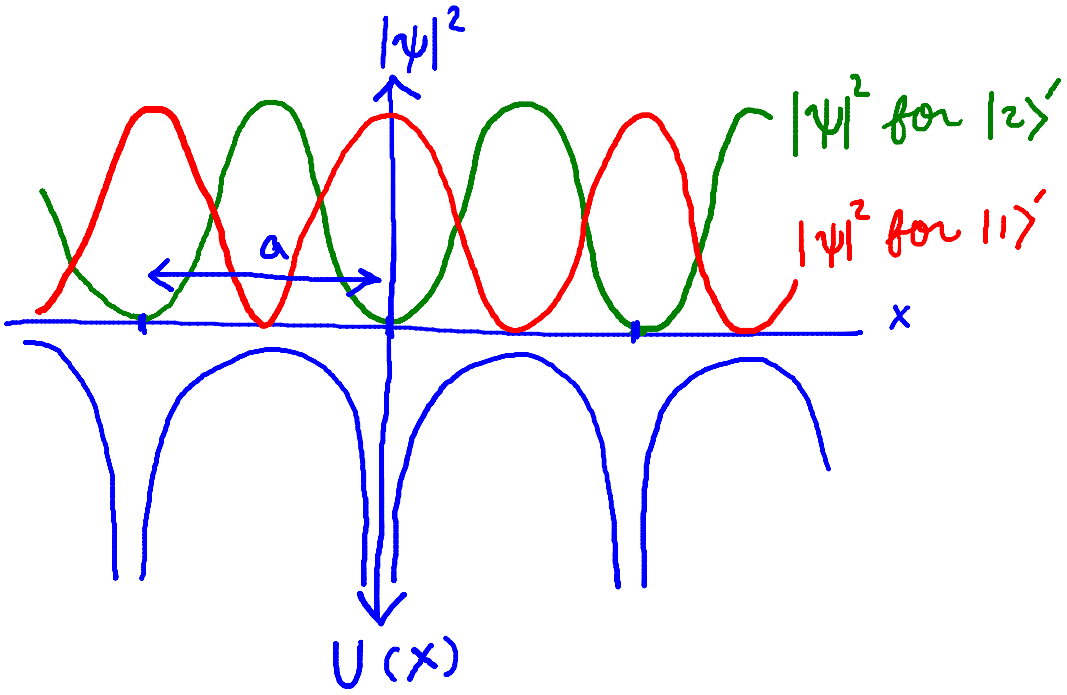
$$h = \begin{pmatrix} E^{(0)} + U_0 & U^*_{\frac{2\pi}{a}} \\ U_{\frac{2\pi}{a}} & E^{(0)} + U_0 \end{pmatrix} \begin{pmatrix} 0 & U^* \\ U & 0 \end{pmatrix}$$

Note that $U_0 = \frac{1}{V_c}\int dx\, U_b(x)$, $U_{\frac{2\pi}{a}} = \frac{1}{V_c}\int dx\, U_b(x) e^{-\frac{i2\pi r}{a}}$, and $E^{(0)} \equiv \lambda_{-\frac{\pi}{a}} = \lambda_{\frac{\pi}{a}} = \frac{\hbar^2\pi^2}{2ma^2}$. If $U_b(x)$ has the even-parity (likely), then $U_{2\pi/a}$ is a real number. Let us assume this. Moreover, let us assume that the potential is attractive. This would be the case for simple metals. Then, $U_0 < 0$ and $U_{2\pi/a} < 0$.

The eigenvalues of h are $E^{(0)} + U_0 \pm |U_{2\pi/a}|$. How do eigenstates look? Keeping in mind that $U_{2\pi/a} < 0$, we get

$|1\rangle' = \frac{1}{\sqrt{2}}\,(|1\rangle + |2\rangle)$ for $\epsilon_{k=-\pi/a,1} = E^{(0)} + U_0 - |U_{2\pi/a}|$
$|2\rangle' = \frac{1}{\sqrt{2}}\,(|1\rangle - |2\rangle)$ for $\epsilon_{k=-\pi/a,2} = E^{(0)} + U_0 + |U_{\frac{2\pi}{a}}|$

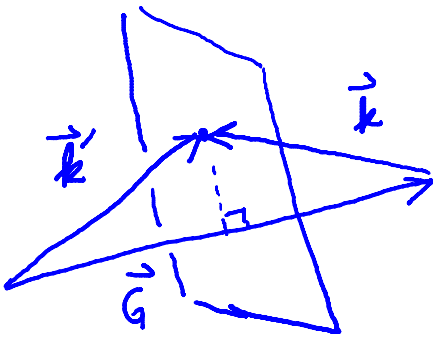
Since $|1\rangle \Rightarrow \frac{1}{\sqrt{V}} e^{-\frac{i\pi}{a}x}$ ($\Rightarrow$ here means "is represented by the wave function") and $|2\rangle \Rightarrow \frac{1}{\sqrt{V}} e^{+\frac{i\pi}{a}x}$, we have $|1\rangle' \Rightarrow \frac{1}{\sqrt{V}}\left(\sqrt{2}\right)\cos\left(\frac{\pi x}{a}\right)$ and $|2\rangle' \Rightarrow \frac{1}{\sqrt{V}}\left(-\sqrt{2}i\right)\sin\left(\frac{\pi x}{a}\right)$.



The above diagram, depicted for a single atom basis, with a positive ion per basis, illustrates the physics at the BZ boundary. For this attractive potential, the lower energy state corresponds to the standing wave maximized at the ion cores, while the higher energy state corresponds to the standing wave maximized between the two adjacent ions. Should the potential be repulsive, then the reverse situation will result.

Why should the standing wave form? This is obvious from the above solution (Bragg diffraction mixing two waves on an equal footing), but it can be viewed in a slightly different way. Recall in the 1D phonon problem that the group velocity is zero since the dispersion curve becomes flat at the BZ boundary. Indeed, the same reason applies in this case as well (see the last diagram -- red curves -- in the previous lecture note).

In higher dimensions, similar things happen. Whenever the Bragg *diffraction* condition is satisfied, i.e. $\vec{k}' - \vec{k} = \vec{G}$ *and* $|\vec{k}'| = |\vec{k}|$, the non-degenerate perturbation theory is valid, and barring accidental cases when $U_{\vec{G}} = 0$ (for a symmetry reason, e.g.) there will be a "level repulsion" in energy between the two levels, one pushed down by $-|U_{\vec{G}}|$ and the other up by $|U_{\vec{G}}|$. Notice that the Bragg diffraction condition is satisfied at (the 1st, 2nd, 3rd,…) BZ boundaries (this is why the BZ boundaries are important), so this level repulsion occurs throughout a surface (3D crystals) or throughout a line (2D crystals).



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

$|\vec{k}'| = |\vec{k}|$

$\vec{k}, \vec{k}'$ on perp. bisector plane of $\vec{G}$! (BZ boundary !)

In higher dimensions, a partially standing wave forms in the sense that the group velocity perpendicular to the BZ boundary is zero. In general, the group velocity parallel to the BZ boundary is non-zero. That is $\vec{v}_{\vec{k}} \cdot \vec{G} = 0$ (cf. Homework). This means that a constant energy surface will cut the BZ boundary at the right angle.

Tight Binding (TB) Method

This is synonymous to the LCAO (linear combination of atomic orbitals) method in the context of the band theory. In Quantum Chemistry, this method is used extensively. Also, in Condensed Matter Physics, one often finds that sophisticated and time-consuming band calculation is re-interpreted with a "tight binding fit" to it. The reason is that by doing so, one has a manageable band theory on top of which to add terms that lead to other effects (magnetism, superconductivity, etc.) that are generally outside band theory, thereby defining a fairly complex but hopefully manageable Hamiltonian to solve exotic real world problems.

In any case, the tight binding method consists of building an eigenstate locally, within a given basis, and then let those eigenstates at neighbors interact with each other, producing a full band structure.

We consider the simplest case only. A 1D crystal with one atom per cell, and one atomic 1s orbital per cell. This would be a model for a 1D solid hydrogen.

The Bloch state can be written as (form 4)

$$|\psi_k\rangle = \sum_s e^{iksa}|s\rangle$$

Here, $|s\rangle$ is the Dirac notation for the 1s orbital at the s-th site. The Hamiltonian is $H = T + U = T + U_s + U_{s,rest}$. Here, U_s is the potential due to the ion at the s-th site, while $U_{s,rest}$ is the potential due to the rest of sites.

Note that here we are not using any "band index", n, since by choosing only one atomic orbital, 1s, per basis, there is only one band, by definition. [In a more realistic calculation, there will be quite a few atomic orbitals per basis.]

Thus, by definition, $(T + U_s)|s\rangle = \varepsilon_0|s\rangle$, where $\varepsilon_0 = -13.6$ eV.

The eigenvalue equation is $H|\psi_k\rangle = \epsilon_k|\psi_k\rangle$.

$H|\psi_k\rangle = \sum_s e^{iksa}\left(\varepsilon_0 + U_{s,rest}\right)|s\rangle$. So, the eigenvalue equation can be written as

$$H|\psi_k\rangle = \sum_s e^{iksa}\left(\varepsilon_0|s\rangle + U_{s,rest}|s\rangle\right) = \sum_s e^{iksa}\epsilon_k|s\rangle$$

Now, we will multiply this equation by $\langle s'|$ from the left. Recall from QM that $|s\rangle$ is a exponentially decaying function. So, we expect $\langle s'|s\rangle \to 0$ and $\langle s'\left|U_{s,rest}\right|s\rangle \to 0$ when $|s' - s| \gg 1$. Here we take the minimal model, where we assume that these integrals are non-zero only when $|s' - s| \leq 1$. Of course, $\langle s|s\rangle = 1$. The rest of the parameters are defined as:

$$\langle s \pm 1|s\rangle = \langle s|s \pm 1\rangle = \beta$$
$$\langle s\left|U_{s,rest}\right|s\rangle = -\alpha$$
$$\langle s \pm 1\left|U_{s,rest}\right|s\rangle = \langle s\left|U_{s,rest}\right|s \pm 1\rangle = -t$$

Here, the fact that the matrix element remains the same when transposed comes from the fact that the [conventional] 1s orbital is a real function.

Then, $\langle s'|\sum_s e^{iksa}\left(\varepsilon_0|s\rangle + U_{s,rest}|s\rangle\right) = e^{iks'a}\varepsilon_0(1 + 2\beta\cos ka) + e^{iks'a}(-\alpha - 2t\cos ka)$. And $\langle s'|\sum_s e^{iksa}\epsilon_k|s\rangle = e^{iks'a}\epsilon_k(1 + 2\beta\cos ka)$.

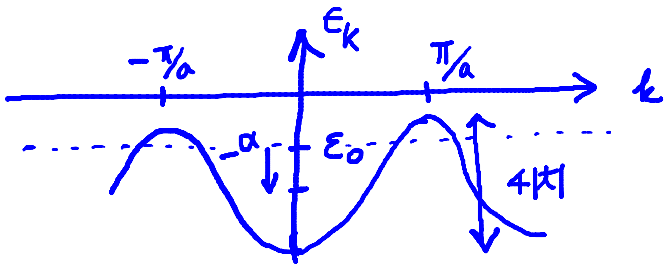
Equating these two terms, we get

$$\epsilon_k = \varepsilon_0 - \frac{\alpha + 2t\cos ka}{1 + 2\beta\cos ka}$$

It is customary to ignore the β term in condensed matter physics. This term tends to be small. Also, note that in writing down form 4 of Bloch's theorem, the wave function $\phi_n(\vec{r}-\vec{R})$ was dubbed the "Wannier orbital." Well, it can be proven rigorously that ϕ_n's at different sites are orthogonal! So, if we are making a tight binding fit of the actual band calculation, then we are implicitly working with the mutually orthogonal Wannier orbital basis and so the β term is strictly zero. [In contrast, in quantum chemistry, it is important to keep all β terms.]

So, with this approximation, we can write

$$\epsilon_k = \varepsilon_0 - \alpha - 2t\cos ka$$



Here the parameter t is particularly important, and it is called a "**hopping amplitude**", a "**hopping integral**" or a "**hopping parameter**," since it is what makes it possible for the electron to hop to the next site and on and on. The band width is given by $|4t|$.

t is a sensitive function of the lattice constant. So, if one imagines starting from the limit of $a \to \infty$, then one starts from the atomic energy level which is absolutely independent of k (dotted line above). As a decreases this energy level is shifted and, more importantly, is broadened into an energy band.

As mentioned above, this model presented here is the simplest tight binding model possible. A more complicated model will include more atomic orbitals, and the number of atomic orbitals that one considers per basis will determine the number of bands. Generally, the multiple bands generated may have overlaps if the hoping integral is large. If the hopping integral is small, then bands will not overlap. When the chemical potential (Fermi energy at zero temperature) happens to lie between non-overlapping bands, then we have, remarkably, insulators and semiconductors!

Metals, Insulators, Semiconductors, Semi-metals

Metals are defined as those materials for which at zero temperature there is at

least one band which is partially filled. Na, Cu, Au, Ag, Al, K, ...

Insulators and semiconductors are those materials for which at zero temperature all bands are full, and the next energy band lies at a finite energy above them. This finite energy is called the energy gap. NaCl, LiF are some of the large band gap insulators. Si, GaAs are some of the small gap insulators, or semi-conductors.

Semi-metals are those materials for which at zero temperature all bands are full, but the next energy band lies without any energy gap. By extension, even if all bands are only nearly full, we call them semi-metals, when the finite temperature behavior is not that different. Bi and C (graphite, graphene) well known semi-metals.

Zone schemes

Different zone schemes can be understood in terms of the plane wave basis. The following statements are easy to understand from the nearly free electron model, but are valid in general, even when the potential energy is large.

Start from the free electron dispersion. Consider the momentum-energy space. Each point $\left(\vec{k}, \lambda_{\vec{k}}\right)$ ($\lambda_{\vec{k}}$ is the free electron energy as before) belongs in a certain BZ. As the potential energy modifies the energy dispersion, it will modify the dispersion relation. However, as they are modified, each BZ still has only one band. This view is called the **extended zone scheme**.

Now, one can bring all dispersions into one BZ, through translation by reciprocal lattice vectors. This view is called the **reduced zone scheme**.

Next, one can repeat the bands of the reduced zone scheme to all other BZs. This scheme is called the **periodic zone scheme**.

Figure 4 of chapter 9 in Kittel nicely summarizes these schemes.

Note that there is no difference in physics between different schemes, but the periodic zone scheme makes it the easiest to describe motions of charge carriers in transport phenomena, as we shall see later.

Fermi surface -- a 2D square crystal example

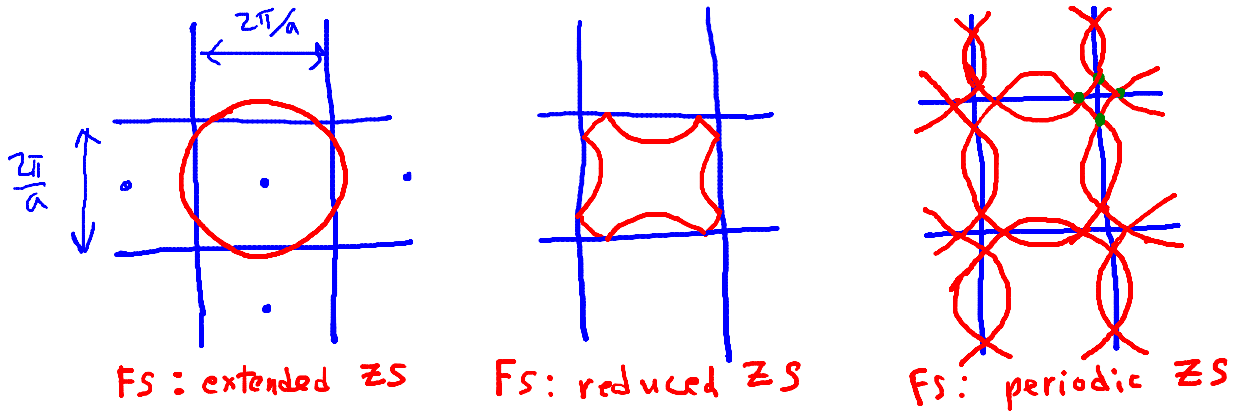
As defined above, metals have at least one un-filled band, and this means that there is a sharp boundary between the occupied states and the unoccupied states. This sharp boundary in the $\vec{k}$ space is the **Fermi surface**. Let us investigate what happens if we have a certain number of electrons per cell in a 2D square crystal.

One thing to note is the following important formula.

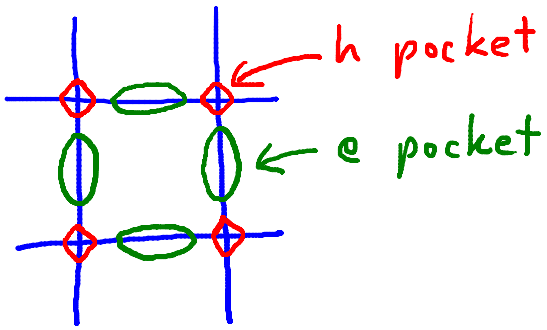
The volume/area/length enclosed by the Fermi surface in the free electron theory = (The volume/area/length of the cell in the $\vec{k}$ space = V_c^*) x (The number of electrons per cell in real space) / 2.

Why is this? (1) Because of the spin degeneracy, each $\vec{k}$ point can accommodate two electrons. (2) There are as many $\vec{k}$ points per cell in the $\vec{k}$ space as there are cells/bases/lattice-points in the $\vec{r}$ space.

With this in mind, the following topologically correct Fermi surface diagram for a 2D square crystal with two electrons per cell can be obtained.



There is more. If one now introduces the crystal potential, then the band degeneracy at the BZ boundaries get lifted. The result, in terms of the FS (Fermi surface) shape, is that the crossing points at which circles intersect each other in the periodic zone scheme (ZS) [four green points in the above figure, and their equivalent points throughout the reciprocal space] disappear (let us call this "avoided crossing mechanism") and **separate Fermi surface pieces** develop (cf. Homework). [This is how Nature seems to avoid the degeneracy.]



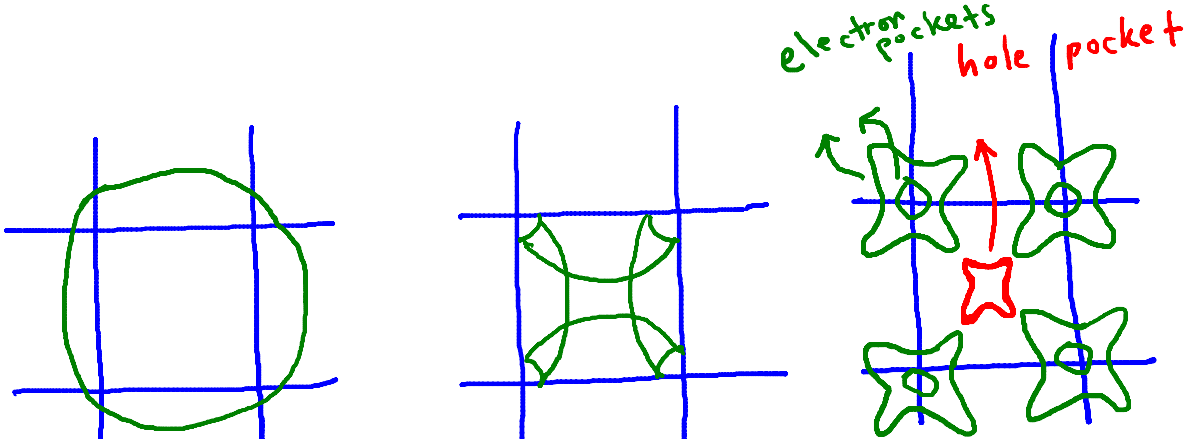
Now we have two distinct FS objects. The reason why the green cigar shaped object is called an **electron pocket** is because if one follows the band dispersion that makes this Fermi surface, that dispersion is occupied within the cigar, but unoccupied outside it. The reason why the small "diamond" piece around the corner point is called a **hole pocket** is the opposite. The dispersion that gives rise to this Fermi surface is occupied outside this diamond, and unoccupied inside it.

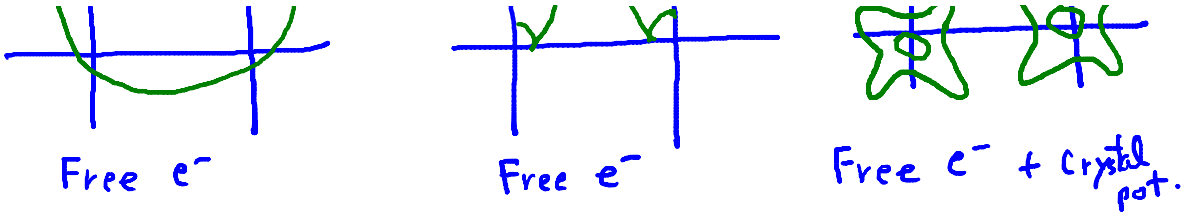
This is the general definition. If a Fermi surface piece forms a closed geometry (closed curve, closed surface), then it is called an electron pocket if its interior is occupied for the band that generates that Fermi surface piece. Otherwise, it is called a hole pocket. This distinction has crucial physical consequences, which we will discuss later.

In the above, the electron pocket or the hole pocket nature can be figured out easily by considering the free electron dispersion for the part of the FS object that is not greatly perturbed by the crystal potential, and the argument of continuity.

The Fermi surface objects above cut the BZ boundary at the right angle, as discussed above. The "avoided crossing mechanism" invoked above can be viewed as a way to satisfy this constraint.

Here is what happens, if there are four electrons per cell in a 2D square crystal.





In homework 7, you will figure out that for the divalent case above, the total area of the hole pocket and the total area of the electron pocket (per cell in the $\vec{k}$ space, of course) are equal to each other. This is the consequence of the particle number conservation. If the crystal potential is large enough, then these Fermi surfaces can completely disappear, giving rise to a situation for an insulator/semi-conductor or a semi-metal. This is true for any case when an even number of electrons are contributed per primitive basis. Note that such large potential cases may be more naturally considered in terms of the tight binding model. On general grounds, both the plane wave basis set and the tight binding basis set are valid basis sets if the set is taken large enough for any given problem. It is just that when the crystal potential is small, the plane wave basis is a better zero-th order approximation, while the tight binding basis is a better zero-th order approximation when the crystal potential is large.

In any case, the discussion of the Fermi surface here leads naturally to the important Wilson's rule.

Wilson's Rule

If there are odd number of electrons per basis, then that crystal must be a metal. This is Wilson's rule. For instance, this rule explains why Na, K, and Al are metals. It also sheds light on why C, Si, Ge are insulators (or semi-metals) without necessarily contradicting the fact that Pb or Te is a metal.

Wilson's rule is a simple consequence of counting, and thus is a quite general rule. Let's see. Each electron energy band has $2N_l$ states available, where N_l is the number of cells/bases/lattice-points in the real space. The factor of 2 comes from the spin degeneracy. So, if there are an even number of electrons per basis, then it is possible to have all bands filled, although this is not necessarily true. On the other hand, if there are an odd number of electrons per basis, then it is necessarily true that there will be at least one band that is left partially filled. Thus, a metal!

Wilson's rule does a good job in most cases, but breaks down completely in many transition metal compounds and rare earth compounds. It also breaks down for solid hydrogen. The reason for this is the neglect of electron-electron correlation in the Bloch-Wilson theory of electron bands, which we have been covering. For this reason, we often refer to the standard band theory as the "one electron" band theory, emphasizing that electron-electron correlations are ignored. On one hand, the one electron band theory has a tremendous success for explaining Si, Ge, GaAs, graphene, diamond, etc., and so its value cannot be under-appreciated. On the other hand, it is very inadequate in describing modern magnetic materials, notably transition metal oxides and rare earth compounds that make your hard drives and your headsets work, as well as most materials that we classify as high temperature superconductors.