

Electron Distribution in Degenerate Semiconductors

When the chemical potential is sufficiently close to one of the band edges that the classical approximation no longer applies, we have to use the full treatment as we did (to some level of approximation) for a Fermi gas in lectures 16 and 17. In particular, our old expression for the number of particles was

$$N = \int_0^{\infty} f(\epsilon) \mathcal{D}(\epsilon) d\epsilon,$$

and μ appearing in $f(\epsilon)$ is adjusted to produce the correct number of particles, and $\mathcal{D}(\epsilon) d\epsilon$ is the number of states with energies in the range $\epsilon \rightarrow \epsilon + d\epsilon$. For a spin 1/2 particle in a box we had

$$\mathcal{D}(\epsilon) = \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{\epsilon}.$$

As we've discussed, to get the right density of states we should replace m by m_e^* . In addition, the energy that appears in the square root in the density of states really came from changing variables from momentum to energy. At the bottom of the conduction band, the electrons presumably have very little momentum. That is, the first state to be filled corresponds to an electron wave which is a standing wave made up of the smallest possible momentum components. The energy ϵ_c is negative and is mostly due the interactions with the atoms (remember, before the atoms formed a crystal, the atomic state that gave rise to the conduction band was a bound electronic state). The upshot of all this is that we should replace the energy in the density of states expression by $\epsilon - \epsilon_c$ and the integral should start at ϵ_c . We have

$$n_e = \frac{N_e}{V} = \frac{1}{2\pi^2} \left(\frac{2m_e^*}{\hbar^2} \right)^{3/2} \int_{\epsilon_c}^{\infty} \frac{\sqrt{\epsilon - \epsilon_c} d\epsilon}{e^{(\epsilon - \mu)/\tau} + 1}.$$

Recall the definition of n_c ,

$$n_c = 2 \left(\frac{m_e^* \tau}{2\pi \hbar^2} \right)^{3/2},$$

from which we get

$$\frac{n_e}{n_c} = \frac{2}{\sqrt{\pi}} \frac{1}{\tau^{3/2}} \int_{\epsilon_c}^{\infty} \frac{\sqrt{\epsilon - \epsilon_c} d\epsilon}{e^{(\epsilon - \mu)/\tau} + 1}.$$

Now let $x = (\epsilon - \epsilon_c)/\tau$ and $\eta = (\mu - \epsilon_c)/\tau$, and we can write the result in dimensionless form

$$\frac{n_e}{n_c} = \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{\sqrt{x} dx}{e^{x - \eta} + 1} = I(\eta),$$

where $I(\eta)$ is called the Fermi-Dirac integral. In the classical limit, where $\eta \ll -1$, the Fermi-Dirac integral reduces to our previous result

$$\frac{n_e}{n_c} = e^\eta = e^{(\mu - \epsilon_c)/\tau}.$$

If this limit doesn't apply, then an excellent approximation is

$$\eta = \log \frac{n_e}{n_c} + \frac{1}{\sqrt{8}} \frac{n_e}{n_c}.$$

Figure 13.4 of K&K shows how well this *Joyce-Dixon* approximation works.

Ionization of Donors and Acceptors

We were working out the electron and hole concentrations in a semiconductor. We found that the concentrations depended on the concentration of donor and acceptor ions. Up to now, we've assumed that these were "givens." Of course, the true givens are the concentrations of the donors and acceptors and an overall equilibrium will be established between donor and acceptor ion concentrations, the concentration of electrons in the conduction band, and the concentration of holes in the valence band.

When a donor is introduced into a semiconductor, the extra electron may stay bound to the donor, or the donor may ionize and contribute the extra electron to the conduction band. When bound to the donor, the energy of the electron is ϵ_d . This must be less than the lower edge of the conduction band ϵ_c (otherwise the donor will always be ionized). However, ϵ_d should not be a lot less than ϵ_c or the donor will hardly ever be ionized and there wasn't much point in introducing this donor in the first place. Thus donor bound states are "in the gap," but close to the conduction band edge. Similarly, acceptor bound states, with energy ϵ_a are in the gap but close to the valence band edge at ϵ_v . Table 13.2 of K&K gives some values of $\Delta\epsilon_d = \epsilon_c - \epsilon_d$ and $\Delta\epsilon_a = \epsilon_a - \epsilon_v$. The values are a few tens of milli-eV.

Next we want to calculate the probability that a donor is ionized or neutral. There are two bound states corresponding to the two spin states of the electron. There is one ionized state and this state has no particles and no energy. Then the grand partition function for this system is

$$\mathcal{Z} = 1 + e^{(\mu - \epsilon_d)/\tau} + e^{(\mu - \epsilon_d)/\tau}.$$

The probabilities that the donor is ionized or neutral are

$$f(\text{D}^+) = \frac{1}{1 + 2e^{(\mu - \epsilon_d)/\tau}},$$

$$f(\text{D}) = \frac{1}{1 + \frac{1}{2}e^{(\epsilon_d - \mu)/\tau}}.$$

In the case of an acceptor, there's one way to ionize it—adding an electron (which means producing the hole in the valence band) gives four covalent bonds with the four neighbors. Since all these electrons are equivalent, there's only one way to produce this state. It

has one extra electron and it has energy ϵ_a . When it's neutral, the added hole (missing electron) has two spin states, and the system has energy 0.

$$\mathcal{Z} = e^{(\mu - \epsilon_a)/\tau} + 1 + 1$$

$$f(A^-) = \frac{e^{(\mu - \epsilon_a)/\tau}}{e^{(\mu - \epsilon_a)/\tau} + 2} = \frac{1}{1 + 2e^{(\epsilon_a - \mu)/\tau}},$$

$$f(A) = \frac{2}{e^{(\mu - \epsilon_a)/\tau} + 2} = \frac{1}{1 + \frac{1}{2}e^{(\mu - \epsilon_a)/\tau}}.$$

To get the concentrations of ions, we multiply by the total concentrations.

$$n_d^+ = n_d f(D^+) = \frac{n_d}{1 + 2e^{(\mu - \epsilon_d)/\tau}},$$

$$n_a^- = n_a f(A^-) = \frac{n_a}{1 + 2e^{(\epsilon_a - \mu)/\tau}},$$

Recall our expressions for n_e and n_h in the classical limit

$$n_e = n_c e^{-(\epsilon_c - \mu)/\tau},$$

$$n_h = n_v e^{-(\mu - \epsilon_v)/\tau}.$$

We now have expressions for four concentrations in terms of givens and one unknown. The givens are the temperature, the concentrations of donors and acceptors, the quantum concentrations of electrons and holes in the conduction and valence bands (which depend on temperature and effective masses), the energies at the band edges, and the donor and acceptor ionization energies (whew!). The unknown is the chemical potential. The chemical potential must be chosen so that the system is electrically neutral. To help with this, note that the total negative charge concentration must equal the total positive charge concentration,

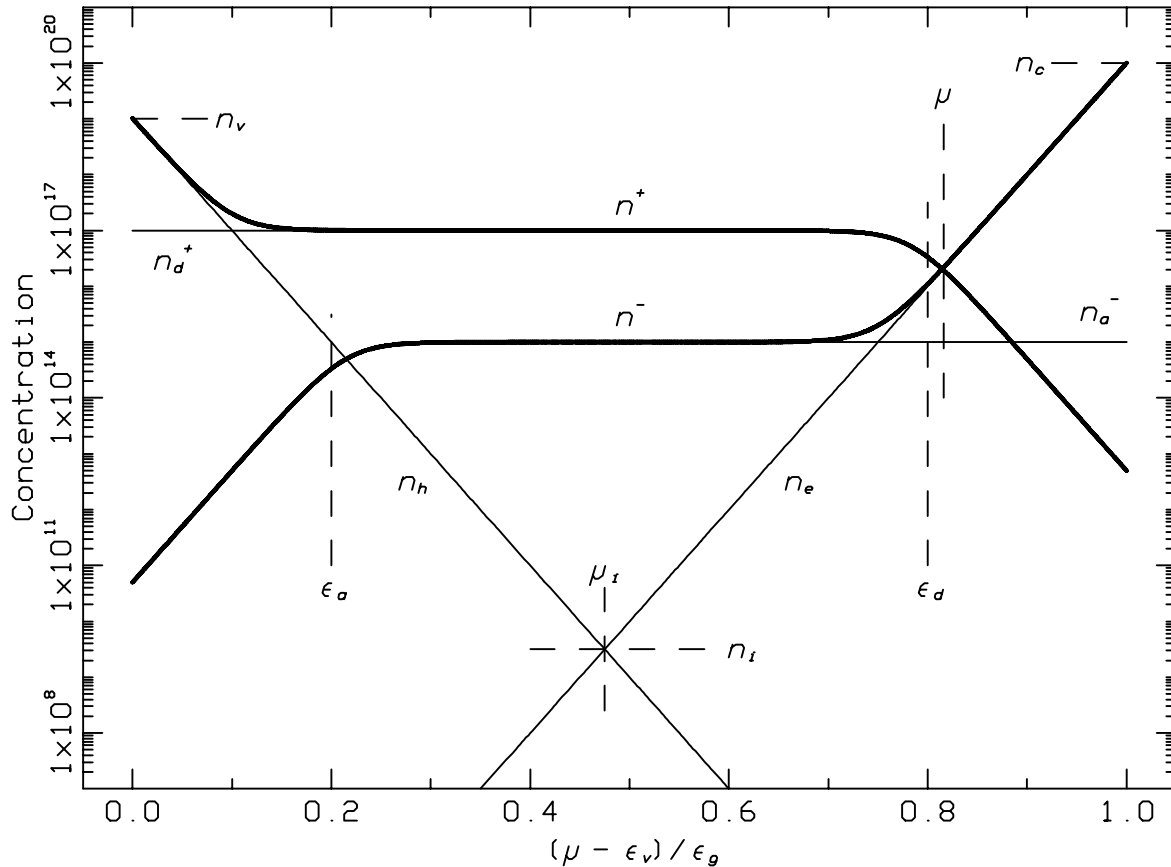
$$n^- = n_e + n_a^- = n_h + n_d^+ = n^+.$$

We can make a semi-log plot of all four concentrations versus μ . The electron and hole concentrations are just straight lines, and the ionized donor and acceptor concentrations are essentially two intersecting straight lines. Then we can plot n^+ and n^- , and where they intersect determines μ and all the concentrations. All this is shown in the figure. The range of μ shown in the figure is from ϵ_v to ϵ_c . Of course, μ must remain a few τ away from either band edge in order that our classical approximation remain valid.

The figure is drawn with μ scaled to run between 0 and 1 at the valence and conduction band edges. The curves have been drawn for $n_d = 10^{17} \text{ cm}^{-3}$, $n_a = 10^{15} \text{ cm}^{-3}$, $n_v = 10^{19} \text{ cm}^{-3}$, and $n_c = 10^{20} \text{ cm}^{-3}$. The donor ionization concentration is flat when $\mu \ll \epsilon_d$ and slopes downward at greater μ . Similarly, the acceptor ion concentration is flat when

$\mu \gg \epsilon_a$ and slopes downward at smaller μ . All the slopes are just $\pm 1/\tau$ which is why all the slopes in the drawing are the same! The temperature has been set so that as μ traverses the gap, $\exp(\mu/\tau)$ changes by 20 orders of magnitude. In other words, the gap width is $20 \times 2.303 \times \tau$. (2.303 is $\log 10$.)

If we really want to consider μ close to the band edges, then we can use the Joyce-



Dixon approximation or the result of numerically integrating the Fermi-Dirac integral. When μ gets to within a few τ of the band edge, the n_e and n_h curves start to flatten.

We can also use the figure to see what happens if the temperature is changed. As an example, suppose the temperature is doubled (actually this might damage your semiconductor, but this is a thought experiment!). The quantum concentrations n_c and n_v increase by a factor of $\tau^{3/2} = 2\sqrt{2}$, so the intersection of the n_e or n_h line with the right or left edges of the diagram moves up by this amount. But the major effect is the change in the slope of the sloping parts of all the curves. All slopes halve for a doubling of the temperature. The n_e or n_h curves go through the new values of n_c or n_v with the new slope, while the n_a^- or n_d^+ curves continue to go approximately through n_a and ϵ_a or n_d and ϵ_d with the new slope. Mentally shifting the curves around can give a qualitative understanding of what happens with a change in temperature.

Electron-Hole Interactions

In our discussion so far, the overall picture has probably been one of a “static” situation. But just as with a gas, which may seem quiescent, but has thermal motions at the molecular level, the situation within a semiconductor is similar. The holes and electrons have thermal motions. Electrons and holes may recombine with donors and acceptors. Similarly, electrons and holes may be ionized from donors and acceptors. Electrons and holes may *recombine* (annihilate!). And an electron-hole pair may be created. Some of these processes can emit or absorb light. For example, the light emitting diode or the CCD detector used in today’s video-cams.

The p - n Junction

Probably the most important thing in semiconductor technology is the p - n junction. Take a piece of donor (n -type) doped semiconductor and place it in contact (thermally and electrically) with a piece of the same kind of semiconductor with acceptor (p -type) doping. What happens?

To start with, the n -type semiconductor has lots of electrons and a Fermi level (μ) near the conduction band edge. The p -type semiconductor has lots of holes and a Fermi level near the valence band edge. When we place them in contact, we have a non-equilibrium situation! This is because the chemical potentials are different. Electrons will flow from regions where their chemical potential is high to regions where their chemical potential is low. Holes flow in the opposite direction (because a high chemical potential for an electron is equivalent to a low chemical potential for a hole).

At the boundary between the two types of semiconductor we have excess electrons (excess negative charge) in the p -type and excess holes (excess positive charge) in the n -type semiconductor. If we imagine that the boundary is a plane, then we have a plane of positive charge next to a plane of equal negative charge. What does this do? It’s like the charge on the plates of a capacitor. It creates an electric field pointing from the n -type to the p -type semiconductor. There is an electric potential difference between the n -type and p -type semiconductors with the n -type at a higher potential.

Now consider following an electron as it travels from the n -type to the p -type semiconductor. When it’s well inside the n -type semiconductor, its intrinsic chemical potential is $\mu_{i,n}$. Here, “intrinsic” does not mean the chemical potential associated with the intrinsic carrier concentration, but rather, the chemical potential calculated ignoring external sources of potential energy, such as the electric potential energy we’ve just been discussing. As the electron traverses the junction and gets well inside the p -type semiconductor, its intrinsic chemical potential drops to $\mu_{i,p}$. On the other hand its electric potential energy rises by $e(V_n - V_p) = e\Delta V$, where V_n and V_p are the electric potentials (voltages) in the n and p -type semiconductors and the charge of an electron is $-e$. The electron traverses the

voltage step from high to low voltage, but since it has a negative charge, it increases its potential energy. As we discussed in lecture 12, any external sources of potential energy must be added to the chemical potential. The system will be in diffusive equilibrium when the total chemical potential is the same everywhere. In other words, when

$$\mu_{i,n} - \mu_{i,p} = e\Delta V ,$$

the electrons and holes stop diffusing across the junction.

What we've just discovered is that a p - n junction spontaneously develops a voltage. Since $\mu_{i,n}$ is close to the conduction band edge and $\mu_{i,p}$ is close to the valence band edge, we expect that the voltage times the electron charge should be the same order as the energy gap. For example, silicon has an energy gap of 1.14 eV at room temperature, so we expect voltage steps across silicon p - n junctions to be about one volt. The exact value depends on the type and amount of doping, as we will now work out.

Suppose one has an n -type semiconductor with no acceptors, suppose also that $n_i \ll n_d \ll n_c$ and that the ϵ_d is close to ϵ_c . Then the electron concentration and the chemical potential are determined by the intersection of the electron exponential with the flat part of the donor curve (that's what all the assumptions are for), and we have

$$n_d = n_e = n_c e^{(\mu - \epsilon_c)/\tau}$$

or

$$\mu_{i,n} = \tau \log \frac{n_d}{n_c} + \epsilon_c .$$

In the p -type conductor, we assume that there are no donors, that $n_i \ll n_a \ll n_v$, and that ϵ_a is close to ϵ_v . Then the hole concentration and the chemical potential are determined by the intersection of the hole exponential with the flat part of the acceptor curve, and

$$n_a = n_h = n_v e^{(\epsilon_v - \mu)/\tau} ,$$

which means

$$\mu_{i,p} = \epsilon_v - \tau \log \frac{n_a}{n_v} .$$

Now we subtract and find

$$e\Delta V = \mu_{i,n} - \mu_{i,p} = \epsilon_g - \tau \log \frac{n_c n_v}{n_d n_a} .$$

One might have $n_c/n_d \approx n_v/n_a \approx 100$, and in this case, the result for silicon at room temperature would be

$$e\Delta V = 1.14 \text{ eV} - 8.62 \times 10^{-5} \text{ eV K}^{-1} \cdot 300 \text{ K} \cdot \log 10000 = 0.90 \text{ eV} .$$

Note that as the temperature goes up, the potential difference gets smaller. The electron and hole concentrations stay at n_d and n_a , but since the quantum concentrations go up with temperature and especially since the slopes on our diagram get shallower with increasing temperature, both chemical potentials move towards the center of the gap.