

Section 12: Semiconductors

Crystal structure and bonding

Semiconductors include a large number of substances of widely different chemical and physical properties. These materials are grouped into several classes of similar behavior, the classification being based on the position in the periodic table of the elements.

The best-known class is the Group IV semiconductors - C (diamond), Si, Ge, - all of which lie in the fourth column of the periodic table. They have been studied intensively, particularly Si and Ge, which have found many applications in electronic devices. The elemental semiconductors all crystallize in the diamond structure. The diamond structure has an fcc lattice with a basis composed of two identical atoms, and is such that each atom is surrounded by four neighboring atoms, forming a regular tetrahedron. Group IV semiconductors are covalent crystals, i.e., the atoms are held together by covalent bonds. These bonds consist of two electrons of opposite spins distributed along the line joining the two atoms. The covalent electrons forming the bonds are hybrid sp^3 atomic orbitals.

Another important group of semiconductors is the Group III-V compounds, so named because each contains two elements, one from the third and the other from the fifth column of the periodic table. The best-known members of this group are GaAs and InSb (indium antimonite), but the list also contains compounds such as GaP, InAs, GaSb, and many others.

These substances crystallize in the zincblende structure which is the same as the diamond structure, except that the two atoms forming the basis of the lattice are now different. Thus, in GaAs, the basis of the fcc lattice consists of two atoms, Ga and As. Because of this structure, each atom is surrounded by four others of the opposite kind, and these latter atoms form a regular tetrahedron, just as in the diamond structure.

The bonding in the III-V compounds is also primarily covalent. The eight electrons required for the four tetrahedral covalent bonds are supplied by the two types of atoms, the trivalent atom contributing its three valence electrons, and the pentavalent atom five electrons. The bonding in this group is not entirely covalent. Because the two elements in the compound are different, the distribution of the electrons along the bond is not symmetric, but is displaced toward one of the atoms. As a result, one of the atoms acquires a net electric charge. Such a bond is called *heteropolar*, in contrast to the purely covalent bond in the elemental semiconductors, which is called *homopolar*.

The distribution of electrons in the bond is displaced toward the atom of higher *electronegativity*. In GaAs, for instance, the As atom has a higher electronegativity than the Ga, and consequently the As atom acquires a net negative charge, whose value is $-0.46e$ per atom (a typical value in Group III-V compounds). The Ga atom correspondingly acquires a net positive charge of $0.46e$. Charge transfer leads to an ionic contribution to the bonding in Group III-V compounds. Their bonding is therefore actually a mixture of covalent and ionic components, although covalent ones predominate in most of these substances.

Band structure

A semiconductor is a solid in which the highest occupied energy band, the *valence band*, is completely full at $T = 0^\circ\text{K}$, but in which the gap above this band is also small, so that electrons may

be excited thermally at room temperature from the valence band to the next-higher band, which is known as the *conduction band*. Generally speaking, the number of excited electrons is appreciable (at room temperature) whenever the energy gap E_g is less than 2 eV. The substance may then be classified as a semiconductor. When the gap is larger, the number of electrons is negligible, and the substance is an insulator.

When electrons are excited across the gap, the bottom of the conduction band (CB) is populated by electrons, and the top of the valence band (VB) by holes. As a result, both bands are now only partially full, and would carry a current if an electric field were applied. The conductivity of the semiconductor is small compared with the conductivities of metals of the small number of electrons and holes involved, but this conductivity is nonetheless sufficiently large for practical purposes.

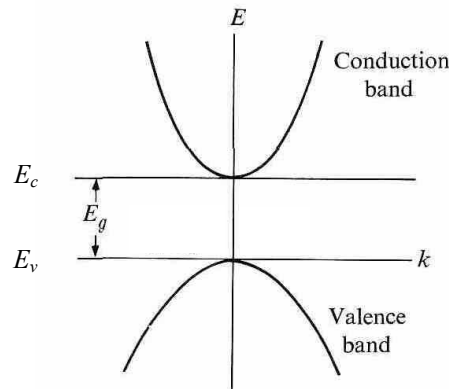


Fig. 2 Band structure in a semiconductor.

The simplest band structure of a semiconductor is indicated in Fig.2. Since we are interested only in the region which lie close to the band gap, where electrons and holes lie, we can ignore a more complex variation of the energy bands far away from the gap.

The energy of the CB has the form

$$E_c(k) = E_c + \frac{\hbar^2 k^2}{2m_e} \quad (1)$$

where k is the wave vector and m_e the effective mass of the electron. The energy E_g represents the energy gap. The zero-energy level is chosen to lie at the top of the VB.

The energy of the VB (Fig. 2) may be written as

$$E_v(k) = E_v - \frac{\hbar^2 k^2}{2m_h} \quad (2)$$

where m_h is the effective mass of the hole which is positive. (Because of the inverted shape of the VB, the mass of an electron at the top of the VB is negative, but the mass of a hole is positive.)

Within this simple picture of the semiconductor, the primary band-structure parameters are thus the electron and hole masses m_e and m_h , and the band gap E_g . Table 1 gives these parameters for

various semiconductors. Note that the masses differ considerably from the free-electron mass. In many cases they are much smaller than the free-electron mass. The energy gaps range from 0.18 eV in InSb to 3.7 eV in ZnS. The table also shows that the wider the gap, the greater the mass of the electron.

Table 1. Band Structure parameters of Semiconductors

Group	Crystal	$E_g(\text{eV})$	m_e/m	m_h/m
IV	C	5.3		
IV	Si	1.1	0.19	0.16
IV	Ge	0.7	0.08	0.04
III-V	GaAs	1.4	0.07	0.09
III-V	GaP	2.3	0.12	0.50
III-V	InSb	0.2	0.01	0.18
II-VI	ZnS	3.6	0.40	5.41
II-VI	ZnSe	2.7	0.1	0.6
II-VI	CdSe	1.7	0.13	0.45

The energy gap for a semiconductor varies with temperature, but the variation is usually slight. That a variation with temperature should exist at all can be appreciated from the fact that the crystal, when it is heated, experiences a volume expansion, and hence a change in its lattice constant. This, in turn, affects the band structure, which is a sensitive function of the lattice constant.

The band structure in Fig.2 is the simplest possible structure. Band structures of real semiconductors are somewhat more complicated, as we shall see, but for the present the simple structure will suffice for our purposes.

Intrinsic Semiconductors

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. *Intrinsic semiconductors* are semiconductors in which the number of carries and the conductivity is not influenced by impurities. Intrinsic conductivity is typical at relatively high temperatures in highly purified specimens.

In order to determine the number of carriers, we need some of the basic results of statistical mechanics. The most important result in this regard is the Fermi-Dirac (FD) distribution function

$$f(E) = \frac{1}{e^{(E-\mu)/k_B T} + 1} \quad (3)$$

This function, gives the probability that an energy level E is occupied by an electron when the system is at temperature T .

The function is plotted versus E in Fig.3. 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 temperature raises the overall energy of the system.

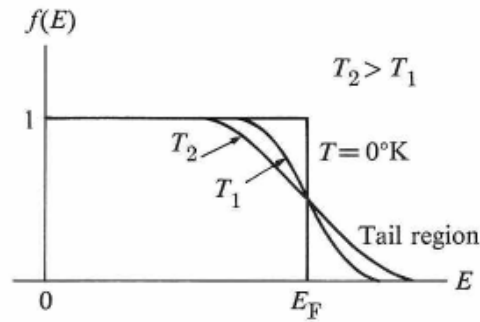


Fig. 3 The Fermi-Dirac distribution function.

We will see later that the Fermi level in intrinsic semiconductors lies close to the middle of the band gap. Therefore we can represent the distribution function and the conduction and valence bands of the semiconductor as shown in Fig.4.

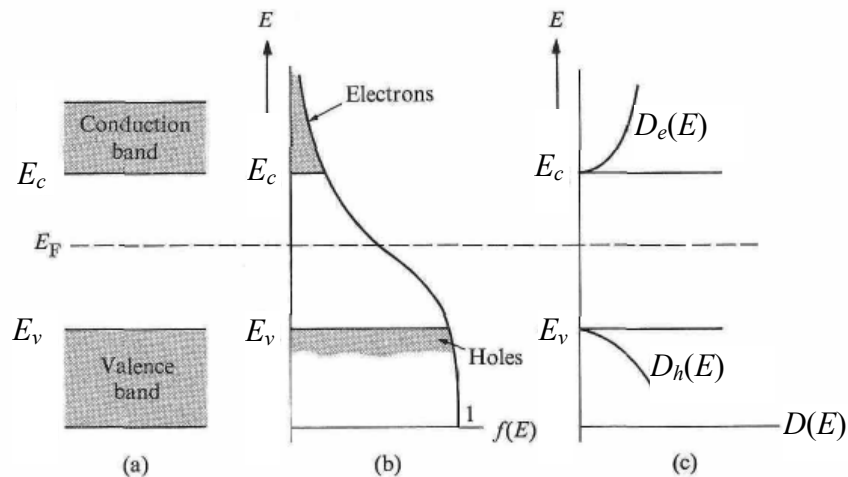


Fig. 4 (a) Conduction and valence bands, (b) The distribution function, (c) Density of states for electrons and holes.

First we calculate the concentration of electrons in the CB. The number of states in the energy range $(E, E + dE)$ is equal to $D_e(E)dE$, where $D_e(E)$ is the density of electron states. Since each of these states has an occupation probability $f(E)$, the number of electrons actually found in this energy range is equal to $f(E)D_e(E)dE$. The concentration of electrons throughout the CB is thus given by the integral over the conduction band

$$n = \int_{E_c}^{\infty} f_e(E) D_e(E) dE \quad (4)$$

where E_c is the bottom the conduction band, as shown in Fig. 4.

The band gap in semiconductors is of the order of 1eV, which is much larger than kT . Therefore $(E-\mu) \gg k_B T$ and we can neglect the unity term in the denominator of the distribution function (3), so that

$$f_e(E) \approx e^{-(E-\mu)/k_B T} . \quad (5)$$

The density of states for the conduction band is given by

$$D_e(E) = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2} . \quad (6)$$

Note that $D_e(E)$ vanishes for $E < E_c$, and is finite only for $E > E_c$, as shown in Fig.4.

When we substitute equations for $f(E)$ and $D_e(E)$ into Eq. (4), we obtain

$$n = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} e^{\mu/k_B T} \int_{E_c}^{\infty} (E - E_c)^{1/2} e^{-E/k_B T} dE . \quad (7)$$

By changing the variable, and using the result

$$\int_0^{\infty} x^{1/2} e^{-x} dx = \frac{\sqrt{\pi}}{2} \quad (8)$$

one can readily evaluate the integral in (7). The electron concentration then reduces to the expression

$$n = 2 \left(\frac{m_e k_B T}{2\pi \hbar^2} \right)^{3/2} e^{(\mu - E_c)/k_B T} \quad (9)$$

The electron concentration is still not known explicitly because the Fermi energy μ 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. Assuming that the Fermi level lies close to the middle of the band gap, i.e. $(\mu-E) \gg k_B T$ for the valence band, we find for the distribution function of holes

$$f_h(E) = 1 - \frac{1}{e^{(E-\mu)/k_B T} + 1} = \frac{1}{e^{(\mu-E)/k_B T} + 1} \approx e^{-(\mu-E)/k_B T} . \quad (10)$$

The density of states for the holes is

$$D_h(E) = \frac{1}{2\pi^2} \left(\frac{2m_h}{\hbar^2} \right)^{3/2} (E_v - E)^{1/2} , \quad (11)$$

where E_v is the energy of the valence band edge. Proceeding in a similar fashion as we did for electrons we find for the concentration of holes in the valence band

$$p = \int_{-\infty}^{E_v} f_h(E) D_h(E) dE = 2 \left(\frac{m_h k T}{2\pi \hbar^2} \right)^{3/2} e^{(E_v - \mu)/k_B T} \quad (12)$$

The electron and hole concentrations have thus far been treated as independent quantities. For intrinsic semiconductors 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,

$$n = p \quad (13)$$

and

$$(m_e)^{3/2} e^{(\mu - E_c)/k_B T} = (m_h)^{3/2} e^{(E_v - \mu)/k_B T} \quad (14)$$

We obtain, then, for the Fermi energy

$$\mu = \frac{E_v + E_c}{2} + \frac{3}{4} k_B T \ln \frac{m_h}{m_e} \quad (15)$$

The second term on the right of (15) is very small compared with the first, and the energy level is close to the middle of the energy gap. 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.

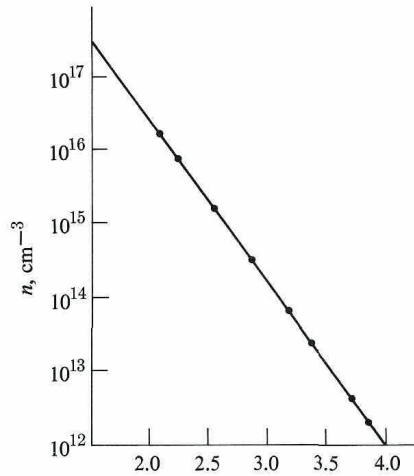


Fig. 5 Electron concentration n versus $1/T$ in Ge (in units 10^{-3}K^{-1}).

The concentration of electrons may now be evaluated explicitly by using the above value of μ . Substitution of Eq. (15) into Eq. (9) yields

$$n = 2 \left(\frac{k_B T}{2\pi \hbar^2} \right)^{3/2} (m_e m_h)^{3/4} e^{-E_g/2k_B T} \quad (16)$$

where $E_g = E_c - E_v$ is the band gap. 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. Figure 5 is a plot of $\log n$ versus $1/T$. The curve is a straight line of slope equal to $(-E_g/2k_B)$. [The $T^{3/2}$ -dependence in (16) is so weak in comparison with the exponential dependence that the former may be disregarded for the purpose of this discussion.)

For Ge at room temperature $n \sim 10^{14}$ electrons/cm³, a typical value of carrier concentration in semiconductors.

Note that the expression (16) also gives the hole concentration, since $n = p$.

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*.

Impurity states

A pure semiconductor has equal numbers of both types of carriers, electrons and holes. In most applications, however, one needs specimens which have one type of carrier only, and none of the other. By doping the semiconductor with appropriate impurities, one can obtain samples which contain either electrons only or holes only.

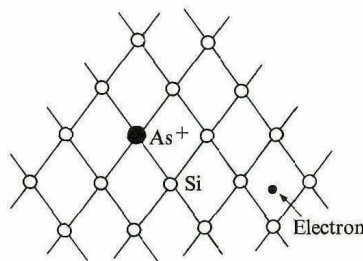


Fig. 6 An As impurity in a Si crystal. The extra electron migrates through the crystal.

Consider, for instance, a specimen of Si which has been *doped* by As. The As atoms (the impurities) occupy some of the lattice sites formerly occupied by the Si host atoms. The distribution of the impurities is random throughout the lattice. But their presence affects the solid in one very important respect: The As atom has valence 5 while Si has valence 4. Of the five electrons of As, four participate in the tetrahedral bond of Si, as shown in Fig. 6. The fifth electron cannot enter the bond, which is now saturated, and hence this electron detaches from the impurity and is free to migrate through the crystal as a conduction electron, i.e., the electron enters the CB. The impurity is

now actually a positive ion, As^+ (since it has lost one of its electrons), and thus it tends to capture the free electron, but we shall show shortly that the attraction force is very weak, and not enough to capture the electron in most circumstances.

The net result is that the As impurities contribute electrons to the CB of the semiconductors, and for this reason these impurities are called *donors*. Note that the electrons have been created *without* the generation of holes.

When an electron is captured by an ionized donor, it orbits around the donor much like the situation in hydrogen. We can calculate the binding energy by using the familiar Bohr model. However, we must take into account the fact that the coulomb interaction here is weakened by the screening due to the presence of the semiconductor crystal, which serves as a medium in which both the donor and ion reside. Thus the coulomb potential is now given by

$$V(r) = -\frac{e^2}{\epsilon r}, \quad (17)$$

where ϵ is the reduced dielectric constant of the medium. The dielectric constant $\epsilon = 11.7$ in Si, for example, shows a substantial decrease in the interaction force. It is this screening which is responsible for the small binding energy of the electron at the donor site.

Using this potential in the Bohr model, we find the binding energy, corresponding to the ground state of the donor, to be

$$E_d = -\frac{e^4 m_e}{2\epsilon^2 \hbar^2}. \quad (18)$$

Note that the effective mass m_e has been used rather than the free mass m . The binding energy of the hydrogen atom, which is equal to 13.6 eV. The binding energy of the donor is reduced by the factor $1/\epsilon^2$, and also by the mass factor m_e/m , which is usually smaller than unity. Using the typical values $\epsilon \sim 10$ and $m_e/m \sim 0.1$, we find that the binding energy of the donor is about 10^{-3} of the hydrogen energy, i.e., about 0.01 eV. This is indeed the order of the observed values.

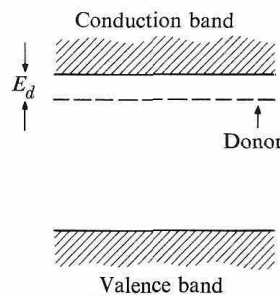


Fig. 7 The donor level in a semiconductor.

The donor level lies in the energy gap, very slightly below the conduction band, as shown in Fig.7. Because the level is so close to the CB, almost all the donors are ionized at room temperature, their electrons having been excited into the CB.

It is instructive to evaluate the Bohr radius of the donor electron. Straightforward adaptation of the Bohr result leads to

$$r_d = \varepsilon \frac{m}{m_e} a_0. \quad (19)$$

where a_0 is the Bohr radius, equal to 0.53 \AA . The radius of the orbit is thus much larger than a_0 , by a factor of 100, if we use the previous values for ε and m_e . A typical radius is thus of the order of 50 \AA . Since this is much greater than the interatomic spacing, the orbit of the electron encloses a great many host atoms, and our picture of the lattice acting as a continuous, polarizable dielectric is thus a plausible one.

Since the donors are almost all ionized, the concentration of electrons is nearly equal to that of the donors. Typical concentrations are about 10^{15} cm^3 . But sometimes much higher concentrations are obtained by heavy doping of the sample, for example, 10^{18} cm^3 or even more.

Acceptors

An appropriate choice of impurity may produce holes instead of electrons. Suppose that the Si crystal is doped with Ga impurity atoms. The Ga impurity resides at a site previously occupied by a Si atom, but since Ga is trivalent, one of the electron bonds remains vacant (Fig. 8). This vacancy may be filled by an electron moving in from another bond, resulting in a vacancy (or hole) at this latter bond. The hole is then free to migrate throughout the crystal. In this manner, by introducing a large number of trivalent impurities, one creates an appreciable concentration of holes, which lack electrons.

The trivalent impurity is called an *acceptor*, because it accepts an electron to complete its tetrahedral bond. The acceptor is negatively charged, by virtue of the additional electron it has entrapped. Since the resulting hole has a positive charge, it is attracted by the acceptor. We can evaluate the binding energy of the hole at the acceptor in the same manner followed above in the case of the donor. Again this energy is very small, of the order of 0.01 eV . Thus essentially all the acceptors are ionized at room temperature.

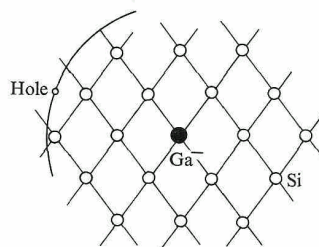


Fig. 8 A Ga impurity in a Si crystal. The extra hole migrates through the crystal.

The acceptor level lies in the energy gap, slightly above the edge of the VB, as shown in Fig. 9. This level corresponds to the hole being captured by the acceptor. When an acceptor is ionized (an electron excited from the top of the VB to fill this hole), the hole falls to the top of the VB, and is now a free carrier. Thus the ionization process, indicated by upward transition of the electron on the energy scale, may be represented by a downward transition of the hole on this scale.

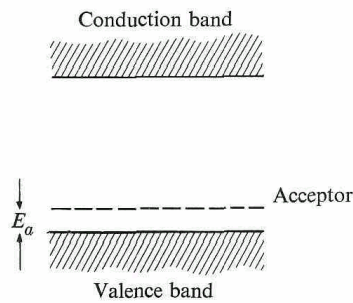


Fig. 9. The acceptor level in a semiconductor.

Semiconductor statistics

Semiconductors usually contain both donors and acceptors. Electrons in the CB can be created either by 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.

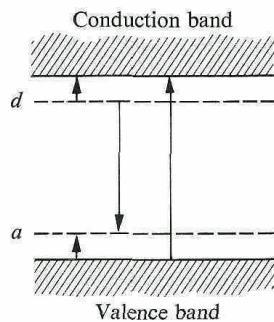


Fig. 10 The various electronic processes in a semiconductor (see text).

Finding the concentrations of carriers, both electrons and holes, taking all these processes into account, is quite complicated. We shall treat a few special cases, 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. In this region $n=p$. 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

$$n \gg N_d, N_a \quad (20)$$

Since n 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 (i.e. the binding energy) is quite small, all the donors are essentially ionized, their electrons going into the CB. Therefore, to a good approximation,

$$n = N_d \quad (21)$$

A semiconductor in which $n \gg p$ is called an *n-type semiconductor* (n for negative). Such a sample is characterized, as we have seen, by a great concentration of electrons.

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

$$p = N_a \quad (22)$$

i.e., all the acceptors are ionized. Such a material is called a *p-type semiconductor*. It is characterized by a preponderance of holes.

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. The temperature at which freeze-out takes place is $E_d \sim kT$, which gives a temperature of about 100°K .

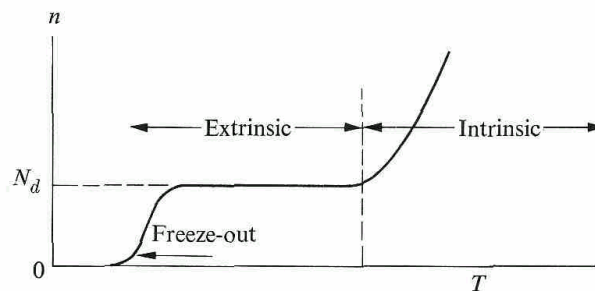


Fig. 11 Variation of electron concentration n with temperature in an n-type semiconductor.

The variation of the electron concentration with temperature in an *n-type* sample is indicated schematically in Fig. 11.

Electrical conductivity; Mobility

Electrical conductivity is, of course, the quantity of primary interest in semiconductors. Both electrons and holes contribute to electric current. Assume first that a sample is strongly n-type and contains only one type of carrier: electrons. The conductivity can be treated according to the free-electron model:

$$\sigma_e = \frac{ne^2\tau_e}{m_e} \quad (23)$$

where m_e is an effective mass and τ_e is the lifetime of the electron. To estimate the value for σ_e , we substitute $n = 10^{14} \text{ cm}^{-3}$, which is eight orders in magnitude less than that in metals, and $m_e = 0.1m$. This leads to $\sigma_e \sim 10^{-7} (\mu\text{ohm}\cdot\text{cm})^{-1}$, which is a typical figure in semiconductors. Although this is many orders of magnitude smaller than the value in a typical metal, where $\sigma_e \sim 1(\mu\text{ohm}\cdot\text{cm})^{-1}$, the conductivity in a semiconductor is still sufficiently large for practical applications.

Semiconductor physicists often use another transport coefficient: *mobility*. The mobility μ_e is defined as the proportionality coefficient between the electron drift velocity and the applied electric field, i.e.

$$|\mathbf{v}_e| = \mu_e E, \quad (24)$$

where $|\mathbf{v}_e|$ is the absolute value of the velocity. Taking into account that $\mathbf{j}_e = -en_e\mathbf{v}_e$ and $\mathbf{j}_e = \sigma_e\mathbf{E}$, we find that

$$\mu_e = \frac{e\tau_e}{m_e}. \quad (25)$$

As defined, the mobility is a measure of the rapidity of the motion of the electron in the field. The longer the lifetime of the electron and the smaller its mass, the higher the mobility.

We can now express electrical conductivity in terms of mobility. We can write

$$\sigma_e = ne\mu_e \quad (26)$$

indicating that σ_e is proportional to μ_e . A typical value for μ_e may be obtained by substituting $\sigma_e = 10^{-7} (\mu\text{ohm}\cdot\text{cm})^{-1}$ and $n = 10^{14} \text{ cm}^{-3}$ in Eq.(26). This yields

$$\mu_e \sim 10^3 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1} \quad (27)$$

What we have said about electrons in a strongly n-type substance can be carried over to a discussion of holes in a strongly p-type substance. The conductivity of the holes is given by

$$\sigma_h = \frac{pe^2\tau_h}{m_h} = pe\mu_h \quad (28)$$

where μ_h is the hole mobility.

Let us now treat the general case, in which both electrons and holes are present. When a field is applied, electrons drift opposite to the field and holes drift in the same direction as the field. The currents of the two carriers are additive, however, and consequently the conductivities are also. Therefore

$$\sigma = \sigma_e + \sigma_h \quad (29)$$

i.e., both electrons and holes contribute to the currents. In terms of the mobilities, one may write

$$\sigma = ne\mu_e + pe\mu_h \quad (30)$$

The carriers' concentrations n and p maybe different if the sample is doped, as discussed before. And one or the other of the carriers may dominate, depending on whether the semiconductor is n - or p -type. When the substance is in the intrinsic region, however, $n = p$, and Eq. (30) becomes

$$\sigma = ne(\mu_e + \mu_h) \quad (31)$$

where n the intrinsic concentration. Even now the two carriers do not contribute equally to the current. The carrier with the greater mobility - usually the electron - contributes the larger share.

Dependence on temperature

Conductivity depends on temperature, and this dependence is often pronounced. Consider a semiconductor in the intrinsic region. Its conductivity is expressed by (31). But in this situation the concentration n increases exponentially with temperature, as may be recalled from (16). We may write the conductivity in the form

$$\sigma = F(T)e^{-E_g/2kT} \quad (32)$$

where $F(T)$ is a function which depends only weakly on the temperature. (This function depends on the mobilities and effective masses of the carriers.) Thus conductivity increases exponentially with temperature as shown in Fig.12.

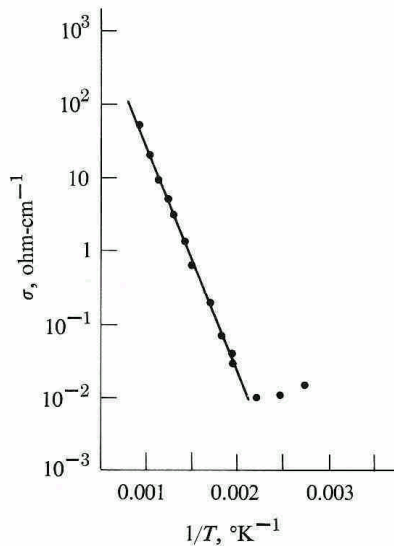


Fig. 12 Conductivity of Si versus $1/T$ in the intrinsic range.

This result can be used to determine the energy gaps in semiconductors. In the early days of semiconductors this was the standard procedure for finding the energy gap. Nowadays, however, the gap is often measured by optical methods.

When the substance is not in the intrinsic region, its conductivity is given by the general expression (30). In that case the temperature dependence of the conductivity on T is not usually as strong as indicated above. To see the reason for this, suppose that the substance is extrinsic and strongly n -type. The conductivity is

$$\sigma = ne\mu_e. \quad (33)$$

But the electron concentration n is now a constant equal to N_d , the donor (hole) concentration. And any temperature dependence present must be due to the mobility of electrons or holes.

Mobility versus temperature: scattering mechanisms

Mobility of electrons (or holes) varies with temperature. In n -type semiconductor

$$\mu_e = \frac{e\tau_e}{m_e} = \frac{el_e}{m_e v_e}. \quad (34)$$

Since the lifetime of the electron, or its collision time, varies with temperature, its mobility also varies with temperature. Normally, both lifetime and mobility diminish as the temperature rises.

The relaxation time is given by $\tau_e = l_e/v_e$, where l_e is the mean free path of the electron and v_e is the drift velocity. The velocity of electrons is different depending on their location in the conduction band. Electrons at the bottom of the conduction band in a semiconductor obey the classical statistics and not the highly degenerate Fermi statistics prevailing in metals. The higher electrons are in the band, the greater their velocity.

We can evaluate the conductivity by assuming that v_e is the average velocity. The average velocity can be estimated using the procedure of the kinetic theory of gases:

$$\frac{1}{2} m_e v_e^2 = \frac{3}{2} kT. \quad (35)$$

This introduces a factor of $T^{-1/2}$ dependence in the mobility:

$$\mu_e = \frac{el_e}{m_e^{1/2} (3kT)^{1/2}}. \quad (36)$$

The mean free path l_e also depends on the temperature, and in much the same way as it does in metals. l_e is determined by the various collision mechanisms acting on the electrons. These mechanisms are the collisions of electrons with thermally excited phonons and collisions with impurities. At high temperatures, at which collisions with phonons is the dominant factor, l_e is inversely proportional to temperature, that is, $l_e \sim T^{-1}$. In that case, mobility varies as $\mu_e \sim T^{3/2}$. Figure 13 shows this for Ge.

Another important scattering mechanism in semiconductors is that of *ionized* impurities. When a substance is doped, the donors (or acceptors) lose their electrons (or holes) to the conduction band. The impurities are thus ionized, and are quite effective in scattering the electrons (holes). At high

temperatures this scattering is masked by the much stronger phonon mechanism, but at low temperatures this latter mechanism becomes weak and the ionized-impurity scattering gradually takes over.

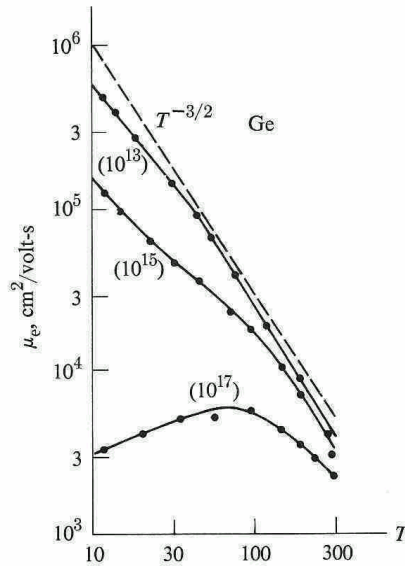


Fig. 13 Electron mobility versus T in Ge. The dashed curve represents the pure phonon scattering; numbers in parentheses refer to donor concentrations.

Band structure of real semiconductors

So far, we have assumed the simplest possible band structure, namely, a conduction band of a standard form, centered at the origin, $k = 0$, and a valence band of a standard inverted form, also centered at the origin. Such a simple structure is applicable for elucidating many observed phenomena, but it does not represent the actual band structures of many common semiconductors. Only when one uses the actual band structure is it possible to obtain a quantitative agreement between experiments and theoretical analysis.

A material whose band structure comes close to the ideal structure is GaAs (Fig. 14). The conduction band has a minimum at the origin $k = 0$, and the region close to the origin is well represented by a quadratic energy dependence, $E(k) = \frac{\hbar^2 k^2}{2m_e}$, where $m_e = 0.072m$. Since the electrons are most likely to populate this region, one can represent this band by a single effective mass.

Note, however, that as k increases, the energy $E(k)$ is no longer quadratic in k , and those states may no longer be represented by a single, unique effective mass. In particular that the next-higher energy minimum occurs along the $[100]$ direction. The dependence of energy on k in the neighborhood of this *secondary* minimum is quadratic, and hence an effective mass may be defined locally, but its value is much greater than that of the primary minimum (at the center). The actual value is $0.36m$. Due to cubic symmetry there are six equivalent secondary minima, or *valleys*, in all along the $[100]$ directions.

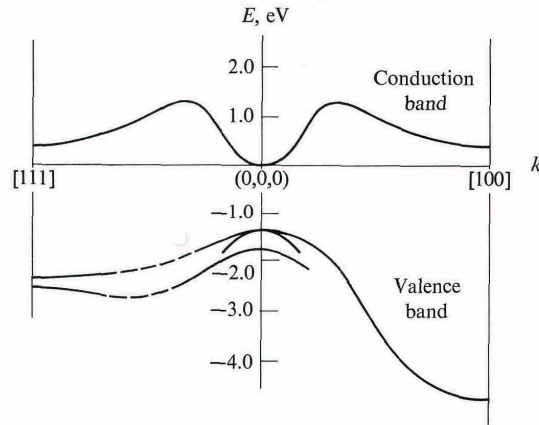


Fig. 14 Band structure of GaAs plotted along the [100] and [111] directions.

These secondary valleys do not play any role under most circumstances, since the electrons usually occupy only the central, or primary, valley. In such situations, these secondary valleys may be disregarded altogether. There are also other secondary valleys in the [111] directions, as shown in Fig. 14. These are higher than the [100] valleys, and hence are even less likely to be populated by electrons.

The valence band is also illustrated in Fig.14. Here it is composed of three closely spaced subbands. Because the curvatures of the bands are different, so are the effective masses of the corresponding holes. One speaks of *light holes* and *heavy holes*.

Other III-V semiconductors have band structures quite similar to that of GaAs.

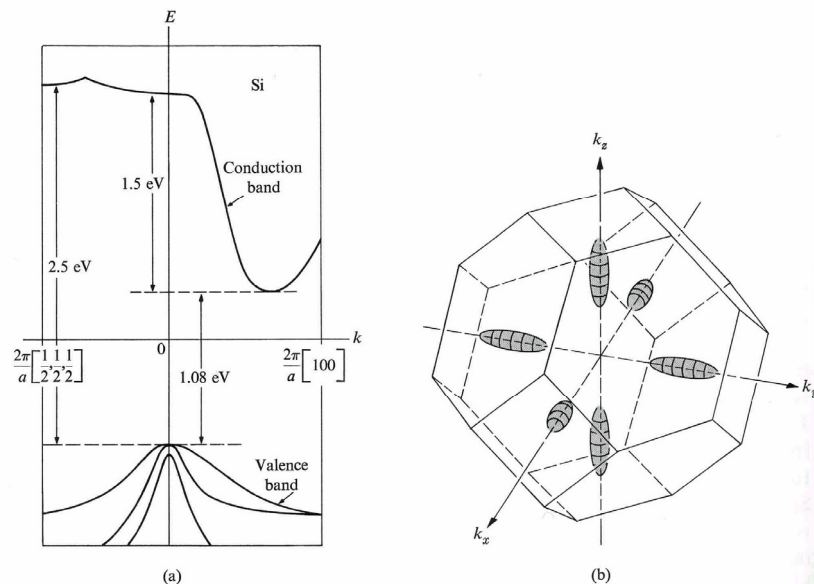


Fig.15 (a) Band structure of Si plotted along the [100] and [111] directions, (b) Ellipsoidal energy surfaces corresponding to primary valleys along the [100] directions.

Figure 15a shows the band structure of Si. An interesting feature is that the conduction band has its lowest (primary) minimum not at $k=0$. The minimum lies along the [100] direction, at about 0.85 the distance from the center to the edge of the zone. Note that the bottom of the conduction does not lie directly above the top of the valence band. This type of semiconductors is known as *indirect gap semiconductors*. These should be distinguished from *direct gap semiconductors* such as GaAs. Because of the cubic symmetry, there are actually six equivalent primary valleys located along the [100] directions. These are illustrated in Fig. 15b. The energy surfaces at these valleys are composed of elongated ellipsoidal surfaces of revolution, whose axes of symmetry are along the [100] directions. There are two different effective masses which correspond to these surfaces: the *longitudinal* and the *transverse* effective masses. The longitudinal mass is $m_l = 0.97m$, while the two identical transverse masses are $m_t = 0.19m$. The mass anisotropy ratio is about 5.

The valence band in silicon is represented by three different holes (Fig.15a). One of the holes is heavy ($m_h = 0.5m$), and the other two are light. The energy gap in Si, from the top of the valence band to the bottom of the conduction band, is equal to 1.08 eV. The fact that the bottom of the conduction does not lie directly above the top of the valence band, is irrelevant to the definition of the band gap.

Excitons

An electron and a hole may be bound together by their attractive coulomb interaction, just as an electron is bound to a proton to form a neutral hydrogen atom. The bound electron-hole pair is called an *exciton*, Fig.16. An exciton can move through the crystal and transport energy; it does not transport charge because it is electrically neutral. It is similar to positronium, which is formed from an electron and a positron.

Excitons can be formed in every insulating crystal. All excitons are unstable with respect to the ultimate recombination process in which the electron drops into the hole.

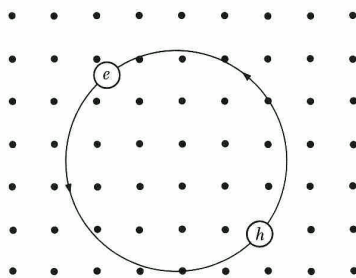


Fig.16a An exciton is a bound electron-hole pair, usually free to move together through the crystal. The exciton shown is weakly bound, with an average electron-hole distance large in comparison with a lattice constant.

The binding energy of the exciton can be measured by optical transitions from the valence band, by the difference between the energy required to create an exciton and the energy to create a free electron and free hole, Fig.17.

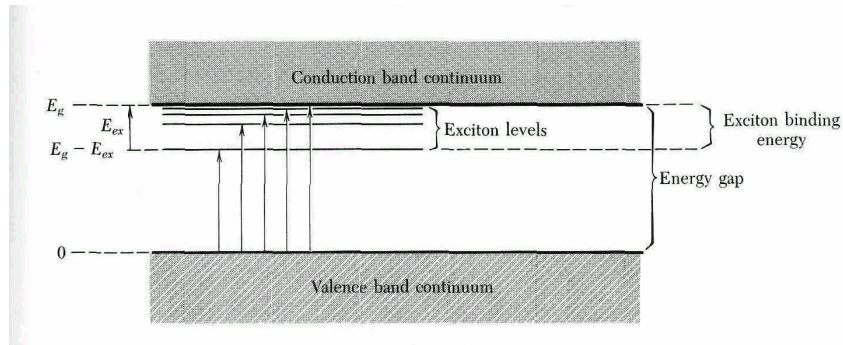


Fig.17 Energy levels of an exciton. Optical transitions from the top of the valence band are shown by the arrows; the longest arrow corresponds to the energy gap. The binding energy of the exciton is E_{ex} , referred to a free electron and free hole.

Energy levels of an exciton can be calculated as follows. Consider an electron in the conduction band and a hole in the valence band. The electron and hole attract each other by the Coulomb potential

$$V(r) = -\frac{e^2}{\epsilon r}, \quad (37)$$

where r is the distance between the particles and ϵ is the appropriate dielectric constant. There will be bound states of the exciton system having total energies lower than the bottom of the conduction band.

The problem is the hydrogen atom problem if the energy surfaces for the electron and hole are spherical and nondegenerate. The energy levels are given by

$$E_n = E_c - \frac{e^4 \mu}{2\epsilon^2 \hbar^2 n^2}. \quad (38)$$

Here n is the principal quantum number and μ is the reduced mass:

$$\frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_h} \quad (39)$$

formed from the effective masses of the electron and hole.

The exciton ground state energy is obtained on setting $n = 1$ in (38); this is the ionization energy of the exciton.