Modern Physics (PHY 3305) Lecture Notes

Solid-State Physics: The Theory of Semiconductors (Ch. 10.6-10.8)

SteveSekula, 30 March 2010 (created 29 March 2010)

Review

no tags

- We applied our model of the solid (using N square wells in a 1-D lattice) to understand several phenomena at a quantum level:
  - We learned that a solid contains bands with N levels, spaced by gaps where there are no allowed energies
  - The gaps occur when you reach an energy level, where the corresponding wavelength of the wave function satisfies $k = n\pi/a$, where a is the spacing of the ions in the lattice. The allowed k are determined by the nature of the lattice, but when k reaches $n\pi/a$ then a gap starts.
  - Conduction and insulation can then be explained by how bands are filled
    - if the conduction electrons fill all energy levels up to a gap, then it is very hard to apply a reasonable electric field to move electrons across the gap. This is an insulator.
    - if the conduction electrons fill only part of the available levels in a band, then they can easily be moved to higher levels (higher k, and thus higher momentum) by a reasonable electric field, this is a conductor.

Insulators, Conductors, and *Semi-Conductors*

(see slides)

- Which is a better conductor?
  - Lithium or Beryllium?
  - Lithium has three electrons. The first two are down in the 1S state, while the third is up in the 2S state. If there are N Lithium atoms, there are N states per band. The 1S band has 2N electrons in it, and the 2S band has N electrons in it, it's half-filled. Lithium should be a good conductor.

  - Beryllium has 4 electrons. For N atoms, there are 2N electrons in the 1S state and 2N electrons in the 2S state. Beryllium should be a bad conductor, but this in reality is not true. That's because the 2P state in Beryllium is close enough to the 2S state that when bands form, they overlap. So the 2S+2P state is only partially filled in Beryllium, and it's actually quite a good conductor.

Solids are complicated. Our simple square well is a bad model at some point, and Beryllium is a good example. To really understand solids, you need to apply the full correct model of the atom with all of its quantum numbers and energy level solutions.

- A semi-conductor is a material that is an insulator at zero temperature, but has a small band gap so that at T>0 some electrons migrate across the gap and can conduct in the next highest band.
  - an insulator with a band gap below 2 eV is a semi-conductor.
  - typical semi-conductors are silicon (1.1 eV) and germanium (0.7 eV).

Semiconductor Theory

A critical component in electronics is amplification, without which a large number of applications would not be possible. Electronic switching is related to amplification (think of switching as "all on" or "all off" - a two state amplifier). Being able to change state is critical to switching, and at the heart of that ability lie semiconductors.

Semiconductors, as applied to electronic switching and amplification, are indispensible. The DIODE is the simplest example of a semiconductor application. It differs from other components (resistors, capacitors, transformers) in a key way:

- Current can only flow one way through a diode

The key to this remarkable property is that the diode controls the flow of both negative and positive charge carriers. While it is true that the ions in a semiconductor lattice do not move, the *net response to an electric field can be precisely the same as freely moving positive charges*. These effective positive charge carriers are called HOLES.

Holes

In normal circuits, we learn about "conventional current" - where the negative charges moving against the electric field are the same as positive charges moving in the opposite direction. For most purposes, these two views are indistinguishable. However, this is not the same as "holes" in semiconductor theory. In conventional current, you can tell the difference between negative and positive charge carriers using the Hall Effect (application of a magnetic field to the conductor). However, in semiconductors, the holes behave the same whether the field applied is electric or magnetic.

The origin of holes is easily understood from our picture of bands and gaps. Consider an electron at the top of the valence band in a semiconductor (see slides). At temperatures above zero, there may be enough thermal energy to promote the electron across the gap, into the conduction band. This leaves a vacancy in a state behind in the valence band - a "hole".

The newly promoted electron is free to move around in the conduction band, but the HOLE is also mobile in the conduction band.

Holes "float". Seeking the lowest possible total energy at a given temperature, if a hole appears lower down in the valence band, electrons will "fall down" into the vacancy, which has the effect of moving the hole up the band to the top. (see slides)

While valence electrons are not free, the hole - the unoccupied state - is free to move in the valence band just as an electron is free to move in the conduction band.

Differences between holes and electrons:

- Holes seek higher energy states, electrons seek lower energy states.
- Holes are free to move in the valence band, while electrons are free to move in the conduction band.
- In moving around the band, the holes motion is in the opposite direction of an electron's motion

Holes have an "effective mass" that differs from the electron mass. At the top of the band, a hole has a positive effective mass. This allows it to participate in conduction along with the electrons, as it can respond to an external force (an electric field, for instance). Holes in this position will move with the electric field, while conduction electrons will flow again the electric field. Holes and electrons BOTH contribute their own current.

Intrinsic vs. Extrinsic Semiconductors

These are *intrinsic* semiconductors that we have been describing - those with pure atomic structure which behave as semiconductors. Extrinsic semiconductors are created by "doping" the crystal - that is, adding impurities to the crystal whose net effect is to increase the number of charge carriers. The doping material either has a large majority of conduction electrons or valence holes, and gives MANY MORE charge carriers to the semiconductor than is available in the pure crystal.

- n-type extrinsic semiconductors
    - typically dope an intrinsic semiconductor with phosphorous or arsenic, which are "pentavalent" - that is, they have 5 valence electrons.
    - typical doping: 1 in every 10^5 atoms is a dopant.
    - the valence band in the intrinsic semiconductor is totally full
    - Each impurity atom adds 1 additional electron, but not into the conduction band of the intrinsic semiconductor; each dopant adds new positive charge to the lattice, as well as negative charge. The net effect is to create a set of states just below the conduction band, called the DONOR STATES. These are NOT a new band, just a new set of states each containing one electron from one of the dopant atoms.
    - At T=0, the additional elections sit in the donor states, about 0.05 eV below the conduction band
    - At T>0, those electrons jump the gap and are now free to conduct in the conduction band.
    - N-type gets it's name not just from the fact that additional electrons are added to the system, but also because the majority participants in conduction are electrons, not holes.

- p-type extrinsic semiconductors
    - complimentary to the n-type.
    - Add a trivalent atom to the intrinsic semiconductor, such as gallium or aluminum.
    - the missing electrons, combined with the added positive charge of the dopant ions, alters the band structure of the intrinsic semiconductor slightly again.

  - the result is that the unfilled states from the dopant are slightly ABOVE the valence band. They are ACCEPTOR STATES.
  - Thermal excitation readily bumps electrons from the valence band to the acceptor states, leaving behind holes in the valence band which can conduct.
  - p-type extrinsic semiconductors are characterized by the fact that VACANCIES in energy levels (holes) are the majority participants in conduction.
  - Vacated donor states do NOT become holes (there is nothing to easily move into them - the lower gap is still large).

There is, in both n-type and p-type extrinsic semiconductors, always some thermal excitation across the gap separating the valence and conduction bands. These create in each the "minority carriers" - in n-type, holes in the valence band that conduct and in p-type electrons in the conduction band that conduct. However, the donor and acceptor states create far more of those specific carriers, and are the "majority carriers."

What makes doped semiconductors so useful is the control that we can exercise over their exact electrical properties via impurity type and concentration.

Example: the Diode

In a diode, the distribution of charge carriers brings about a physical asymmetry in the material that allows current to flow in one direction but NOT in the other.

The diode is a sandwich of p-type and n-type extrinsic semiconductor - a "p-n junction" is the area of contact between them.

Behaviors:

- when the diode is FORWARD BIASED, electrons are injected into the n-type semiconductor, while its existing electrons conduct out the other side into the p-type semiconductor, filling the holes there. On the p-type side, holes are added into one side and combine with electrons at the junction. The net flow of electrons into n-type and the net flow of holes into p-type causes a sustained current through the circuit.
- when the voltage is inverted and the diode is *REVERSE-BIASED*, now holes are draw away from the junction and electrons are drawn away from the junction. The junction becomes a region depleted of charge carriers, and no current can flow through the system.

At the boundary between the two (the junction), the bands don't line up. When the semiconductors are brought together, it's like joining two columns of water of different height by opening a valve between them. They seek a common equilibrium level, and settle there. Likewise, at the junction electrons in the n-type and holes in the p-type semiconductor intermingle until they reach equilibrium - in this case, until the Fermi energy on each side of the junction is the same. This corresponds to the lining up of thje donor and acceptor energy levels, and now there will be no net flow of electrons and holes.

When a voltage is applied (see slides), this causes the bands to move further apart at the region in-between is depleted of charge carriers even further (reverse bias) or to match up so that the bands on either side of the junction coincide (forward-bias). Minority carriers are always present, and even reversed biased the p-n junction has a tiny current due to this; but it's essentially zero for the purposes of most electronic applications.

When electron-hole recombination occurs in the forward-biased case, energy is released as the electron leaves the conduction band for a hole in the valence band. This energy can result in heat, or if the design is tuned, in light (LED).

QUESTION: what is the band gap in a red LED?
ANSWER: red light has a wavelength of about 700 nm.
$E = hc/\lambda = 2.8 \times 10^{-19} J = 1.8 eV$.

Example: The Transistor

A double sandwich of n-p-n doped semiconductor. The p-type is the base, the n-types are the emitter and the collection (they typically have different doping). The sandwich causes the emitter-base junction to be forward biased while the base-collector junction is reverse-biased. While the electrons entering the base are minority carriers there, and a small current leaves the bias output, the majority of electrons slide down into the conduction band of the collector, resulting in a huge current out the collector. emitter-collector current can be 100 times larger than the emitter-base current. This is the amplification, then. A small input signal (emitter-base) can be amplified into a large emitter-collector signal. (see slides)

For their 1948 discovery of the transistory, John Bardeen, Walter Brattain, and William Shockley were awarded the 1956 Nobel prize in physics.