Notes for Lecture 6

Electrons, Holes

In the previous lecture, we discussed the valence band and the conduction band in a semiconductor. We also discussed how at $T = 0$ the valence band is full while the conduction band is empty. Here, we begin to consider excited states. In an excited state, "electrons" and "holes" come into play.

6.1 Crystal Momentum and Angular Frequency

As explained in the previous LN, $\hbar\vec{k}$ inside a crystal is called the "crystal momentum." In the 1D crystal example, this means that k and $k+2\pi n/a$ are completely equivalent. The physical origin of this equivalence is the existence of all Bragg diffracted waves in an eigenstate for an electron in a crystal. We have[1]

$$\vec{p} = \hbar\vec{k} \qquad \textbf{Crystal Momentum} \tag{6.1}$$

On the energy side,

$$\varepsilon(\vec{k}) = \hbar\omega(\vec{k}) \qquad \textbf{Energy Quantum} \tag{6.2}$$

[1]How about the case of an electron in free space? That can be formally considered as the limit of lattice constants $\to 0$. In such a limit, the crystal momentum becomes the momentum.

These two important relations are almost[2] nothing else than De Broglie relation and Planck quantum, but now they are used for electrons in a crystal. They are indeed applicable to any quantum particle in a crystal.

As noted above, the energy ε or the frequency ω is a function of wave vector $\vec{k}$. This function is often referred to as the **dispersion relation**.

6.2 Wave Packet and Group Velocity

In textbooks, we often deal with a perfectly monochromatic wave of the standard form

$$A \ \exp\left[i(\vec{k}\cdot\vec{r}-\omega t)\right]$$

Here, A is the amplitude, $\vec{k}$ is the wave vector, and ω is the angular frequency ($2\pi\nu = 2\pi/T$, where T is the period).

This type of ideal wave is easy to deal with in theory, and so a lot of textbook discussions depend on it. In reality, it does not exist, and it should be taken with a grain of salt. It is a useful idealization. In particular, one should keep in mind that, between the phase velocity and the group velocity (see definitions below), the latter is the more physically meaningful one, as far as the "movement" or "transport" of any information by the wave.

You can skip reading the next two paragraphs – they are physically rich, but are not really essential for following my notes – if you simply realize/accept that a wave packet is a real thing, but a monochromatic wave is not.

Why does such an ideal wave not exist? First, one can ask where does any wave come from? Take the wave of photons – i.e. the wave of light. Where do they come from? One might say that they existed from time zero. It probably has. More significantly for our thinking, though, one can note that when quantum transitions occur in atoms/molecules/solids/nuclei/all-kinds-of-stuff, we expect photons to be a common by-product. Therefore, *photons and other matters must be thought of as in a mutual equilibrium state, through exchanges of energy.* Now, take the process of two atomic levels involved in emitting a photon. An electron is in the high energy level, and it makes a transition to the low energy level, emitting a photon of energy, which is precisely the difference of the two energy levels. Sounds like we can get a perfectly

[2] "Almost" because we have now generalized the momentum to the crystal momentum.

monochromatic wave! This is false, of course, because of the Heisenberg uncertainty principle, $\Delta E \Delta t \gtrsim \hbar$. How so? Recall that this Heisenberg uncertainty principle for energy and time means the following. If a quantum system has a finite width in energy, then its state is not constant[3] – it will either decay or oscillate at a time scale given by Δt. And the reverse is also true. If a system is observed to decay or oscillate with a certain time scale, then its energy must be uncertain by $\Delta E \gtrsim \hbar/\Delta t$. This uncertainty in energy means, for a wave, an uncertainty in wave vector, due to the dispersion relation: $\varepsilon(\vec{k})$.

Second, note that all sciences are based on experiments. Whatever we write down must be falsifiable by experiments before it can be accepted to, or rejected from, the body of scientific knowledge. This poses a practical problem. For experiments on a solid, we invariably deal with finite size samples/devices. Say the typical dimension of the sample is L (length). Then, it follows that $\Delta k \gtrsim 1/L$. Which means a wave packet! In practice, when we consider conduction of current, heat etc., from one side of the sample to another side of the sample, we are preparing wave packets whose size L_2 is actually smaller than L. Thus, Δk for our wave packet will be even greater $\sim 1/L_2 > 1/L$. As another example, consider a wave of photons. One might think that if we have a really precise instrument, then we can monochromatize the light at a single frequency. However, this runs into problems – a "really precise instrument" would be an instrument that is characterized by an infinite size grating, an infinite size crystal, or a zero-size aperture. These are impossible to make (due to practical reasons or fundamental quantum mechanics).

So, either fundamentally or empirically, we are bound to accept that any real wave must be a wave packet.

For a wave packet, the following is well-known[4].

> Consider a wave packet consisting of, i.e. which is a superposition of, sinusoidal waves ($e^{i(kx-\omega t)}$) with a distribution of wave vector values. Let us assume that the distribution is centered at k and has a small width, Δk. Then, such a wave packet has a size of $\Delta x \gtrsim 1/\Delta k$ in the real space, and propagates at the **group velocity**
>
> $$v_g = \frac{d\omega}{dk}$$

For notational convenience, the above statement is made for 1D. The generaliza-

[3] It is not in a "stationary state" in the QM language. In QM, a stationary state is defined as an eigenstate of H, the Hamiltonian operator.

[4] The proof for the group velocity part is left for a homework problem.

tion to higher dimensions is straightforward. In general, the group velocity is given by $\vec{v}_g = \nabla_{\vec{k}}\,\omega(\vec{k})$. Here, $\nabla_{\vec{k}}$ means the gradient with respect to the wave vector.

In the above, $\Delta x \gtrsim 1/\Delta k$ is the Heisenberg uncertainty principle.

In any case, the group velocity can be quite different from the phase velocity, ω/k. Even in sign! The group velocity is the physical velocity at which the signal/information (whether it is current, heat, particle number, etc.) propagates.

6.3 Semi-Classical EOM

While Newton's laws are invalid in atomic scales, *some* behaviors of atomic systems *can* be conveniently described with an equation that "looks like" Newton's laws. When such a handy equation is discovered, we call it a "semi-classical equation of motion." However, an extreme care must be practiced not to misinterpret the result. "Semi-classical" means that it is neither classical nor quantum. It is something in between: it has to be kept in mind that both quantum physics and classical physics are mixed in such an equation.

The thing about these "semi-classical" equations/rules (recall the Bohr-Sommerfeld semi-classical quantization rule from LN 2) is that they are extremely attractive ... to us, big creatures for whom classical mechanics is a natural while contemplating quantum mechanics gives us a lot of wrinkles on the forehead.

So, here is a very useful **semi-classical equation of motion (EOM)** for an electron in a crystal:

$$\hbar\frac{d\vec{k}}{dt} = \vec{F} \tag{6.3}$$

Here, $\vec{F}$ is the force. If the electron is subjected to an electric field ($\vec{E}$), then $\vec{F} = -e\vec{E}$. If the electron is subjected to a magnetic field ($\vec{B}$), then $\vec{F} = -\frac{e}{c}\vec{v}_g \times \vec{B}$.

The following three paragraphs are advanced and can be skipped on first reading.

The derivation of the above equation requires knowing QM pretty well, and is given at `http://griffin.ucsc.edu/teaching/11Q1-155/Lecture%2016.pdf`, if you are interested and well-prepared with QM[5].

[5]The derivation given there is more detailed than anything that you can find in a standard solid

Here, however, it suffices to note that the above EOM is based on a set of assumptions. (1) The electron wave is a wave packet with Δk such that $\Delta x \sim 1/\Delta k \gg$ lattice constant(s). (2) However, Δx must be much smaller than the wave length of the external field (E field, $\vec{B}$ field, or any force field). Namely, the external field can be an AC field, but its wave length must be much longer than the wave packet size.

The first assumption ensures that we can talk meaningfully about the electron belonging, and moving, in one band. If this assumption is broken, then we would be talking of an atomically localized wave function. Clearly, the wave function will have too large a Δk in that case and a description of the time dependence in terms of the time dependence of the average value of k only will become problematic. The second assumption means that from the external field point of view, the electron is a point particle to a good approximation. This is where the "classical" part of the "semi-classical EOM" comes in.

Note that the above equation 6.3 is a much more general version than Eq. (2.3) of the textbook. As we shall see later, Eq. (2.3) of the book is a limited expression that is valid, *only* for a semiconductor at the top of the valence band or at the bottom of the conduction band, whereas my equation above is valid for *any* crystal and for any part of any band in any material.

Eq. 6.3 should be read as follows: the rate at which the crystal momentum changes is given by the force. Thus, except the momentum is replaced by the crystal momentum, $\hbar\vec{k}$, this sounds a lot like classical mechanics.

Thinking further, one might wonder whether the crystal momentum has any relationship to the group velocity, as the momentum is, in classical mechanics, mass times velocity. The answer is basically no. The crystal momentum is *not* mass times $\vec{v}$ except under special circumstances. Having said this, the "special circumstances" turn out to be for top of the valence band and bottom of the conduction band: these are the most important states for semiconductor devices! In these cases, *if* the crystal momentum is carefully referenced and *if* the *effective* mass (m^*) is properly defined, then the "crystal momentum" *is* equal to the classical momentum $m^*\vec{v}_g$. We will learn what this statement actually means in the next lecture.

Now, let us apply the above equation to some interesting typical situations – and learn the *new* concepts of "electron" and "hole."

state textbook

6.4 Electrons and Holes

So far, we have been considering the ground state, mostly. However, now that we have the semi-classical EOM, we can start considering what happens when a crystal is perturbed by an external field.

Basically, knowing what electrons do in response to an external field should be enough to know what happens. In the band theory, what we need to know is how each electron responds to the perturbation, and how many electrons there are.

What is an "external field"? For instance, we might connect our sample to a battery, shine light on a sample, or join two samples to make a "junction device." As the result of the perturbation by such an external field, electrons will be "excited" a little bit, and interesting things may happen.

In this view, it appears as though we actually do *not* need to invent any "new particle" such as hole. In a mathematical sense, this is a true statement, but in a physical sense, this is not a true statement. In a fundamental physical sense, we actually do need to invent a hole in a semiconductor. *A hole is a physical entity that is separate from an electron.* We will see why in this and the next lecture. While we do that, we will also see, by the way, that the "electron" in a crystal is really *not* the electron that we know[6]!

6.4.1 Nearly Empty Band

This would be the case of the conduction band in a semiconductor at a very low temperature.

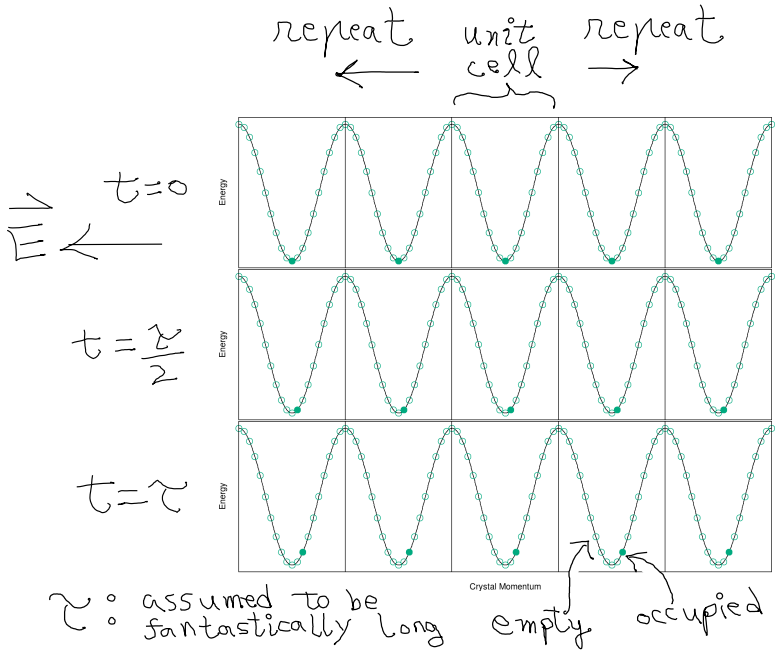
For illustration, we consider one electron in a band. For visualization, we consider a one dimensional crystal consisting of only 20 bases. We have 20 quantized k values per unit cell, as shown in the figure.

Suppose we apply a DC electric field $\vec{E}$. According to Eq. 6.3, the electron will change its k value according to $dk/dt = -eE/\hbar$. We assume that the electric field is applied to the left[7]. Then, the crystal momentum k will increase over time. We assume that after a finite time τ, the **relaxation time**, the electron loses its gained momentum and comes back to $k = 0$. This "relaxation" is caused by the scattering that the electron experiences with impurity, lattice imperfections, lattice vibrations

[6]The "electron" and the "hole" in a crystal are simple examples of the so-called "quasi-particles."

[7]In class, the field might have been drawn incorrectly, pointing to the right. Please correct your note, if this was the case!

("phonons"), and other electrons. It is the source of resistivity (inverse of conductivity), and, subsequently, the energy dissipation through Joule heating[8].



The current that this electron carries will be proportional to the average velocity that it gains during the relaxation time τ. This velocity is the so-called **drift velocity**.

The drift velocity is positive in the figure above. As the electron charge is negative, the current will flow to the left, i.e. in the same direction as the $\vec{E}$ field.

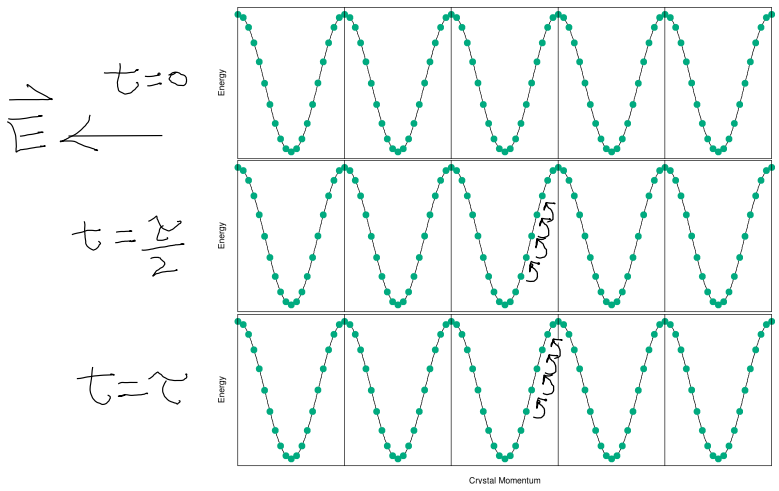
Note that in the above figure the magnitude of τ is greatly exaggerated to make the effect of τ easily visible. For real materials, τ is very small and the change of k during τ is minute compared to the unit cell.

6.4.2 Full Band

What if a band is completely filled?

With the same assumptions as above, except for the band filling, then, we have the following figure.

[8]Note that if $\tau \to \infty$, the current will oscillate due to the oscillating nature of the band dispersion. This is the so-called "Bloch oscillation." For common solids, this oscillation is not observable due to the very small τ value.



What is interesting in this case is that, while each electron moves to the next available k point after $\tau/2$, there is absolutely no change in the overall state! In particular, at any given time, there is no net momentum, or no net current. The latter can be defined precisely in terms of the **current density** $\vec{J}$, which is defined as[9]

$$\vec{J} = \sum_k (-e)\vec{v}_g/V \tag{6.4}$$

Here, V is the volume of the system, and $\vec{v}_g$ is the group velocity, which depends on k. $\sum_k$ is the sum over all k values in precisely one unit cell.

In the above figure, the group velocity is the slope of the curve divided by $\hbar$, because $v_g = d\omega/dk$ and $\omega(k) = \varepsilon(k)/\hbar$.

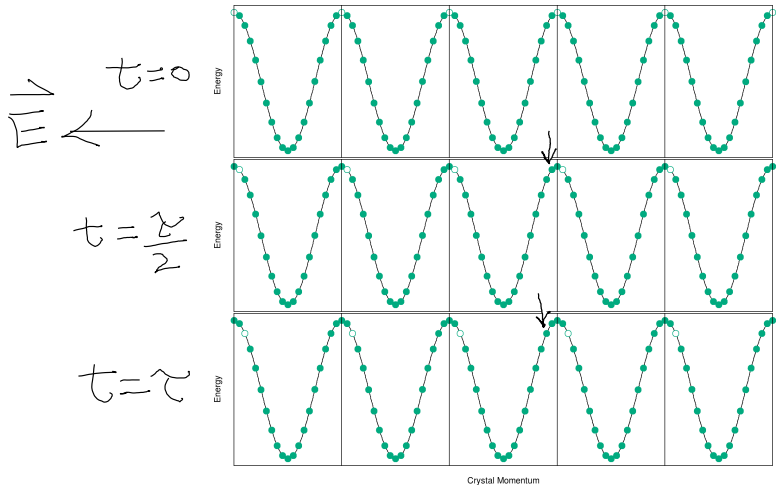
Due to symmetry[10]. $\sum_k \vec{v}_g = 0$, and thus $\vec{J} = 0$! This is the precise reason why a completely filled band does not lead to any electrical conduction! This is the reason why a material for which all bands are filled is defined as an insulator, in the band theory.

[9] A sign might have been dropped in classroom discussion of this formula. Please check your notebook and make a correction, if necessary. Also, here I use J in preference to j.

[10] Note that, in the figure, the sum over one unit cell is the sum over the range $(-\pi/a, \pi/a]$, excluding $-\pi/a$ (a vertical line) but including π/a (the next vertical line). So, the "symmetry" does not hold for one value of k, π/a, as its negative partner $-k$ is not included in the sum. This is no problem, however, since the group velocity at this special value is zero, as can be seen easily from the fact that the group velocity is given by the slope of the energy dispersion curve divided by $\hbar$.

6.4.3 Nearly Full Band

Now having considered a full band, let us take one electron out.



Here, we have a finite current! This is because the current is now given by

$$\vec{J} = \sum_{k'} (-e)\vec{v}_g / V \tag{6.5}$$

Here, I use $\sum_{k'}$ to mean all k values in a unit cell, except one. Clearly the distribution of electrons is not symmetric in the second ($t = \tau/2$) and the third ($t = \tau$) panels. The average of the instant currents for these two times will be the current of this system, as we assume that τ is the relaxation time (which means that right after $t = \tau$, the system relaxes back to the first panel situation – i.e. the equilibrium distribution).

Notice that in the second panel or the third panel, the k value marked by a vertical arrow (and the corresponding k values in all other unit cells) corresponds to a finite group velocity state occupied by an electron but the $-k$ state for it is not occupied. Namely, this k value is responsible for giving rise to a non-zero group velocity sum and, thus, a non-zero current. Note that the group velocity for this k value is positive. Therefore, again, the current flows to the left, as expected.

In the next lecture, we will see why this description in terms of electrons is less natural than the description of this situation in terms of a hole.