Modern Physics (PHY 3305) Lecture Notes

Solid State Physics: Bonding, Bands, and Gaps (Ch. 10.1, 10.4-10.5)

SteveSekula, 23 March 2010 (created 22 March 2010)

Review

no tags

We have now covered the basic ingredients and conclusions of Statistical Mechanics:

1. The behavior of large systems is probabilistic, and predictions become more accurate as the number of participants in the system increases.
2. A wide class of physics can be approached assuming that particles are non-/weakly interacting, while still able to exchange energy
3. Distinguishable particles occupy states, n, with energy E_n with probabilities given by the Boltzmann Distribution
4. Identical particles occupy states, n, with energy E_n with probabilities given by either the *Bose-Einstein* (Bosons) or *Fermi-Dirac* (Fermions) distributions.
5. We "built" a model of a metal (fermion electron gas) from Coulomb potentials in one dimension and argued that, to a good approximation, this looks like N conduction electrons in a 1-D finite square well.
    a. Electrons in the gas will tend to fall down into the lowest energy state possible in the well owing to the "coldness" of room temperature compared to a metal's energy levels. This makes only the electrons near the top of the filled energy levels participate in things like the photoelectric effect, and explains a wide variety of phenomena (e.g. the work function).
6. Photons in a gas enclosed in a box are exactly the "blackbody radiation" problem, and we saw qualitatively that the blackbody energy distribution per unit frequency contains the *Bose-Einstein* probability.

Class Participatory Discussion on Model Building

Begin by building a two-atom molecule. Students will be expected to go to the board and draw the potential for a single atom, the potential for two atoms far away, then in close proximity. Relate this to a simpler model: two finite square-well potentials.

Discuss the wave functions in two separated single wells, and then what happens to the wave functions as the wells become closer. Start with the ground state, work up to the first excited state (n=2). What happens to the wave functions for each conduction electron in each atomic well?

Discuss how energy levels bunch and that there begins the formation of "bands" - energy levels closely spaced, with large gaps in between the groupings.

Add more atoms (make it three). What happens to each of the original energy levels? Discuss how the number of splittings scales with the number of atoms bound together. Now we really have "bands" - groups of very closely spaced energy levels separated by much larger zones with no energy levels allowed. Bands and band-gaps are the nomenclature to be introduced here.

PhET Demonstration:
http://phet.colorado.edu/simulations/sims.php?sim=Double_Wells_and_Covalent_Bonds

What lessons do we take from this exercise?

- The finite square well offers a fertile ground to begin thinking hard about how to model, say, a metal at the quantum level
- The proximity of the wells causes an electron to be "shared" between them, with some wave functions having nodes and some having anti-nodes in the "forbidden region" between the atoms.
- Combining N atoms causes each single energy level of a finite square well to "split" into a "band" of N closely spaced levels, centered on the energy of the original finite square well level.

Valence Electrons

DISCUSSION: what are valence electrons? How do they help us predict the properties of an atom?
GUIDE: they are the outermost electrons in an atom (ones in the highest energy states). When the highest energy states are only partially filled by valence electrons, such atoms are HIGHLY REACTIVE as they are prone to give up those electrons to form bonds. When the highest energy states are nearly completely filled by valence electrons, they are HIGHLY REACTIVE and seek to gain completion of their valence shell. A complete valence shell leads to chemically inert atoms.

Bonding and Energy

Bonding occurs when a joint energy level has a LOWER energy than the corresponding "nth" energy level of either well individually. This is a simplified view, but fairly accurate. It's energetically favorable for the two atoms to bind than to remain on their own. Recall from stat. mech. that overall, a system of N atoms will seek the lowest possible energy state as it has the most corresponding micro-states.

- EXERCISE: Ask the students to think about whether H_2 or He_2 are possible by considering the energy of the system when the 2/4 electrons fill the lowest molecular energy states. Compare the average energies to the energy of the singular atomic ground state.
    - CONCLUSIONS: there are "bonding" states, which means their energy is lower than the atomic nth-level energy, and "anti-bonding" states, whose energy is higher. He_2 is not favorable because it leads to no improvement in the energy of the pair. He_2^+ is possible, even with 1 electron in the anti-bonding state, because the average energy is lower than the atomic nth level.

The Covalent Bond

We now understand the "covalent bond" we learned about in chemistry. In this bond, two electrons are shared between a pair of atoms. This bond it very strong because the electrons' wave function probability is maximal in the space between the two atoms (the charge density is maximal between the two, and the two protons are strongly attracted to each others' electrons).

The Covalent Bond in Modern Society

Covalent bonds are a "hot topic" in American society right now. Does anybody know why?

- Lead the discussion: it has to do with food safety. (c.f. http://twit.tv/kiki1) Specifically, it has to do with plastics.
- The answer: Perfluorocarbons. They are highly inert chemically, and have a number of (potential) applications including as a means to replace blood in an emergency (they carry gases very well) and as a means to treat decompression sickness. They are used in manufacturing (e.g. in the application of teflon to pans), and several

PFC derivatives have been found in human blood samples taken all over the globe. *PFCs* are potent greenhouse gases, and are regulated by the Kyoto Protocol. Their health effects on humans are not understood, but they are difficult for the body to handle and remove (they are biochemically inert, and the covalent CF bond cannot be broken by the body - it's the strongest organic chemical bond in nature).

Solids

All atoms at sufficiently low temperature or high pressure form solids, and a large number of those are crystal lattices.

- DISCUSSION: what is a "crystal lattice"?
- GUIDE: a lattice is a regularly spaced arrangement of atoms in well-defined locations. The crystal occurs when the lattice is made from repeating sub-units of geometric atomic arrangement.
    - There are also "amorphous solids" where the bond angles and lengths are so irregular as to make an accountable relationship between atoms unmeasurable. Glass and rubber are good examples of these.

Geometric considerations lead to the conclusion that there are 14 lattice types in 3-D, or ways in which atoms can be arranged in a unique, geometrical pattern. Common examples in nature:

- simple cubic
- body-centered cube (first column of periodic table, transition metals such as iron, chromium, and tungsten)
- face-centered cube (copper, silver, gold, and noble gases, sodium-chloride where each *Na-Cl* pair replaces an atom)
- hexagonal closest-packed (helium, magnesium, zinc, titanium)

QUESTION: why can't atoms be in ANY location in the cube?

Solids are classes not just by geometry but by the type of bond that coheres the solid:

- Covalent Solid: diamond - strong bond means strong properties: hardness, poor electrical conduction, high melting point
- Ionic Solid: when atoms with nearly filled shells meet atoms with weakly bound valence electrons, the former may seize electrons from the latter. Strong electrostatic attraction is what forms the bond, and it's quite strong. Materials like this have a high melting point and are relatively hard. The exchange of electrons leads to nearly filled noble-gas-like states, meaning few/no electrons free for conduction. Geometry is determined by whatever configuration leads to the lowest electrostatic energy. This bond relies on an imbalance between two atoms, so ELEMENTS do NOT form ionic bonds.
- Metallic Solid: When a bonding is unable to join the valence electrons of one atom to all atoms around it, this leaves extra valence electrons and forms a metallic solid. The extra valence electrons are shared by all atoms, a state which leads to a lowest overall energy for the solid. Shared by all atoms, valence electrons move with little resistance around the solid, and form an electron gas. The attraction that holds a metallic solid together is between the positive ions and the pervasive electron gas. Since the bond that forms the metal is not between ions forming the solid, metals are soft (malleable) and possess lower melting points than other solids (e.g. covalent solids). Valence electrons free to move about the solid are called CONDUCTION ELECTRONS. They readily respond to an electric field. Metallic solids are excellent conductors.
- Molecular Solid: When molecules bonded in a "noble" way (no unused valence electrons and shells filled), they can still form a solid through a molecular bond. This happens due to an induced dipole in the electron clouds of two molecules, leading to a very weak bond (the London force). These form a low temperature, and even then are very soft. Molecules with strong innate dipole moments (e.g. water) will form comparatively stronger bonds. The dipole-dipole attraction in water is responsible for at least some of its odd properties, such as its high melting point. They all lack free electrons and are poor conductors.

Bonding is, in general, quite complex, and is usually a combination of at least a few of the above. Some properties of solids are not so easily explained, like malleability. What about conduction?

Even covalent and ionic bonded solids show variation in properties. Consider the "tetravalent" atoms - those with 4 valence electrons. Carbon (Z=6), silicon (Z=14), germanium (Z=32), and tin (Z=50). Carbon (diamond) is a terrible conductor, yet tin is an excellent conductor, with progression of conductivity in between. If these are covalent bonded structures, why do their properties vary?

The answer is that even in covalent and ionic solids, as in metallic solids, we have to think of the valence electrons as being shared across the solid as a gas. Why, then, are some solids in this class good and bad conductors? The answer has to do with how energy levels are structured in a solid: bands and gaps.

Energy Bands

PhET Demonstration:
http://phet.colorado.edu/simulations/sims.php?sim=Band_Structure

The number of conduction electrons often does not equal the number of valence electrons. But, if nonmetals (bad conductors) are to also be treated as an electron gas, what makes them insulate? ANSWER: the well is not a simple well - it has energy levels, closely spaced (bands) and separated by band gaps - large regions where there are no allowed quantized energy levels.

Why are the energy levels closely spaced? Some have more kinetic, and some more potential (anti-nodes in the classically forbidden regions), energy. This tradeoff causes close spacing of the levels, and when you reach a next-lower state where the nodes and anti-nodes REVERSE, then the energy jumps and you have a gap.

Use the *PhET* demonstration with four wells, and compare ψ_8 to 7,6,5, and then to 4. See the big jump in where the nodes/anti-nodes are located?

Energy Bands at large N

When the number of atoms is large (N), we can draw several conclusions:

- N-levels appear at each single-atom energy level, closely spaced.
- The jump between bands at n and n+1 occurs sharply when $k = n\pi/a$, where a is the inter-atomic spacing and $k = 2\pi/\lambda = p/\hbar$.
- Because N is so large, each band is essentially a continuum of energies (very closely spaced quantized energies).
- The band width depends not on the number of atoms but their spacing. Note that at too small a separation, the atoms tend to repel and that is something we've ignored. In general, atoms try to achieve a spacing that lowers overall energy.

-

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.