

Chapter 14

RECTIFIERS AND TRANSISTORS

14-1. Rectifying properties of a barrier layer between two metals

Although in solid-state rectifiers one usually employs one semiconducting contact, semiconduction itself is not essential for the rectification process. This may be illustrated by considering two metals of different work function separated by a thin vacuum gap. As we have seen in Sec. 9-10, the Fermi levels of the two metals must coincide in thermal equilibrium, leading to the situation depicted in Fig. 14-1a; the metal of low work function acquires a positive surface charge, the other acquires a negative surface charge. The total potential drop across the gap is

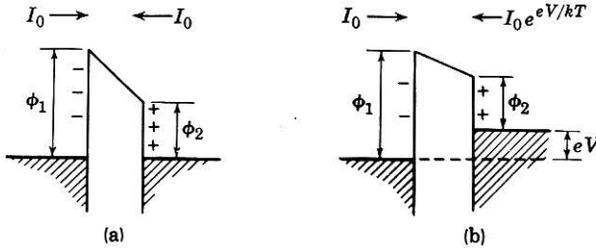


Fig. 14-1. (a) Shows the equilibrium between two metals of different work functions separated by a thin vacuum gap. In (b) a forward voltage is applied (metal 2 negative).

equal to $(\phi_1 - \phi_2)/e$. It is convenient to consider this situation a dynamic equilibrium in which the electronic current from 1 to 2 is equal to that from 2 to 1. Let us denote this current density by I_0 . Suppose now metal 2 is made negative with respect to 1 by applying an external voltage smaller than the voltage drop $(\phi_1 - \phi_2)/e$. The energy levels of 2 are then raised relative to those in 1, and the situation corresponding to Fig. 14-1b results. The current $I_{1 \rightarrow 2}$ is still equal to I_0 because the barrier viewed from the position of metal 1 has not changed. On the other hand, the potential energy hill as viewed from metal 2 is lowered by an amount eV , which makes the probability for an electron to cross the hill larger by a factor $e^{eV/kT}$. Hence the net electron current is

$$I_f = I_0 (e^{eV/kT} - 1) \quad (14-1)$$

Similarly, if the applied voltage has such polarity as to make metal 2

positive with respect to 1, we have again $I_{1 \rightarrow 2} = I_0$. However, the current from $2 \rightarrow 1$ is now $I_0 e^{-eV/kT}$, yielding a net current of

$$I_r = I_0(1 - e^{-eV/kT}) \quad (14-2)$$

where I_f and I_r are referred to, respectively, as the forward current and the reverse current. Now I_f increases exponentially with the voltage applied in the forward direction; I_r , on the other hand, saturates rapidly to the low value I_0 . The current-voltage characteristic of the contact is similar to that given in Fig. 14-8 and can be used for rectifying purposes.

14-2. The Schottky theory of a metal-semiconductor contact

A simple theory for the contact between a metal and a semiconductor has been developed by Schottky.¹ It leads to the formation of a physical

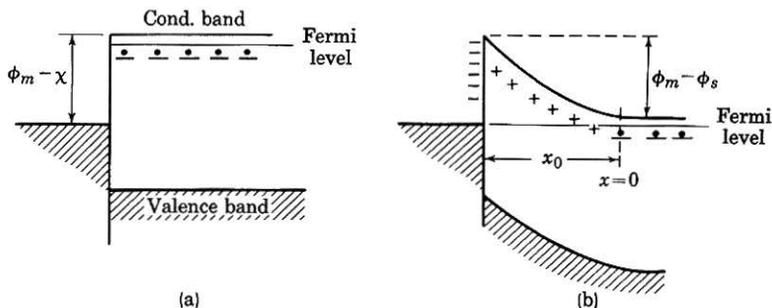


Fig. 14-2. (a) Refers to a metal-semiconductor contact not yet in equilibrium; χ is the electron affinity of the semiconductor. In (b) equilibrium has been established by the formation of a Schottky layer; the total potential drop is $(\phi_m - \phi_s)/e$.

barrier layer at the metal-semiconductor interface as explained below. Such barriers must be distinguished from chemical barrier layers which may be present between the metal and semiconductor as a result of chemical preparation.

To explain the nature of the physical barrier layer, consider an ideal contact between a metal of work function ϕ_m and an n -type semiconductor with an electron affinity χ .² Before equilibrium has been established, the energy band scheme may be represented by Fig. 14-2a. According to Sec. 12-5, the effective work function of the semiconductor is given by the energy difference between its Fermi level and the vacuum level; let this difference be ϕ_s . Thus, if $\phi_s < \phi_m$, electrons will flow from the

¹ W. Schottky, *Z. Physik*, **113**, 367 (1939).

² The electron affinity is defined as the energy required to transfer an electron from the bottom of the conduction band to vacuum.

semiconductor into the metal. Consequently, the metal acquires a negative surface charge and the semiconductor charges up positively. Now, because the density of donors is relatively small, the donors will be come ionized over a region which extends into the semiconductor, i.e., a space charge rather than a surface charge is created (see Fig. 14-2b). The thickness of the barrier layer thus formed may be estimated as follows: Let us assume that all donors in the region between $x = 0$ and $x = x_0$ in Fig. 14-2b are ionized. The potential energy of an electron ϕ in this region is then determined by the Poisson equation

$$d^2\phi/dx^2 = (4\pi/\epsilon)n_d e^2 \quad (14-3)$$

where n_d is the donor concentration and ϵ the dielectric constant. Taking $\phi = 0$ at $x = 0$ and $\phi = \phi_m - \phi_s$ at $x = x_0$, one readily finds that the thickness x_0 of the barrier must satisfy the relation

$$\phi_m - \phi_s = (2\pi/\epsilon)n_d e^2 x_0^2 \quad (14-4)$$

Thus, for a given value of the required potential energy drop, x_0 varies as $n_d^{-1/2}$. A few examples may be given here for $\phi_m - \phi_s = 1$ eV and $\epsilon = 10$.

n_d	10^{15}	10^{17}	10^{19}	per cm ³
x_0	10^{-4}	10^{-5}	10^{-6}	cm

It is observed that an externally applied voltage changes the potential drop across the barrier and hence results in a change in the thickness of the barrier. For example, if the semiconductor is made positive, the thickness increases and is determined by

$$\phi_m - \phi_s + eV = (2\pi/\epsilon)n_d e^2 x_0^2 \quad (14-5)$$

For an applied voltage in the opposite direction, x_0 decreases. Note that the barrier thickness increases with increasing dielectric constant. If the dielectric constant of the material is known, the thickness of the barrier can be determined from capacitance measurements with a small ac signal for a given bias. To a first approximation the equivalent circuit of the contact may be represented by a voltage-dependent capacitor in parallel with a nonlinear resistor, the combination being in series with the bulk resistance of the semiconductor. Changes as indicated by (14-5) can indeed be observed.

The above model is admittedly simplified and neglects, for example, the influence of the image force; for high dielectric constants (≥ 10) the image force has little influence. Also, the influence of surface states has been neglected; these may play an important role.

The Schottky barrier layer forms an essential factor in the theory of rectifying contacts as we shall see below. It is left to the reader to discuss the barrier formed at a metal-*p*-type-semiconductor contact.

14-3. Single-carrier theories of rectification

As an example of the rectifying properties of a metal-semiconductor contact we represent in Fig. 14-3 the current-voltage characteristic of a metallic point contact on p -type AlSb.³ In general, the forward current is observed under the following circumstances: for n -type material the semiconductor should be negative, for p -type material the semiconductor should be positive.

In the conventional theories of rectification it is assumed that either electrons or holes take part in the current flow across the barrier; i.e., they are single-carrier theories. In order to explain certain properties of Ge rectifiers, a two-carrier theory has been developed. In the present section we shall consider only the single-carrier case.

The tunnel theory. The oldest theory of rectification was developed in 1932 by Wilson and Nordheim.⁴ These authors assumed that for an n -type semiconductor the electrons crossed the barrier by tunnel effect, i.e., the carriers penetrate through rather than cross over the potential barrier. Such a mechanism of course requires thin barrier layers ($\sim 10^{-7}$ cm); as we have seen above, the barrier is usually considerably thicker. The strongest objection against the tunnel theory is, however, that it predicts the wrong sign for rectification. For example, the reader can readily convince himself that for a metal- n -type-semiconductor contact, the tunnel theory predicts the forward current when the metal is negative. This is simply because the number of electrons available for tunneling in the metal is much larger than that in the conduction band of the semiconductor. The tunnel theory will therefore not be discussed here, although it should be kept in mind that under special circumstances tunneling may well occur.

The Mott-Schottky theory for thick barriers. A new theory of rectification in which it is assumed that the carriers surmount the potential barrier by thermal excitation was proposed by Mott⁵ and further developed by Schottky.⁶ Mott was particularly concerned with selenium and with Cu-Cu₂O rectifiers. In these rectifiers, as a result of the chemical way in

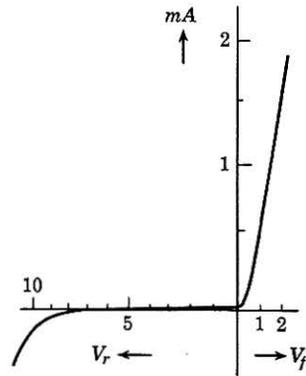


Fig. 14-3. The current in mA as function of forward and reverse voltage (in volts) for a point-contact AlSb rectifier.

[After Welker, ref. 3]

³ H. Welker, *Z. Naturforsch.*, **8a**, 248 (1953).

⁴ H. Wilson, *Proc. Roy. Soc. (London)*, **136**, 487 (1932); L. W. Nordheim, *Z. Physik*, **75**, 434 (1932).

which they are prepared, the density of donors (or acceptors) is very small near the metal and gradually increases to a constant value as one moves into the semiconductor. In some instances one produces deliberately an insulating layer between the metal and the semiconductor by chemical means.

As an idealized model for this type of system we shall consider the following case.⁵ An n -type semiconductor contains n_d donors per cm^3 . The semiconductor is separated from the metal by a layer of 10^{-4} – 10^{-5} cm thick of the same material but without donor levels. We shall further

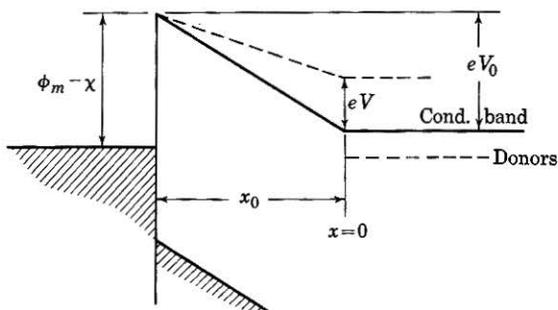


Fig. 14-4. Metal-insulator-semiconductor contact; the voltage drop across the insulating layer in equilibrium is $V_0 = (\phi_m - \phi_s)/e$. With an applied forward voltage V , the voltage drop is $(V_0 - V)$ as indicated by the dashed conduction band.

assume that any potential difference between the metal and semiconductor exists essentially across the insulating layer, the field strength in the layer being constant (see Fig. 14-4).⁶ Since the thickness of the layer x_0 is large compared with the mean free path for scattering of the electrons by lattice vibrations, the electron current through the layer is due to (1) the electric field and (2) diffusion. Let E represent the field strength for a unit negative charge, and let I be the electronic current density.

We may then write

$$I = n(x)e\mu E - D \frac{dn}{dx} \quad (14-6)$$

where $n(x)$ is the density of electrons. The diffusion coefficient D in terms of the mobility μ is given by the Einstein relation $D = \mu kT/e$. Integrating (14-6), one thus obtains

$$n(x) = I/\mu e E + C e^{eEx/kT} \quad (14-7)$$

where C is a constant. In order to calculate the current I , we make use

⁵ N. F. Mott, *Proc. Roy. Soc. (London)*, **A171**, 27 (1939).

⁶ W. Schottky, *Z. Physik*, **118**, 539 (1942); also, W. Schottky and E. Spenke, *Wiss. Veröffentl. Siemens Werken*, **18**, 225 (1939). These references also take account of the space charge region in the semiconductor.

of the following boundary conditions: for $x = 0$, the density of electrons is equal to that in the conduction band in the bulk semiconductor $n(0)$. Also, the density of electrons for $x = x_0$ in the absence of an external field must be equal to

$$n(x_0) = n(0)e^{-(\phi_m - \phi_s)/kT} = n(0)e^{-eV_0/kT} \quad (14-8)$$

where ϕ_m and ϕ_s are the work functions of the metal and insulator, respectively; the total voltage drop is then $V_0 = (\phi_m - \phi_s)/e$. We shall assume that $n(x_0)$ is not influenced by the current flow resulting from an external field, although this is only approximately true. The first boundary condition leads to $C = n(0) - I/\mu eE$. Substituting this into (14-7) and applying the second boundary condition in the form (14-8), one obtains for the current density,

$$I = \mu eE \frac{n(0)e^{eEx_0/kT} - n(x_0)}{e^{eEx_0/kT} - 1} \quad (14-9)$$

Now when V is the applied voltage in the forward direction as indicated in Fig. 14-4, $Ex_0 = -(V_0 - V)$; also, as long as we are interested in cases for which the potential barrier is large compared with kT , the denominator in (14-9) reduces to -1 . Thus (14-9) may be written in the form

$$I(V) = [\mu e(V_0 - V)/x_0]n(x_0)(e^{eV/kT} - 1) \simeq A(e^{eV/kT} - 1) \quad (14-10)$$

where A is approximately constant. The form of the current-voltage characteristic is thus essentially the same as (14-1); for the reverse current one obtains a relation similar to (14-2). For the type of rectifiers for which this theory is developed (thick barriers), it is generally in accord with the experimental results; refinements involving the image force may be found in footnotes 5 and 6.

*The diode theory.*⁷ In germanium and silicon rectifiers the barrier layer thickness is of the order of 10^{-6} cm, i.e., comparable with the mean free path of the carriers. In that case, the diffusion theory cannot be applied, and the so-called diode theory has been developed. In this theory it is assumed that collisions of carriers with the lattice are absent and the problem reduces to that of two thermionic emitters facing each other. We leave it to the reader to show that the electronic current density from the semiconductor to the metal in this case is given by

$$\left(\frac{1}{2}\right)n(0)e\langle v \rangle e^{-e(V_0 - V)/kT} \quad (14-11)$$

where $\langle v \rangle = (2kT/\pi m)^{1/2}$ is the average thermal velocity of the electrons in the conduction band of the semiconductor; the other symbols have

⁷ See, for example, H. C. Torrey and C. A. Whitmer, *Crystal Rectifiers*, McGraw-Hill, New York, 1948, p. 81.

the same meaning as above. The electronic current density from the metal to the semiconductor is obtained by putting $V = 0$ in expression (14-11), since for $V = 0$ the two currents are equal and opposite. The resultant current is then

$$I(V) = (\frac{1}{2})n(x_0)e\langle v \rangle (e^{eV/kT} - 1) = A'(e^{eV/kT} - 1) \quad (14-12)$$

Comparing this with (14-10), it is observed that the form of the two expressions is essentially the same. However, $\langle v \rangle$ may be considerably smaller than μE , leading to $A' \ll A$. For example, at room temperature $\langle v \rangle \simeq 10^7$ cm sec $^{-1}$; for a barrier of 10^{-6} cm and a voltage drop of 1 eV, $E \simeq 10^6$ volts per cm, and with $\mu \simeq 10^3$ cm per volt sec, we obtain $\mu E \simeq 10^9$ cm sec $^{-1}$. It should be remarked here that V represents the voltage across the barrier, i.e., it is equal to the applied voltage minus the voltage drop across the bulk semiconductor.

Although the diode theory has been applied in the past to interpret the rectifying properties of germanium and silicon diodes, a number of observations remained unexplained. For example, according to the above theory, the magnitude of the currents should depend strongly on the work function of the metal because $n(x_0)$ is proportional to $\exp(-eV_0/kT)$ and V_0 is determined by $(\phi_m - \phi_s)$. Thus, for a variation of 0.5 eV in the work function of the metals used, the currents should vary by a factor of $\sim 10^8$. Experiments indicate variations by a factor of 10 or less for different metal points. The origin of this discrepancy will be discussed in the next section.

14-4. Surface states on semiconductors

In 1948 Shockley and Pearson reported the following relatively simple but crucial experiment.⁸ Consider a thin layer of n -type germanium on an insulating support. Opposite the layer and separated from it is a metal plate, the system as a whole forming a parallel plate condenser. When the metal plate is made negative relative to the germanium layer, a negative charge is induced in the latter, which, if it were free to move, should enhance the conductivity of the layer. For example, if the applied field is 3×10^4 volts cm $^{-1}$ the induced charge per cm 2 corresponds to about 3×10^{10} electrons. On the other hand, if the layer is 5000 Å thick and contains 10^{15} electrons per cm 3 , the number of electrons per cm 2 of the layer without external field is 5×10^{10} . It should thus be possible to measure this effect. The experiments indicated, however, that only about one-tenth of the total induced charge contributed to the increase in conductivity. It was proposed by Bardeen that the immobile fraction of the induced charge resides in electronic states at the surface of the

⁸ W. Shockley and G. L. Pearson, *Phys. Rev.*, **74**, 232 (1948).

material.⁹ Such states, which may lie within the normally forbidden region, may arise partly as a consequence of the sudden departure from periodicity of the potential at the surface or in part from adsorbed atoms. In other words, the simple band picture which one normally employs for the bulk properties is in general not applicable close to the surface. Thus a certain number of these surface states may be occupied without giving rise to an excess surface charge. When the material is placed opposite a positively charged metal, more surface states may be filled by the induced charge.

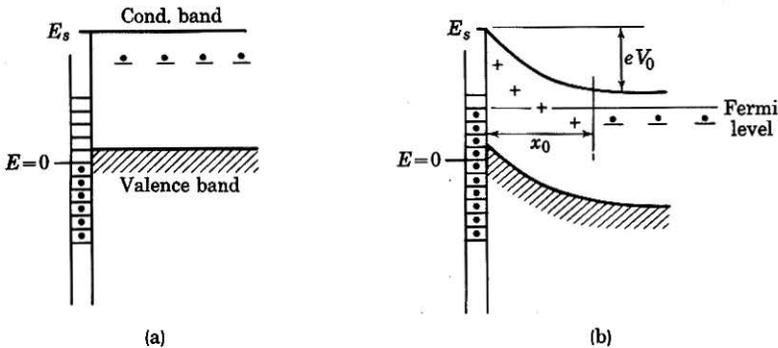


Fig. 14-5. In (a) there is no equilibrium; a number of surface states are filled but no surface charge exists. In (b) equilibrium is established, leading to a surface charge equal to the space charge extending over x_0 .

The existence of surface states has an important consequence for the electron distribution near the surface of a neutral germanium crystal. The reason is that the Fermi level associated with the surface states should coincide with that of the bulk material. Thus, in the absence of a net surface charge, let the surface states of an n -type Ge crystal be filled up to the level $E = 0$ as indicated in Fig. 14-5a. Let the conduction band be located at E_s . Evidently, electrons in the conduction band will tend to fill up more surface states until a potential drop V_0 is built up such that the highest filled surface state coincides with the Fermi level of the bulk material (see Fig. 14-5b). Let the density of surface states in the vicinity of $E = 0$ be equal to n_s per cm^2 per electron volt. The neutrality of the crystal then requires that per cm^2 the number of ionized donors extending over a thickness x_0 must equal the excess number of electrons in surface states. Thus, if we assume that the bottom of the conduction band practically coincides with the Fermi level in n -type Ge, we may write

$$n_d x_0 = n_s (E_s - eV_0) \quad (14-13)$$

⁹ J. Bardeen, *Phys. Rev.*, **71**, 717 (1947).

According to (14-4) we also have

$$eV_0 = 2\pi n_d e^2 x_0^2 / \epsilon \quad (14-14)$$

so that x_0 may be eliminated, leading to

$$n_s^2 = n_d \epsilon e V_0 / 2\pi e^2 (E_s - eV_0)^2 \quad (14-15)$$

We note that for very small values of n_s , the voltage drop V_0 is very small because a small number of extra electrons will bring the Fermi level at the surface up to that of the bulk material. For very large values of n_s , V_0 becomes approximately equal to E_s/e . According to Shockley and Pearson, n_s is of the order of $5 \times 10^{13} \text{ cm}^{-2} \text{ volts}^{-1}$.

From what has been said above it follows that due to the presence of surface states, a layer of depleted conductivity is formed below the surface. Under these circumstances, the space charge layer is a property of the material itself and not particularly sensitive to the work function of a metal which may be brought in contact with the surface. This argument has been used by Bardeen to explain the fact that the properties of point-contact germanium rectifiers are rather insensitive to the metal used.¹⁰ The presence of surface states also play a role in the interpretation of contact potential measurements across n - p junctions.¹¹

14-5. The two-carrier theory of rectification

In Sec. 14-3 we have seen that if V is the applied voltage in the forward direction the forward current of a rectifying contact should be given by

$$I = A[e^{x(V-Ir)} - 1] \quad (14-16)$$

where r is the bulk resistance of the rectifier. Although this formula is in agreement with results obtained for germanium point-contact rectifiers up to about 0.2 or 0.3 volts, deviations occur at higher voltages; these deviations are such that they require a decrease in the resistance r . This difficulty has been explained by Bardeen and Brattain in terms of the model represented in Fig. 14-6.¹² They assume that as a result of surface states the Fermi level in n -type material crosses the surface near the top of the valence band. From Fig. 14-5b it may be seen that this is possible if the density of surface states per unit energy interval is sufficiently large. Under these circumstances, the concentration of holes in the valence band near the surface will be larger than the concentration of electrons in the conduction band. Hence a thin layer of the n -type material

¹⁰ J. Bardeen, *Phys. Rev.*, **71**, 717 (1947); see also S. Benzer, *J. Appl. Phys.*, **20**, 804 (1949).

¹¹ W. H. Brattain, "Semiconducting Materials," *Proc. Reading Conference*, Butterworths Scientific Publications, 1951, p. 37.

¹² J. Bardeen and W. H. Brattain, *Phys. Rev.*, **74**, 230, 231 (1948).

will become p -type. Suppose now that the semiconductor is made negative relative to a metal in contact with it. This will lead to an increase in the electronic current from the semiconductor into the metal but at the same time a hole current will begin to flow from the surface into the semiconductor. In other words, two types of carriers contribute to the current. This has the effect of decreasing the apparent value of r in (14-16), because r is based on electronic conductivity only. Because of lack of space it is not possible to discuss the quantitative aspects of the two-carrier theory for point contacts here.¹³ It will be evident that this model also explains the hole injection into n -type material by metals biased in the forward direction, as referred to in Sec. 14-6.

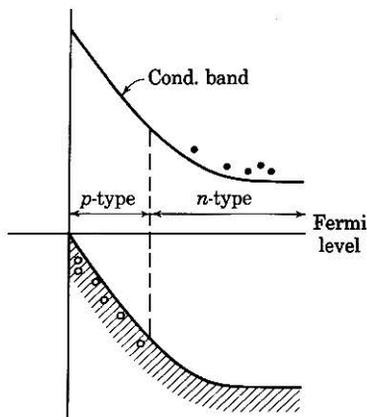


Fig. 14-6. As a result of surface states, the valence band at the surface lies close to the Fermi level; this gives rise to a thin p -type layer near the surface of n -type material.

14-6. The p - n junction rectifier

When a piece of p -type material is in contact with an n -type region, one speaks of a p - n junction. Such junctions may be made in several ways; in germanium they have been produced by converting part of an n -type region into p -type by heating or by nuclear bombardment. In other cases, these junctions are formed during the growth of single crystals as a result of segregation of impurities. In general, the acceptor concentration and the donor concentration will not change abruptly at the junction, but for simplicity we shall assume this to be the case (Fig. 14-7a). The Fermi level in the bulk p -type is located close to the top of the valence band; the Fermi level in the n -type region lies close to the bottom of the conduction band. As a consequence, the situation of Fig. 14-7b is unstable; electrons will flow from n to p and holes from p to n until two space charge regions are established, producing a voltage drop V_0 (see Fig. 14-7c). The space charge in the n -region results from ionized donors, that in the p -region from ionized acceptors. The voltage drop V_0 is approximately equal to the width of the forbidden gap. For an ideal case, assuming a simple variation of donor and acceptor

¹³ J. Bardeen and W. H. Brattain, *Phys. Rev.*, **75**, 1208 (1949); J. A. Swanson, *J. Appl. Phys.*, **25**, 314 (1954); for a discussion of the reverse characteristics of high inverse voltage point contacts on Ge, see J. H. Simpson and H. L. Armstrong, *J. Appl. Phys.*, **24**, 25 (1953).

concentration at the junction, the potential may be calculated in a way similar to that used in a metal-semiconductor contact (Sec. 14-2).

It is convenient to consider the equilibrium situation in the absence of an external field as a dynamic one. Thus, there must be a certain hole current I_{h_0} flowing from P to n and an equal but opposite one from

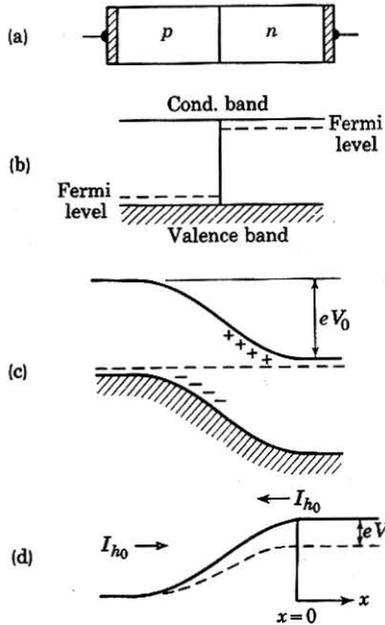


Fig. 14-7. (a) p - n junction; (b) non-equilibrium energy band scheme; (c) equilibrium energy band scheme with space charge regions; (d) full curve represents the potential energy for a hole in equilibrium and the two compensating hole currents I_{h_0} are indicated; the dashed curve represents the potential energy for a hole when the junction is biased in the forward direction; in that case the hole current to the left remains I_{h_0} , the hole current to the right equals $I_{h_0} \exp(eV/kT)$.

n to p (Fig. 14-7d). The same is true for the equilibrium electron current I_{e_0} . This implies that there must be a certain concentration of holes in the n -region as well as a certain concentration of electrons in the p -region. This is a result of the continuous thermal creation of electron-hole pairs, the creation being compensated by recombination. For example, if g is the rate of production of pairs and n_{e_0} and n_{h_0} are the equilibrium concentrations anywhere, we must have $rn_{e_0}n_{h_0} = g$, where r is the recombination coefficient. Thus in either region, $n_{e_0}n_{h_0} = g/r$, which is constant at a given temperature. If one assumes that the ratio g/r is independent of the donor or acceptor concentration, g/r must also

be equal to n_i^2 , where n_i is the density of carriers in intrinsic material. Under these circumstances

$$n_e n_{h_0} = g/r = n_i^2 \tag{14-17}$$

Suppose now that a negative voltage $-V$ is applied to the n -region. If we assume that the voltage drop is essentially across the space-charge region, the hole current from n to p is still I_{h_0} . However, holes going

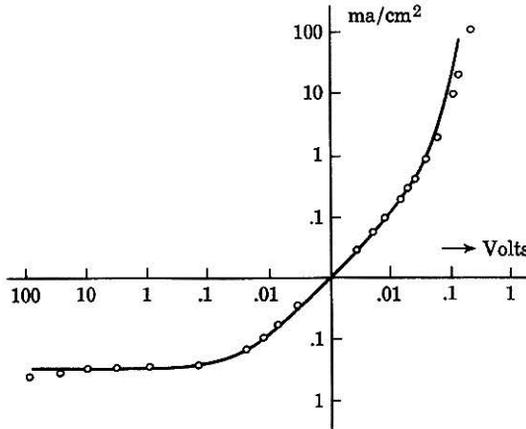


Fig. 14-8. Current-voltage characteristic of a p - n junction. The circles are experimental points, the curve is theoretical. [After W. Shockley, ref. 14]

from p to n have to climb a smaller potential hill (Fig. 14-7d) and will give rise to a current $I_{h_0} \exp(eV/kT)$. Hence

$$I_h(V) = I_{h_0}(e^{eV/kT} - 1) \tag{14-18}$$

For the electron current one finds a similar expression and the total forward current across the junction should be

$$I_f(V) = (I_{h_0} + I_{e_0})(e^{eV/kT} - 1) \tag{14-19}$$

For positive voltages applied to the n -type region the reverse current is obtained,

$$I_r = (I_{h_0} + I_{e_0})(1 - e^{-eV/kT}) \tag{14-19a}$$

In Fig. 14-8 we have represented the experimental points obtained for a p - n junction characteristic;¹⁴ the fully drawn curve is the theoretical one. The agreement is very good indeed.

¹⁴ W. Shockley, *Proc. IRE*, **40**, 1289 (1952).

Let us now consider the rectification process in some more detail; we shall discuss only the hole current because a similar reasoning may be given for the electron current. We shall use the following symbols:

- V_0 = equilibrium potential drop
- $-V$ = voltage applied to n -region
- n_0 = equilibrium density of holes in bulk n -region
- $n_h(x)$ = actual density of holes in n -region
- n_p = density of holes in p -region
- $n(x) = n_h(x) - n_0$ = density of excess holes in n -region

The point $x = 0$ indicated in Fig. 14-7d corresponds to the point where the derivative of the potential vanishes. The general equations governing the motion of holes are (13-32) and (13-33). If in these equations we put $g_h = 0$ (no external pair generation) and assume the electric field to be negligible for $x > 0$, we obtain for the steady state ($\partial n/\partial t = 0$) in that region,

$$D_h d^2 n/dx^2 = n/\tau_h \quad (14-20)$$

where τ_h is the life time of holes in the n -region. The solution of this equation is

$$n(x) = n(0)e^{-x/L_h} \quad \text{where} \quad L_h^2 = D_h \tau_h \quad (14-21)$$

where L_h is the diffusion length of the holes. Thus the excess hole density in the n -region decreases by a factor $1/e$ over a distance L_h . According to equation (13-33) the hole current density diffusing across the junction is equal to

$$I_h = -eD_h(\partial n_h/\partial x)_{x=0} = (eD_h/L_h)n(0) \quad (14-22)$$

In order to find an expression for $n(0)$ in terms of the applied voltage, we make use of the fact that according to Boltzmann,

$$n(0) = n_h(0) - n_0 = n_p e^{-(V_0 - V)/kT} - n_p e^{-eV_0/kT} \quad (14-23)$$

From the last two equations it then follows that

$$I_h = \frac{eD_h}{L_h} n_0 (e^{eV/kT} - 1) \quad (14-24)$$

Comparing this with (14-18) it is observed that the equilibrium current $I_{h_0} = eD_h n_0/L_h = en_0 L_h/\tau_h$. This result has a simple physical interpretation: n_0/τ_h represents the number of holes recombining per second in equilibrium, and hence also represents the rate of creation of holes. The created holes diffuse about and recombine at an average distance L_h from their point of origin. Therefore the holes diffusing across the barrier are essentially those created within a range L_h on the right of $x = 0$

(Fig. 14-7d). For electrons in Ge, the diffusion constant $D \simeq 100 \text{ cm}^2/\text{sec}$ and a typical lifetime is 10^{-4} sec . This gives a diffusion length of the order of 0.1 cm.

The correctness of the diffusion theory for the rectifying p - n junction has been tested further by using junctions as a photoconductive device. For example, let photons of sufficient energy to create electron-hole

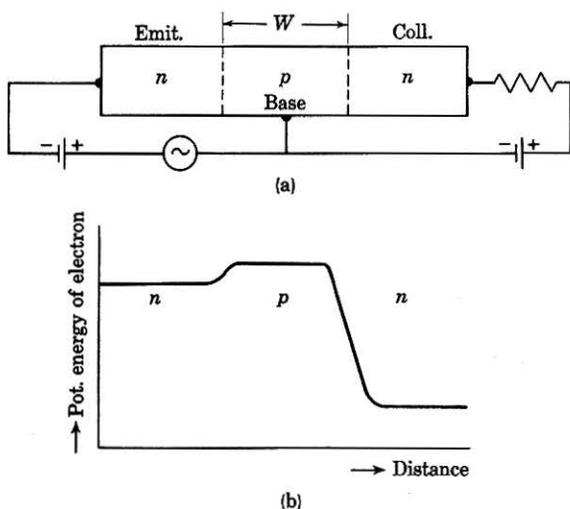


Fig. 14-9. (a) Represents schematically an n - p - n junction transistor with the emitter biased in the forward direction, the collector in the reverse direction. Actually, the p -region is much narrower than indicated. In (b) the potential energy of an electron is represented for the biased transistor.

pairs be incident on the n -type region at a distance x from $x = 0$. According to what has been said above, the current response should vary as $\exp(-x/L)$ and this has indeed been verified experimentally by Goucher and coworkers.¹⁵ Also, the value of L so obtained is consistent with the one required by the rectifier equation (14-24). When the light is incident at the junction itself, the electron and hole are separated by the strong field at the junction and a current of one electron per absorbed photon may be obtained.

14-7. Transistors

An n - p - n junction transistor is built up of two n -type regions separated by a thin layer of weakly p -type material. It is mainly this type of transistor which will be discussed below. The same reasoning applies to p - n - p

¹⁵ F. S. Goucher *et al.*, *Phys. Rev.*, **78**, 816 (1950); **81**, 637 (1951).

junction transistors. When the junction transistor is used as an amplifier, one of the n - p junctions is biased in the forward direction, the other in the reverse direction, as indicated in Fig. 14-9a. The former is called the emitter because the corresponding n -type region emits electrons into the p -region (the base); these electrons are collected at the junction with the reverse bias (the collector). The discussion below deals with the reasons for the amplifying action of the transistor.

Let W be the width of the p -type base region and D_e the diffusion coefficient of electrons in the base. The time required for an electron to cross the base, if it stays "alive" during the crossing, is equal to $t = W^2/2D_e$; this follows from the elementary theory of diffusion. Thus the probability for an electron to recombine with a hole during the crossing of the base is given by

$$t/\tau_e = W^2/2D_e\tau_e = W^2/2L_e^2 \quad (14-25)$$

where τ_e is the lifetime of electrons and L_e is their diffusion length in the base material. In most cases, $t/\tau_e \ll 1$ because the width of the base is small compared with L_e ; we shall assume this to be the case in the remainder of this section. In other words, we shall assume that all electrons emitted by the emitter are collected by the collector.

Let us now consider the current flow between the emitter and the base. The total current is made up of two parts: (a) a hole current I_h from the base into the emitter; (b) an electron current I_e from the emitter into the base. The ratio of these currents is important for the amplifying action of the transistor, and we shall now show that $I_e/I_h = L_h\sigma_e/W\sigma_h$, where σ_e and σ_h are the conductivities of the emitter and base regions, and L_h is the diffusion length of holes in the emitter region.

According to (14-24), the hole current is given by

$$I_h = \frac{eD_h}{L_h}n_{h_0}(e^{eV/kT} - 1) \quad (14-26)$$

where V is the applied voltage between base and emitter and n_{h_0} is the equilibrium concentration of holes in the emitter region. Because of what follows it is important to realize that I_h is determined by the diffusion length of holes in the emitter region. In the same way, the electronic current from emitter to base is determined by the diffusion length of electrons in the base region. However, because the width of the base region is $W \ll L_e$, one should use W rather than L_e for the electronic current. Hence

$$I_e = \frac{eD_e}{W}n_{e_0}(e^{eV/kT} - 1) \quad (14-26a)$$

where n_{e_0} is the equilibrium density of electrons in the base. From (14-26a) and (14-26) it then follows that

$$I_e/I_h = D_e n_{e_0} L_h / D_h n_{h_0} W \quad (14-27)$$

We may now apply expression (14-17) to the base and emitter regions, giving

$$n_{e_0} = n_i^2/n_h \quad \text{and} \quad n_{h_0} = n_i^2/n_e \quad (14-28)$$

where n_i is the density of carriers in intrinsic material, n_h is the density

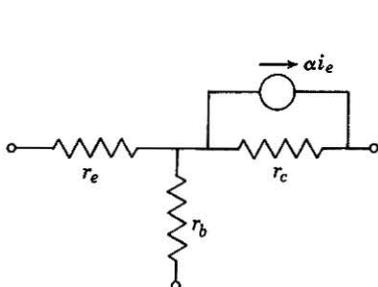


Fig. 14-10. Equivalent circuit of the transistor.

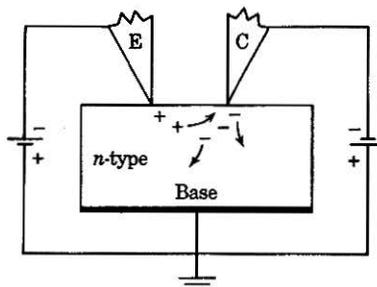


Fig. 14-11. A point-contact transistor.

of holes in the base, and n_e is the density of electrons in the emitter. Substituting into (14-27) we obtain

$$I_e/I_h = D_e L_h n_e / D_h W n_h = L_h \sigma_e / W \sigma_h \quad (14-29)$$

Here we have made use of the fact that in general $\sigma = ne\mu$ and $\mu = De/kT$, where μ is the mobility.

Suppose now that the base potential is altered; this will give rise to a change in the hole current from the base to the emitter, and at the same time, to a change in the electron current from the emitter into the base. However, the latter is collected completely by the collector, and it thus follows that the current gain is simply given by (14-29); the current gain so obtained may be 100 or more. Other factors also favor a high gain. The collector impedance is very high because of the reverse bias; it is evident from the reverse junction characteristic in Fig. 14-8 that for voltages larger than a few times kT/e , the collector current is essentially independent of the bias, i.e., the impedance would approach infinity. Actual collector impedances are of the order of 10^6 ohms or higher. Furthermore, the resistance r_e of the emitter is very low; in fact, it follows from the forward characteristic that

$$r_e = (kT/eI)$$

where I is the emitter current. For $I = 1$ ma, this gives at room temperature, $r_e = 25$ ohms.

From the above discussion one arrives at the equivalent circuit represented in Fig. 14-10. Here r_e is the emitter junction resistance, r_b represents the resistance of the thin p -type base region, and r_c is the resistance of the collector junction. These resistances are, of course, functions of the bias voltages. The collecting action may be represented by a current generator αi_e , where i_e is the emitter alternating current and α is the fraction of the emitter current collected by the collector. For a good junction transistor, α is nearly unity, as was assumed above.

A point-contact transistor (called type A) is represented schematically in Fig. 14-11. The emitter and collector in this case are metallic points pressed on the surface of a small die of n -type material. The base contact is simply a large area contact at the bottom of the die. The emitter is positive relative to the n -type material, and thus injects holes into the germanium. The holes diffuse towards the collector under influence of the electric field. This hole current adds to the electron current flowing from the collector into the germanium as a result of the reverse bias of the collector. At the same time the presence of the holes near the collector enhances the electronic current. The ratio of the collector current increase to the emitter current increase is again denoted by α ; in point-contact transistors α is therefore larger than unity (see table below). Since the collector current flows through the high collector impedance, whereas the emitter current is injected through the low emitter impedance, one also obtains a voltage gain. Some typical values for a point-contact type A transistor and a junction transistor are given below; the resistances are in ohms.

	<i>Junction</i>	<i>Point-contact</i>
r_e	25	125
r_b	200	75
r_c	5×10^6	2×10^4
α	0.95-0.99	2-3

REFERENCES

- J. Bardeen and W. H. Brattain, *Phys. Rev.*, **75**, 1208 (1949); *Bell System Tech. J.*, **28**, 239 (1949).
- J. S. Blakemore, A. E. De Barr, and J. B. Gunn, "Semiconductor circuit Elements," *Repts. Progr. Phys.*, **16**, 160 (1953).
- H. K. Henisch, *Metal Rectifiers*, Oxford, New York, (1949).
- Proc. IRE.*, **40**, (1952) (transistor issue).
- W. Shockley, *Electrons and Holes in Semiconductors*, Van Nostrand, New York, 1950.
- H. C. Torrey and C. A. Whitmer, *Crystal Rectifiers*, McGraw-Hill, New York, 1948.

PROBLEMS

14-1. Two metallic surfaces with work functions of 3 and 4 eV are separated by a gap of 10 \AA . Calculate the surface charge density in equilibrium at room temperature in terms of a number of electrons.

14-2. A metal with a work function of 3 eV is in contact with a semiconductor with an electron affinity of 1 eV; the semiconductor contains 10^{16} donors per cm^3 close to the conduction band. Calculate the capacitance of the barrier layer per cm^2 for zero applied voltage, when the dielectric constant is 12. Do the same problem for a reverse bias of 5 volts.

14-3. Consider a block of semiconducting material with a large area contact on one of its faces; the opposite face has a small circular point contact of radius a . Show that the bulk or spreading resistance of the system is $r = 1/4\sigma a$, where σ is the conductivity of the semiconductor.

14-4. Consider an idealized p - n junction in which the acceptor concentration is constant for $x < 0$, and the donor concentration is constant for $x > 0$. Find an expression for the barrier thickness in terms of the acceptor and donor concentrations and the forbidden energy gap; assume that the donor and acceptor levels lie very close, respectively, to the conduction and valence bands. Also discuss the variation of the barrier thickness with an applied voltage.

14-5. Repeat Problem 14-4 for a junction consisting of a p -type region containing N acceptors per cm^3 and an n -type region containing N donors per cm^3 , the two regions being separated by a transition region in which the concentrations vary linearly with x . Assume that the transition region is large compared with the physical barrier layer.

14-6. Consider a p - n junction with an area of 0.25 cm^2 in which the current is carried mainly by holes. Given that for small forward voltages the junction resistance is 800 ohms, calculate the density of holes in the n -region if the life time of the holes in this region is 10^{-4} sec and their mobility is $1800 \text{ cm}^2 \text{ volt}^{-1} \text{ sec}^{-1}$.