

Topic 4: Carrier Concentrations in Intrinsic and Extrinsic Semiconductors

Preface

In this set of notes we will continue to formulate quantitative expressions for carrier concentrations in intrinsic and extrinsic semiconductors. I have chosen to separate this section from Topic 3 to keep the former more conceptual. We will begin with the simple expression, $n = p = n_i$, and manipulate it to discover several fundamental expressions within semiconductor physics.

Intrinsic Review

We showed in the last set of notes that an intrinsic semiconductor by definition has an equivalent hole and electron concentration. That is $n = p = n_i$, where n_i is the concentration of holes and electrons in an intrinsic material. The derivation concluded with the claim that hole and electron state densities must be equal. That derivation can also be used to express the Fermi level in terms of material parameters.

$$\begin{aligned} n &= p \\ \Rightarrow \frac{N_v}{N_c} &= e^{\frac{2E_F - E_c - E_v}{kT}} \\ kT \ln\left(\frac{N_v}{N_c}\right) &= 2E_F - E_c - E_v \\ \therefore E_F &= \frac{kT}{2} \ln\left(\frac{N_v}{N_c}\right) + \frac{E_g}{2} \end{aligned}$$

Choosing to solve for the Fermi level results in an expression-based entirely on material parameters or measurable constants. The thermal voltage, bandgap energy, and density of states are all dependent on the material. For context, the density of states for both holes and electrons is on the order of 10^{19} in silicon. The thermal voltage is approximately 26 milli-eV (meV), and lets say N_v/N_c varies by a factor of 10; $\ln(N_v/N_c) \approx \ln(10) = 2.3$. Approximately 60 meV is somewhat less than half the bandgap; 560 meV in silicon. Hence, we often approximate the Fermi level as the intrinsic level in an intrinsic semiconductor. Note that a ratio of 10 for the state densities is not realistic but shows that the Fermi level does not vary much from the middle of the bandgap in an extreme case.

Typically we use the quantity n_i to represent the carrier concentration in an intrinsic semiconductor but we lack an expression for it. We can make a seemingly arbitrary decision to multiply n with p and see what happens. For context, we choose to multiply them because we can an expression in terms of n_i . The expression $np = n_i^2$ is so fundamental to semiconductor physics we gave it the name, **Mass action law**. By the end of these notes, you will see why it is so fundamental.

$$\begin{aligned} n &= p = n_i \\ \Rightarrow np &= n_i^2 \\ n_i^2 &= N_c N_v e^{\frac{E_c - E_F}{-kT}} \cdot e^{\frac{E_F - E_v}{-kT}} \\ &= N_c N_v e^{\frac{E_c - E_F + E_F - E_v}{-kt}} \\ &= N_c N_v e^{\frac{E_c - E_v}{-kT}} \\ &= N_c N_v e^{\frac{E_g}{-kT}} \end{aligned}$$

Taking the square root of an exponential term halves the argument. Hence, our final expression for n_i^2 is:

$$n_i = \sqrt{N_c N_v} e^{\frac{E_g}{-2kT}} \approx 1.5 \cdot 10^{10} \text{ cm}^{-3} \quad (1)$$

If you were to plug the various constants into the equation above, it will likely not be equal to n_i . This discrepancy is due to the way we express state density. It is beyond the scope of this course but you are free to theorize why this may occur.

Extrinsic Semiconductors

Extrinsic semiconductors compose the bulk of our study. To be honest, there is not much we can do with just intrinsic semiconductors. Allowing non-equal carrier concentrations provides an opportunity to delve into the fundamentals driving many modern-day devices such as diodes, transistors, MOS devices, and such. However, before we can delve into the analysis we need to define some terminology and assumptions. We will not delve into why some of these assumptions apply at this moment, but we hope they will become evident throughout our analysis.

Assumptions:

- **Mass Action Law:** $np = n_i^2$. For extrinsic semiconductors, the mass action law still applies. We will not prove this fact, but throughout our discussion, we will show that this is a necessity.
- **Atomic Charge Neutrality:** From chemistry and electrostatics we know that individual atoms are neutral and that these atoms become ions when a valence electron is added or removed from the valence shell.
- **Charge Neutrality:** $N_D + p = N_A + n$. A chunk of semiconducting material must be neutral. There may be regions of non-neutral charge within the material but from a macroscopic view the material is neutral.
- **Total Dopant Ionization:** All doping impurities are assumed to be ionized.

Terminology:

- **Donor Concentration (N_d) :** The concentration of donor ions in the lattice in cm^{-3} .
- **Acceptor Concentration (N_a) :** The concentration of acceptor ions in the lattice in cm^{-3} .
- **Charge Density (ρ) :** The net charge density in a given region in $\frac{C}{\text{cm}^3}$.
- **Majority Carrier:** The charge carrier, hole or electron, which composes the bulk of the free charge in the material.
- **Minority Carrier:** The charge carrier, hole or electron, which is not very present within the material

Beginning with charge neutrality, the expression above may seem incorrect. If we create an n-type semiconductor with donor ions, why do we treat them as a positive charge? The answer lies in the word definition of donor and acceptor ions. When we dope with donors, Ex. Phosphorous, we insert an electron into the lattice. If this electron were to gain enough energy to leave the influence of the host atom, a positively charged ion would be left behind. The inverse is true for acceptor ions, Ex. Boron, where an electron would enter the influence of atom resulting in a negatively charged ion.

Without going into extensive details, dopant ions result in the formation of additional levels within the energy band diagram. Additional bands are formed near the conduction and valence bands for donor and acceptor ions, respectively. The figure below is provided in Chapter 1, Fig. 20, of (Sze S. M., 2012) which visualizes these dopant bands. Boron and phosphorous are approximately 0.045 eV away from the valence and conduction bands, respectively. Implying that only 0.045 eV is necessary to release the charge carrier and convert the dopant into an ion. We may assume that every dopant atom is ionized because the thermal voltage is 0.026 eV. If the temperature were closer to freezing, this assumption would not be valid. For our purposes, we may assume that every dopant atom is ionized, having released it's charge carrier to the lattice.

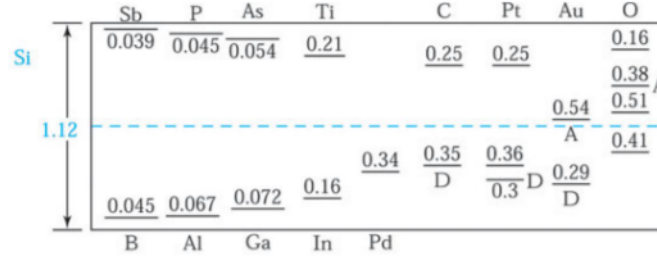


Figure 1: Silicon Energy Band Diagram with dopant ionization levels.

Extrinsic Analysis:

With the formality and definitions aside, we have two expressions to begin our analysis: 1. $np = n_i^2$ and 2. $N_D + p = N_A + n$. We begin by substituting $p = \frac{n_i^2}{n}$:

$$\begin{aligned}
 N_D + p &= N_A + n \\
 \Rightarrow N_D + \frac{n_i^2}{n} &= N_A + n \\
 nN_D + n_i^2 &= nN_A + n^2 \\
 \Rightarrow n^2 + n(N_A - N_D) - n_i^2 &= 0
 \end{aligned}$$

This final equation should be similar, it is a quadratic equation. Applying the quadratic formula gives us the following expression for the electron concentration:

$$n = \frac{N_D - N_A}{2} \pm \frac{\sqrt{(N_D - N_A)^2 + 4n_i^2}}{2} \quad (2)$$

I encourage you to repeat the steps replacing n with $p = \frac{n_i^2}{p}$. Doing so results in the following expression:

$$p = \frac{N_A - N_D}{2} \pm \frac{\sqrt{(N_A - N_D)^2 + 4n_i^2}}{2} \quad (3)$$

Take a moment to note what this equation represents term-by-term, and to acknowledge some of the mathematical “trickery” used. We can interchange $(N_A - N_D)^2 = (N_D - N_A)^2$ within the radical. The electron and hole concentration is dependent on the difference in dopant concentration with the intrinsic concentration. What does the \pm represent? Is it possible to have a negative concentration for either n or p ? What happens if N_A is much larger than N_D ? What happens if N_A and N_D are equal? These are the types of questions that this course focuses on and will be the focus on examinations.

Starting with the plus-minus symbol, there is **no** physical meaning for a negative concentration. Charge carriers are either present or not, so it will be up to you to determine when it is appropriate, or if it appropriate, to take the plus or minus term. Regardless of the values for n and p , it must obey the mass action law which you can use as a self-check.

Ex. 1:

Determine the doping concentrations for a piece of silicon at room temperature doped with 10^{16} acceptor atoms. $T = 300K$, $n_i = 1.5 \cdot 10^{10} \text{ cm}^{-3}$, and $N_A = 10^{16} \text{ cm}^{-3}$

Solution: The solution is quite straightforward, we just need to substitute the given values into Eq. (2) and (3). Since N_D is not given, it is assumed to be zero.

$$\begin{aligned}
 n &= \frac{0 - 10^{16}}{2} \pm \frac{\sqrt{(0 - 10^{16})^2 + 4 \cdot 1.5 \cdot 10^{10}}}{2} \\
 \Rightarrow n &\approx \frac{-10^{16} \pm 10^{16}}{2} \\
 \therefore n &= -22500 \text{ cm}^{-3} \text{ or } -10^{16} \text{ cm}^{-3}
 \end{aligned}$$

$$\begin{aligned}
p &= \frac{10^{16} - 0}{2} \pm \frac{\sqrt{(10^{16} - 0)^2 + 4 \cdot 1.5 \cdot 10^{10}}}{2} \\
\Rightarrow p &\approx \frac{10^{16} \pm 10^{16}}{2} \\
\therefore p &= 10^{16} \text{ cm}^{-3} \text{ or } 22,500 \text{ cm}^{-3}
\end{aligned}$$

We need to ask ourselves if these two answers mean anything. I encourage you to think through these questions before looking at the answers. Is it possible for there to be zero concentration of either carrier type? Can we have a negative carrier concentration? Why does the radical simplify to N_A ? You also need to be aware of your calculator's limitations. Calculating one of the values for n may result in a zero instead of -22,500.

I mentioned earlier we cannot have a negative concentration, we also cannot have zero concentration. Note, one of the values for n is not actually zero even though it is approximately zero in the second step. Remember $np = n_i^2$ and n_i^2 is a positive constant. If either concentration is zero or negative the mass action law is violated. As for the radical, $(10^{16})^2 = 10^{32}$ whereas $4n_i^2 = 9 \cdot 10^{20}$. The first term in the radical is several orders of magnitude larger than the intrinsic term, hence, we can make an approximation ignoring the intrinsic concentration. **Note**, this is dependent on the magnitude of doping concentration. If we have a lightly doped semiconductor, this may not be a valid approximation.

The only valid term above then is the hole concentration of $p = N_A = 10^{16}$. We still need to address what the electron concentration is. Since equation 2 gave us nonsense for an answer, we turn to the mass action law.

$$\begin{aligned}
n &= \frac{n_i^2}{p} \\
&= \frac{2.25 \cdot 10^{20}}{10^{16}} \\
\therefore n &= 22,500 \text{ cm}^{-3}
\end{aligned}$$

Note that the hole concentration is much larger than the electron concentration due to this doping; $n \lll p$. There are several orders of magnitude between concentrations, but it is not valid to approximate the hole concentration as zero. I may sound like a broken record, but I am reiterating this because it is important. If the mass action law is violated, there was an error in your calculation. I suggest keeping a list of assumptions throughout a problem because not every assumption we make now will be apply to every problem we encounter.

From this moment, we can devise many different types of questions. What happens if we swap N_A and N_D in the prior example? What happens if $N_A = N_D$? What happens if $N_A \neq N_D$ but $N_A \neq 0$ and $N_D \neq 0$? What will happen if either $N_A < n_i$ or $N_D < n_i$. I will leave these examples for you but make a brief comment on the second question. If $N_A = N_D$, $n = p = n_i^2$. That is, **equal doping concentrations produces a semiconductor no different from an intrinsic semiconductor**. This is due to recombination, a later topic in the course, but the electrons and holes will continually recombine into bonds, preventing the carriers from contributing to a substantial current. The semiconductor is still extrinsic but behaving intrinsically.

Fermi Level Confusion:

Recall from prior notes that we have two main energy levels within the conduction and valence bands; the intrinsic level and the Fermi level. There is often confusion over which value to use in equations such as:

$$n = N_c e^{\frac{E_c - E_F}{-kT}}$$

We simply replace the Fermi level with the intrinsic level for intrinsic semiconductors. By definition, these two terms represent the same probability of 50% electron probability. However, there is a clever way of reorganizing the formula into something more useful to us.

$$n = N_c e^{\frac{E_c - E_F}{-kT}}$$

$$\begin{aligned}
&= N_c e^{\frac{E_c + E_i - E_i + E_F}{-kT}} \\
&= N_c e^{\frac{E_c - E_i}{-kT}} \cdot e^{\frac{E_F - E_i}{kT}} \\
\therefore n &= n_i e^{\frac{E_F - E_i}{kT}}
\end{aligned}$$

This derivation is valid because we know the value of n in an intrinsic semiconductor; $n = n_i$. The argument of the exponent is no different if we add and subtract a constant from it. Using the intrinsic level instead of the Fermi level, we relate $n_i = N_c e^{\frac{E_c - E_i}{-kT}}$. I encourage you to find the equation for the hole concentration using this process. The final expressions are:

$$n = n_i e^{\frac{E_F - E_i}{kT}} \quad (4)$$

$$p = n_i e^{\frac{E_i - E_F}{kT}} \quad (5)$$

$$E_F - E_i = kT \ln\left(\frac{n}{n_i}\right) \quad (6)$$

$$E_i - E_F = kT \ln\left(\frac{p}{n_i}\right) \quad (7)$$

Eq. (6) and (7) are rearranged from Eq. (4) and (5) solving for the energy between the Fermi and intrinsic level. A good way of remembering the exponential terms of Eq. (4) and (5) is by how the Fermi level changes with doping type. N-type materials have an electron concentration above the intrinsic concentration. To increase n from n_i the exponential must have a positive argument. Therefore, the Fermi level must be closer to the conduction band for $E_F - E_i > 0$ to be true. The inverse is true for P-type material. With a Fermi level closer to the valence band, $E_i - E_F > 0$ is required to increase the hole concentration.

Ex. 2:

Using the doping concentration from Example 1 sketch the energy band diagram of the system. Explicitly show the energy between each neighboring energy level.

Solution:

Plugging in the hole and electron concentrations from Ex 1 we may find the energy difference between Fermi and intrinsic levels.

$$\begin{aligned}
E_F - E_i &= 8.617 \cdot 10^{-5} \frac{eV}{K} \cdot 300K \ln\left(\frac{22500}{1.5 \cdot 10^{10}}\right) = -0.347 eV \\
E_i - E_F &= 8.617 \cdot 10^{-5} \frac{eV}{K} \cdot 300K \ln\left(\frac{10^{16}}{1.5 \cdot 10^{10}}\right) = 0.347 eV
\end{aligned}$$

It is no coincidence that our two values are equal in magnitude. Combining this information into a sketch, we get the following figure.

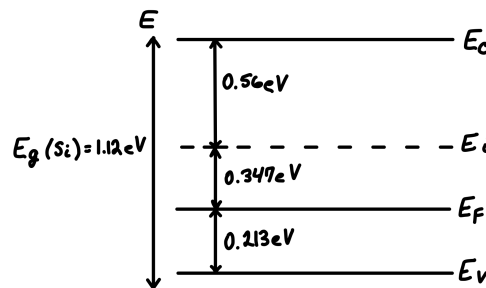


Figure 2: Ex. 2 Energy Band Diagram

I hope this figure accurately demonstrates the large amount of information encoded within carrier concentration. Knowing the carrier concentration allows us to determine how much energy is needed to jump between various energy levels, the type of the semiconductor, and doping intensity. If I were to show you an energy band diagram with E_F above E_i you could immediately determine it is an N-type

material. I have also hand-written the diagram to show what will be expected when we ask you to draw an energy band diagram on tests and quizzes.

In general, we can abstract this idea a little bit further. We notice that $E_i - E_F$ is a weak function of carrier concentration; since we have an $\ln(\frac{p}{n_i})$ or $\ln(\frac{n}{n_i})$ term. The slow growth of the natural logarithm explains why typical doping concentrations are 6 to 9 orders of magnitude above the intrinsic concentration; 10^{16} to 10^{21} cm^{-3} . To generate a substantial current, it is necessary to introduce many, many charge carriers. With this knowledge we may ask ourselves if we can dope beyond E_c or E_v ? Yes, it is possible to dope beyond the conduction and valence bands. We call these devices **degenerate semiconductors** and will not be a large focus in this course. In short, moving the Fermi level beyond either band results in conductor-like behavior. The extreme quantity of either holes or electrons allows substantial current to flow through the device. All the devices we will study are **non-degenerate**, meaning the Fermi level is bounded within the bandgap of the material.

Temperature Dependence:

Some of you may ask what about temperature? Eq. (6) and (7) are linear with respect to temperature. We discussed in the introduction that if we heat a semiconductor to outrageously high temperatures substantial current is possible. The figure below plots the electron concentration over various temperatures. The plot is provided in Chapter 1, Fig. 25, of the text (Sze S. M., 2012) but I have included it here for your convenience.

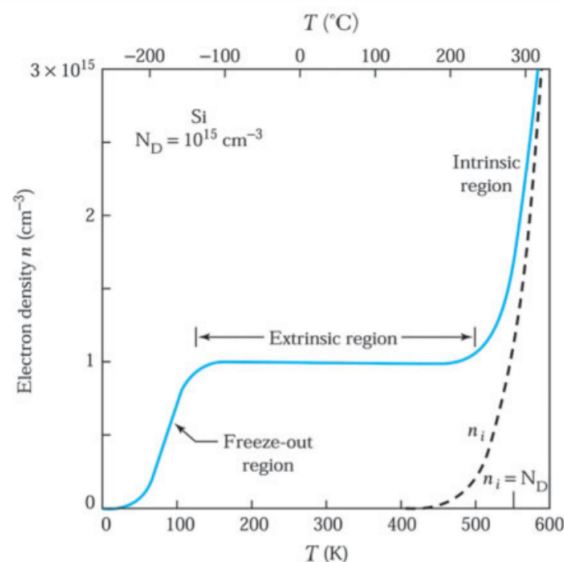


Figure 3: Electron concentration as a function of temperature

On the horizontal axis is temperature, in Kelvin. On the vertical axis is the electron concentration scaled by a factor of $3 \cdot 10^{15}$. There are two distinct curves in the plot, a blue curve, and a dashed curve. The blue curve represents the analysis of extrinsic semiconductors. At low temperatures, the electron concentration falls to zero due to the lack of thermal energy needed to ionize dopant atoms. The flat, middle, section represents the temperature range where the material behaves extrinsically. As the temperature increases drastically, we model the device as intrinsic. It does not matter what the doping concentration is because the thermal energy going into breaking bonds will generate more charge carriers than the impurities. Hence, the material may be modeled by an intrinsic material. The dashed curve represents an intrinsic material. Note how the two curves converge at larger temperatures.

Conclusion

In this set of notes we shifted our focus to extrinsic semiconductors and how doping affects carrier concentrations. We developed an expression for carrier concentrations from the mass action law and charge neutrality. These equations will sometimes lie to us if we just plug in numbers. Therefore, we

need to keep our focus on the physics that represents these equations to determine when a formula gives us nonsense.

From these carrier concentrations it is possible to quantify the difference in Fermi levels from the intrinsic level. As mentioned, the knowledge of carrier concentration encodes significant information regarding the energy band diagram. I highly recommend going through these derivations again, on your own and without your notes. Doing so will cement your understanding and better prepare you for quizzes and exams. As well as provide a more intuitive understanding of the physics within the device.

References

Sze S. M., L. M. K. ((2012)). Semiconductor devices: Physics and technology 3rd edition. Wiley.