14

Carrier concentrations

The *activation energy* of impurities will be frequently used in this chapter. It is useful to recall the interdependence of free energy, internal energy, enthalpy, entropy, and activation energy. To do so, consider the electronic ionization of an impurity, for example a donor

$$D^0 \leftrightarrow D^+ + \mathrm{e} \ . \qquad (14.1)$$

The effective work necessary to accomplish the ionization process at a constant temperature and pressure equals to the change of ***Gibbs free energy*** of the system (Kittel and Kroemer, 1980; Reif, 1965). In thermodynamics, Gibbs free energy G is defined as

$$G = H - T\,S \ , \qquad (14.2)$$

$$H = E + P\,V \qquad (14.3)$$

where H is the reaction enthalpy, S the entropy, E the internal energy, and PV the product of pressure and volume of the system. The change in Gibbs free energy occurring during the donor ionization process of Eq. (14.1) at a constant temperature T is then given by

$$\Delta G = \Delta H - T\Delta S \ , \qquad (14.4)$$

$$\Delta H = \Delta E + P\,\Delta V \qquad (14.5)$$

where constant temperature and constant pressure is assumed. Gibbs free energy is the proper energy to be used in a Boltzmann factor or Fermi function (see Sect. on *semiconductor statistics*). The change in volume of the system occurring during chemical reactions can be quite significant. However, the change in volume during the electronic reaction of Eq. (14.1) is very small since the valence electron configuration does not change. The change in volume can therefore be neglected. In this chapter, the change in entropy as well as the mechanical work ($P\,\Delta V$) will be neglected. In this case, it is $\Delta G \approx \Delta H \approx \Delta E$. The energy required for the ionization reaction of Eq. (14.1) is the difference in internal energy, *i. e.* the difference in energy of states occupied by the electron before and after the ionization process. The change in free energy for donors can then be written as $\Delta G \approx \Delta H \approx \Delta E = E_C - E_D = E_d$, that is, the ionization energy equals the donor level energy relative to the bottom of the conduction band. The enthalpy and the entropy of ionization of centers in semiconductors were further considered by Thurmond (1975) and by Van Vechten and Thurmond (1976a, 1976b). The authors made simple estimates of the entropy of ionization of coulombic, isoelectronic, and vacancy-type defects in semiconductors by considering the effect of localized and free-carrier charge distributions upon the lattice modes. The empirical values of these entropies are observed as the temperature variation of the corresponding ionization levels (*i. e.* the term $T\,\Delta S$ in Eq. 14.4). The change in entropy during the ionization reaction of Au-related levels in Si was considered by Lang *et al*. (1980), who differentiated between the entropy change due to electronic degeneracy and due to atomic

© E. F. Schubert

vibrational changes. The authors showed that the change in entropy can be a small fraction (10 %) of the ionization enthalpy.

Typical densities of free carriers in semiconductors range from 10^{15} cm^{-3} to 10^{20} cm^{-3}. It is impossible to describe the energies or velocities of those carriers individually. An alternative to the individual characterization of particles is the *statistical* description of a carrier system. The statistical description uses *probabilities* of velocities or energies rather than knowing these quantities for all individual carriers. Thus, the statistical treatment represents a simplification. The derivation of the energy distribution function treats the carrier system as an *ideal gas*, for example a gas of oxygen molecules. The ideal gas is assumed to have only *elastic* collisions between atoms or molecules. Furthermore, the energy of the gas molecules is assumed to be purely translational *kinetic*. Since these properties are applied to the electron or hole system, those systems are frequently referred to as *electron-gases* or *hole-gases*.

The free carrier concentration in semiconductors depends on a number of parameters such as the doping concentration, impurity activation energy, temperature, and other parameters. Given the results of the previous sections on the density of states and the distribution functions, the carrier concentration can now be calculated. In the calculation intrinsic, extrinsic, and compensated semiconductors will be considered.

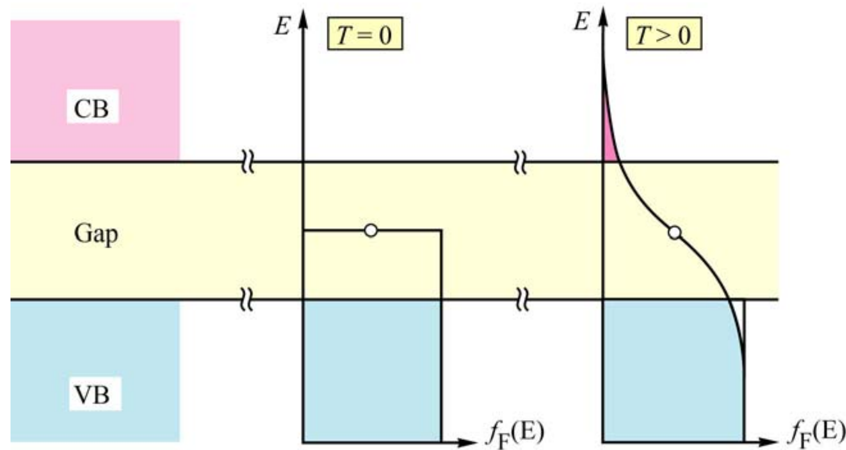

Fig. 14.1. Carrier distributions in the conduction band (CB) and valence band (VB) and Fermi distribution function in an intrinsic semiconductor at T = 0 K and at finite temperatures $T > 0$ K.

14.1 Intrinsic semiconductors

The carrier concentration of pure, undoped semiconductors is determined by thermal excitation of electrons from the valence band to states in the conduction band. An intrinsic semiconductor has a filled valence band and an empty conduction band at zero temperature. This property is the very definition of semiconductors. The band diagram along with the $T = 0$ K Fermi distribution function is shown in ***Fig***. 14.1.

As the temperature increases, a small fraction of electrons in the valence band is excited into the conduction band. Thus the number of holes (unoccupied states) p in the valence band coincides with the number of electrons n in the conduction band. Semiconductors for which $n = p$ are called *intrinsic*. The condition that the concentration of electrons coincides with the concentration of holes requires that the Fermi energy be within the forbidden gap. The position of the Fermi energy in the gap is visualized in ***Fig***. 14.1. The electron and hole concentrations are given by

$$n = p \ , \tag{14.6}$$

$$N_c \, F_{1/2}(\eta_F) = N_v \, F_{1/2}(\eta_F) \ . \tag{14.7}$$

© E. F. Schubert

Since the Fermi energy is within the forbidden gap, *i. e.* many values of kT below the conduction band and many values of kT above the valence band, simpler Boltzmann statistics can be used instead of Fermi–Dirac statistics. Equation (14.7) then simplifies to

$$N_c \exp\left(-\frac{E_C - E_F}{kT}\right) = N_v \exp\left(-\frac{E_F - E_V}{kT}\right) . \quad (14.8)$$

Rearrangement of the equation and the definition of the gap energy $E_g = E_C - E_V$ yields for the Fermi energy of an *intrinsic* semiconductor

$$E_F = E_V + \frac{1}{2} E_g + \frac{kT}{2} \ln \frac{N_v}{N_c} . \quad (14.9)$$

In this equation, $E_V + (1/2)\,E_g$ represents the mid-gap energy. Since the logarithmic function changes weakly with N_v / N_c, the Fermi energy of an intrinsic semiconductor is approximately at mid-gap. The temperature dependence of the intrinsic Fermi energy is weak due to the (weak) logarithmic dependence of the Fermi energy on the temperature. Using Boltzmann statistics the Fermi energy allows us to determine the ***intrinsic carrier concentration***, n_i, of electrons and holes in an undoped semiconductor.

$$\boxed{n_i = \sqrt{N_v\, N_c}\, \exp\left(-\frac{E_g}{2\,kT}\right)} \quad (14.10)$$

According to this equation the intrinsic carrier concentration increases exponentially with temperature. In addition, the effective density of states have the comparatively weak temperature dependence of $N_{c,v} \propto T^{3/2}$. The intrinsic carrier concentration is of special importance. Calculating the product of electron and hole concentration for *any* (non-degenerate) Fermi level using Boltzmann statistics yields

$$n\, p = n_i^2 = N_v\, N_c \exp\left(-\frac{E_g}{kT}\right) . \quad (14.11)$$

Thus the product $n\,p$ is a constant at a given temperature and, since the result does not depend on the Fermi level, is independent of the doping concentration. The intrinsic carrier concentrations of GaAs and InP are shown as a function of temperature in ***Fig***. 14.2.

© E. F. Schubert

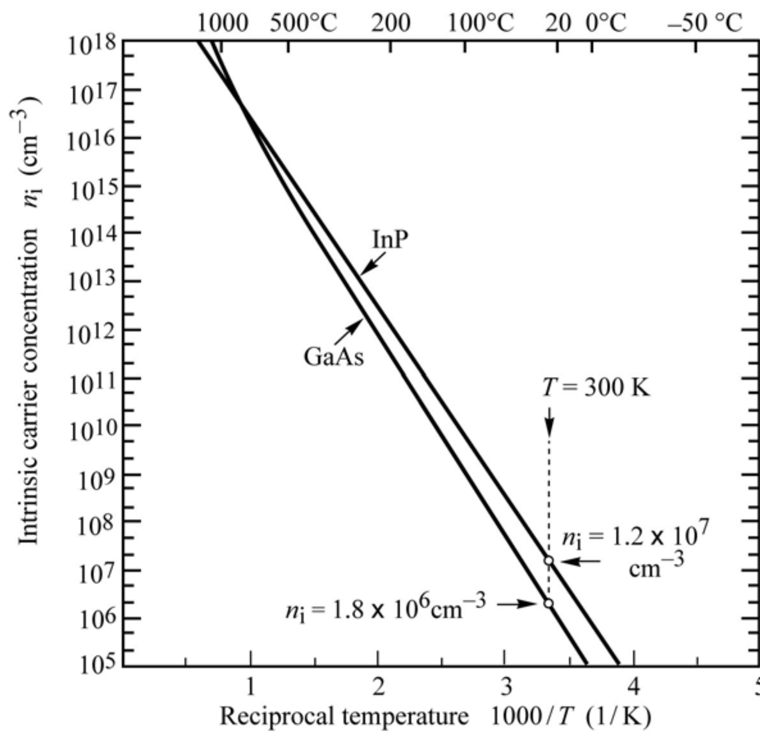

Fig. 14.2. Intrinsic carrier concentrations of GaAs and InP as a function of temperature. The slopes of the curves are proportional to the bandgap energy (after Thurmond, 1975; Laufer *et al*., 1980).

14.2 Extrinsic semiconductors (single donor species)

Substitutional donors and acceptors have an excess or a deficit electron in their outer electron shell, respectively, as compared to the replaced lattice atom. ***Donors*** have one excess electron which can be *donated* to the conduction band. ***Acceptors*** have one less electron than the replaced lattice atom and can *accept* an electron from the filled valence band of the semiconductor, thereby creating a *hole*. Here we consider donors and acceptors being represented by an energy state close to the conduction band edge (donor) or close to the valence band edge (acceptor), as shown in ***Fig***. 14.3. We will next investigate the free carrier concentration as a function of temperature in a semiconductor with donor impurities of *one* chemical species. The donor concentration is assumed to be N_D.

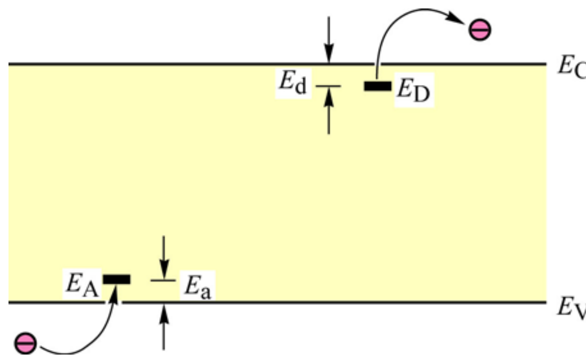

Fig. 14.3. Energy levels of acceptors and donors in the semiconductor band diagram. A donor (acceptor) level at energy E_D (E_A) has an ionization energy E_d (E_a).

The charge state of donors is *neutral* when occupied by an electron and positively charged if the electron is excited to the conduction band. The total concentration of donors is the sum of neutral donor concentration and ionized donor concentration, *i. e.*

$$N_D = N_D^0 + N_D^+ . \tag{14.12}$$

© E. F. Schubert

The energy of the donor impurity state is denoted as E_D. The donor energy is frequently given with respect to the conduction band edge, that is

$$E_d = E_C - E_D \, . \tag{14.13}$$

The probability of occupation of an acceptor or donor follows Fermi – Dirac statistics. Consequently, the concentration of neutral donors, *i. e.* donors occupied by an electron is

$$N_D^0 = N_D \, f_F(E_D) \tag{14.14}$$

where $f_F(E_D)$ is the value of the Fermi–Dirac distribution at the energy of the donor. With $N_D^+ = N_D - N_D^0$ one obtains the concentration of ionized donors

$$N_D^+ = N_D \left[1 - f_F(E_D)\right] = N_D \left[1 - \left(1 + \frac{1}{g} \exp\left(\frac{E_D - E_F}{kT}\right)\right)^{-1}\right] = N_D \left[1 + g \, \exp\left(\frac{E_F - E_D}{kT}\right)\right]^{-1} \tag{14.15}$$

where g is the ground-state degeneracy of the donor. The value of the ground-state degeneracy in GaAs is $g = 2$ for hydrogen-like donors since the donor can donate one electron of either spin (see Chap. 1). The ground-state degeneracy of acceptors in GaAs is $g = 4$, since the acceptor can accept electrons of either spin from the heavy-hole and the light-hole valence band (see Chap. 1). Note that Eq. (14.15) is limited to concentrations below the Mott transition (see Chap. 1). Above the Mott transition, impurities cannot bind charge carriers, *i. e.* donors and acceptors cannot be in the neutral charge state.

If a semiconductor has one carrier type dominating due to doping, the other carrier type has an extremely small equilibrium concentration. If, for example, GaAs is doped with $N_D = 10^{17}\,\text{cm}^{-3}$ donors and $n \approx 10^{17}\,\text{cm}^{-3}$, the hole-concentration inferred from Eq. (14.10) at 300 K is $p = n_i^2 / n = 3.2 \times 10^{-3}\,\text{cm}^{-3}$. Thus, there are approximately 3 holes in 1000 cm^3 of this n-type semiconductor. The very small concentration of the ***minority carrier*** allows us to completely neglect minority carriers in many semiconductor structures. Such semiconductor devices are called ***majority carrier devices***.

Charge neutrality is maintained in a doped semiconductor and has to be taken into account in addition to Fermi – Dirac statistics. Since minority carriers can be neglected, the free carrier concentration coincides with the ionized dopant concentration. If we restrict ourselves to n-type semiconductors, then

$$n = N_D^+ \, . \tag{14.16}$$

We now consider the semiconductor at low temperatures, when most electrons occupy donor states. Then Boltzmann statistics can be used for the occupation of conduction band states according to

$$n = N_c \exp\left(-\frac{E_C - E_F}{kT}\right) . \tag{14.17}$$

© E. F. Schubert

If Fermi–Dirac statistics are used for the occupation of the donor level according to Eq. (14.15), one obtains a quadratic equation for the free carrier concentration

$$n^2 - \frac{1}{g} N_D N_c e^{-E_d/kT} + \frac{1}{g} n N_c e^{-E_d/kT} = 0 \, . \tag{14.18}$$

At low temperatures the free carrier concentration, n, is much smaller than the donor concentration, N_D. Thus, the third term of the quadratic equation is much smaller than the second term. The free carrier concentration is given by

$$\boxed{n \approx \left(\frac{1}{g} N_D N_c \right)^{1/2} \exp\left(- \frac{E_d}{2kT} \right)} \tag{14.19}$$

where the ground state degeneracy for donors is $g = 2$.

Equation (14.19) was first obtained by de Boer and van Geel (1935) by the method described here. The formulas can also be obtained by minimizing the free energy change due to thermal excitation of electrons from donor states to conduction band states (Mott and Gurney, 1940). At higher temperatures all donors become ionized. The carrier concentration is then constant $n = N_D^+ = N_D$ and independent of temperature. This temperature regime is called the ***saturation regime***.

As the temperature is increased even further, the *intrinsic* carrier concentration n_i increases and at sufficiently high temperatures assumes values comparable or higher than the dopant concentration. For most technologically useful semiconductors, the crossover from the saturation to the ***intrinsic regime*** occurs at temperatures much higher than room temperature. The three temperature regimes *(i)* thermal ionization regime *(ii)* saturation regime and *(iii)* intrinsic regime are shown schematically in ***Fig***. 14.4 along with the associated activation energies.

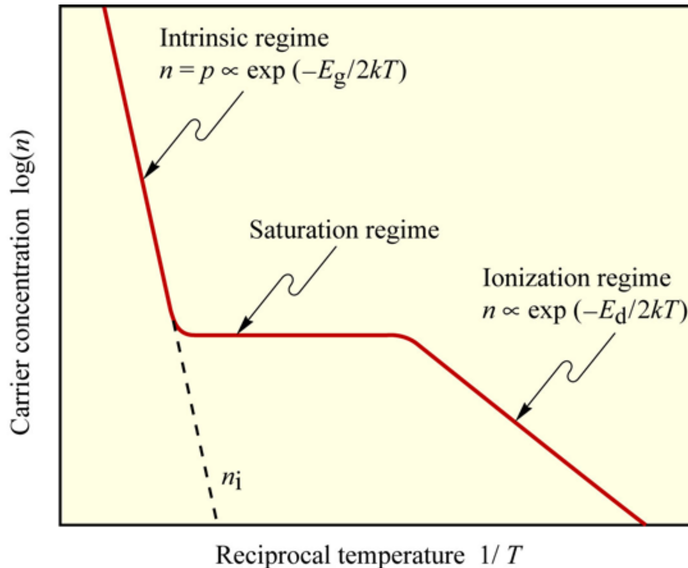

Fig. 14.4. Carrier concentration as a function of reciprocal temperature for an uncompensated n-type semiconductor. The donor is assumed to form one level at energy E_d below the conduction band edge. Three regimes, namely (*i*) the ionization regime, (*ii*) the saturation regime, and (*iii*) the intrinsic regime and their corresponding activation energies are indicated (after Smith, 1986).

The thermal ionization energy of a donor can be obtained from the slope of n versus reciprocal temperature. Rearrangement of Eq. (14.19) yields

$$E_d = -2k \frac{d(\ln n)}{d(1/T)} \tag{14.20}$$

© E. F. Schubert

which allows one to determine E_d directly from the temperature dependent carrier concentration. The change in carrier concentration with increasing temperatures also implies a continuously changing Fermi level in the semiconductor. Consider an n-type semiconductor. At low temperatures the donor levels are filled, while the conduction band is empty. Thus, the Fermi level must be slightly above the donor level. As the temperature increases, the Fermi distribution becomes smeared out, as conduction band states become filled and donor states become unoccupied. Simultaneously the Fermi level moves deeper into the forbidden gap. At still higher temperatures, the Fermi level approaches the (near) mid-gap level and the semiconductor becomes intrinsic.

14.3 Extrinsic semiconductors (two donor species)

In the following, the free carrier concentration as a function of temperature is investigated in a semiconductor with two different species of donor impurities. It is assumed that the two donors form two different energy levels in the gap of the semiconductor. The two types of donor levels can originate from two different chemical species (*e. g.* Sn and Te donors in GaAs).

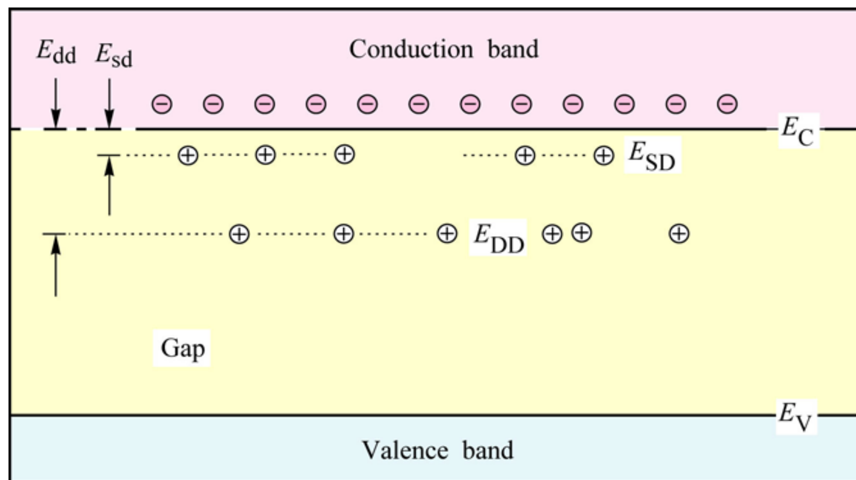

Fig. 14.5. Band diagram of a semiconductor with two species of donors namely a shallow donor with energy E_{SD} (ionization energy E_{sd}) and a deep donor with energy E_{DD} (ionization energy E_{dd}).

If both donor types have very similar thermal ionization energies, there is no necessity to differentiate between the two types of donors. It is therefore assumed that the two types of donors have markedly different ionization energies. In particular, we assume that one of the donors is relatively shallow and the other one is relatively deep. The band diagram of a semiconductor with two types of donors with energies E_{SD} and E_{DD} is shown in ***Fig***. 14.5. If the shallow and deep donor concentrations are given by N_{SD} and N_{DD}, respectively, then the free carrier concentration is given by

$$n = N_{SD}^{+} + N_{DD}^{+} . \tag{14.21}$$

The carrier concentration in the conduction band is given by the Boltzmann distribution

$$n = N_c \exp\left(-\frac{E_C - E_F}{kT}\right) \tag{14.22}$$

while Fermi–Dirac statistics is assumed for the donor levels (see Eq. 14.15).

© E. F. Schubert

$$N_{SD}^{+} = N_{SD} - \frac{N_{SD}}{1 + \frac{1}{g} \exp\left(\frac{E_{SD} - E_F}{kT}\right)} , \tag{14.23}$$

$$N_{DD}^{+} = N_{DD} - \frac{N_{DD}}{1 + \frac{1}{g} \exp\left(\frac{E_{DD} - E_F}{kT}\right)} \tag{14.24}$$

where E_{SD} and E_{DD} are the energies of the donor states. We assume that $E_{sd} << E_{dd}$ and $E_{dd} << E_g$, where $E_{sd} = E_C - E_{SD}$ and $E_{dd} = E_C - E_{DD}$.

All donors are neutral at very low temperatures. As temperature increases, shallow donors will donate their electrons to the conduction band, until all shallow donors are ionized. As the temperature is further increased, the deep donors start to become ionized until all deep donors are ionized. At even higher temperatures the intrinsic carrier concentration exceeds the dopant concentration and the semiconductor becomes intrinsic. In the following, the ionization regimes of the shallow and the deep donor are investigated.

At low temperatures when both types of donors are neutral, the Fermi energy is higher than the shallow donor energy. Then the energy difference between the Fermi energy and the deep donor energy is relatively large and according to Eq. (14.24), the deep donor can be considered as neutral, *i. e.* $N_{DD}^{+} = 0$. The carrier concentration is then given by

$$n = N_{SD}^{+} = N_{SD} - \frac{N_{SD}}{1 + \frac{1}{g} \exp\left(\frac{E_{SD} - E_F}{kT}\right)} . \tag{14.25}$$

Using Boltzmann statistics for the conduction band one obtains the quadratic equation

$$n^2 - \frac{1}{g} N_{SD} N_c \exp(-E_{sd}/kT) + \frac{1}{g} n N_c \exp(-E_{sd}/kT) = 0 \tag{14.26}$$

which is identical to the single donor equation Eq. (14.18). Thus, the low temperature solution is

$$\boxed{n \approx \left(\frac{1}{2} N_{SD} N_c\right)^{1/2} \exp\left(-\frac{E_{sd}}{2kT}\right)} \tag{14.27}$$

where the ground-state degeneracy is assumed to be $g = 2$. Ionization of the shallow donor continues until all shallow donors are ionized, *i. e.* $n = N_{SD}^{+} = N_{SD}$.

As the temperature is increased further, deep donors become ionized. Using Boltzmann statistics for the conduction band (Eq. 14.22) and Fermi–Dirac statistics for the deep donor (Eq. 14.24) one obtains the quadratic equation

$$(N_{DD}^{+})^2 + N_{DD}^{+}\left(N_{SD}^{+} + \frac{1}{g} N_c e^{-E_{dd}/kT}\right) - \frac{1}{g} N_{DD} N_c e^{-E_{dd}/kT} = 0 . \tag{14.28}$$

For $N_{SD}^{+} >> (g^{-1}) N_c \exp(-E_{dd}/kT)$ one obtains

© E. F. Schubert

$$(N_{DD}^{+})^2 + N_{SD}^{+} N_{DD}^{+} \approx \frac{1}{g} N_{DD} N_c e^{-E_{dd}/kT} . \tag{14.29}$$

Since the free carrier concentration is the sum of ionized deep and shallow donor concentration ($n = N_{SD}^{+} + N_{DD}^{+}$) the equation can be written as

$$n(n - N_{SD}^{+}) \approx \frac{1}{2} N_{DD} N_c e^{-E_{dd}/kT} \tag{14.30}$$

where the donor ground-state degeneracy is assumed to be $g = 2$.

At the elevated temperatures considered here, the shallow donor is ionized ($N_{SD}^{+} = N_{SD}$) and therefore the slope of the carrier density with respect to temperature follows the proportionality

$$\boxed{n\left(n - N_{SD}\right) \propto e^{-E_{dd}/kT}} \tag{14.31}$$

Note that this equation is significantly different from the simple relation $n \propto \exp(-E_d / kT)$, which would lead to incorrect results if applied to a semiconductor with two types of donors. Equation (14.31) was applied to shallow and deep Si donors in $Al_xGa_{1-x}As$ (Schubert and Ploog, 1984).

The ionization of the deep donor continues until shallow and deep donors are ionized, which corresponds to the saturation regime. At even higher temperatures the intrinsic carrier concentration increases above the dopant concentration and the semiconductor becomes intrinsic. The carrier concentration is shown in ***Fig***. 14.6 for a semiconductor containing two different donor species as a function of temperature. The different saturation and ionization regimes along with their activation energies are indicated in the figure. Note that the different ionization regimes discussed above may not be as clearly distinguishable if the difference between E_{sd} and E_{dd} is small.

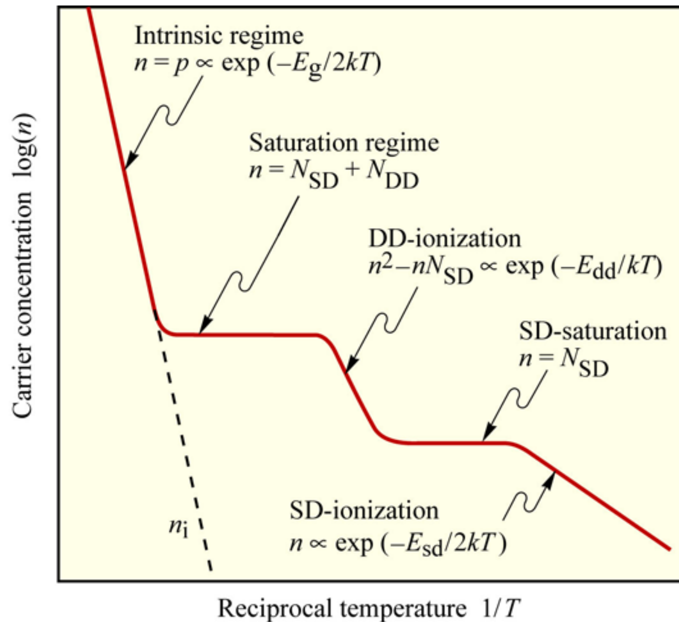

Fig. 14.6. Carrier concentration as a function of reciprocal temperature for a semiconductor with two different dopant species, namely a shallow donor (SD) and a deep donor (DD). At low temperatures the shallow donor becomes ionized and saturates when $n = N_{SD}$. At higher temperatures the deep donor becomes ionized and saturates ($n = N_{SD} + N_{DD}$). At even higher temperatures the semiconductor becomes intrinsic ($n = n_i$). The activation energies of each regime are indicated.

© E. F. Schubert

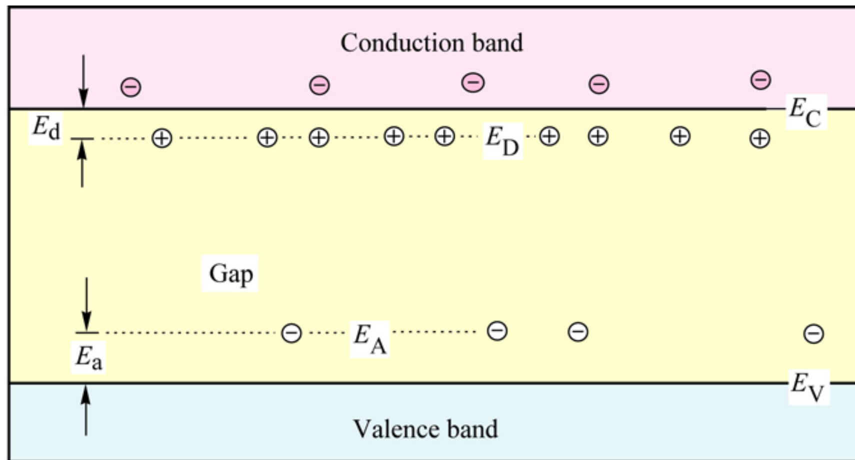

Fig. 14.7. Band diagram of a lightly compensated n-type semiconductor.

14.4 Compensated semiconductors

A partially compensated semiconductor contains dopant atoms of one type (n- or p-type) and, in addition, a smaller number of dopants of the other type. The band diagram of a partially compensated n-type semiconductor with a small number of acceptors is shown in ***Fig***. 14.7. The free carrier concentration is given by

$$n \; = \; N_D^+ - N_A^- \; . \tag{14.32}$$

Electrons from donors prefer to occupy lower-energy acceptor states at all temperatures. Thus, even for low temperatures ($T \to 0\,\text{K}$) *all* acceptors and *some* donors are ionized. Fermi–Dirac statistics for the occupation of the donor state (see Eq. 14.14) and Eq. (14.32) yield

$$n + N_A^- \; = \; N_D^+ \; = \; N_D - \frac{N_D}{1 + \frac{1}{g}\exp\left(\frac{E_D - E_F}{kT}\right)} \; . \tag{14.33}$$

Using Boltzmann statistics for the free electron concentration in the conduction band allows one to eliminate the Fermi energy. One obtains the quadratic equation

$$n^2 + n\left(N_A + \frac{1}{g} N_c \, e^{-E_d/kT}\right) - \frac{1}{g} N_c \, e^{-E_d/kT} \left(N_D - N_A\right) \; = \; 0 \tag{14.34}$$

where acceptors are assumed to be ionized at all temperatures, *i. e.* $N_A = N_A^-$. The solution of the quadratic equation is, for $g = 2$:

$$n \; = \; \frac{-1}{2}\left(N_A + \frac{N_c}{2} e^{-E_d/kT}\right) + \frac{1}{2}\left[\left(N_A + \frac{N_c}{2} e^{-E_d/kT}\right)^2 + 2N_c e^{-E_d/kT}\left(N_D - N_A\right)\right]^{\frac{1}{2}} \tag{14.35}$$

At low temperatures when $(1/2)\, N_c \exp(E_d / kT) << N_A$ the equation simplifies to

© E. F. Schubert

$$n = -\frac{1}{2} N_A + \frac{1}{2} N_A \left[1 + 2 N_c \, e^{-E_d / kT} \left(\frac{N_D - N_A}{N_A{}^2} \right) \right]^{1/2} . \qquad (14.36)$$

An approximate solution of the equation can be found by applying $(1 + x)^{1/2} \approx 1 + (1/2)\, x$ (for $x << 1$) to the square root term of the equation.

$$n \approx \frac{1}{2} N_c \frac{N_D - N_A}{N_A} e^{-E_d / kT} \qquad (14.37)$$

Note that the temperature dependence of the carrier concentration as a function of temperature has a different slope as compared with the uncompensated semiconductor (see Eq. 14.19). The slopes are different by a factor of two for the compensated and uncompensated case.

As the temperature is further increased, $N_D >> (1/2) N_c \exp(-E_d / kT) >> N_A$ and Eq. (14.35) simplifies to

$$n \approx \left(\frac{1}{2} N_c \, N_D \right)^{1/2} \exp\left(-\frac{E_d}{2\,k\,T} \right) \qquad (14.38)$$

which is identical to the uncompensated case given by Eq. (14.19).

Even further increase of the temperature results in fully ionized donors. The free carrier concentration is then given by $n = N_D^+ - N_A^- = N_D - N_A$. The two different slopes for the carrier concentration vs. temperature have indeed been observed experimentally (Morin, 1959). The carrier concentration versus temperature of a compensated semiconductor is schematically illustrated in ***Fig***. 14.8, where the different regimes are indicated.

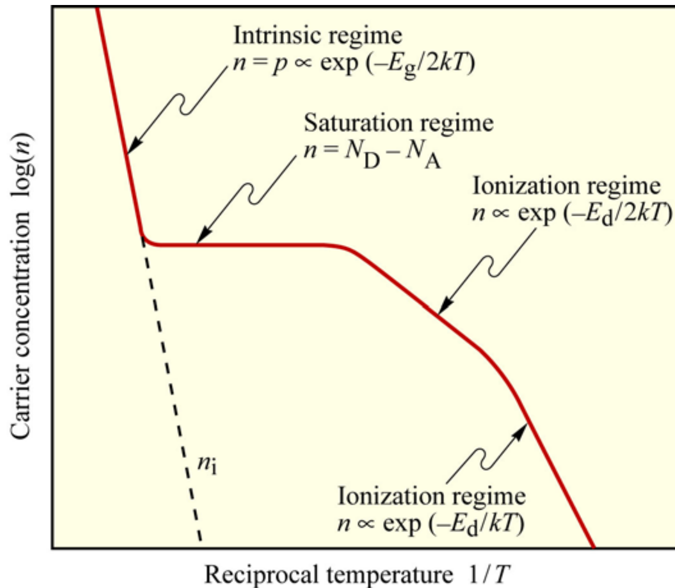

Fig. 14.8. Carrier concentration as a function of reciprocal temperature for a partially compensated n-type semiconductor. The ionization regime is characterized by two different activation energies, namely E_d and $E_d/2$.

© E. F. Schubert