ENERGY BANDS; FERMI LEVEL;
ELECTRONS AND HOLES

The application of the Fermi-Dirac distribution to electrons in semiconductors is central to the design and operation of all semiconductor devices, and thus to much of modern electronics. We treat below those aspects of the physics of semiconductors and semiconductor devices that are parts of thermal physics. We assume that the reader is familiar with the basic ideas of the physics of electrons in crystalline solids, as treated in the texts on solid state physics and on semiconductor devices cited in the general references. We assume the concept of energy bands and of conduction by electrons and holes. Our principal aim is to understand the dependence of the all-important concentrations of conduction electrons and of holes upon the impurity concentration and the temperature.

A semiconductor is a system with electron orbitals grouped into two energy bands separated by an energy gap (Figure 13.1). The lower band is the valence band and the upper band is the conduction band.* In a pure semiconductor at \( \tau = 0 \) all valence band orbitals are occupied and all conduction band orbitals are empty. A full band cannot carry any current, so that a pure semiconductor at \( \tau = 0 \) is an insulator. Finite conductivity in a semiconductor follows either from the presence of electrons, called conduction electrons, in the conduction band or from unoccupied orbitals in the valence band, called holes.

Two different mechanisms give rise to conduction electrons and holes: Thermal excitation of electrons from the valence band to the conduction band, or the presence of impurities that change the balance between the number of orbitals in the valence band and the number of electrons available to fill them.

We denote the energy of the top of the valence band by \( e_v \), and the energy of the bottom of the conduction band by \( e_c \). The difference

\[ e_g = e_c - e_v, \]  

is the energy gap of the semiconductor. For typical semiconductors \( e_g \) is between 0.1 and 2.5 electron volts. In silicon, \( e_g \approx 1.1 \) eV. Because \( \tau \approx 1/40 \) eV at room

* We treat both bands as single bands; for our purposes it does not matter that both may be groups of bands with additional gaps within each group.
Figure 13.1 Energy band structure of a pure semiconductor or insulator. The electron orbitals occur in bands which extend through the crystal. At $\tau = 0$ all orbitals up to the top of the valence band are filled, and the conduction band is empty. The energy interval between the bands is called the energy gap.

temperature, we usually have $\varepsilon_g \gg \tau$. Substances with a gap of more than about 2.5 eV are usually insulators. Table 13.1 gives the energy gaps for selected semiconductors, together with other properties needed later.

Let $n_e$ denote the concentration of conduction electrons and $n_h$ the concentration of holes. In a pure semiconductor the two will be equal:

$$n_e = n_h,$$

if the crystal is electrically neutral.

Most semiconductors as used in devices have been intentionally doped with impurities that may become thermally ionized in the semiconductor at room temperature. Impurities that give an electron to the crystal (and become positively charged in the process) are called donors. Impurities that accept
Table 13.1 Band structure data of some important semiconductors

<table>
<thead>
<tr>
<th>Energy gaps at 300 K</th>
<th>Quantum concentrations of electrons and holes at 300 K</th>
<th>Density-of-states effective masses, in units of the free electron mass</th>
<th>Dielectric constants, relative to vacuum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_F$, eV</td>
<td>$n_e$, cm$^{-3}$</td>
<td>$n_h$, cm$^{-3}$</td>
<td>$m_e^* / m_0$</td>
</tr>
<tr>
<td>Si</td>
<td>$1.14$</td>
<td>$2.7 \times 10^{19}$</td>
<td>$1.1 \times 10^{19}$</td>
</tr>
<tr>
<td>Ge</td>
<td>$0.67$</td>
<td>$1.0 \times 10^{19}$</td>
<td>$5.2 \times 10^{18}$</td>
</tr>
<tr>
<td>GaAs</td>
<td>$1.43$</td>
<td>$4.6 \times 10^{17}$</td>
<td>$1.5 \times 10^{19}$</td>
</tr>
<tr>
<td>InP</td>
<td>$1.35$</td>
<td>$4.9 \times 10^{17}$</td>
<td>$6.9 \times 10^{18}$</td>
</tr>
<tr>
<td>InSb</td>
<td>$0.18$</td>
<td>$4.6 \times 10^{16}$</td>
<td>$6.2 \times 10^{18}$</td>
</tr>
</tbody>
</table>

an electron from the valence band (and become negatively charged in the process) are called acceptors.

Let $n_d^+$ be the concentration of positively charged donors and $n_a^-$ the concentration of negatively charged acceptors. The difference

$$\Delta n = n_d^+ - n_a^-$$

is called the net ionized donor concentration. The electrical neutrality condition becomes

$$n_e - n_h = \Delta n = n_d^+ - n_a^-,$$

which specifies the difference between electron and hole concentrations.

The electron concentration may be calculated from the Fermi-Dirac distribution function of Chapter 6:

$$f_e(\varepsilon) = \frac{1}{\exp[(\varepsilon - \mu)/\tau] + 1},$$

where $\mu$ is the chemical potential of the electrons. The subscript $e$ refers to electrons. In semiconductor theory the electron chemical potential is always called the Fermi level. Further, in semiconductor theory the character $\mu$ is almost always reserved for the electron and hole mobilities, and the Fermi level is designated by $\varepsilon_F$ or by $\zeta$. To avoid confusion with the Fermi energy of a metal which we designated as $\varepsilon_F$ and which stands for the Fermi level in the limit $\tau \to 0$, we shall maintain our previous usage of the letter $\mu$ for the chemical potential at any temperature.
Chapter 13: Semiconductor Statistics

Given $\mu$ and $\tau$, the number of conduction electrons is obtained by summing the distribution function $f_e(\varepsilon)$ over all conduction band orbitals:

$$N_e = \sum_{CB} f_e(\varepsilon). \quad (6)$$

The number of holes is

$$N_h = \sum_{VB} [1 - f_h(\varepsilon)] = \sum_{VB} f_h(\varepsilon), \quad (7)$$

where the summation is over all valence band orbitals. Here we have introduced the quantity

$$f_h(\varepsilon) \equiv 1 - f_e(\varepsilon) = \frac{1}{\exp[(\mu - \varepsilon)/\tau] + 1}, \quad (8)$$

which is the probability that an orbital at energy $\varepsilon$ is unoccupied. We say that the unoccupied orbital is "occupied by a hole"; then $f_h(\varepsilon)$ is the distribution function for holes just as $f_e(\varepsilon)$ is the distribution function for electrons. Comparison of (8) with (5) shows that the hole occupation probability involves $\mu - \varepsilon$ where the electron occupation probability involves $\varepsilon - \mu$.

The concentrations $n_e = N_e/V$ and $n_h = N_h/V$ depend on the Fermi level. But what is the value of the Fermi level? It is determined by the electrical neutrality requirement (4), now written as $n_e(\mu) - n_h(\mu) = \Delta n$. This is an implicit equation for $\mu$; to solve the equation we must determine the functional dependences $n_e(\mu)$ and $n_h(\mu)$.

Classical Regime

We assume that both electron and hole concentrations are in the classical regime defined by the requirements that $f_e \ll 1$ and $f_h \ll 1$, as in Chapter 6. This will be true if, as in Figure 13.2, the Fermi level lies inside the energy gap and is separated from both band edges by energies large enough that

$$\exp[-(\varepsilon_c - \mu)/\tau] \ll 1; \quad \exp[-(\mu - \varepsilon_v)/\tau] \ll 1. \quad (9)$$

To satisfy (9) both $(\varepsilon_c - \mu)$ and $(\mu - \varepsilon_v)$ have to be positive and at least a few times larger than $\tau$. Such a semiconductor is called nondegenerate. The inequalities (9) place upper limits on the electron and hole concentrations and are satisfied in many applications. With (9) the two occupation probabilities $f_e(\varepsilon)$ and $f_h(\varepsilon)$ reduce to classical distributions:
Figure 13.2 Occupancy of orbitals at a finite temperature, according to the Fermi-Dirac distribution function. The conduction and valence bands may be represented in terms of temperature-dependent effective numbers $N_c, N_v$ of degenerate orbitals located at the two band edges $e_c, e_v$. The $n_c, n_v$ are the corresponding quantum concentrations.

\[ f_s(e) \simeq \exp\left[-(e - \mu)/\tau\right]; \quad f_v(e) \simeq \exp\left[-(\mu - e)/\tau\right]. \quad (10) \]

We use (6) and (10) to write the total number of conduction electrons in the form

\[ N_c = \sum_{CB} \exp\left[-(e_c - \mu)/\tau\right] = \exp\left[-(e_c - \mu)/\tau\right] \sum_{CB} \exp\left[-(e_c - \mu)/\tau\right], \]

or

\[ N_c = N_c \exp\left[-(e_c - \mu)/\tau\right], \quad (11) \]
where we define

\[ N_e \equiv \sum_{\epsilon_B} \exp\left[-(\epsilon - \epsilon_c)/\tau\right]. \]  

(12)

Here \( \epsilon - \epsilon_c \) is the energy of a conduction electron referred to the conduction band edge \( \epsilon_c \) as origin.

The expression for \( N_e \) has the mathematical form of a partition function for one electron in the conduction band. In Chapter 3 we evaluated a similar sum denoted there by \( Z_1 \), and we can adapt that result to the present problem with an approximate modification for band structure effects. Because of the rapid decrease of \( \exp\left[-(\epsilon - \epsilon_c)/\tau\right] \) as \( \epsilon \) increases above its minimum value at \( \epsilon_c \), only the distribution of orbitals within a range of a few \( \tau \) above \( \epsilon_c \) really matters in the evaluation of the sum in (12). The orbitals high in the band make a negligible contribution. The important point is that near the band edge the electrons behave very much like free particles. Not only are the electrons mobile, which causes the conductivity of the semiconductor, but the energy distribution of the orbitals near the band edge usually differs from that of free particles only by a proportionality factor in the energy and eventually in the sum for \( Z_1 \).

We can arrange for a suitable proportionality factor by use of a device called the density-of-states effective mass. For free particles we calculated the partition function \( Z_1 \) in (3.62), but for zero spin. For particles of spin \( \frac{1}{2} \) the result is larger by a factor of 2, so that (12) becomes

\[ N_e = Z_1 = 2n_0V = 2(m*2\pi\hbar^2)^{3/2}V. \]  

(13)

Numerically, this gives

\[ N_e/V \approx 2.509 \times 10^{19} \times (T/300 \text{K})^{3/2} \text{ cm}^{-3}. \]  

(14)

where \( T \) is in kelvin.

The quantity \( N_e \) for actual semiconductors exhibits the same temperature dependence as (13), but differs in magnitude by a proportionality factor. We express this formally by writing, in analogy to (13),

\[ N_e = 2(m_e^*/2\pi\hbar^2)^{3/2}V. \]  

(15)

where \( m_e^* \) is called the density-of-states effective mass for electrons. Experimental values are given in Table 13.1. The introduction of effective masses is more than a formality. In the theory of electrons in crystals it is shown that
the dynamical behavior of electrons and holes, under the influence of external forces such as electric fields, is that of particles with effective masses different from the free electron mass. The dynamical masses usually are different from the density-of-states masses, however.

We define the quantum concentration $n_e$ for conduction electrons as

$$n_e \equiv N_e/V = 2(m_e^*/\tau/2\pi h^2)^{3/2}. \quad (16)$$

By (11) the conduction electron concentration $n_e \equiv N_e/V$ becomes

$$n_e = n_e \exp[-(e_e - \mu)/\tau]. \quad (17)$$

The earlier assumption (9) is equivalent to the assumption that $n_e \ll n_i$, so that the conduction electrons act as an ideal gas. As an aid to memory, we may think of $N_e$ as arising from $N_i$ orbitals at $e_i$, with the Fermi level at $\mu$. Warning: In the semiconductor literature $n_i$ is invariably called the effective density of states of the conduction band.

Similar reasoning gives the number of holes in the valence band:

$$N_h = \sum_{vb} \exp[-(\mu - \epsilon)/\tau] = N_e \exp[-(\mu - e_i)/\tau], \quad (18)$$

with the definition

$$N_i \equiv \sum_{vb} \exp[-(e_i - \epsilon)/\tau]. \quad (19)$$

We define the quantum concentration $n_h$ for holes as

$$n_h \equiv N_h/V = 2(m_h^*/\tau/2\pi h^2)^{3/2}, \quad (20)$$

where $m_h^*$ is the density-of-states effective mass for holes. By (18) the hole concentration $n_h \equiv N_h/V$ is

$$n_h = n_e \exp[-(\mu - e_i)/\tau]. \quad (21)$$
Like (17), this gives the carrier concentration in terms of the quantum concentration and the position of the Fermi level relative to the valence band edge. In the semiconductor literature \( n_e \) is called the effective density of states of the valence band.

**Law of Mass Action**

The product \( n_en_h \) is independent of the Fermi level so long as the concentrations are in the classical regime. Then

\[
n_en_h = n_en_e \exp\left[-(\varepsilon_e - \varepsilon_f)/\tau\right] = n_en_e \exp(-\varepsilon_g/\tau), \tag{22a}
\]

where the energy gap \( \varepsilon_g \equiv \varepsilon_e - \varepsilon_f \). In a pure semiconductor we have \( n_e = n_h \), and the common value of the two concentrations is called the intrinsic carrier concentration \( n_i \) of the semiconductor. By (22a),

\[
n_i = (n_en_e)^{1/2} \exp(-\varepsilon_g/2\tau). \tag{22b}
\]

The Fermi level independence of the product \( n_en_h \) means that this product retains its value even when \( n_e \neq n_h \), as in the presence of electrically charged impurity atoms, provided both concentrations remain in the classical regime. We may then write (22a) as

\[
n_en_h = n_i^2. \tag{22c}
\]

The value of the product depends only on the temperature. This result is the mass action law of semiconductors, similar to the chemical mass action law (Chapter 9).

**Intrinsic Fermi Level**

For an intrinsic semiconductor \( n_e = n_i \), and we may equate the right-hand sides of (17) and (22b):

\[
n_i \exp\left[-(\varepsilon_e - \mu)/\tau\right] = (n_en_e)^{1/2} \exp(-\varepsilon_g/2\tau). \tag{23}
\]

Insert \( \varepsilon_g = \varepsilon_e - \varepsilon_f \) and divide by \( n_e \exp(-\varepsilon_f/\tau) \):

\[
\exp(\mu/\tau) = (n_i/n_e)^{1/2} \exp[(\varepsilon_e + \varepsilon_f)/2\tau].
\]
We take logarithms to obtain

$$\mu = \frac{1}{2}(\varepsilon_c + \varepsilon_v) + \frac{3}{2} \tau \log(n_i/n_e) = \frac{1}{2}(\varepsilon_c + \varepsilon_v) + \frac{3}{2} \tau \log(m_\text{e}^*/m_\text{v}^*),$$  \hspace{1cm} (24)$$

by use of (16) and (20). The Fermi level for an intrinsic semiconductor lies near the middle of the forbidden gap, but displaced from the exact middle by an amount that is usually small.

**n-TYPE AND p-TYPE SEMICONDUCTORS**

**Donors and Acceptors**

Pure semiconductors are an idealization of little practical interest. Semiconductors used in devices usually have impurities intentionally added in order to increase the concentration of either conduction electrons or holes. A semiconductor with more conduction electrons than holes is called n-type; a semiconductor with more holes than electrons is called p-type. The letters \(n\) and \(p\) signify negative and positive majority carriers. Consider a silicon crystal in which some of the Si atoms have been substituted by phosphorus atoms. Phosphorus is just to the right of Si in the periodic table, hence each P has exactly one electron more than the Si it replaces. These extra electrons do not fit into the filled valence band; hence a Si crystal with some P atoms will contain more conduction electrons and, by the law of mass action, fewer holes than a pure Si crystal. Next consider aluminum atoms. Aluminum is just to the left of Si in the periodic table, hence Al has exactly one electron fewer than the Si it replaces. As a result, Al atoms increase the number of holes and decrease the number of conduction electrons.

Most impurities in the same columns of the periodic table as P and Al will behave in Si just as P and Al behave. What matters is the number of valence electrons relative to Si and not the total number of electrons on the atom. Impurities from other columns of the periodic table will not behave so simply. Similar reasoning can be applied to other semiconductors, for example GaAs. For the present we assume that each donor atom contributes one electron which may enter the conduction band or fill one hole in the valence band. We also assume that each acceptor atom removes one electron, either from the valence band or from the conduction band. These assumptions are called the approximation of fully ionized impurities: all impurities when ionized are either positively charged donors \(D^+\) or negatively charged acceptors \(A^-\).

The electrical neutrality condition (4) told us that

$$\Delta n = n_c - n_h = n_\text{d}^+ - n_\text{a}^-.$$  \hspace{1cm} (25)$$
Because \( n_b = n_i^2/n_e \) from the mass action law, we see that (25) leads to a quadratic equation for \( n_e \):

\[
n_e^2 - n_e \Delta n = n_i^2. \tag{26}
\]

The positive root is

\[
n_e = \frac{1}{2} \left( [\Delta n]^2 + 4n_i^2 \right)^{1/2} + \Delta n, \tag{27a}
\]

and because \( n_b = n_e - \Delta n \) we have

\[
n_b = \frac{1}{2} \left( [\Delta n]^2 + 4n_i^2 \right)^{1/2} - \Delta n. \tag{27b}
\]

Most often the doping concentration is large compared to the intrinsic concentration, so that either \( n_e \) or \( n_b \) is much larger than \( n_i \):

\[
|\Delta n| \gg n_i. \tag{28}
\]

This condition defines an extrinsic semiconductor. The square roots in (27) can then be expanded:

\[
[\Delta n]^2 + 4n_i^2 \right)^{1/2} = |\Delta n| \left[ 1 + (2n_i/\Delta n)^2 \right]^{1/2} \approx |\Delta n| + 2n_i^2/|\Delta n|. \tag{29}
\]

In an \( n \)-type semiconductor \( \Delta n \) is positive and (27) becomes

\[
n_e \simeq \Delta n + n_i^2/\Delta n \simeq \Delta n; \quad n_b \simeq n_i^2/\Delta n \ll n_i. \tag{30}
\]

In a \( p \)-type semiconductor \( \Delta n \) is negative and (27) becomes

\[
n_e \simeq n_i^2/|\Delta n| \ll n_i; \quad n_b \simeq |\Delta n| + n_i^2/|\Delta n| \simeq |\Delta n|. \tag{31}
\]

The majority carrier concentration in the extrinsic limit (28) is nearly equal to the magnitude of \( \Delta n \), while the minority carrier concentration is inversely proportional to \( |\Delta n| \).

**Fermi Level in Extrinsic Semiconductor**

By use of the mass action law we calculated the carrier concentrations without having to calculate the Fermi level first. The Fermi level is obtained from \( n_e \) or \( n_b \) by solving (17) or (21) for \( \mu \).
Figure 13.3 The Fermi level in silicon as a function of temperature, for various doping concentrations. The Fermi levels are expressed relative to the band edges. A small decrease of the energy gap with temperature has been neglected.

\[ \mu = \epsilon_e - \tau \log(n_e/n_i) = \epsilon_v + \tau \log(n_p/n_h). \]  

(32)

We may now use (27) to find \( \mu \) as a function of temperature and doping level \( \Delta n \). Figure 13.3 gives numerical results for Si. With decreasing temperature the Fermi level in an extrinsic semiconductor approaches either the conduction or the valence band edge.

**Degenerate Semiconductors**

When one of the carrier concentrations is increased and approaches the quantum concentration, we may no longer use the classical distribution (10) for that carrier. The calculation of the carrier concentration now follows the treatment of the Fermi gas in Chapter 7. The sum over all occupied orbitals, which is equal to the number of electrons, is written as an integral over the density of states times the distribution function:

\[ N = \int d\epsilon \mathcal{D}(\epsilon) f(\epsilon). \]  

(33)
where for free particles of mass $m$ the density of states is

$$\Phi(\varepsilon) = \frac{V}{2\pi^2} \left(\frac{2m^*}{\hbar^2}\right)^{3/2} \varepsilon^{1/2}. \quad (34)$$

That is, $\Phi(\varepsilon) d\varepsilon$ is the number of orbitals in the energy interval $(\varepsilon, \varepsilon + d\varepsilon)$. To make the transition to conduction electrons in semiconductors we replace $N$ by $n_c V$; $m$ by $m^*$; and $\varepsilon$ by $\varepsilon - \varepsilon_c$. We obtain

$$n_c = \frac{1}{2\pi^2} \left(\frac{2m^*}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{d\varepsilon (\varepsilon - \varepsilon_c)^{1/2}}{1 + \exp(\varepsilon - \mu)/\tau}. \quad (35)$$

Let $x = (\varepsilon - \varepsilon_c)/\tau$ and $\eta = (\mu - \varepsilon_c)/\tau$. We use the definition (16) of $n_c$ to obtain

$$n_c/n_e = I(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{dx x^{1/2}}{1 + \exp(x - \eta)}. \quad (36)$$

The integral $I(\eta)$ in (36) is known as the Fermi-Dirac integral.

*When $\varepsilon_c - \mu \gg \tau$ we have $-\eta \gg 1$, so that $\exp(x - \eta) \approx 1$. In this limit

$$n_c/n_e \approx \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{dx e^{-x} x^{1/2}}{1 + \exp[1/\tau]} = \frac{2}{\sqrt{\pi}} \Gamma(3/2)e^\eta = \exp[(\mu - \varepsilon_c)/\tau], \quad (37)$$

the familiar result for the ideal gas.

In semiconductors the electron concentration rarely exceeds several times the quantum concentration $n_q$. The deviation between the value of $\mu$ from (35) and the approximation (37) then can be expanded into a rapidly converging power series of the ratio $r = n_c/n_e$, called the Joyce-Dixon approximation:* 

$$\eta - \log r \approx \frac{1}{\sqrt{8}} r - \left(\frac{3}{16} - \frac{\sqrt{3}}{9}\right) r^2 + \cdots; \quad r = n_c/n_e, \quad (38)$$

Figure 13.4 compares the exact relation (36) with the approximations (37) and (38).

---

* W. B. Joyce and R. W. Dixon, Appl. Phys. Lett. 31, 354 (1977). If the right side of (38) is written as $\sum A_n r^n$, the first four coefficients are $A_1 = 3.53553 \times 10^{-4}$; $A_2 = -4.95009 \times 10^{-6}$; $A_3 = 1.48386 \times 10^{-8}$; $A_4 = -4.42563 \times 10^{-10}$.
Figure 13.4 Electron concentration as a function of Fermi level, for $\mu$ near or above conduction band edge $e_c$. The dashed curve represents the first term of the Joyce-Dixon approximation (38).
When $n_e$ is no longer small compared to $n_i$, the expression of the mass action law must be modified. In Problem 4 we ask the reader to show that

$$n_p n_h = n_i^2 \exp\left[-n_e/\sqrt{8n_e} + \cdots\right].$$  \hfill (39)

If the gap itself depends on the carrier concentrations, the value of $n_i$ to be used here will depend on concentration.

**Impurity Levels**

The addition of impurities to a semiconductor moves some orbitals from the conduction or valence band into the energy gap, where the orbitals now appear as localized bound states. We consider phosphorous in a silicon crystal. If the P atom has released its extra electron to the Si conduction band, the atom appears as a positively charged ion. The positive ion attracts the electrons in the conduction band, and the ion can bind an electron just as a proton can bind an electron in a hydrogen atom. However, the binding energy in the semiconductor is several orders of magnitude lower, mostly because the binding energy is to be divided by the square of the static dielectric constant, and partly because of mass effects. Table 13.2 gives the ionization energies for column V donors in Si and Ge. The lowest orbital of an electron bound to a donor corresponds to an energy level $\Delta \varepsilon_d = \varepsilon_v - \varepsilon_d$ below the edge of the conduction band (Figure 13.5). There is one set of bound orbitals for every donor.

A parallel argument applies to holes and acceptors. Orbitals are split off from the valence band, as in Figure 13.5. For each acceptor atom there is one set of bound orbitals with an ionization energy $\Delta \varepsilon_a = \varepsilon_v - \varepsilon_a$, of the same order as $\Delta \varepsilon_d$. Ionization energies for column III acceptors in Si are listed in Table 13.2.

In GaAs the ionization energies for all column VI donors except oxygen are close to 6 meV. For zinc, the most important acceptor, $\Delta \varepsilon_a = 24$ meV. Some

<table>
<thead>
<tr>
<th>Table 13.2</th>
<th>Ionization energies of column V donors and column III acceptors in Si and Ge, in meV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Donors</td>
</tr>
<tr>
<td></td>
<td>P</td>
</tr>
<tr>
<td>Si</td>
<td>45</td>
</tr>
<tr>
<td>Ge</td>
<td>12.0</td>
</tr>
</tbody>
</table>
impurities generate orbitals deep inside the forbidden gap, sometimes with multiple orbitals corresponding to different ionization states.

**Occupation of Donor Levels**

A donor level can be occupied by an electron with either spin up or spin down. Hence there are two different orbitals with the same energy. However, the occupations of these two orbitals are not independent of each other: Once the level is occupied by one electron, the donor cannot bind a second electron with opposite spin. As a result, the occupation probability for a donor level is not given by the simple Fermi-Dirac distribution function, but by a function treated in Chapter 5. We write the probability that the donor orbital is vacant, so that the donor is ionized, in a form slightly different from (5.73):

\[
\tilde{f}(D^+) = \frac{1}{1 + \exp[(\mu - \epsilon_d)/\tau]}.
\]  

(40)

Here \( \epsilon_d \) is the energy of a singly occupied donor orbital relative to the origin of the energy. The probability that the donor orbital is occupied by an electron, so that the donor is neutral, is given by (5.74):

\[
f(D) = \frac{1}{1 + \frac{1}{2}\exp[(\epsilon_d - \mu)/\tau]}.
\]

(41)
Chapter 13: Semiconductor Statistics

Acceptors require extra thought. In the ionized condition $A^-$ of the acceptor, each of the chemical bonds between the acceptor atom and the surrounding semiconductor atoms contains a pair of electrons with antiparallel spins. There is only one such state, hence the ionized condition contributes only one term, $\exp[(\mu - \varepsilon_a)/\tau]$, to the Gibbs sum for the acceptor. In the neutral condition $A$ of the acceptor, one electron is missing from the surrounding bonds. Because the missing electron may have either spin up or spin down, the neutral condition is represented twice in the Gibbs sum for the acceptor, by a term $2 \times 1 = 2$.

Hence the thermal average occupancy is

$$f(A^-) = \frac{\exp[(\mu - \varepsilon_a)/\tau]}{2 + \exp[(\mu - \varepsilon_a)/\tau]} = \frac{1}{1 + 2 \exp[(\mu - \varepsilon_a)/\tau]}.$$  \hspace{1cm} (42)

The neutral condition $A$, with the acceptor orbital unoccupied, occurs with probability

$$f(A) = \frac{2}{2 + \exp[(\mu - \varepsilon_a)/\tau]} = \frac{1}{1 + \frac{1}{2} \exp[(\mu - \varepsilon_a)/\tau]}.$$ \hspace{1cm} (43)

The value of $\Delta n \equiv n_a^+ - n_a^-$ is the difference of concentrations of $D^+$ and $A^-$. From (40) or (42) we have

$$n_a^+ = n_a f(D^+) = \frac{n_a}{1 + 2 \exp[(\mu - \varepsilon_a)/\tau]},$$ \hspace{1cm} (44)

$$n_a^- = n_a f(A^-) = \frac{n_a}{1 + 2 \exp[(\mu - \varepsilon_a)/\tau]}.$$ \hspace{1cm} (45)

The neutrality condition (4) may be rewritten as

$$n^+ \equiv n_e + n_a^- = n_h + n_a^+ \equiv n^+.$$ \hspace{1cm} (46)

This expression may be visualized by a logarithmic plot of $n^-$ and $n^+$ as functions of the position of the Fermi level (Figure 13.6). The four dashed lines represent the four terms in (46); the two solid lines represent the sum of all positive and all negative charges. The actual Fermi level occurs where the total positive charges equal the total negative charges.

For $n_a^+ - n_a^- \gg n_i$, as in Figure 13.6, the holes can be neglected; for $n_a^- - n_a^+ \gg n_i$, the electrons can be neglected. If one of the two impurity species can be neglected, the majority carrier concentration can be calculated in closed form. Consider an $n$-type semiconductor with no acceptors. The
neutrality point in Figure 13.6 is now given by the intersection point of the $n^+$ curve with the $n_0$ curve. If the donor concentration is not too high, the intersection will be on the straight portion of the $n_e$ curve, along which the approximation (17) holds. We rewrite this as

\[
\exp(\mu/\tau) = \frac{n_e}{n_i} \exp(\epsilon_e/\tau); \tag{47}
\]

\[
\exp[(\mu - \epsilon_d)/\tau] = \frac{n_e}{n_i} \exp[(\epsilon_e - \epsilon_d)/\tau] = \frac{n_e}{n_e^*}. \tag{48}
\]
where

\[ n_e^* = n_e \exp\left[ -\frac{e_c - e_d}{\tau} \right] = n_e \exp\left( -\Delta e_d/\tau \right) \]  

is the electron concentration that would be present in the conduction band if the Fermi level coincided with the donor level. Here \( \Delta e_d \equiv e_c - e_d \) is the donor ionization energy.

We insert (48) into (44) and set \( n_e = n_d^* \) to obtain

\[ n_e = \frac{n_d}{1 + (2n_e/n_e^*)}, \]  

\[ n_e^2 + \frac{1}{2}n_en_e^* = \frac{1}{2}n_en_e^*. \]  

This is a quadratic equation in \( n_e \); the positive solution is

\[ n_e = \frac{1}{2}n_e^* \left[ 1 + \sqrt{1 + 8(n_d/n_e^*)} \right]. \]  

For shallow donor levels, \( n_e^* \) is large and close to \( n_e \). If the doping is sufficiently weak that \( 8n_d \ll n_e^* \), the square root may be expanded by use of

\[ (1 + x)^{1/2} \approx 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \cdots, \]  

for \( x \ll 1 \). With \( x = 8n_d/n_e^* \) we obtain

\[ n_e \approx n_d - 2n_d^2/n_e^* = n_d(1 - 2n_d/n_e^*). \]  

The second term in the parentheses gives the first order departure from complete ionization. For example, for P in Si at 300 K, we have \( \Delta e_d \approx 1.74 \tau \) from Table 13.2, so that \( n_e^* \approx 0.175n_e \) from (49). If \( n_d = 0.01n_e \), Eq. (54) predicts that 11.4 pct of the donors remain un-ionized. The limit of weak ionization is the subject of Problem 6.

Example: Semi-insulating gallium arsenide. Could pure GaAs be prepared, it would have an intrinsic carrier concentration at room temperature of \( n_i \approx 10^9 \text{cm}^{-3} \). With such a low concentration of carriers, \( 10^{-15} \) less than a metal, the conductivity would be closer to an insulator than to a conventional semiconductor. Intrinsic GaAs would be useful as an insulating substrate on which to prepare thin layers of doped GaAs as needed for devices. There does not exist a technology to purify any substance to \( 10^7 \) impurities per \( \text{cm}^3 \).