Lecture notes: 01

Carrier concentration

Dr. Mohammad A Rashid
November 14, 2020
just.edu.bd/t/rashid

Contents

1 Density of States 1

2 Fermi-Dirac distribution function 4

3 Fermi energy 7

4 The dynamical effective mass 7

5 Carrier concentration in intrinsic semiconductor 10

6 Impurity distribution functions 15

7 Semiconductor statistics 16

References 18

1 Density of States

The density of states for electrons in a band yields the number of states in a certain energy range. This function is important in electronic processes, particularly in transport phenomena. When we denote the density-of-states function by $g(E)$, it is defined by the relation

$$g(E)\mathrm{d}E = \text{number of electron states per unit volume in the energy range } (E, E+\mathrm{d}E).$$

Consider a cube of semiconductor crystal with length L on each side. The electron waves

in the crystal are standing waves. In the x direction, the wavelength is

$$\lambda_x = \frac{L}{n_x} \tag{1}$$

n_x is equal to $1, 2, 3 \ldots$ The wavelength is related to the electron momentum in the x direction, p_x, through the de Broglie relationship.

$$p_x = \pm \frac{h}{\lambda_x} \tag{2}$$

h is the Planck's constant and $+$ and $-$ represent the momentum in the x and the $-x$ directions. Since λ_x only takes on a set of discrete values, so does p_x.

$$p_x = \pm \frac{n_x h}{L} \tag{3}$$

The increment between the allowable p_xs is h/L . Similarly p_y and p_z can only take on discrete values with increments of h/L. Fig. 1 shows a three dimensional space with axes p_x, p_y, and p_z. Allowed energy levels occupy points separated from one another by h/L in p_x, p_y, and p_z. There are two allowed states (the factor of 2 accounting for the two spin directions) for every cube of h^3/L^3 volume in the momentum space. Each state therefore occupies a volume of $h^3/2L^3$.

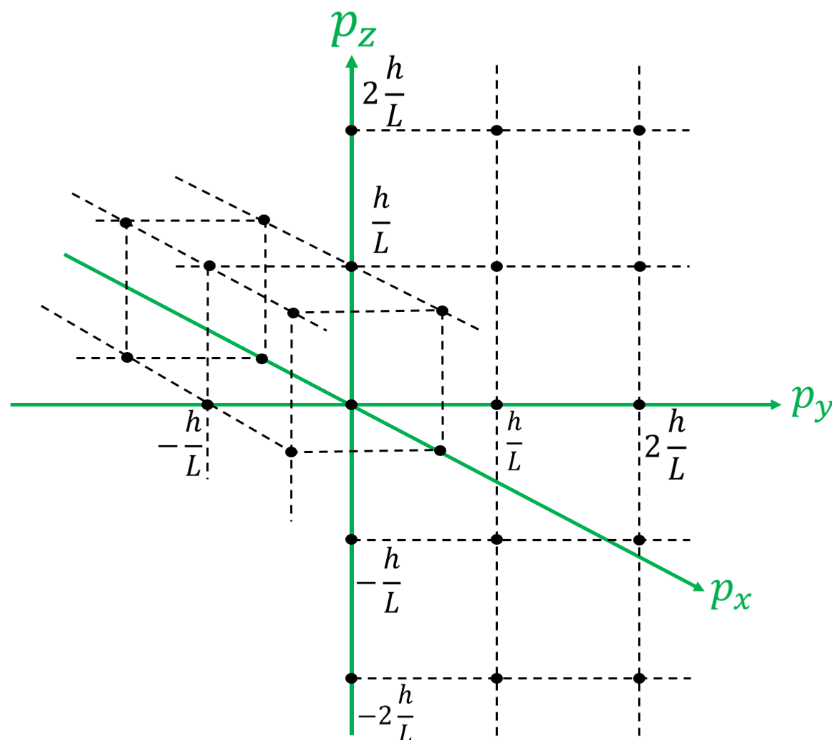

Figure 1: The allowable states in the momentum space from a large uniform 3-D grid. Only a few grid points are shown for simplicity [1].

Figure 2 shows the same momentum space as Fig. 1 but in a very much zoomed-out scale. The allowable states are now semi-continuous. Nonetheless, each allowed state

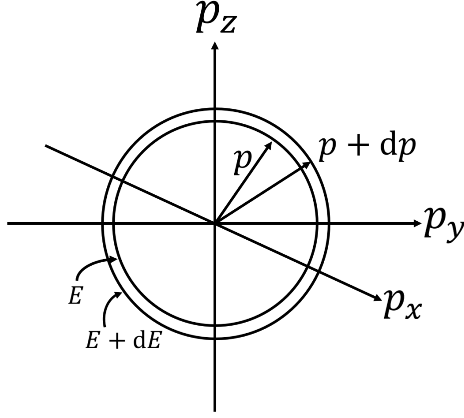

Figure 2: Each sphere in the momentum space represents a constant-energy surface [1].

still occupies a volume of $h^3/2L^3$. A sphere in this space represents a constant total momentum, p, and therefore a constant kinetic energy, E.

$$E = \frac{p^2}{2m^*} \tag{4}$$

where m^* is the electron effective mass. Using (4) twice, we find that

$$\frac{\mathrm{d}E}{\mathrm{d}p} = \frac{p}{m^*} = \frac{\sqrt{2m^*E}}{m^*} = \sqrt{\frac{2E}{m^*}}. \tag{5}$$

According to (5), two spheres that differ in energy $\mathrm{d}E$ have two radii that differ by

$$\mathrm{d}p = \sqrt{\frac{m^*}{2E}}\mathrm{d}E. \tag{6}$$

The volume of the shell that is between these two spheres is the surface area times $\mathrm{d}p$.

$$\text{Volume} = 4\pi p^2 \mathrm{d}p = 4\pi(2m^*E)\mathrm{d}p = 8\pi m^*\sqrt{\frac{m^*E}{2}}\mathrm{d}E, \tag{7}$$

where we have first used (4) and then (6). The number of states contained in this shell between E and $E + \mathrm{d}E$ is the volume of the shell divided by $h^3/2L^3$.

$$8\pi m^*\sqrt{\frac{m^*E}{2}} \times \frac{2L^3}{h^3}\mathrm{d}E. \tag{8}$$

Therefore the number of states per unit volume (the volume of the sample is L^3) per unit energy is

$$g(E) = \frac{8\pi m^*\sqrt{2m^*E}}{h^3}. \tag{9}$$

Using $\hbar = h/2\pi$, reduced Planck's constant, we have

$$\boxed{g(E) = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{3/2} E^{1/2}.} \tag{10}$$

Here E is the kinetic energy of the electron. This shows that $g(E) \sim E^{1/2}$, which means that the curve $g(E)$ has a parabolic shape (Fig. 3). The function $g(E)$ increases with E because, as we see from Fig. 2, the larger the energy the greater the radius, and hence the volume of the sphere, and consequently the larger the number of states lying within it. Also note that $g(E) \sim m^{*3/2}$. That is, the larger the mass the greater the density of states.

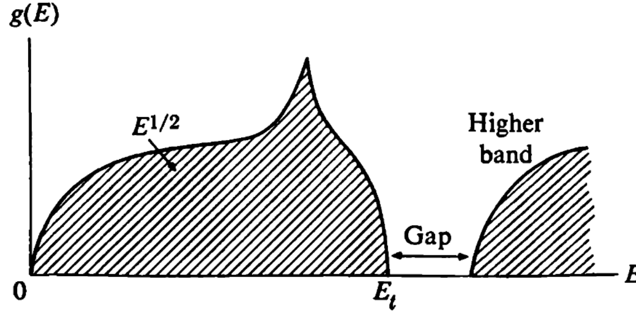

Figure 3: The density of states [2].

2 Fermi-Dirac distribution function

The Fermi-Dirac distribution function, also called Fermi function, provides the probability of occupancy of energy levels by Fermions. Fermions are half-integer spin particles, which obey the Pauli exclusion principle. The Pauli exclusion principle postulates that only one Fermion can occupy a single quantum state. Therefore, as Fermions are added to an energy band, they will fill the available states in an energy band just like water fills a bucket. The states with the lowest energy are filled first, followed by the next higher ones. At absolute zero temperature ($T = 0$ K), the energy levels are all filled up to a maximum energy, which we call the Fermi level. No states above the Fermi level are filled. At higher temperature, one finds that the transition between completely filled states and completely empty states is gradual rather than abrupt.

Electrons are Fermions. Therefore, the Fermi function provides the probability that an energy level at energy, E, in thermal equilibrium with a large system, is occupied by an electron. The system is characterized by its temperature, T, and its Fermi energy, E_F.

To derive the Fermi-Dirac distribution function, we start from a series of possible energies, labeled E_i. At each energy, we can have g_i possible states and the number of states that are occupied equals $g_i f_i$, where f_i is the probability of occupying a state at energy E_i. We also assume that the number of possible states is very large, so that the discrete nature of the states can be ignored.

The number of possible ways - called configurations - to fit $g_i f_i$ electrons in g_i states, given the restriction that only one electron can occupy each state, equals:

$$W_i = \frac{g_i!}{(g_i - g_i f_i)!(g_i f_i)!}. \tag{11}$$

This equation is obtained by numbering the individual states and exchanging the states rather than the electrons. This yields a total number of $g_i!$ possible configurations. However since the empty states are all identical, we need to divide by the number of permutations between the empty states, as all permutations cannot be distinguished from each other and can therefore only be counted once. In addition, all the filled states are indistinguishable from each other, so we need to divide also by all permutations between the filled states, namely $(g_i f_i)!$.

The number of possible ways to fit the electrons in the number of available states is called the multiplicity function.

The multiplicity function for the whole system is the product of the multiplicity functions for each energy E_i:

$$W_i = \prod_i W_i = \prod_i \frac{g_i!}{(g_i - g_i f_i)!(g_i f_i)!}. \tag{12}$$

Using Stirlings approximation:

$$\ln n! = n \ln n - n + O(\ln n) \approx n \ln n - n \tag{13}$$

one can eliminate the factorial signs in (12), yielding:

$$\ln W = \sum_i W_i = \sum_i [g_i \ln g_i - g_i(1 - f_i)\ln(g_i - g_i f_i) - g_i f_i \ln(g_i f_i)]. \tag{14}$$

The total number of electrons in the system equals N and the total energy of those N electrons equals U. These system parameters are related to the number of states at each energy, g_i, and the probability of occupancy of each state, f_i, by:

$$N = \sum_i g_i f_i \tag{15}$$

and

$$U = \sum_i E_i g_i f_i. \tag{16}$$

According to the basic assumption of statistical thermodynamics, all possible configurations are equally probable. The multiplicity function provides the number of configurations for a specific set of occupancy probabilities, f_i. The multiplicity function sharply peaks at the thermal equilibrium distribution since this is the most likely distribution of the system and must therefore be associated with the largest number of - equally probable - configurations. The occupancy probability in thermal equilibrium is therefore obtained by finding the maximum of the multiplicity function, W, while keeping the total energy and the number of electrons constant.

For convenience, we maximize the logarithm of the multiplicity function instead of the multiplicity function itself. According to the Lagrange method of undetermined multipliers, we must maximize the following function:

$$\ln W - a \sum_j g_j f_j - b \sum_j E_j g_j f_j \tag{17}$$

where a and b need to be determined. The maximum of the multiplicity function, W, is obtained from:

$$\frac{\partial}{\partial(g_i f_i)}\left[\ln W - a\sum_j g_j f_j - b\sum_j E_j g_j f_j\right] = 0, \tag{18}$$

which can be solved, yielding:

$$\ln\frac{g_i - g_i f_i}{g_i f_i} - a - bE_i = 0 \tag{19}$$

or

$$f_i = f_{FD}(E_i) = \frac{1}{1+\exp(a+bE_i)} \tag{20}$$

which can be written in the following form

$$f_{FD}(E_i) = \frac{1}{1+e^{(E_i-E_F)/\beta}} \tag{21}$$

with $\beta = 1/b$ and $E_F = -a/b$. The symbol E_F was chosen since this constant has units of energy and will be the constant associated with this probability distribution.

Taking the derivative of the total energy, (16) one obtains:

$$\mathrm{d}U = \sum_i E_i \mathrm{d}(g_i f_i) + \sum_i g_i f_i \mathrm{d}E. \tag{22}$$

Using the Lagrange equation, this can be rewritten as:

$$\mathrm{d}U = \beta \mathrm{d}(\ln W) + \sum_i g_i f_i \mathrm{d}E + E_F \mathrm{d}N. \tag{23}$$

Any variation of the energies, E_i, can only be caused by a change in volume, so that the middle term can be linked to a volume variation $\mathrm{d}V$.

$$\mathrm{d}U = \beta \mathrm{d}(\ln W) + \left[\sum_i g_i f_i \frac{\mathrm{d}E}{\mathrm{d}V}\right]\mathrm{d}V + E_F \mathrm{d}N. \tag{24}$$

Comparing this to the thermodynamic identity:

$$\mathrm{d}U = T\mathrm{d}Sp\mathrm{d}V + \mu \mathrm{d}N \tag{25}$$

one finds that $\beta = kT$ and $S = k\ln W$, where k is a constant that must be determined . The energy, E_F, equals the energy associated with the particles, namely the electrochemical potential, μ.

The Fermi-Dirac distribution function then becomes:

$$f_{FD}(E) = \frac{1}{1+e^{(E-E_F)/kT}}. \tag{26}$$

The constant k is determined to be the Boltzmann constant $k_B = 1.380649 \times 10^{23}$ J·K^{-1}. In terms of k_B Fermi-Dirac distribution function is written as

$$\boxed{f_{FD}(E) = \frac{1}{e^{(E-E_F)/k_BT} + 1}.} \tag{27}$$

The Fermi function has a value of one for energies, which are more than a few times k_BT below the Fermi energy. It equals 1/2 if the energy equals the Fermi energy and decreases exponentially for energies which are a few times k_BT larger than the Fermi energy. While at $T = 0$ K the Fermi function equals a step function, the transition is more gradual at finite temperatures and more so at higher temperatures. This function is plotted in Figure 4 at 150, 300 and 600 K.

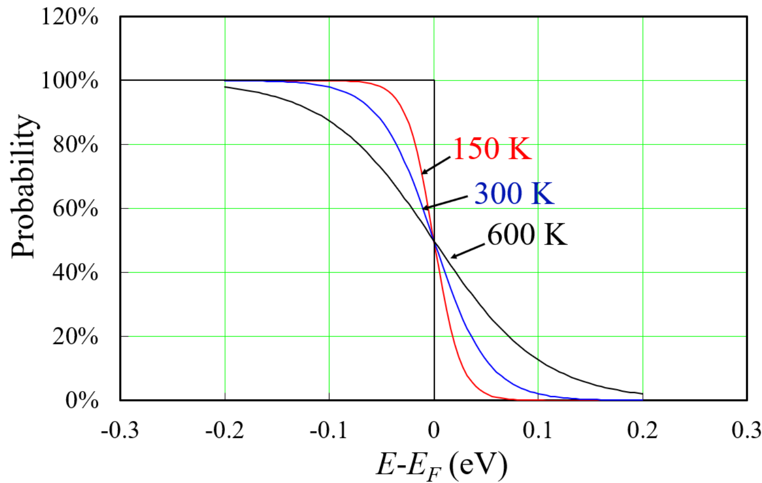

Figure 4: The Fermi function at three different temperatures [3].

3 Fermi energy

Fermi energy is often defined as the highest occupied energy level of a material at absolute zero temperature. In other words, all electrons in a body occupy energy states at or below that body's Fermi energy at 0 K. The concept of the Fermi energy is a crucially important concept for the understanding of the electrical and thermal properties of solids. At higher temperatures a certain fraction, characterized by the Fermi function, will exist above the Fermi level. The Fermi level plays an important role in the band theory of solids. In doped semiconductors, p-type and n-type, the Fermi level is shifted by the impurities, illustrated by their band gaps (Fig. 5). The Fermi level is referred to as the electron chemical potential in other contexts.

4 The dynamical effective mass

An electron in crystal may behave as if it had a mass different from the free electron mass m_0. There are crystals in which the effective mass of the carriers is much larger

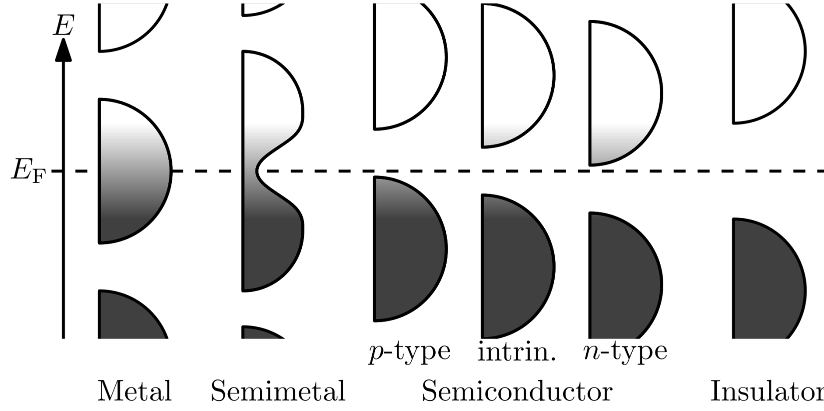

Figure 5: Filling of the electronic states in various types of materials at equilibrium. Here, height is energy while width is the density of available states for a certain energy in the material listed. The shade follows the FermiDirac distribution (black = all states filled, white = no state filled). In metals and semi-metals the Fermi level E_F lies inside at least one band. In insulators and semiconductors the Fermi level is inside a band gap; however, in semiconductors the bands are near enough to the Fermi level to be thermally populated with electrons or holes [wikipedia].

or much smaller than m_0. The effective mass may be anisotropic, and it may even be negative. The important point is that the electron in a periodic potential is accelerated relative to the lattice in an applied electric or magnetic field as if its mass is equal to an effective mass.

When an electric field $\mathscr{E}$ is applied to a crystal, the electron undergoes an acceleration. This can be calculated as follows: Since acceleration is the time derivative of velocity, we have

$$a = \frac{\mathrm{d}v}{\mathrm{d}t}, \tag{28}$$

where we have chosen to treat the one-dimensional case.

The de Broglie wavelength λ of a particle of momentum p is determined by the relation

$$\lambda = \frac{h}{p}. \tag{29}$$

Therefore,

$$p = \frac{h}{\lambda} = \frac{h}{2\pi}\frac{2\pi}{\lambda} = \hbar k, \tag{30}$$

where $k = 2\pi/\lambda$ is the wave vector and $\hbar = h/2\pi$ is the reduced Planck's constant. Hence the velocity of the electron is in terms of the wave vector is

$$v = \frac{\hbar k}{m^*}, \tag{31}$$

where the electronic mass m_e is replaced by m^*, the effective mass. Since the velocity is a function of the wave vector k, and consequently the eq. (28) may be rewritten as

$$a = \frac{\mathrm{d}v}{\mathrm{d}k}\frac{\mathrm{d}k}{\mathrm{d}t}. \tag{32}$$

Using the (30) we can write the kinetic energy of the electron as

$$E = \frac{p^2}{2m^*} = \frac{\hbar^2 k^2}{2m^*}. \tag{33}$$

The first derivative of the above equation with respect to k gives

$$\frac{\mathrm{d}E}{\mathrm{d}k} = \frac{2\hbar^2 k}{2m^*} = \frac{\hbar k}{m^*}\hbar = \frac{p}{m^*}\hbar = v\hbar, \tag{34}$$

or

$$v = \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k}. \tag{35}$$

Now, because of the electric field $\mathscr{E}$, electron in the crystal experiences a force $F = -e\mathscr{E}$ and hence a change in its energy. The rate of absorption of energy by the electron is

$$\frac{\mathrm{d}E(k)}{\mathrm{d}t} = -e\mathscr{E} \cdot v, \tag{36}$$

where the term on the right is clearly the expression for the power absorbed by a moving object. If we write

$$\frac{\mathrm{d}E(k)}{\mathrm{d}t} = \frac{\mathrm{d}E}{\mathrm{d}k}\frac{\mathrm{d}k}{\mathrm{d}t}, \tag{37}$$

and use (35) and (36) we find the simple relation

$$\hbar\frac{\mathrm{d}k}{\mathrm{d}t} = -e\mathscr{E} = F. \tag{38}$$

Hence with the help of (35), (38) we get from (32)

$$a = \frac{\mathrm{d}}{\mathrm{d}k}\left(\frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k}\right) \cdot \frac{F}{\hbar} = \frac{1}{\hbar^2}\frac{\mathrm{d}^2E}{\mathrm{d}k^2}F. \tag{39}$$

This has the same form as Newton's second law, provided we define a *dynamical effective mass* m^* by the relation

$$\boxed{m^* = \hbar^2 \Big/ \left(\frac{\mathrm{d}^2E}{\mathrm{d}k^2}\right).} \tag{40}$$

Thus, insofar as the motion in an electric field is concerned, the electron behaves like a free electron whose effective mass is given by (40).

The mass m^* is inversely proportional to the curvature of the band; where the curvature is large –that is, $\frac{\mathrm{d}^2E}{\mathrm{d}k^2}$ is large –the mass is small; a small curvature implies a large mass (Fig. 6).

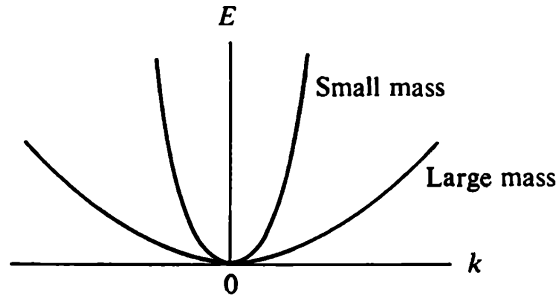

Figure 6: The inverse relationship between the mass and the curvature of the energy band [2].

The effective mass m^* is positive near the bottom of the band, where the curvature is positive. But near the top, where the band curvature is negative, the effective mass is also negative. The fact that the effective mass is different from the free mass is due to the effect of the lattice force on the electron (Fig. 7).

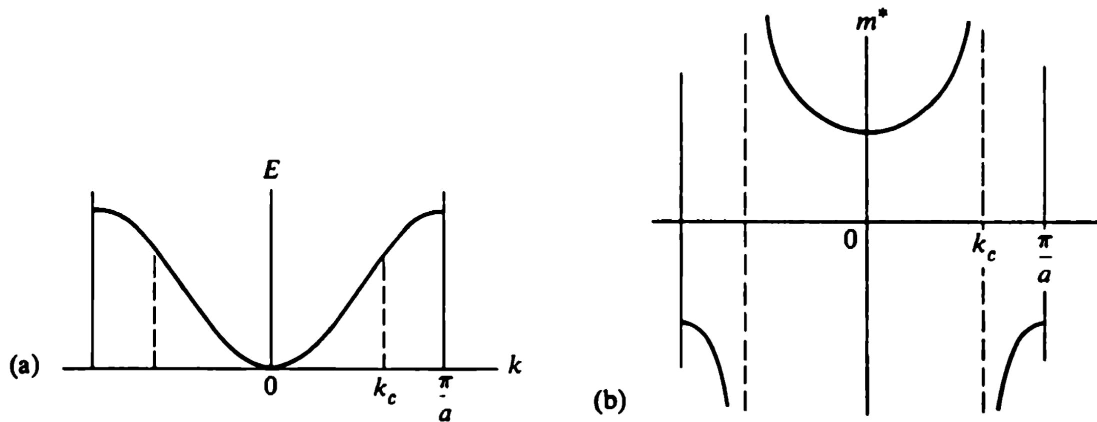

Figure 7: (a) The band structure, and (b) the effective mass m^* versus k [2].

5 Carrier concentration in intrinsic semiconductor

In the field of semiconductors, electrons and holes are usually referred to as free carriers, or simply carriers, because it is these particles which are responsible for carrying the electric current. The number of carriers is an important property of a semiconductor, as this determines its electrical conductivity. In order to determine the number of carriers, we need to use the the Fermi-Dirac (FD) distribution (eq. (27)),

$$f(E) = \frac{1}{e^{(E-E_F)/k_B T} + 1}, \tag{41}$$

which gives the probability that an energy level E is occupied by an electron when the system is at temperature T (Fig. 8) .

Here we see that, as the temperature rises, the unoccupied region below the Fermi level E_F becomes longer, which implies that the occupation of high energy states increases as the temperature is raised, a conclusion which is most plausible, since increasing the

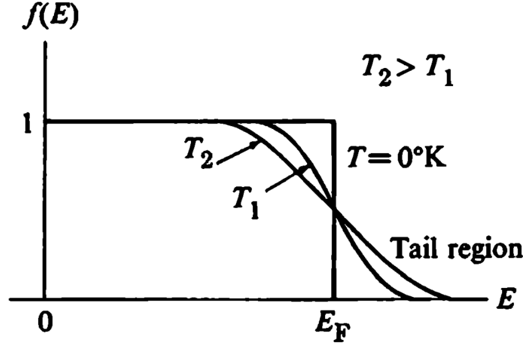

Figure 8: The Fermi-Dirac distribution function [2].

temperature raises the overall energy of the system. Note also that $f(E) = \frac{1}{2}$ at the Fermi level ($E = E_F$) regardless of the temperature. That is, the probability that the Fermi level is occupied is always equal to one-half.

In semiconductors it is the tail region of the FD distribution which is of particular interest. In that region the inequality $(E - E_F) \gg k_B T$ holds true, and one may therefore neglect the term unity in the denominator of (41). The FD distribution then reduces to the form

$$f(E) = e^{E_F/k_B T}\, e^{-E/k_B T}, \tag{42}$$

which is the familiar Maxwell-Boltzmann, or classical, distribution. This simple distribution therefore suffices for the discussion of electron statistics in semiconductors.

We can calculate the concentration of electrons in the CB in the following manner. The number of states in the energy range $(E, E + \mathrm{d}E)$ is equal to $g_e(E)\mathrm{d}E$, where $g_e(E)$ is the density of electron states (eq. (10)). Since each of these states has an occupation probability $f(E)$, the number of electrons actually found in this energy range is equal to $f(E)g_e(E)\mathrm{d}E$. The concentration of electrons throughout the CB is thus given by the integral over the band

$$n = \int_{E_{c1}}^{E_{c2}} f(E) g_e(E)\mathrm{d}E, \tag{43}$$

where E_{c1} and E_{c2} are the bottom and top of the band, respectively, as shown in Fig. 9(a) and the distribution function is shown in Fig. 9(b). Note that the entire CB falls in the tail region. Thus we may use the Maxwell-Boltzmann function for $f(E)$ in (43).

The density of states is given by

$$g_e(E) = \frac{1}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2}(E - E_g)^{1/2}, \tag{44}$$

where the zero-energy level has been chosen to lie at the top of the VB. Thus $g_e(E)$ vanishes for $E < E_g$, and is finite only for $E > E_g$, as shown in Fig. 9(c).

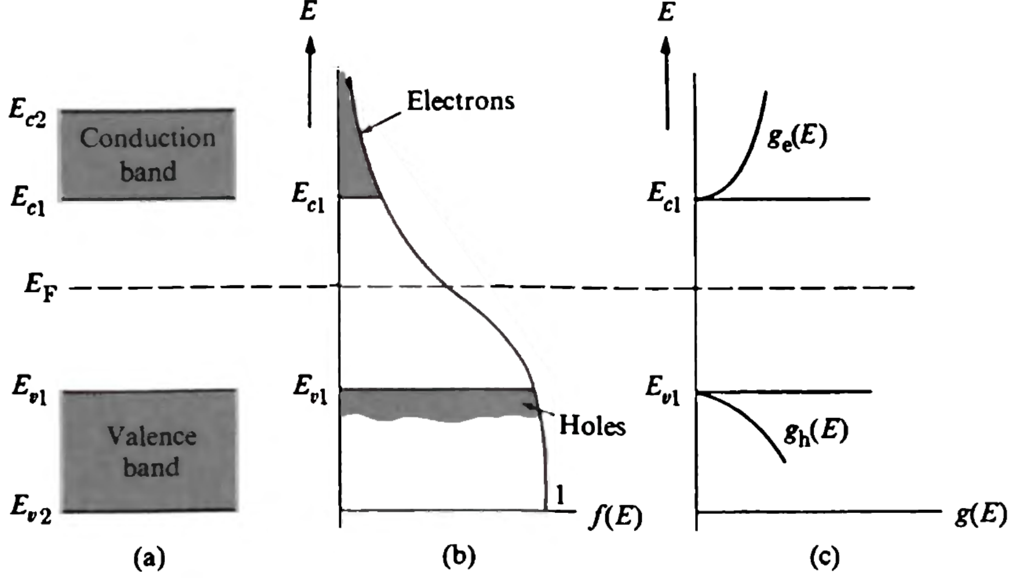

Figure 9: (a) Conduction and valence bands, (b) The distribution function, (c) Density of states for electrons and holes: $g_e(E)$ and $g_h(E)$ [2].

When we substitute for $f(E)$ and $g_c(E)$ into (43), we obtain

$$\begin{aligned} n &= \frac{1}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2} e^{E_F/k_BT}\int_{E_g}^{\infty}(E-E_g)^{1/2}e^{-E/k_BT}\mathrm{d}E \\ &= \frac{1}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2} e^{(E_F-E_g)/k_BT}\int_{E_g}^{\infty}(E-E_g)^{1/2}e^{-(E-E_g)/k_BT}\mathrm{d}E. \end{aligned} \tag{45}$$

For convenience, the top of the CB has been set equal to infinity. Since the integrand decreases exponentially at high energies, the error introduced by changing this limit from E_{c2} to ∞ is quite negligible. In order to evaluate the integration we assume $E - E_g = (k_BT)x$ to have

$$\int_{E_g}^{\infty}(E-E_g)^{1/2}e^{-(E-E_g)/k_BT}\mathrm{d}E = (k_BT)^{3/2}\int_0^{\infty}x^{1/2}e^{-x}\mathrm{d}x. \tag{46}$$

Now using the result

$$\int_0^{\infty}x^{1/2}e^{-x}\mathrm{d}x = \frac{\pi^{1/2}}{2}, \tag{47}$$

we find the electron concentration to be

$$\boxed{n = 2\left(\frac{m_ek_BT}{2\pi\hbar^2}\right)^{3/2} e^{(E_F-E_g)/k_BT}} \tag{48}$$

or

$$\boxed{n = N_c e^{(E_F-E_g)/k_BT}.} \tag{49}$$

Here $N_c \equiv 2\left(\frac{m_e k_B T}{2\pi\hbar^2}\right)^{3/2}$ is the effective density of states of the conduction band.

The electron concentration is still not known explicitly because the Fermi energy E_F is so far unknown. Essentially the same ideas employed above may also be used to evaluate the number of holes in the VB. The probability that a hole occupies a level E in this band is equal to $1 - f(E)$, since $f(E)$ is the probability of electron occupation. Thus the probability of hole occupation f_h is

$$f_h = 1 - f(E). \tag{50}$$

Since the energy range involved here is much lower than E_F, the FD function of (41) must be used rather than (42). Thus

$$f_h = 1 - \frac{1}{e^{(E-E_F)/k_B T} + 1} = \frac{1}{e^{(E_F-E)/k_B T} + 1} \simeq e^{(E-E_F)/k_B T}, \tag{51}$$

where the approximation in the last expression follows as a result of the inequality $(E_F - E) \gg k_B T$. The validity of this inequality in turn can be seen by referring to Fig. 9(b), which shows that $E_F - E$ is of the order of $E_g/2$, which is much larger than $k_B T$ at room temperature.

The density of states for the holes is

$$g_h = \frac{1}{2\pi^2}\left(\frac{2m_h}{\hbar^2}\right)^{3/2}(-E)^{1/2}, \tag{52}$$

which is appropriate for an inverted band. Note that the term $(-E)$ in this equation is positive, because the zero-energy level is at the top of the VB, and the energy is measured positive upward and negative downward from this level.

The hole concentration is thus given by

$$p = \int_{-\infty}^{0} f_h(E)\, g_h(E)\mathrm{d}E. \tag{53}$$

When we substitute for $f_h(E)$ and $g_h(E)$ from the above equations and carry out the integral as in the electron case, we obtain

$$\boxed{p = 2\left(\frac{m_h k_B T}{2\pi\hbar^2}\right)^{3/2} e^{-E_F/k_B T}.} \tag{54}$$

or

$$\boxed{p = N_v e^{-E_F/k_B T}.} \tag{55}$$

Here $N_v \equiv 2\left(\frac{m_h k_B T}{2\pi\hbar^2}\right)^{3/2}$ is the effective density of states of the valance band.

The electron and hole concentrations have thus far been treated as independent quantities. The two concentrations are, in fact, equal, because the electrons in the CB are due

to excitations from the VB across the energy gap, and for each electron thus excited a hole is created in the VB. Therefore

$$n = p. \tag{56}$$

If we substitute the values of n and p into the above equation, we obtain an equation involving the only unknown, E_F. The solution of this equation is

$$E_F = \frac{1}{2}E_g + \frac{3}{4}k_B T \ln\left(\frac{m_h}{m_e}\right). \tag{57}$$

Since $k_B T \ll E_g$ under usual circumstances, the second term on the right of (57) is very small compared with the first, and the energy level is close to the middle of the energy gap, i.e.

$$\boxed{E_F \approx \frac{1}{2}E_g.} \tag{58}$$

This is consistent with earlier assertions that both the bottom of the CB and the top of the VB are far from the Fermi level. The fact that the Fermi level falls in the energy gap –the forbidden region –poses no difficulties. This level is a theoretical concept and no electrons need be present there.

The concentration of electrons may now be evaluated explicitly by using the above value of E_F. From (57) we have

$$\frac{E_F}{k_B T} = \frac{E_g}{2k_B T} + \ln\left(\frac{m_h}{m_e}\right)^{3/4}, \tag{59}$$

or

$$e^{E_F/k_B T} = e^{E_g/2k_B T}\left(\frac{m_h}{m_e}\right)^{3/4}. \tag{60}$$

Substitution of this into (48) yields

$$n = 2\left(\frac{m_e k_B T}{2\pi\hbar^2}\right)^{3/2} e^{E_g/2k_B T}\left(\frac{m_h}{m_e}\right)^{3/4} e^{-E_g/k_B T}, \tag{61}$$

or

$$\boxed{n = 2\left(\frac{k_B T}{2\pi\hbar^2}\right)^{3/2} (m_e m_h)^{3/4} e^{-E_g/2k_B T}.} \tag{62}$$

The important feature of this expression is that n increases very rapidly –exponentially –with temperature, particularly by virtue of the exponential factor. Thus as the temperature is raised, a vastly greater number of electrons is excited across the gap. (This can be visualized by recalling that as the temperature is raised, the tail of the FD distribution in the CB becomes longer, and more states are occupied in this band.)

We can estimate the numerical value of n by substituting the values $E_g = 1$ eV, $m_e = m_h = 9.1 \times 10^{-31}$ kg, and $T = 300$ K. Which gives $n \simeq 10^{15}$ electrons/cm^3, a typical value of carrier concentration in semiconductors. Note that the expression (62) also gives the hole concentration, since $n = p$. This fact immediately suggest that,

$$n = p = n_i \tag{63}$$

and

$$\boxed{n_i = \sqrt{N_c N_v} e^{-E_g/k_B T}.} \tag{64}$$

Therefore, n_i is called the *intrinsic carrier concentration*.

Our discussion of carrier concentration in this section is based on the premise of a pure semiconductor. When the substance is impure, additional electrons or holes are provided by the impurities. In that case, the concentrations of electrons and holes may no longer be equal, and the amount of each depends on the concentration and type of impurity present. When the substance is sufficiently pure so that the concentrations of electrons and holes are equal, we speak of an *intrinsic semiconductor*. That is, the concentrations are determined by the intrinsic properties of the semiconductor itself. On the other hand, when a substance contains a large number of impurities which supply most of the carriers, it is referred to as an *extrinsic semiconductor*.

6 Impurity distribution functions

The distribution function of impurities differs from the Fermi-Dirac distribution function even though the particles involved are still Fermions. The difference is due to the fact that an ionized donor energy level still contains one electron with either spin. The donor energy level cannot be empty since this would leave a doubly positively charged atom, which would have an energy different from that of the singly ionized donor level. The distribution function for donors therefore differs from the Fermi function and is given by:

$$f_{\text{doner}}(E_d) = \frac{1}{\frac{1}{2} e^{(E_d - E_F)/k_B T} + 1} \tag{65}$$

The distribution function for acceptors differs also because of the different possible ways to occupy the acceptor level. The neutral acceptor contains no electrons. The ionized acceptor contains one electron, which can have either spin, while the doubly negatively charged state is not allowed since this would require a different energy. This restriction would yield a factor of 2 in front of the exponential term. In addition, one finds that most commonly used semiconductors have a two-fold degenerate valence band, which causes this factor to increase to four, yielding:

$$f_{\text{acceptor}}(E_a) = \frac{1}{4 e^{(E_a - E_F)/k_B T} + 1} \tag{66}$$

7 Semiconductor statistics

Semiconductors usually contain both donors and acceptors. Electrons in the CB can be created either by interband thermal excitation or by thermal ionization of the donors. Holes in the VB may be generated by interband excitation, or by thermal excitation of electrons from the VB into the acceptor level. And in addition, electrons may fall from the donor levels to the acceptor level. Figure 10 indicates these various processes.

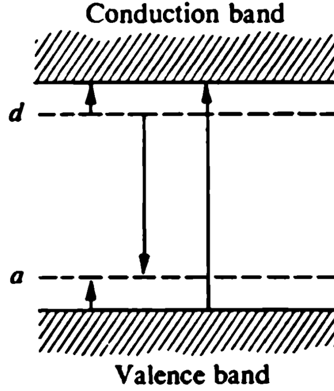

Figure 10: The various electronic processes in a semiconductor [2].

Finding the concentrations of carriers–both electrons and holes –under the most general of circumstances, taking all these processes into account, is quite complicated. Instead of giving the details of such general calculations here, we shall treat a few special cases instead, ones which are often encountered in practice. Two regions may be distinguished, depending on the physical parameters involved: The *intrinsic* and the *extrinsic* regions.

The intrinsic region

The concentration of carriers in the intrinsic region is determined primarily by thermally induced interband transitions. Consequently we have, to a good approximation,

$$n = p. \tag{67}$$

In that case, we find the carrier concentrations as

$$\boxed{n_i = 2\left(\frac{k_B T}{2\pi\hbar^2}\right)^{3/2}(m_e m_h)^{3/4} e^{-E_g/2k_B T},} \tag{68}$$

where $n_i = n = p$ is the intrinsic concentration. The intrinsic region obtains when the impurity doping is small. When we denote the concentrations of donors and acceptors by N_d and N_a, the requirement for the validity of the intrinsic condition is

$$n_i \gg (N_d - N_a) \tag{69}$$

The reason for this condition is readily understandable. There are N_d electrons at the donor level, but of these a number N_a may fall into the acceptors, leaving only $N_d - N_a$

electrons to be excited from the donor level into the conduction band. When condition (69) is satisfied, the ionization of all these remaining impurities is not sufficient to appreciably affect the number of electrons excited thermally from the VB. The semiconductor may then be treated as a pure sample, and the influence of impurities disregarded. This is precisely what we did in obtaining (68).

Since n_i increases rapidly with temperature, the intrinsic condition becomes more favorable at higher temperatures. All semiconductors, in fact, become intrinsic at sufficiently high temperatures (unless the doping is unusually high).

The extrinsic region

Quite often the intrinsic condition is not satisfied. For the common dopings encountered, about 10^{15} cm^{-3}, the number of carriers supplied by the impurities is large enough to change the intrinsic concentration appreciably at room temperature. The contribution of impurities, in fact, frequently exceeds those carriers that are supplied by interband excitation. When this is so, the sample is in the *extrinsic region*.

Two different types of extrinsic regions may be distinguished. The first occurs when the donor concentration greatly exceeds the acceptor concentration, that is, when $N_d \gg N_a$. In this case, the concentration of electrons may be evaluated quite readily. Since the donor's ionization energy is quite small, all the donors are essentially ionized, their electrons going into the CB. Therefore, to a good approximation,

$$n = N_d \tag{70}$$

The hole concentration is small under this condition. To calculate this concentration, we multiply (48) and (54) and find

$$np = 4\left(\frac{k_B T}{2\pi\hbar^2}\right)^3 (m_e m_h)^{3/2} e^{-E_g/k_B T}. \tag{71}$$

Note that the troublesome Fermi energy has disappeared from the right side. Thus the product np is independent of E_F, and hence of the amount and type of doping; the product np depends only on the temperature. We also see from comparison with (68) that the right side is equal to n_i^2 [which is reasonable, since (71) is also valid in the intrinsic region, in which case the left side is equal to n_i^2]. We may thus write

$$np = n_i^2. \tag{72}$$

This equation means that, if there is no change in temperature, the product np is a constant, independent of the doping. If the electron concentration is increased, by varying the doping, the hole concentration decreases, and vice versa.

When the doping is primarily of the donor type, $n \simeq N_d$, as shown by (70). According to (72), the concentration

$$p = \frac{n_i^2}{N_d}. \tag{73}$$

Since we are in the extrinsic region, $n_i \ll N_d$, and hence $p \ll N_d = n$. Thus the concentration of electrons is much larger than that of holes.

A semiconductor in which $n \gg p$ is called an *n-type semiconductor* (n for negative); this terminology dates back to the early days of semiconductors. Such a sample is characterized, as we have seen, by a great concentration of electrons (donors). (For a strongly n-type sample, $n \gg p$, while for a weakly n-type sample, $n \gtrsim p$.)

The other type of extrinsic region occurs when $N_a \gg N_d$, that is, the doping is primarily by acceptors. Using an argument similar to the above, one then has

$$p \simeq N_a, \tag{74}$$

i.e., all the acceptors are ionized. The electron concentration, which is small, is given by

$$n = \frac{n_i^2}{N_a}. \tag{75}$$

Such a material is called a *p-type semiconductor*. It is characterized by a preponderance of holes (acceptors).

In discussing ionization of donors (and acceptors), we assumed that the temperature is sufficiently high so that all of these are ionized. This is certainly true at room temperature. But if the temperature is progressively lowered, a point is reached at which the thermal energy becomes too small to cause electron excitation. In that case, the electrons fall from the CB into the donor level, and the conductivity of the sample diminishes dramatically. This is referred to as freeze-out, in that the electrons are now "frozen" at their impurity sites. We can estimate the temperature at which freeze-out takes place from the equation $E_d \simeq k_B T$, which gives a temperature of about 100 K.

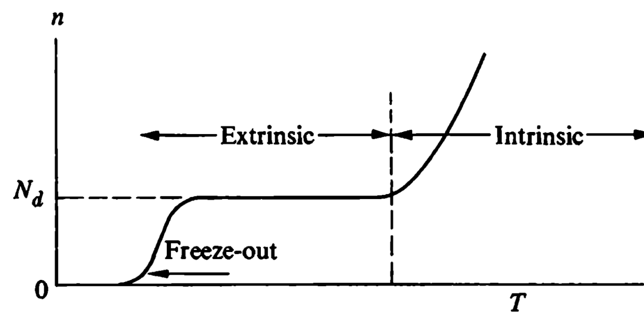

Figure 11: Variation of electron concentration with T in an n-type semiconductor [2].

The variation of the electron concentration with temperature in an n-type sample is indicated schematically in Fig. 11, in which the various regions –freeze-out, extrinsic, intrinsic –are clearly marked.

References

1. Modern Semiconductor Devices for Integrated Circuits by Chenming Hu
2. Elementary Solid State Physics by M. Ali Omar
3. Principles of Semiconductor Devices by Bart Van Zeghbroeck