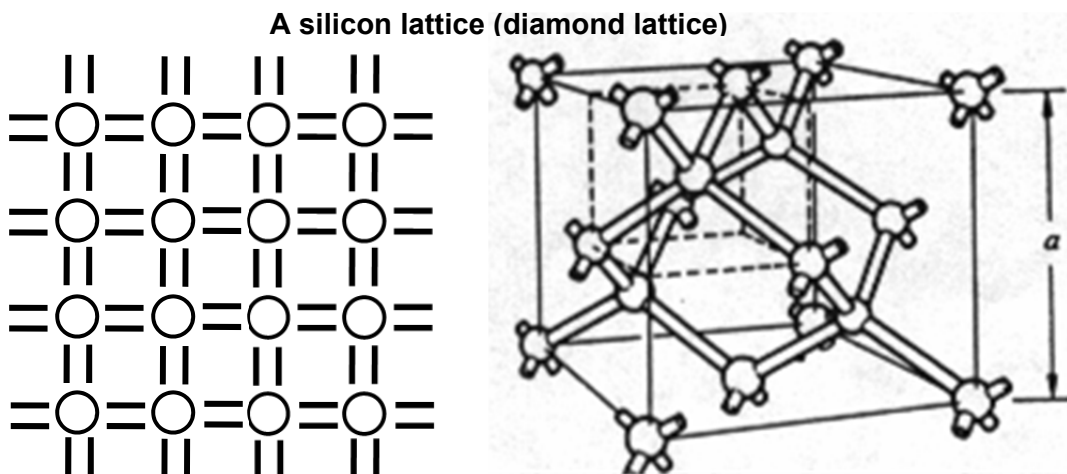# Chapter 1

Review of Basic Semiconductor Physics

1.1 Semiconductors

This review is not meant to teach you semiconductor physics–only to refresh your memory. Most semiconductors are formed from elements from groups II, III, VI, V, VI of the periodic table. The most commonly used semiconductor is silicon or Si. In a Si crystal each Si atom forms a covalent bond with 4 other Si atoms. Si has 4 electrons in its valence (or outer most shell) and therefore it can bond with 4 other Si atoms.

A cartoon depiction of Si crystal is then as shown in the Figure.

**A silicon lattice (diamond lattice)**

In a Si crystal each Si atom bonds with 4 other Si atoms in a tetrahedral geometry, as shown. This structure is called a "diamond Lattice" (since diamond crystals consisting of C atoms also have the same structure). The diamond lattice is essentially an FCC lattice (face centered cubic) with a single-atom basis. The lattice constant 'a' is also shown in the figure. Note that 'a' is not the actual distance between the nearest Si atoms. 'a' is the length of one side of the diamond unit cell (not the wigner-seitz cell) that has the cubic symmetry. Semiconductors are also formed by combining elements from group III and group V of the periodic table. This is possible since group III elements have 3 electrons in their outer most shell and group V elements have 5 electrons in their outermost shell. So a III-V covalent bond is possible. Most common III-V semiconductors are GaAs and InP .

A GaAs lattice (zincblende lattice)
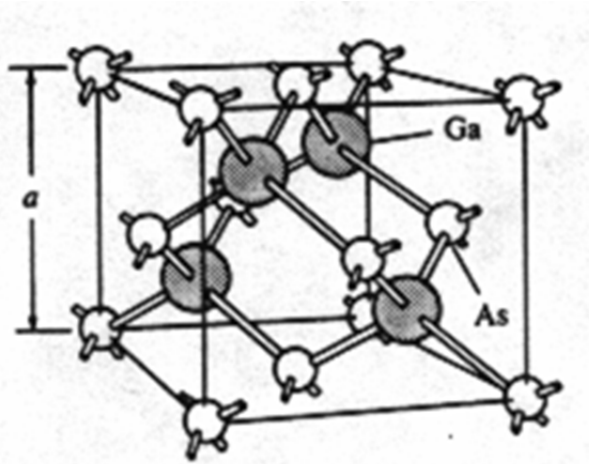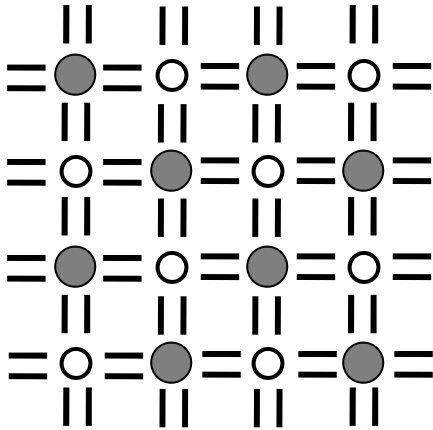
Each Ga atom is surrounded by 4 As atoms and each As atom is surrounded by 4 Ga atoms in a tetrahederal geometry. GaAs lattice is an example of "zinc blende lattice". The difference between zinc blende and diamond lattices is that in diamond lattice all atoms are the same. InP also has a zinc blende lattice. GaAs and InP are examples of "compound semiconductors". Si, C, and Ge are examples of "elemental semiconductors". Not all compound semiconductors have the zinc blende lattice. For example, III-Nitrides (e.g GaN , AlN, InN) can also have the wurtzite lattice structure show below.

A GaN lattice (wurtzite lattice)
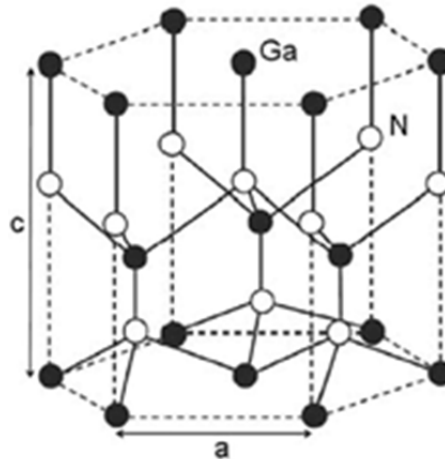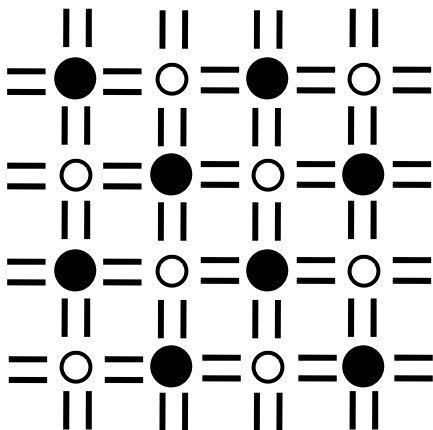
Just like the zinc blende lattice is a FCC lattice with a single-atom basis, the wurtzite lattice is a HCP lattice (hexagonal close packed) with a single-atom basis. For ideal HCP lattice $\frac{c}{a} = \sqrt{\frac{8}{3}}$. There is one thing common in zinc blende and wurtzite lattices; both have tetrahederal coordination.

Group II elements and group VI elements also combine to give II-VI compound semiconductor like ZnSe, CdTe, CdSe, ZnO etc. Most of these have zine blende or wurtzite lattices (but some do have "rock salt" lattic structures). Most of the IV–VI semiconductors (e.g PbS, PbSe, PbTe) called "lead salts" have the "rock salt" structure (similar to a NaCl crystal).

1.2 Semiconductor Bandstructure

In a solid the electronic energy levels are obtained by solving the Schrodinger Equation

$$\left[\frac{-\hbar^2}{2m} \nabla^2 + V(\vec{r}) \right] \psi(\vec{r}) = E\,\psi(\vec{r}) \qquad (1)$$

where $V(\vec{r})$ is the periodic potential from the atoms sitting on the lattice sites. The solutions of (1) can be written as,

$$\psi(r) = \psi_{n,\vec{k}}(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} u_{n,\vec{k}}(\vec{r})$$

The eigenfunctions $\psi_{n,\vec{k}}(\vec{r})$ are called Bloch functions and satisfy,

$$\left[\frac{-\hbar^2}{2m} \nabla^2 + V(\vec{r}) \right] \psi_{n,\vec{k}}(\vec{r}) = E_n(\vec{k}) \psi_{n,\vec{k}}(\vec{r})$$

The vector $\vec{k}$ can take values belonging to the first Brillouin zone (FBZ) and 'n' takes integral values. Therefore, all the possible energy levels of the solid can be labeled by the set of values $\{n,\vec{k}\}$. If one plots $E_n(\vec{k})$ as a function of $\vec{k}$ for different integral values of n one obtains the bandstructure of the solid. The first Brillouin zone (FBZ) corresponding to a FCC lattice (or a diamond or a zinc blende lattice) is shown below.

FBZ of an FCC lattice
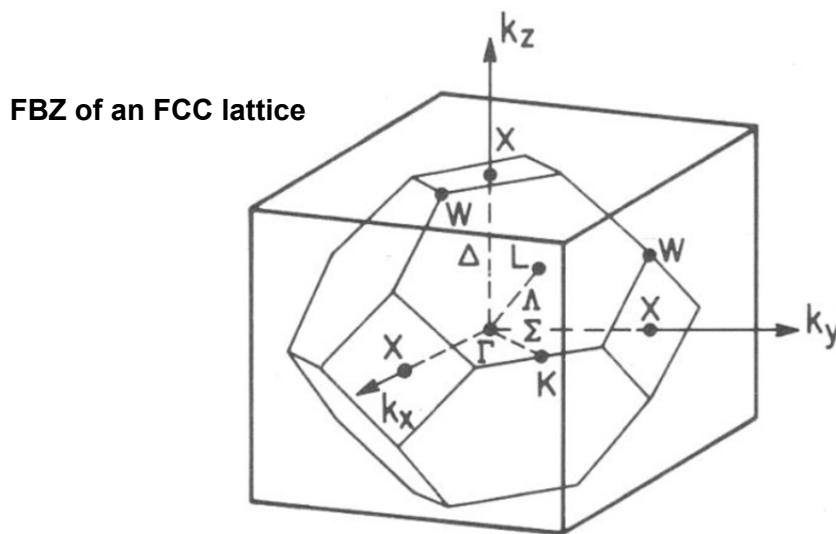
The bandstructures of Si, Ge, and $GaAs$ are shown below.

Germanium bandstructure   **Silicon bandstructure**   **GaAs bandstructure**
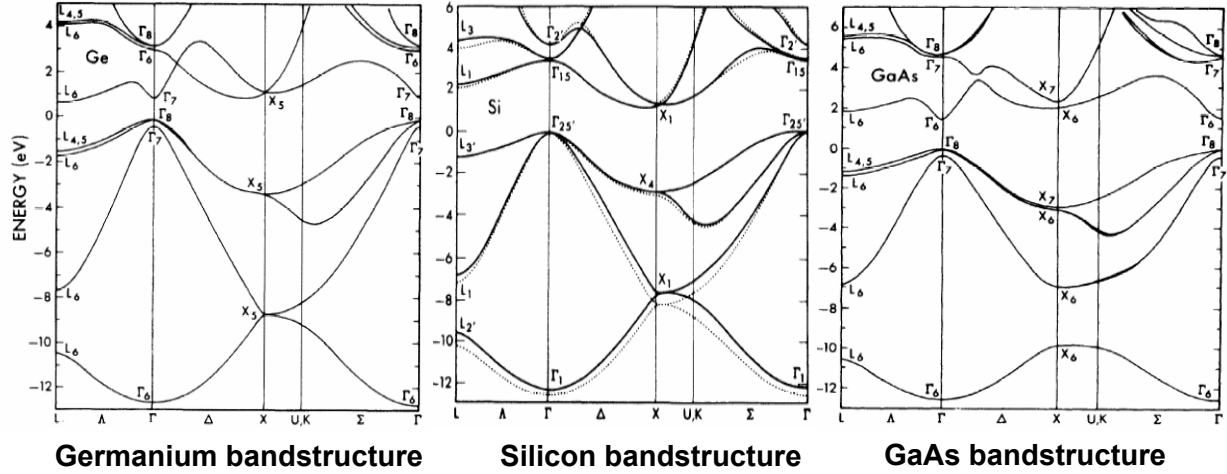
Only few chosen bands are shown. A particular feature of all semiconductor is that electrons in semiconductors fill all the low lying energy bands (called the valence bands). There are four valence bands, but only the highest three are shown in the figure. The highest energy in the valence bands is denoted by E_v. In pure semiconductors the conduction bands are all empty on electrons. The lowest energy in the conduction bands is denoted by E_c. There are also four conduction bands and all four are shown in the figure. The difference $E_c - E_v = E_g$ is called the band gap of the semiconductor.

Near the bottom of the lowest conduction band and the top of the highest valence band one may Taylor expand the energy $E_n(\vec{k})$. Assuming isotropic parabolic bands, conduction band dispersion near the band bottom can be written as,

$$E_c(\vec{k}) = E_c + \frac{\hbar^2}{2m_e}(\vec{k} - \vec{K}_c) \cdot (\vec{k} - \vec{K}_c)$$

And for the valence band one can write,

$$E_v(\vec{k}) = E_v - \frac{\hbar^2}{2m_h}(\vec{k} - \vec{K}_v) \cdot (\vec{k} - \vec{K}_v)$$

where m_e and m_h are electron and hole effective masses and the vectors $\vec{K}_c$ and $\vec{K}_v$ are the locations in k-space of conduction band minimum and valence band maximum. $\vec{K}_v = 0$ for all semiconductors that we will consider. $\vec{K}_c = 0$ for most III-V and II-VI semiconductors. Semiconductors for which $\vec{K}_c = \vec{K}_v$ are called "direct gap" or just "direct (e.g. $GaAs$, InP, GaN, $ZnSe$, $CdSe$, ZnO). Semiconductors for which $\vec{K}_c \neq \vec{K}_v$ are called "indirect gap" or just "indirect" (e.g. Si, Ge, C, SiC, GaP, $AlAs$). As we will see later in the course, all optically active semiconductors are direct gap.

When $\vec{K}_c = \vec{K}_v = 0$, and assuming isotropic parabolic bands,

Conduction band: $E_c(\vec{k}) = E_c + \dfrac{\hbar^2 k^2}{2m_e}$

Valence band: $E_v(\vec{k}) = E_v - \dfrac{\hbar^2 k^2}{2m_h}$

More generally, one can write for the conductions band with minimum at $\vec{K}_c$ (assuming parabolic bands),

$$E_c(\vec{k}) = E_c + \frac{\hbar^2}{2}(\vec{k} - \vec{K}_c) \cdot M_e^{-1} \cdot (\vec{k} - \vec{K}_c)$$

where M_e is the effective mass matrix,

$$M_e = \begin{bmatrix} m_{xx} & m_{xy} & m_{xz} \\ m_{yx} & m_{yy} & m_{yz} \\ m_{zx} & m_{zy} & m_{zz} \end{bmatrix}$$

Physical considerations demand that M_e be symmetric. Similarly, for the valence band one get,

$$E_v(\vec{k}) = E_v - \frac{\hbar^2}{2}(\vec{k} - \vec{K}_v) \cdot M_h^{-1} \cdot (\vec{k} - \vec{K}_v)$$

1.3 Counting Electronic States in Semiconductors

In a solid of volume V, the number of energy levels is one band in volume $d^3\vec{k}$ of the FBZ is

$2 \times V \dfrac{d^3\vec{k}}{(2\pi)^3}$ (The multiplies 2 accounts for the two spin states). So all summations of the from $\displaystyle\sum_{\vec{k} \in \text{FBZ}} \cdots$

where values of $\vec{k}$ are restricted to the FBZ can be replaced by the integral,

$$\sum_{\vec{k}} \cdots \rightarrow V \int_{\text{FBZ}} \frac{d^3\vec{k}}{(2\pi)^3} \cdots$$

The number of energy levels per band in a crystal of volume V is given by,

$$2 \sum_{\vec{k} \in \text{FBZ}} = 2V \int_{\text{FBZ}} \frac{d^3 k}{(2\pi)^3} = \frac{2V}{(2\pi)^3} \times \text{volume of FBZ}$$

But,

$$\text{volume of FBZ} = \frac{(2\pi)^3}{\text{volume of the primitive unit cell}}$$

The number of energy levels per band in a crystal of volume V is then given by,
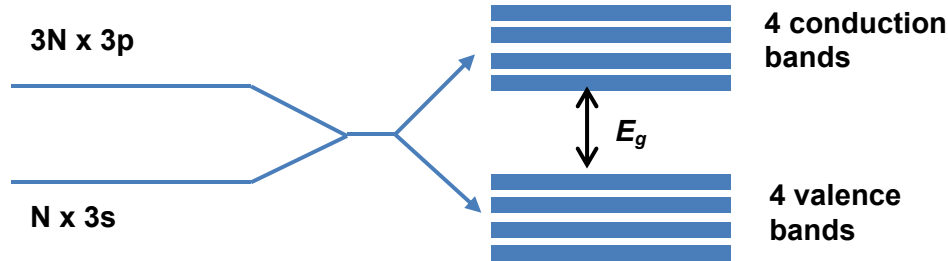
$$2 \frac{V}{\text{volume of the primitive cell}} = 2\left\{\text{number of primitive cells in the crystal}\right\}$$

Therefore each primitive cell contributes two states or energy levels to each band.

1.4 Linear Combination of Atomic Orbitals Approach to Energy Bands

Linear combination of atomic orbitals is another way to understand energy band formation in semiconductors. In semiconductors, the atomic states of the outermost shell (e.g the single $3s$ and the three $3p$ in a Si atom, and the single $4s$ and the three $4p$ in a Ga atom and the same in a As atom) combine or hybridize with the states of the neighboring atoms to result in the four valence bands and the four conditions bands, as shown in the figure below.

Energy band formation in a crystal of N Silicon atoms

In this hybridization process the total number of energy levels of all the atoms is conserved. Suppose N Si atoms from a crystal then the total number of energy levels before hybridization is $2 \times 4N$. Now lets find the total number of energy levels in the resulting crystal. As found earlier, the number of energy levels per band in a crystal of volume V is then given by,

$$2 \frac{V}{\text{volume of the primitive cell}} = 2 \left\{ \text{number of primitive cells in the crystal} \right\}$$

So the total number of energy levels in eight bands (four conduction bands and four valence bands) is,

$$2 \times \left\{ \begin{array}{l} \text{\# of primitive cells} \\ \text{in the crystal} \end{array} \right\} \times \left\{ \begin{array}{l} \text{\# of} \\ \text{bands} \end{array} \right\}$$

Since each Si primitive cell has two Si atoms (diamond lattice is an FCC lattice with a two-atom basis) we get,

$$2 \times \left\{ \begin{array}{l} \text{\# of primitive cells} \\ \text{in the crystal} \end{array} \right\} \times \left\{ \begin{array}{l} \text{\# of} \\ \text{bands} \end{array} \right\} = 2 \times \frac{N}{2} \times 8 = 4N$$

And we get the same answer as before.

1.5 Properties of Semiconductor Alloys

Other than elemental and compound semiconductors, semiconductor alloys also exist and are extremely useful. For example $Si_{1-x} Ge_x$ is a binary allows of Si and Ge and the lattice of $Si_{1-x} Ge_x$ consists of x fraction of Ge atoms and $(1-x)$ fraction of Si atoms arranged randomly. On the other hand SiC is a compound semiconductor with zinc blende Lattice.

Binary Alloys:
$Si_{1-x} Ge \Rightarrow$ Alloy of two elemental semiconductors consists of x fraction of Ge atoms and $(1-x)$ fraction of Si atoms

Ternary Alloys:

$Al_x Ga_{1-x}As$ $\Rightarrow$ Alloy of two compound semiconductors $AlAs$ and $GaAs$ with x fraction of $AlAs$ and $(1-x)$ fraction of $GaAs$, also written as: $x\,AlAs(1-x)GaAs$

$in_x Ga_{1-x}As$ $\Rightarrow$ Alloy of two compound semiconductors $inAs$ and $GaAs$ with x fraction of $inAs$ and $(1-x)$ fraction of $GaAs$ also written as: $x\,inAs(1-x)GaAs$

Quaternary Alloys:
$In_{1-x}Ga_xAs_yP_{1-y} \Rightarrow (1-x)y\,InAs(1-x)(1-y)InP\,x(1-y)GaP\,xy\,GaAs$

Here we will discuss what happens when we make an alloy of two semiconductors (whether elemental or compound) and how do the properties of the alloy differ from those of its constituents.

1.5.1 Vegards law:
Vegard's law says the lattice constant of an alloy is a weighted sum of the lattice constants of each of its constituents, and the weight aligned to each constituent is equal to its fraction is the alloy.

Example: For binary alloys like Si_xGe_{1-x}
$$a[Si_xGe_{1-x}] = x\,a[Si] + (1-x)a[Ge]$$

Example: For a ternary alloy like $Al_xGa_{1-x}As$
$$a[Al_xGa_{1-x}As] = x\,a[AlAs] + (1-x)a[GaAs]$$

Example: For a quaternary alloy like $In_{1-x}Ga_xAs_yP_{1-y}$
$$a[In_{1-x}Ga_xAs_yP_{1-y}] = (1-x)Y\,a[InAs] + (1-x)(1-y)a[InP]$$
$$+ x(1-y)a[GaP] + xy\,a[GaAs]$$

What about other material parameter like dielectric constants, effective masses, band gaps, etc? The linear rule rays that you average all quantities just like the lattice constant is averaged according to the Vegards law. If you don't know any better, the linear rule law can be a good first approximation. But it does not always work very well for quantities other than the lattice constant. For example, the band gap of $Al_xGa_{1-x}As$ at the Γ point is given more accurately by:
$$E_g(\Gamma)[Al_xGa_{1-x}As] = a + bx + c\,x^2$$
where,
$$b = 1.087 \text{ eV}$$
$$c = 0.438 \text{ eV} \left\{ at\,300K \right\}$$
$$a = 1.42 \text{ eV}$$
The values of a, b, c are determined experimentally. Similarly, for $Ga_xIn_{1-x}As$ one has,
$$E_g(\Gamma)[Ga_x\,In_{1-x}As] = a + bx + cx^2$$
where
$$b = 0.7 \text{ eV}$$
$$c = 0.4 \text{ eV} \left\{ at\,300K \right\}$$
$$a = 0.324 \text{ eV}$$
In designing alloys one has to be careful. $GaAs$ is a direct gap semiconductor (conduction band minimum and valence band maximum occur at the Γ point). $AlAs$ is as indirect gap semiconductor

(conduction band minimum is at the X point and valence band maximum is at the Γ point). The Alloy $Al_xGa_{1-x}As$ must therefore be direct gap for small values of x and indirect gap for large values of x. At some value of x the transition from direct to indirect gap occurs. This is shown in the Figure. For the effective masses, the linear rule works better if the inverse effective masses are averaged, and provided the effective masses refer to the same point in k-space in the FBZ. For Eeample, the electron effective mass m_e in $Ga_xIn_{1-x}As$ is,

$$\frac{1}{m_e[Ga_xIn_{1-x}As]} = \frac{x}{m_e[GaAs]} + \frac{1-x}{m_e[InAs]}$$

Here, all effective masses are at the Γ point in the conduction band. For the dielectric constants, and refractive indices, the linear rule can be hopelessly wrong, especially if the wavelength at which these are desired is close to the bandgap of any one of the constituents in the alloy. As the course proceeds we will consider many different examples.

Bandgaps for $Al_xGa_{1-x}As$
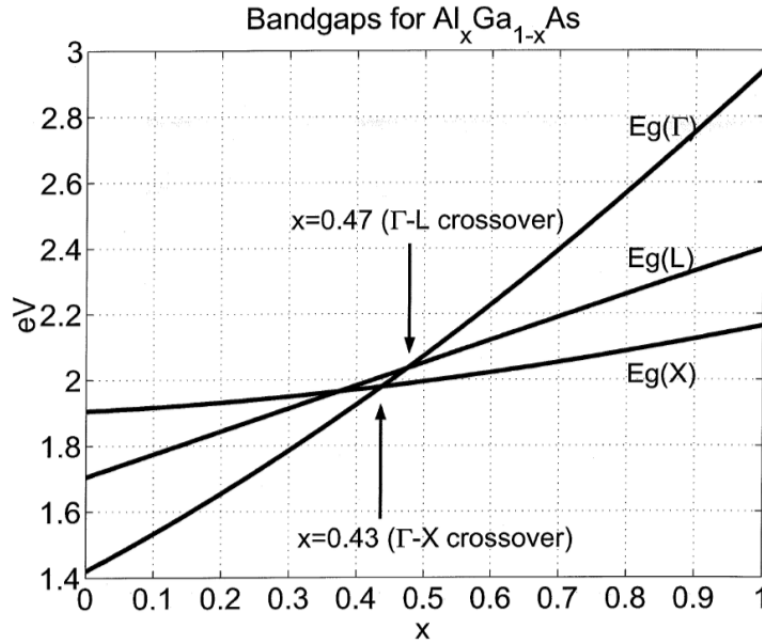
1.6 Density of States in Energy

We know that the number of allowed energy states in volume $d^3\vec{k}$ of the reciprocal space is $2 \times V \dfrac{d^3\vec{k}}{(2\pi)^3}$ in one energy band. Suppose we want to find out the number of allowed energy states in an interval dE of energy in a band. Suppose a simple isotropic, parabolic conduction band with energy dispersion given by,

$$E(\vec{k}) = E_c + \frac{\hbar^2 k^2}{2m_e} \quad \{k^2 = \vec{k} \cdot \vec{k}\}$$

The equal energy surfaces in reciprocal space or k-space are spherical shells (i.e. all states on a shell have the same energy since the $E(\vec{k})$ vs $\vec{k}$ relation is isotropic). Suppose the thickness of the shell is dk. Then the number of states in the shell of radius k is,

$$2 \times V \times \frac{4\pi k^2}{(2\pi)^3} dk$$

Since,

$$E(\vec{k}) = E_c + \frac{\hbar^2 k^2}{2m_e}$$

$$\Rightarrow dE = \frac{\hbar^2}{m_e} k \, dk$$

A spherical shell of thickness dk in k-space corresponds to interval dE in energy space. The number of states in the shell is,

$$2 \times V \times \frac{4\pi k^2}{(2\pi)^3} dk$$

$$= \frac{V}{\pi^2}\left((E - E_c)\frac{2m_e}{\hbar^2}\right)^{\frac{1}{2}} \frac{m_e}{\hbar^2} dE$$

$$= V \frac{\sqrt{2}}{\pi^2}\left(\frac{m_e}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E - E_c} \, dE$$

The number of states in energy interval dE is therefore,

$$V \frac{\sqrt{2}}{\pi^2}\left(\frac{m_e}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E - E_c} \, dE = V \, g_c(E) dE$$

where $g_c(E)$ is called the conduction band density of states and represents the number of states per unit energy interval in the band per unit crystal volume.

For conduction band:

$$g_c(E) = \begin{cases} \dfrac{\sqrt{2}}{\pi^2}\left(\dfrac{m_e}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E - Ec} & \text{for} \quad E \geq E_c \\ 0 & \text{for} \quad E < E_c \end{cases}$$

For valence band:
Similarly, for valence band with an isotropic parabolic effective hole mass m_h,

$$g_v(E) = \begin{cases} \dfrac{\sqrt{2}}{\pi^2}\left(\dfrac{m_h}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E_v - E} & \text{for} \quad E \leq E_v \\ 0 & \text{for} \quad E > E_v \end{cases}$$

Now consider a more complicated example of a non-isotropic but parabolic conductions band with dispersion given by,

$$E(\vec{k}) = E_c + \frac{\hbar^2 k_x^2}{2m_x} + \frac{\hbar^2 k_y^2}{2m_y} + \frac{\hbar^2 k_z^2}{2m_z}$$

Now how do we find $g_c(E)$? Equal energy surfaces are now elliptical shells. Define,

$$q_x = k_x\left(\frac{m}{m_x}\right)^{\frac{1}{2}} \quad q_y = k_y\left(\frac{m}{m_y}\right)^{\frac{1}{2}} \quad q_z = k_z\left(\frac{m}{m_2}\right)^{\frac{1}{2}}$$

$$\Rightarrow E(\vec{k}) \to E(\vec{q}) = E_c + \frac{\hbar^2}{2m}\left(q_x^2 + q_y^2 + q_z^2\right) = E_c + \frac{\hbar^2 q^2}{2m}$$

Equal energy surfaces in q-space are now spherical shells. We have simply scaled the coordinates. Volume element $\dfrac{d^3\vec{q}}{(2\pi)^3}$ is q-space corresponds to volume element $\sqrt{\dfrac{m}{m_x m_y m_z}}\,\dfrac{d^3\vec{k}}{(2\pi)^3}$ in k-space.

Number of states in volume $d^3\vec{k}$ in k-space is,

$$2\times V \times \frac{d^3\vec{k}}{(2\pi)^3} = 2\times V \times \frac{\sqrt{m_x m_y m_z}}{m^{3/2}}\frac{d^3\vec{q}}{(2\pi)^3}$$

Number of states in volume $d^3\vec{q}$ in q-space is,

$$2\times V \times \frac{\sqrt{m_x m_y m_z}}{m^{3/2}}\frac{d^3\vec{q}}{(2\pi)^3}$$

Number of states in spherical shell of radius q in q-space is,

$$2\times V \times \frac{\sqrt{m_x m_y m_z}}{m^{3/2}}\frac{4\pi q^2}{(2\pi)^3}\,dq = V\frac{\sqrt{2}}{\pi^2}\frac{\sqrt{m_x m_y m_z}}{\hbar^3}\sqrt{E-E_c}\,dE \qquad \left\{\text{since } E(\vec{q}) = E_c + \frac{\hbar^2 q^2}{2m}\right\}$$

Therefore, the density of states is,

$$g_c(E) = \frac{\sqrt{2}}{\pi^2}\frac{\sqrt{m_x m_y m_z}}{\hbar^3}\sqrt{E-E_c}$$

$$= \frac{\sqrt{2}}{\pi^2}\left(\frac{m_{de}}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{E-E_c}$$

where the density of states effective mass m_{de} for electrons is

$$m_{de} = (m_x m_y m_z)^{\frac{1}{3}}$$

We can write the conduction band density of states as follows,

$$g_c(E) = \begin{cases} \dfrac{\sqrt{2}}{\pi^2}\left(\dfrac{m_{de}}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{E-E_c} & \text{for} \quad E \geq E_c \\ 0 & \text{for} \quad E < E_c \end{cases}$$

1.7 Occupation Statistics

1.7.1 The Fermi-Dirac Distribution Function:
The probability of an electron occupying a state of energy E in the crystal is given by the Fermi-Dirac distribution function,

$$f(E) = \frac{1}{1 + e^{(E-E_f)/KT}}$$

where E_f is the Fermi level (or the chemical potential) and K is the Boltzmann constant.

Example:
Consider isotropic parabolic conduction band with the dispersion,

$$E(\vec{k}) = E_c + \frac{\hbar^2 k^2}{2m_e}$$

At zero temperature all the valence bands are occupied by electrons and all the conduction bands are empty. At any non-zero temperature electrons can be thermally excited from the valence band into the conduction band. We need to find the electron density at a non-zero temperature. The electron density can be written as,

$$n = 2 \times \int \frac{d^3 \vec{k}}{(2\pi)^3} \ f\left(E\left(\vec{k}\right)\right)$$

$$= \int_{-\infty}^{\infty} dE \ g_c(E) f(E) dE$$

$$= \int_{E_c}^{\infty} dE \ \frac{\sqrt{2}}{\pi^2}\left(\frac{m_e}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E - E_c} \ f(E)$$

When the Fermi level E_f is much below the conduction band edge, i.e. $E_c - E_f \gg KT$, then the Fermi distribution function can be approximated by an exponential,

$$f(E) = \frac{1}{1 + e^{(E-E_f)/KT}} \approx e^{-(E-E_f)/KT}$$

This approximation, called the Maxwell-Boltzmann approximation, does not always work. It only works when the electron density is very small. With this approximation we get,

$$n = \int_{E_c}^{\infty} dE \ \frac{\sqrt{2}}{\pi^2}\left(\frac{m_e}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E - E_c} \ e^{\frac{-(E-E_f)}{kT}}$$

$$= N_c \ e^{(E_f - E_c)/KT}$$

Where,

$$N_c = 2\left(\frac{m_e \ kT}{2\pi \ \hbar^2}\right)^{\frac{3}{2}}$$

is called the effective density of states of the conduction band. More generally (when $E_c - E_f \gg kT$ is not true) and the band is not isotropic (but is still parabolic),

$$n = \int_{E_c}^{\infty} dE \ \frac{\sqrt{2}}{\pi}\left(\frac{m_{de}}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E - E_c} \ \frac{1}{1 + e^{(E-E_f)/KT}}$$

$$= N_c \ \frac{2}{\sqrt{\pi}} F_{1/2}\left(\frac{E_f - E_c}{kT}\right) \qquad \left\{ N_c = 2\left(\frac{m_{de} \ kT}{2\pi \ \hbar^2}\right)^{\frac{3}{2}} \right.$$

where $F_{1/2}(x)$ is called the Fermi function and is defined as,,

$$F_{1/2}(x) = \int_0^{\infty} \frac{\sqrt{y}}{1 + e^{y-x}} \ dy \qquad \left\{ \approx \frac{\sqrt{\pi}}{2} e^x \ \text{ for } \ x \ll -1 \right\}$$

Example: We can also find the hole density in a semiconductor at non-zero temperature. The probability of a hole at energy E is given by $1 - f(E)$. Consider a parabolic and isotropic valence band given with energy dispersion given by,

$$E\left(\vec{k}\right) = E_v - \frac{\hbar^2 k^2}{2m_h}$$

The hole density in the valence band is,

$$p = 2\int \frac{d^3\vec{k}}{(2\pi)^3} \left(1 - f\left(E(\vec{k})\right)\right)$$

$$P = \int_{-\infty}^{\infty} dE \; g_v(E) \; (1 - f(E))$$

$$= \int_{-\infty}^{E_v} dE \; \frac{\sqrt{2}}{\pi}\left(\frac{m_h}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{E_v - E} \; (1 - f(E))$$

$$= N_v \; \frac{2}{\sqrt{\pi}} F_{1/2}\left(\frac{E_v - E_f}{kT}\right) \qquad \left\{ N_v = 2\left(\frac{m_h kT}{2\pi \hbar^2}\right)^{\frac{3}{2}} \right.$$

For $E_f - E_v \gg KT$, when Maxwell-Boltzmann approximation works,

$$p \approx N_v \; e^{(E_v - E_f)/KT}$$

For parabolic but not isotropic valence band N_v is given by,

$$N_v = 2\left(\frac{m_{dh} KT}{2\pi\hbar^2}\right)^{\frac{3}{2}}$$

where m_{dh} is the density of states effective mass for holes.

1.7.2 Intrinsic Semiconductors:

Semiconductors are either doped (n-type or p-type) or they are undoped (intrinsic). In intrinsic semiconductors the number of electrons and holes must be equal (i.e. $n = p$). Assume Maxwell-Boltzmann approximation.

$$n = p$$

$$\Rightarrow N_c e^{\frac{E_f - E_c}{KT}} = N_v e^{\frac{E_v - E_f}{KT}}$$

$$\Rightarrow E_f = \frac{E_c + E_v}{2} + \frac{KT}{2} \ln\left(\frac{N_v}{N_c}\right)$$

If $N_v \sim N_c$ then E_f is in the center of the band gap and therefore Maxwell-Boltzmann approximation works. Also,

$$np = N_c N_v e^{\frac{-(E_c - E_v)}{KT}} = N_c N_v e^{\frac{-E_g}{KT}}$$

Since,

$$n = p \Rightarrow n = p = \sqrt{N_c N_v} \, e^{\frac{-E_g}{2KT}} = n_i$$

The intrinsic electron and hole concentration is therefore related to the semiconductor bandgap.

1.7.3 Doped semiconductors:

Semiconductors can be doped n-type or p-type by introducing donor or acceptor impurities, respectively. Each donor atom contributes an energy state with energy E_d right below the conduction band minimum, as shown in the Figure. The electron in the donor atom can get enough thermal energy (at room
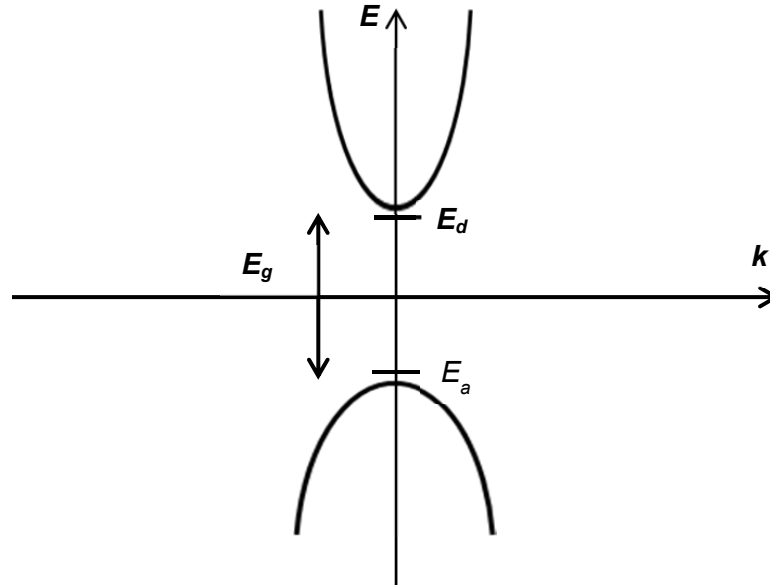
temperature) to get into the conduction band and move away leaving behind a positively charged ionized donor atom. Each donor atom can get ionized with a probability given by,

$$\frac{1}{1+g_d\, e^{(E_f-E_d)/KT}}$$

where $g_d \approx 2$. If the donor concentration is N_d, then the ionized donor concentration is,

$$N_d^+ = \frac{N_d}{1+g_d e^{(E_f-E_d)/KT}}$$

Similarly, an acceptor impurity atom contributes an energy state with energy E_a right below the valence band maximum, as shown in the Figure. And electron in the valence band can get enough thermal energy (at room temperature) to move up from the valence band into the acceptor energy level thereby negatively charging the acceptor impurity atom and leaving behind a hole in the valence band. Conversely, one can say that at room temperature the hole in the acceptor energy level can move down into the valence band and move away leaving behind a negatively charged ionized acceptor atom. Each donor impurity can get ionized with a probability given by,

$$\frac{1}{1+g_a e^{(E_a-E_f)/KT}}$$

Where E_a is the acceptor binding energy and $g_a \approx 2$. The ionized acceptor concentration is then,

$$N_a^- = \frac{N_a}{1+g_a e^{(E_a-E_f)/KT}}$$

Each ionized donor (acceptor) atom is singly positively (negatively) charged. Overall change neutrality implies,

$$q\left(p-n+N_d^+ - N_a^-\right) = 0$$

If Maxwell-Boltzmann statistics apply then,

$$np = n_i^2$$

$$\frac{n_i^2}{n} - n + N_d^+ - N_a^- = 0$$

$$n^2 - n\left(N_d^+ - N_a^-\right) - n_i^2 = 0$$

The solution is,

$$n = +\frac{\left(N_d^+ - N_a^-\right)}{2} + \sqrt{\frac{\left(N_d^+ - N_a^-\right)^2}{4} + n_i^2}$$

$$p = -\frac{\left(N_d^+ - N_a^-\right)}{2} + \sqrt{\frac{\left(N_d^+ - N_a^-\right)^2}{4} + n_i^2}$$

N-type Semiconductor:

$$N_d^+ \gg N_a^- \quad \text{and} \quad N_d^+ \gg n_i^2$$

$$\Rightarrow n \approx N_d^+ \quad \text{and} \quad p = \frac{n_i^2}{N_d^+} \ll n$$

In a n-type semiconductors since,

$$n = N_c \, e^{(E_f - E_c)/KT} = N_d^+$$

$$\Rightarrow E_f - E_c = kT \, \ln\left(\frac{N_d^+}{N_c}\right)$$

Therefore, E_f shifts close to the conductions band in an n-type semiconductor.

P-type Semiconductor:

$$N_a^- \gg N_d^+ \quad \text{and} \quad N_a^- \gg n_i^2$$

$$\Rightarrow p \approx N_a^- \quad \text{and} \quad n = \frac{n_i^2}{N_a^-} \ll p$$

In a p-type semiconductors since,

$$p = N_v e^{(E_v - E_f)/KT} = N_a^-$$

$$\Rightarrow E_v - E_f = KT \ln\left(\frac{N_a^-}{N_v}\right)$$

Therefore, E_f shifts close to the valence band in a p-type semiconductor.

1.7.4 Degenerate Doping:

When doping N_d (or N_a) is much larger than N_c (or N_v) the Maxwell-Boltzmann approximation does not work. In this case electron (or hole) density is found by solving the non-linear equations for the Fermi level,

$$n = N_c \frac{2}{\sqrt{\pi}} F_{\frac{1}{2}}\left(\frac{E_f - E_c}{KT}\right) = N_d^+ = \frac{N_d}{1 + g_d e^{(E_f - E_d)/KT}} \qquad (n\text{-type})$$

or

$$p = N_v \frac{2}{\sqrt{\pi}} F_{\frac{1}{2}}\left(\frac{E_v - E_f}{KT}\right) = N_a^- = \frac{N_a}{1 + g_a e^{(E_a - E_f)/KT}} \qquad (p\text{-type})$$

and then computing n (or p).

1.8 Electron and Hole Transport in Semiconductors

1.8.1 Electron and Hole Current Densities:

Electron current density $\vec{J}_e$ (units: Amps/cm^2) has drift and diffusive components,

$$\vec{J}_e(\vec{r}) = q n(\vec{r}) \mu_e \vec{E}(\vec{r}) + q D_e \nabla n(\vec{r})$$

The electron mobility is μ_e (units: cm^2/volt-sec) and the electron diffusion constant is D_e (units: cm^2/sec). When Maxwell-Boltzmann approximation applies, the mobility can be related to the diffusion constant by the Einstein relation,

$$\frac{\mu_e}{D_e} = \frac{q}{KT}$$

Similarly, the hole current density is,

$$\vec{J}_h(\vec{r}) = q p(\vec{r}) \mu_h \vec{E}(\vec{r}) - q D_h \nabla p(\vec{r})$$

And the corresponding Einstein relation is,

$$\frac{u_h}{D_h} = \frac{q}{KT}$$

1.8.2 Particle Number Conservation Equations:

Since electron number in the conduction band is conserved (unless electrons recombines with a hole) we have:

$$\frac{\partial n(\vec{r})}{\partial t} - \frac{1}{q} \nabla \cdot \vec{J}_e(\vec{r}) = G_e(\vec{r}) - R_e(\vec{r})$$

For holes we have,

$$\frac{\partial p(\vec{r})}{\partial t} + \frac{1}{q} \nabla \cdot \vec{J}_n(\vec{r}) = G_h(\vec{r}) - R_h(\vec{r})$$

For electron-hole generation or recombination,

$$G_h(\vec{r}) = G_e(\vec{r})$$

and,

$$R_h(\vec{r}) = R_e(\vec{r})$$

1.8.3 Poisson Equation:

The electric field inside a semiconductor is given by the Poisson equation,

$$\nabla \cdot \varepsilon(x) \vec{E}(\vec{r}) = q\left(p(\vec{r}) - n(\vec{r}) + N_d^+(\vec{r}) - N_a^-(\vec{r})\right)$$

1.8.4 Five Shockley Equations:

The following five equation, named after William Shockley, form the backbone of basic semiconductor device analysis,

$$\vec{J}_e(\vec{r}) = q\, n(\vec{r})\mu_e \vec{E}(\vec{r}) + q\, D_e \nabla n(\vec{r})$$

$$\vec{J}_h(\vec{r}) = q\, p(\vec{r})\mu_h \vec{E}(\vec{r}) - q\, D_h \nabla p(\vec{r})$$

$$\frac{\partial n(\vec{r})}{\partial t} - \frac{1}{q}\nabla \cdot \vec{J}_e(\vec{r}) = G_e(\vec{r}) - R_e(\vec{r})$$

$$\frac{\partial p(\vec{r})}{\partial t} + \frac{1}{q}\nabla \cdot \vec{J}_h(\vec{r}) = G_h(\vec{r}) - R_h(\vec{r})$$

$$\nabla \cdot \varepsilon(\vec{r})\vec{E}(\vec{r}) = q\big(p(\vec{r}) - n(\vec{r}) + N_d^+(\vec{r}) - N_a^-(\vec{r})\big)$$

When solved with proper boundary conditions, these equations determine the electron-hole dynamics in most semiconductor devices.

1.8.5 Electron-Hole Recombination-Generation and Minority Carrier Lifetimes:

There are different mechanisms by which electrons and holes recombine. We will go into the details later in the course. Here we give a simple qualitative expression for the case of minority carriers. In a p-type semiconductor, electrons are the minority carriers and holes are the majority carriers and $n \ll p$. In a n-type semiconductor, holes are the minority carriers and electrons are the majority carriers, and $p \ll n$. In equilibrium, generation and recombination rates are always equal. They differ only in non-equilibrium situations. We consider a non-equilibrium situation in a p-type semiconductor in which the electron and hole concentrations are slightly perturbed from their equilibrium values given by n_o and p_o, respectively ($n_o \approx n_i^2/p_o$). In this case, one can write the net recombination rate as,

$$R_e - G_e = \frac{n - n_o}{\tau_e}$$

where τ_e is the minority carrier (i.e. electron) lifetime. In equilibrium, $n = n_o$, and $R_e = G_e$, as it must be. If holes are the minority carriers, then one can write,

$$R_h - G_h = \frac{p - p_o}{\tau_h}$$

In equilibrium $p = p_o$ and $R_h = G_h$. In intrinsic semiconductors the expressions for recombination-generation are more complicated and we will examine them later in the course.

1.8.6 Bands and Band Diagrams in Real Space:

Recall that the total energy of an electron in a crystal is a sum of kinetic and potential energies, and this total energy $(\text{K.E} + \text{P.E})$ is what is plotted in $E(\vec{k})$ vs $\vec{k}$ band diagrams. Now suppose an electric field is present inside a semiconductor. The electrostatic potential $\phi(\vec{r})$ associated with an electric field $\vec{E}(\vec{r})$ satisfies

$$\vec{E}(\vec{r}) = -\vec{\nabla}\phi(\vec{r})$$

In the presence of an electric field the potential energy of the electrons gets an additional term $-q\phi(\vec{r})$. This addition shifts the energies of all the bands in real space as a function of position. In particular, the conduction band minimum E_c in real space is given by the equation

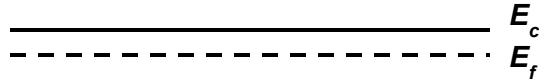$$E_c(\vec{r}) = E_c(\vec{r}_o) - q[\phi(\vec{r}) - \phi(\vec{r}_o)]$$

and similarly.

$$E_v(\vec{r}) = E_v(\vec{r}_o) - q\big[\phi(\vec{r}) - \phi(\vec{r}_o)\big]$$

$$\Rightarrow \vec{E}(\vec{r}) = \frac{1}{q}\vec{\nabla}E_c(\vec{r}) = \frac{1}{q}\nabla\vec{E}_v(\vec{r})$$

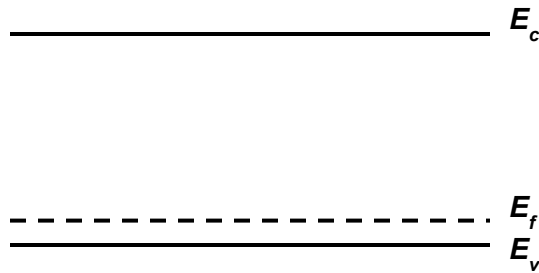Band diagrams (in real space) are plots of the lowest conduction band energy E_c, highest valence band energy E_v, and the Fermi level E_f in real space. Some examples are shown below.
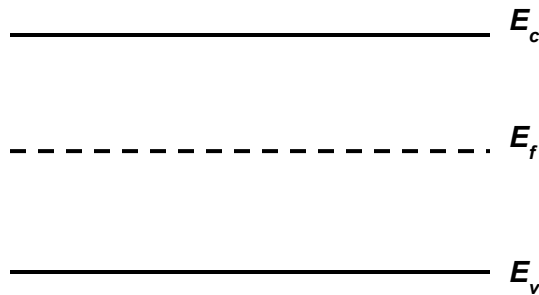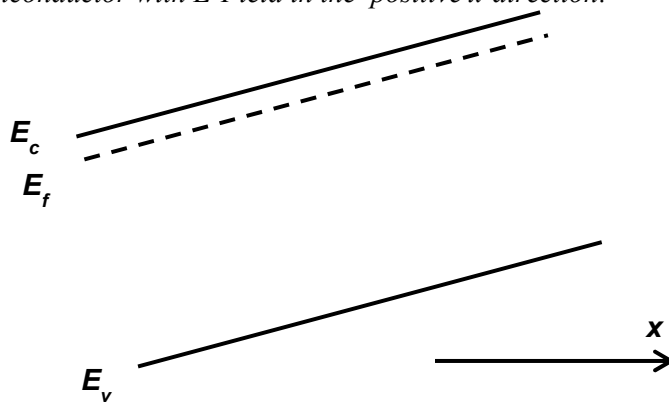
N-type Semiconductor without E-Field:

E_c

E_f

E_v

P-type Semiconductor without E-Field:

E_c

E_f

E_v

Intrinsic Semiconductor without E-Field:

E_c

E_f

E_v

N-Type Semiconductor with E-Field in the positive x-direction:

E_c

E_f

E_v

x

Note that since $E_f(\vec{r}) - E_c(\vec{r})$ is independent of position, electron density is uniform. Hole density is also uniform (why?). The situation depicted in the figure would result if we take a n-type semiconductor, put metal contacts on its two ends, and apply a voltage from the external circuit, as shown below.
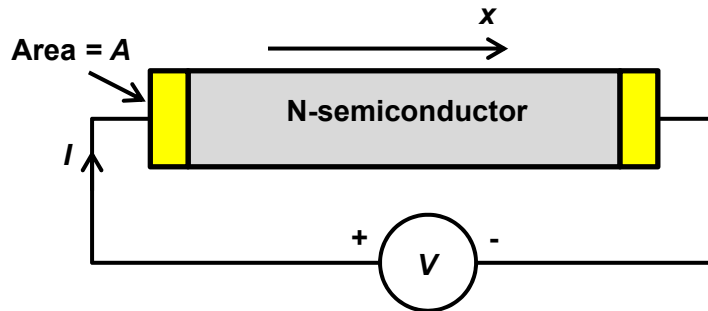
Let us find the current using Schockley's equations,

$$J_e = qn\mu_e\,E + q\,D_e\,\frac{\partial n}{\partial x}$$

$$J_h = qp\mu_h\,E - q\,D_h\,\frac{\partial p}{\partial x}$$

$$J_t = J_e + J_n$$

$$\Rightarrow J_t = (qn\mu_e + qp\mu_h)E \qquad \left\{\begin{array}{l} n(x) = \text{constant} = n \\ p(x) = \text{constant} = p \end{array}\right\}$$

$$\Rightarrow J_t L = (qn\mu_e + qp\mu_h)EL = (qn\mu_e + qp\mu_h)V$$

$$\Rightarrow I = AJ_t$$

$$\Rightarrow I = \frac{V}{R}$$

where,

$$R = \frac{L}{A\sigma} \qquad \{\sigma = qn\mu_e + qp\mu_h = \sigma_e + \sigma_h$$

No surprise here; the n-doped semiconductor acts like a conductor with conductivity that is the sum of the electron and hole conductivities.

1.8.7 Fermi-level in Equilibrium:
In equilibrium, the Fermi level, being the chemical potential, must have the same value at all locations in the device, i.e. the Fermi level is a straight horizontal line in the band diagram in equilibrium. Fermi level can change with position only in non-equilibrium situations.