# 8.231 Physics of Solids I — Fall 2017
# Lecture Note 14

Huitao Shen

November 9, 2017

## 1 Impurity Levels in Semiconductors

### 1.1 Energy Level in the Hydrogen Atom

The motion of electron in the hydrogen atom is described by the Schrödinger equation

$$\left(-\frac{\hbar^2\nabla^2}{2m}-\frac{e^2}{4\pi\varepsilon_0 r}\right)\psi(\mathbf{r})=E\psi(\mathbf{r}). \tag{1}$$

Its energy levels (for bound states) are

$$E_n=-\frac{me^4}{2(4\pi\varepsilon_0\hbar)^2}\frac{1}{n^2}=-\frac{1}{4\pi\varepsilon_0}\frac{e^2}{2a_0}\frac{1}{n^2}\approx\frac{-13.6\,\mathrm{eV}}{n^2}, \tag{2}$$

where

$$a_0=\frac{4\pi\varepsilon_0\hbar^2}{me^2}\approx 0.53\text{Å}, \tag{3}$$

is the Bohr radius.

### 1.2 Hydrogenic Impurities in Semiconductors

Let us consider the silicon, whose band structure is given below:

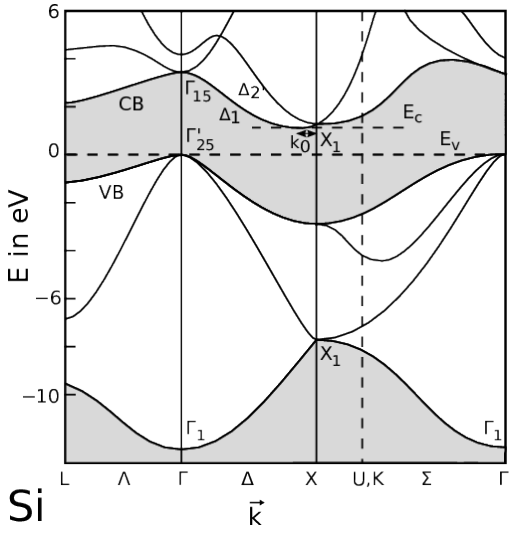


Figure 1: The silicon band structure calculated by the pseudopotential method.

It has a band gap of $E_g = E_c - E_v \approx 1.1\,\text{eV}$. It is helpful to remember the temperature - energy conversion rule: $1\,\text{eV} \sim 10\,\text{K}$. This means in order for the silicon to be conductive due to the thermal excitations, the temperature should be of order $10^4\,\text{K}$, which is way beyond the room temperature. Similarly, in order for the thermal motion to ionize hydrogen atoms, the temperature needs to be even higher (around $10^5\,\text{K}$).

In order for the silicon to be a good semiconductor at room temperature, it is necessary to dope the silicon with impurity atoms. Based on the electronic property of the impurity atoms, they can be classified into

- Donors, which have more valence electrons than the silicon, such as P and As.
- Acceptors, which have fewer valence electrons than the silicon, such as B and Al.

Suppose the silicon is doped with P. P will substitute Si in the crystal structure. The excessive valence electron in P that cannot form $\sigma$ bonds with adjacent Si, along with the $P^+$ ion, can be effectively treated as a hydrogenic atom, with modified effective mass and the dielectric constant. The hydrogenic atom has energy level

$$E_n = -\frac{m^* e^4}{2(4\pi\varepsilon_0\varepsilon\hbar^2)}\frac{1}{n^2}. \tag{4}$$

For silicon, the effective mass is $m^* = 0.2m$ and the dielectric constant $\varepsilon = 11.8$. It can be estimated that the bounding energy is only $E_D = 19.5\,\text{meV}$, instead of $13.6\,\text{eV}$ for the real hydrogen. This is the energy of the localized level below the conduction band. The physical origin of this weak binding is due to the screening effect of the Coulomb potential due to the remaining valence electrons in $P^+$ and other Si atoms in the lattice. Note that this energy is in the order of room temperature. The Bohr radius can also be estimated to be

$$a_D = \frac{4\pi\varepsilon_0\varepsilon\hbar^2}{m^* e^2} = \frac{\varepsilon m}{m^*}a_0 \approx 3\,\text{nm}. \tag{5}$$

Compared with the Bohr radius for the real hydrogen, this effective hydrogenic atom is quite large because of the loose binding. This means with large doping density $n > 10^{-6}$, these impurity orbitals will overlap and form a shallow impurity band called donor band below the conduction band.

Similar argument can be made for acceptors doping. There will be an empty acceptor band above the valence band.

The semiconductor doped with donors are called n-type. "n" represents "negative", meaning the main charge carriers (majority carrier) are of negative charge, which are electrons. The holes are minority carriers. The semiconductor doped with acceptors are called p-type. "p" represents "positive", meaning the main charge carriers are of positive charge, which are holes. It is instructive to use the flat band diagram to represent the energy levels of the donor band and the acceptor band.

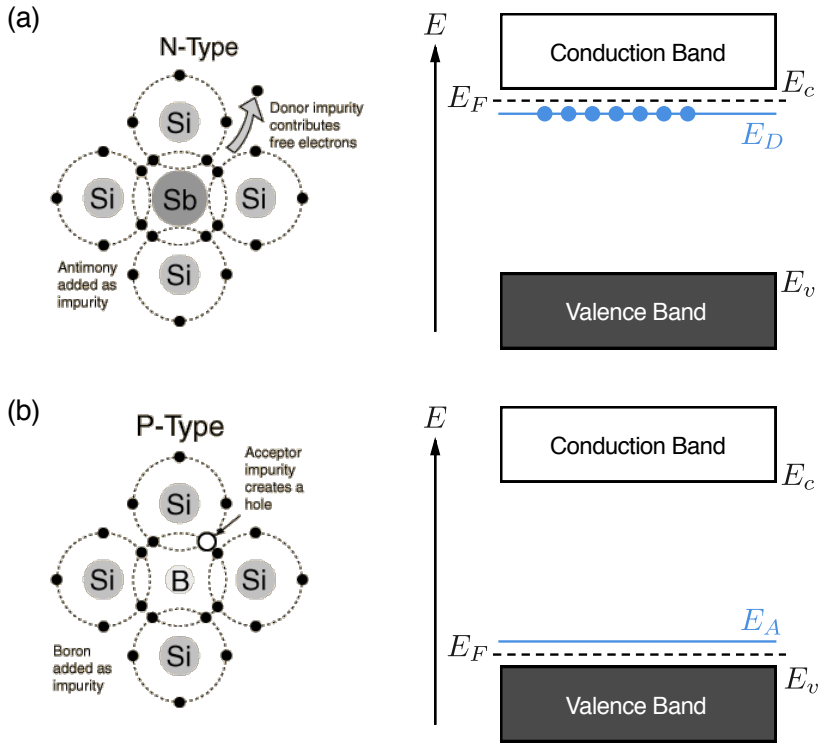


Figure 2: Different types of semiconductors with its flat band diagram. (a) n-type; (b) p-type.

## 2 Interface Effect

Consider a junction made by two pieces of materials.

- Homojunction. The two materials have the same band structure. For example, p-Si and n-Si.
- Heterojunction. The two materials have different band structures. For example,
  - Semiconductor-semiconductor (S-S): GaAs and AlAs;
  - Metal-semiconductor (M-S): Ag and n-Si;
  - ...

All these junctions play a vital role in lasers and transistors. Indeed, they lay the foundation for the modern information society. In the following, we will consider the most simple homojunction: the p-n junction formed between p-Si and n-Si. The underlying physics for more complicated heterojunctions are essentially the same.

### 2.1 p-n Junction with Zero Bias

Let us first consider p-Si and n-Si before touching each other.

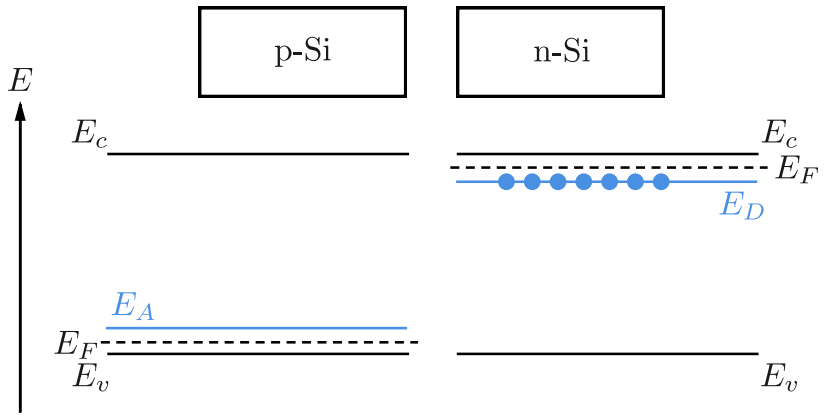


Figure 3: p-Si and n-Si before touching each other with their flat band diagram.

The Fermi energies in the two semiconductors are not equal. Therefore, when they contact, the chemical potential difference will induce electron diffusion from n-Si to p-Si and hole diffusion from p-Si to n-Si. Near the interface, some diffused electrons and holes recombine to become charge neutral. Therefore, there will be a thin layer where there is hardly any mobile charge carriers. This is called the "depletion region" in the p-n junction. Because the electrons diffuse out of n-Si, net positive charges will accumulate in the n-Si side of the interface. Similarly, there will be net negative charges in the p-Si side of the interface. These net charges result in a built-in electric field near the interface, pointing from n-Si to p-Si. Due to this electric field, electrons drift back into the n-Si and holes drift back into the p-Si. In other words, the built-in voltage is an energy barrier blocking the electron diffusion from n-Si to p-Si, and hole diffusion from p-Si to n-Si. The final thermal equilibrium state is obtained due to the balance of the diffusion and drift motion. This process is schematically demonstrated in the following figure:

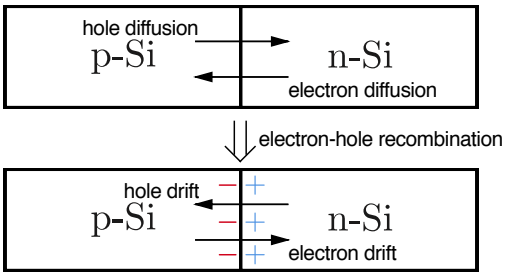


Figure 4: Diffusion and drift motion of electrons and holes in the p-n junction.

The equilibrium configuration of charge distribution, electric field distribution, electric potential and the local band structure are shown in the following figure. From a band structure point of view, the built-in electric field will bend the bands so that the Fermi energies are equal.

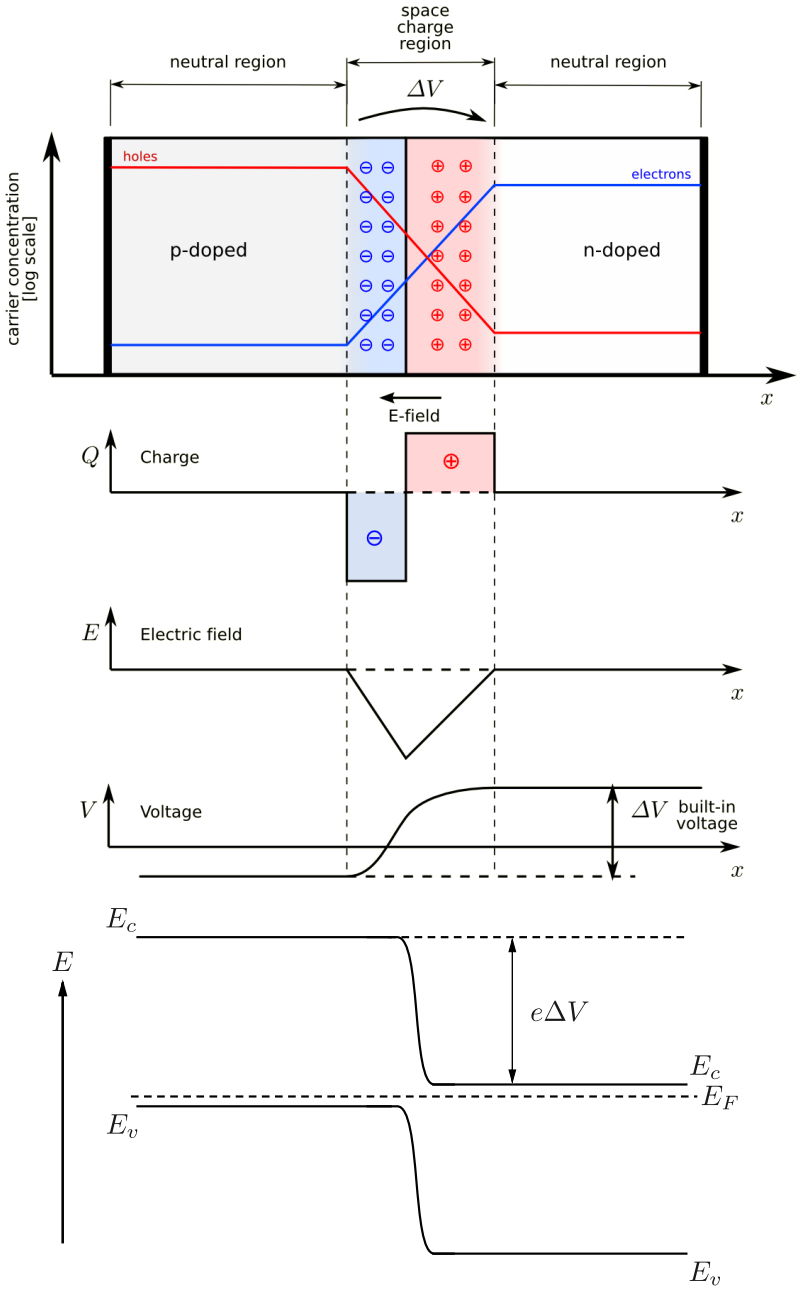


Figure 5: The equilibrium configuration of charge distribution, electric field distribution, electric potential and local band structure in a p-n junction.

Let us try to relate the electron and hole density in the junction to the built-in voltage.

As discussed above, the current comes from the diffusion $J_{\text{dif}} = -D\nabla n$ (Fick's law) and drift motion $J_{\text{dri}} = \sigma E$ (Ohm's law) of electrons and holes. At thermal equilibrium, there is no net charge current. Let us first consider the current of holes. For simplicity, we consider a one-dimensional model.

$$J_h = \sigma E - eD_h \frac{dp}{dx} = ep\mu_h E - eD_h \frac{dp}{dx} = 0, \tag{6}$$

where we $p(x)$ is the hole density. $\mu_h$ and $D_h$ are the mobility and the diffusion constant for the hole. Note that the carrier mobility $\mu$ for drift velocity $v$ is defined as $v = \mu E$, which characterizes how quickly the carrier can move through a metal or semiconductor, when pulled by an electric field.

We first prove the so called Einstein relation. Because the doped hole density is very small and the temperature is not very high, the probability that a state in the valence band is occupied by a hole is usually much smaller than 1. In this case, the Pauli exclusion principle is not effective and the occupation of holes satisfies Maxwell-Boltzmann relation:

$$f_h(E) = 1 - f_e(E) = \frac{1}{\exp\left(-\frac{E-E_F}{k_BT}\right) + 1} \sim \exp\left(\frac{E - E_F}{k_BT}\right). \tag{7}$$

Therefore, the local density of states of holes is

$$p(x) = A \exp\left(-\frac{eV(x)}{k_BT}\right). \tag{8}$$

Then Eq. (6) can be rewritten as

$$0 = p\mu_h E - D_h \frac{dp}{dx} = p\mu_h E - D_h \frac{dp}{dV}\frac{dV}{dx} = \left(\mu_h - \frac{eD_h}{k_BT}\right) \Rightarrow \frac{D_h}{\mu_h} = \frac{k_BT}{e}. \tag{9}$$

We then integral over $x$ through the depletion region in Eq. (6):

$$\Delta V = -\int E(x)dx = -\frac{D_h}{\mu_h}\int \frac{1}{p}\frac{dp}{dx}dx = -\frac{D_h}{\mu_h}\ln\frac{p_n^0}{p_p^0} = -\frac{k_BT}{e}\ln\frac{p_n^0}{p_p^0}. \tag{10}$$

It is instructive to rewrite it into the following form:

$$p_n^0 = p_p^0 e^{-e\Delta V/(k_BT)}. \tag{11}$$

This equation is often called "Boltzmann equation".

A parallel analysis can be made for electrons. Denote the electron density as $n(x)$. We have

$$n_p^0 = n_n^0 e^{-e\Delta V/(k_BT)}. \tag{12}$$

## 2.2 p-n Junction with Bias

The Boltzmann relation relates the equilibrium carrier density with the built-in potential. We are now going to discuss the non-equilibrium case.

In the non-equilibrium case, we are mainly interested in the minority carriers, because the number variation of majority carriers are relatively too few to be significant. That is to say, we are interested in electrons in p-Si and holes in n-Si.

When the density of minority carriers somehow becomes non-equilibrium, the system tends to relax to equilibrium. In fact, the thermal equilibrium is a dynamical equilibrium. Every instant, there is equal number particle-hole pair generation and recombination. In the non-equilibrium case, suppose there are more holes than the equilibrium. Denote the excessive hole density as $\Delta p \equiv p - p^0$. The recombination rate is $\Delta p/\tau$, where $\tau$ is the electron-hole pair lifetime.

Apart from the electron-hole recombination, the inhomogeneity of carrier density will also drive a diffusion within the semiconductor, which is governed by Fick's law. (In principle, there may also be drift motion if there is electric field. However, note the Einstein relation $D_h/\mu_h = k_BT/e$. At high temperature the diffusion dominates over drift motion. ) Therefore, the continuity equation of the carrier density becomes:

$$\frac{\partial p}{\partial t} + \nabla \cdot \mathbf{J_h} = -\frac{\Delta p}{\tau_h}. \tag{13}$$

Consider the steady state,

$$\frac{d}{dx}\left(D_h \frac{dp}{dx}\right) - \frac{\Delta p}{\tau_h} = 0. \tag{14}$$

The solution is

$$\Delta p = Ae^{-x/L_h} + Be^{x/L_h}, \tag{15}$$

where $L_h = \sqrt{D_h \tau_h}$ is the diffusion length. With boundary condition $\Delta p(x = 0) = P$ (excessive holes keep produced at $x = 0$) and $\Delta p(x = +\infty) = 0$, we have $\Delta p = Pe^{-x/L_h}$.

Now we add bias voltage to the p-n junction. The density distribution becomes non-equilibrium. There are two scenarios:

- Forward bias. Positive voltage to p-Si. The bias potential cancels the built-in potential.
- Reverse bias. Positive voltage to n-Si. The bias potential enhances the built-in potential.

Let us first consider the forward bias and compute the density of minority carriers. The build-in potential $\Delta V$ is now $\Delta V - V$, where $V$ is the biased potential. In this way, the potential barrier for the diffusion is lowered and the holes will diffuse back into n-Si. The hole density in n-Si is approximately

$$p_n = p_p^0 e^{-e(\Delta V - V)/(k_B T)}. \tag{16}$$

The excessive density is

$$\Delta p_n = p_n - p_n^0 = p_n^0 (e^{eV/(k_B T)} - 1). \tag{17}$$

The depletion region is usually very narrow. We can approximately fix it to be the boundary condition for the excessive hole diffusion at $x = 0$, which is the junction interface. Insert Eq. (15), we have

$$\Delta p_n = p_n^0 (e^{eV/(k_B T)} - 1) e^{-x/L_h}, \tag{18}$$

where $x > 0$ represents the n-Si. The diffusion current is then

$$J_h = -eD_h \frac{d\Delta p_n}{dx} = \frac{eD_h}{L_h} p_n^0 (e^{eV/(k_B T)} - 1) e^{-x/L_h}. \tag{19}$$

The diffusion current of electrons in p-Si can be computed similarly:

$$J_e = eD_h \frac{d\Delta n_p}{dx} = \frac{eD_e}{L_e} n_p^0 (e^{eV/(k_B T)} - 1) e^{x/L_h}, \tag{20}$$

where $x < 0$. Therefore, the total current at the interface $x = 0$ is

$$J = J_h + J_e = \left(\frac{eD_h}{L_h} + \frac{eD_e}{L_e}\right)(e^{eV/(k_B T)} - 1) \equiv J_0 (e^{eV/(k_B T)} - 1). \tag{21}$$

It is not hard to show for reverse bias, the total current takes the same form while just reversing the sign of $V$. This is the current-voltage characteristics for a diode:

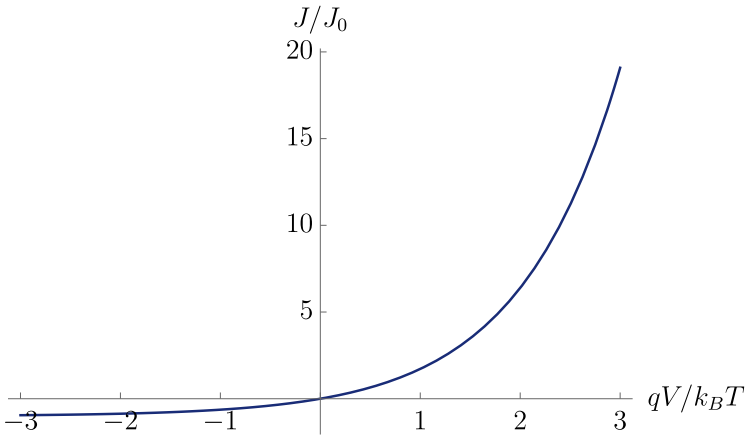


Figure 6: Current-voltage characteristics for a diode.

Diodes primarily conduct in one direction. This property can be used to, say, implement logic circuits.

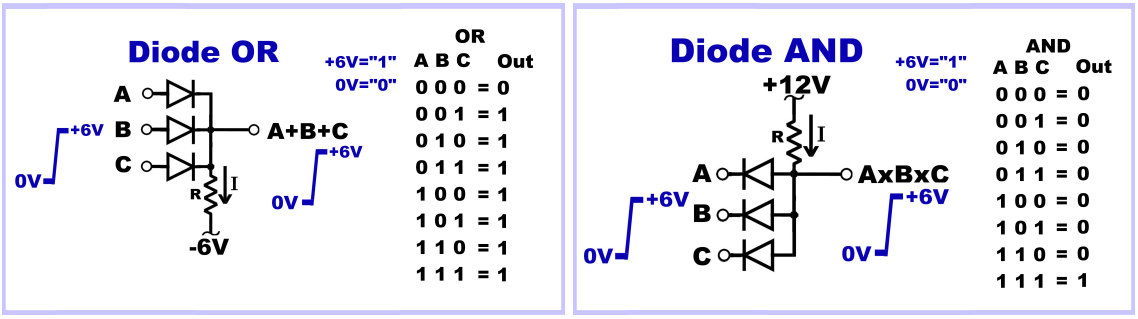


Figure 7: Diode implementation of OR and AND gates.

Note that while diode logic has the advantage of simplicity, the lack of an amplifying stage in each gate limits its application.

- Not all logical functions can be implemented in diode logic alone; only the non-inverting logical AND and logical OR functions can be realized by diode gates.
- If several diode logic gates are cascaded, the voltage levels at each stage are significantly changed, so one-stage is normally used.

In modern computers, transistors are used to form logic gates. Typical transistors involve metal-oxide-semiconductor junction (MOS). The analysis of electronic properties of transistors are similar to the p-n junction.