Physics of Semiconductor Devices, 3rd Edition
by S. M. Sze and Kwok K. Ng

3

Metal-Semiconductor Contacts

3.1 INTRODUCTION

The earliest systematic investigation on metal-semiconductor rectifying systems is generally attributed to Braun, who in 1874 noted the dependence of the total resistance of a point contact on the polarity of the applied voltage and on the detailed surface conditions.[1] The point-contact rectifier in various forms found practical applications beginning in 1904.[2] In 1931, Wilson formulated the transport theory of semiconductors based on the band theory of solids.[3] This theory was then applied to metal-semiconductor contacts. In 1938, Schottky suggested that the potential barrier could arise from stable space charges in the semiconductor alone without the presence of a chemical layer.[4] The model arising from this consideration is known as the Schottky barrier. In 1938, Mott also devised a more appropriate theoretical model for swept-out metal-semiconductor contacts that is known as the Mott barrier.[5] These models were further enhanced by Bethe in 1942 to become the thermionic-emission model which accurately describes the electrical behavior.[6] The basic theory, the historical development, and the device technology of rectifying metal-semiconductor contacts can be found in References 7–11.

Because of their importance in direct current and microwave applications and as intricate parts of other semiconductor devices, metal-semiconductor contacts have been studied extensively. Specifically, they have been used as photodetectors, solar cells, as the gate electrode of the MESFET, etc. Most importantly, the metal contact on heavily doped semiconductor forms an ohmic contact that is required for every semiconductor device in order to pass current in and out of the device.

3.2 FORMATION OF BARRIER

When metal makes contact with a semiconductor, a barrier is formed at the metal-semiconductor interface. This barrier is responsible for controlling the current conduction as well as its capacitance behavior. In this section, we consider the basic energy-band diagrams leading to the formation of the barrier height and some effects that can modify the value of this barrier.

3.2.1 Ideal Condition

We will first consider the ideal case without surface states and other anomalies. Figure 1a shows the electronic energy relations of a high work-function metal and an n-type semiconductor which are not in contact and are in separate systems. If the two are allowed to communicate with each other, for example by an external wire connection, charge will flow from the semiconductor to the metal and thermal equilibrium is established as a single system. The Fermi levels on both sides will line up. Relative to the Fermi level in the metal, the Fermi level in the semiconductor is lowered by an amount equal to the difference between the two work functions.

The work function is the energy difference between the vacuum level and the Fermi level. This quantity is denoted by $q\phi_m$ for the metal, and is equal to $q(\chi + \phi_n)$ in the semiconductor, where $q\chi$ is the electron affinity measured from the bottom of the conduction band E_C to the vacuum level, and $q\phi_n$ is the energy difference between E_C

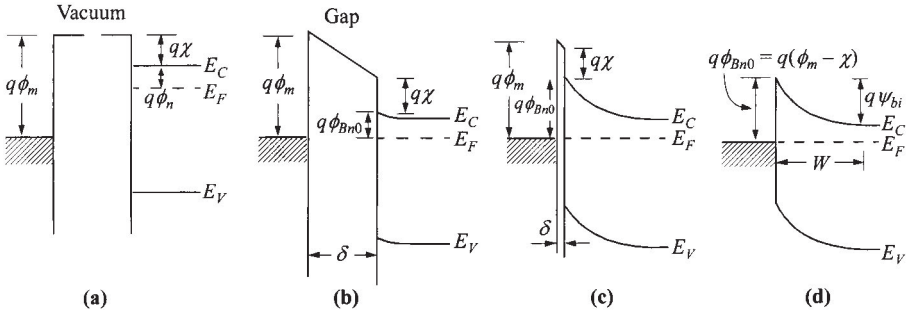


Fig. 1 Energy-band diagrams of metal-semiconductor contacts. Metal and semiconductor (a) in separated systems, and (b) connected into one system. As the gap δ (c) is reduced and (d) becomes zero. (After Ref. 7.)

and the Fermi level. The potential difference between the two work functions $\phi_m - (\chi + \phi_n)$ is called the contact potential. As the gap distance δ decreases, the electric field in the gap increases and an increasing negative charge is built up at the metal surface. An equal and opposite charge (positive) must exist in the semiconductor depletion region. The potential variation within the depletion layer is similar to that in one side of a p-n junction. When δ is small enough to be comparable to the interatomic distances, the gap becomes transparent to electrons, and we obtain the limiting case, as shown on the far right (Fig. 1d). It is clear that the limiting value of the barrier height $q\phi_{Bn0}$ is given by

$$q\phi_{Bn0} = q(\phi_m - \chi). \tag{1}$$

The barrier height is simply the difference between the metal work function and the electron affinity of the semiconductor. Conversely, for an ideal contact between a metal and a p-type semiconductor, the barrier height $q\phi_{Bp0}$ is given by

$$q\phi_{Bp0} = E_g - q(\phi_m - \chi). \tag{2}$$

Thus, for any given semiconductor and metal combination, the sum of the barrier heights on n-type and p-type substrates is expected to be equal to the bandgap, or

$$q(\phi_{Bn0} + \phi_{Bp0}) = E_g. \tag{3}$$

In practice, however, simple expressions for the barrier heights as given by Eqs. 1 and 2 are never realized experimentally. The electron affinities of semiconductors and the work functions of metals have been established. For metals, $q\phi_m$ is of the order of a few electron volts (2 – 6 eV). The values of $q\phi_m$ are generally very sensitive to surface contamination. The most reliable values for clean surfaces are given in Fig. 2. The main deviations of experimental barrier heights from the ideal condition are: (1) an unavoidable interface layer, $\delta \neq 0$ as in Fig. 1c, and (2) the presence of interface states. Furthermore, the barrier height can be modified due to image-force lowering. These effects will be discussed in the following sections.

3.2.2 Depletion Layer

The depletion layer of a metal-semiconductor contact is similar to that of the one-sided abrupt (e.g., p^+-n) junction. It is clear from the discussion above that when a metal is brought into intimate contact with a semiconductor, the conduction and valence bands of the semiconductor at the surface are brought into a definite energy relationship with the Fermi level in the metal. Once this relationship is established, it serves as a boundary condition to the solution of the Poisson equation in the semiconductor, which proceeds in exactly the same manner as in a p-n junction. The energy-band diagrams for metals on both n-type and p-type materials are shown, under different biasing conditions, in Fig. 3.

For contacts on n-type semiconductors, under the abrupt approximation that $\rho \approx qN_D$ for $x < W_D$, $\rho \approx 0$ and $\mathscr{E} \approx 0$ for $x > W_D$, where W_D is the depletion width, we obtain

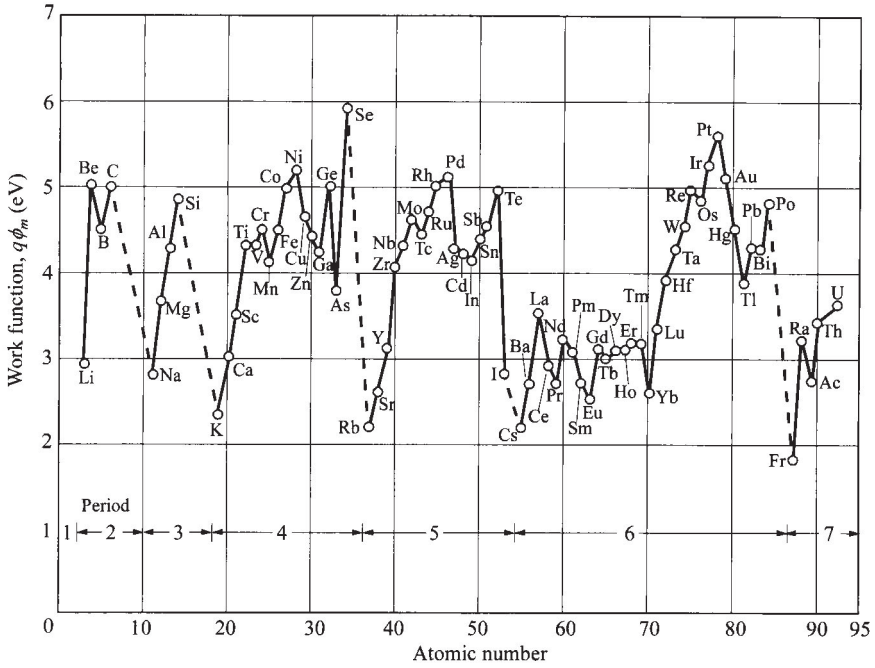


Fig. 2 Metal work function for a clean metal surface in a vacuum versus atomic number. Note the periodic nature of the increase and decrease of the work functions within each group. (After Ref. 12.)

$$W_D = \sqrt{\frac{2\varepsilon_s}{qN_D}\left(\psi_{bi} - V - \frac{kT}{q}\right)} \tag{4}$$

$$|\mathscr{E}(x)| = \frac{qN_D}{\varepsilon_s}(W_D - x) = \mathscr{E}_m - \frac{qN_D x}{\varepsilon_s} \tag{5}$$

$$E_C(x) = q\phi_{Bn} - \frac{q^2 N_D}{\varepsilon_s}\left(W_D x - \frac{x^2}{2}\right) \tag{6}$$

where the term kT/q arises from the contribution of the majority-carrier distribution tail (electrons in n-side, see the footnote on p. 84) and $\mathscr{E}_m$ is the maximum field strength which occurs at $x = 0$:

$$\mathscr{E}_m = \mathscr{E}(x = 0) = \sqrt{\frac{2qN_D}{\varepsilon_s}\left(\psi_{bi} - V - \frac{kT}{q}\right)} = \frac{2[\psi_{bi} - V - (kT/q)]}{W_D}. \tag{7}$$

The space charge Q_{sc} per unit area of the semiconductor and the depletion-layer capacitance C_D per unit area are given by

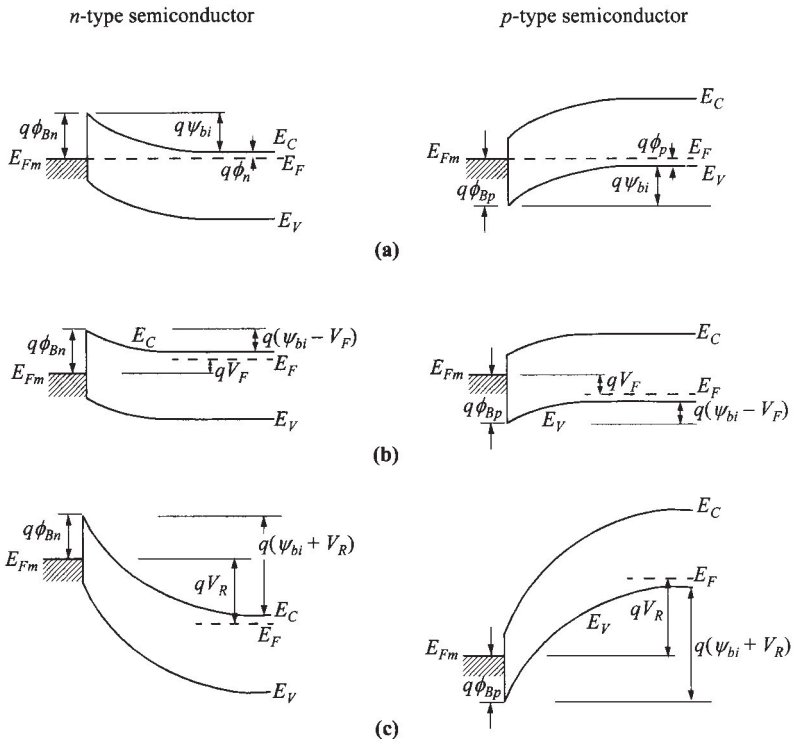


Fig. 3 Energy-band diagrams of metal on n-type (left) and on p-type (right) semiconductors under different biasing conditions. (a) Thermal equilibrium. (b) Forward bias. (c) Reverse bias.

$$Q_{sc} = qN_D W_D = \sqrt{2q\varepsilon_s N_D\left(\psi_{bi} - V - \frac{kT}{q}\right)} \tag{8}$$

$$C_D \equiv \frac{\varepsilon_s}{W_D} = \sqrt{\frac{q\varepsilon_s N_D}{2[\psi_{bi} - V - (kT/q)]}}. \tag{9}$$

Equation 9 can be written in the form

$$\frac{1}{C_D^2} = \frac{2[\psi_{bi} - V - (kT/q)]}{q\varepsilon_s N_D} \tag{10}$$

or

$$N_D = \frac{2}{q\varepsilon_s}\left[-\frac{1}{d(1/C_D^2)/dV}\right]. \tag{11}$$

If N_D is constant throughout the depletion region, one should obtain a straight line by plotting $1/C_D^2$ versus voltage. If N_D is not a constant, the differential capacitance method can be used to determine the doping profile from Eq. 11, similar to the case of a one-sided *p-n* junction as discussed in Section 2.2.1.

The *C-V* measurement can also be used to study deep impurity levels. Figure 4 shows a semiconductor with one shallow donor level and one deep donor level.[13] While all the shallow donors above the Fermi level will be ionized, only deep impurities near the surface are above the Fermi level and ionized, giving a higher effective doping concentration near the interface. In a *C-V* measurement where a small ac signal is superimposed on the dc bias, there will be a frequency dependence on capacitance since the deep impurities can only follow slow signals, i.e. dN_T/dV is absent at high frequencies. Comparing *C-V* measurements at various frequencies can reveal the properties of these deep-level impurities.

3.2.3 Interface States

The barrier heights of metal-semiconductor systems are, in general, determined by both the metal work function and the interface states. A general expression of the barrier height can be obtained on the basis of the following two assumptions:[14] (1) with intimate contact between the metal and the semiconductor, and with an interfa-

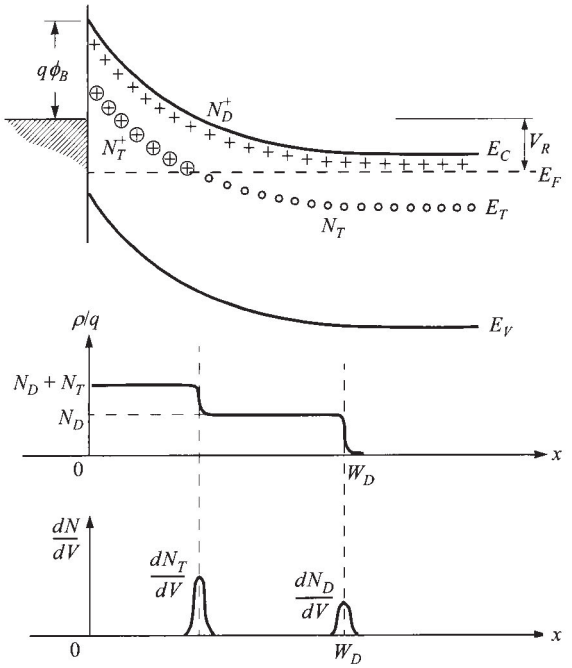
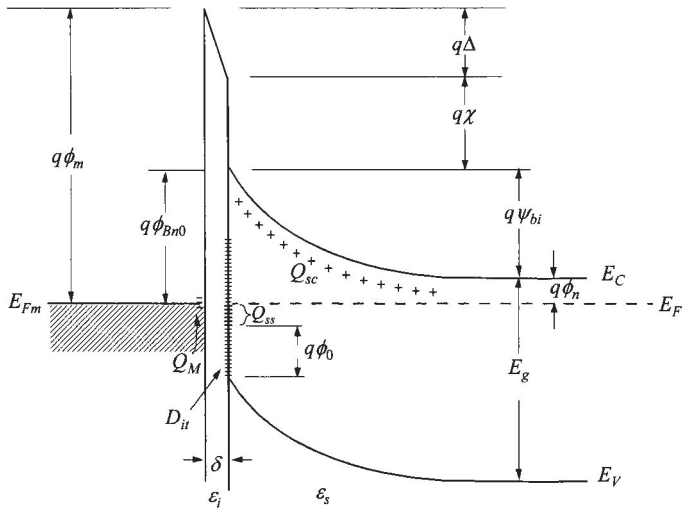


Fig. 4 Semiconductor with one shallow donor level and one deep donor level. N_D and N_T are the shallow donor and deep donor concentration, respectively. (After Ref. 13.)

cial layer of atomic dimensions, this layer will be transparent to electrons but can withstand potential across it, and (2) the interface states per unit area per energy at the interface are a property of the semiconductor surface and are independent of the metal. A more detailed energy-band diagram of a practical metal-n-semiconductor contact is shown in Fig. 5. The various quantities used in the derivation that follows are defined in this figure. The first quantity of interest is the energy level $q\phi_0$ above E_V at the semiconductor surface. It is called the neutral level above which the states are of acceptor type (neutral when empty, negatively charged when full) and below which the states are of donor type (neutral when full of electrons, positively charged when empty). Consequently, when the Fermi level at the surface coincides with this neutral level, the net interface-trap charge is zero.[15] This energy level also tends to pin the semiconductor Fermi level at the surface before the metal contact was formed.



ϕ_m = Work function of metal
ϕ_{Bn0} = Barrier height (without image-force lowering)
ϕ_0 = Neutral level (above E_V) of interface states
Δ = Potential across interfacial layer
χ = Electron affinity of semiconductor
ψ_{bi} = Built-in potential
δ = Thickness of interfacial layer
Q_{sc} = Space-charge density in semiconductor
Q_{ss} = Interface-trap charge
Q_M = Surface-charge density on metal
D_{it} = Interface-trap density
ε_i = Permittivity of interfacial layer (vacuum)
ε_s = Permittivity of semiconductor

Fig. 5 Detailed energy-band diagram of a metal-n-semiconductor contact with an interfacial layer (vacuum) of the order of atomic distance. (After Ref. 14.)

The second quantity is $q\phi_{Bn0}$, the barrier height of the metal-semiconductor contact; it is this barrier that must be surmounted by electrons flowing from the metal into the semiconductor. The interfacial layer will be assumed to have a thickness of a few angstroms and will therefore be essentially transparent to electrons.

We consider a semiconductor with acceptor interface traps (since in this particular example E_F is above the neutral level) whose density is D_{it} states/cm^2-eV, and is a constant over the energy range from $q\phi_0 + E_V$ to the Fermi level. The interface-trap charge density on the semiconductor Q_{ss} is therefore negative and is given by

$$Q_{ss} = -qD_{it}(E_g - q\phi_0 - q\phi_{Bn0}) \qquad \text{C/cm}^2. \tag{12}$$

The quantity in parentheses is simply the energy difference between the Fermi level at the surface and the neutral level. The interface-trap density D_{it} times this quantity yields the number of surface states above the neutral level that are full.

The space charge that forms in the depletion layer of the semiconductor at thermal equilibrium is given as

$$Q_{sc} = qN_D W_D = \sqrt{2q\varepsilon_s N_D\left(\phi_{Bn0} - \phi_n - \frac{kT}{q}\right)}. \tag{13}$$

The total equivalent surface charge density on the semiconductor surface is given by the sum of Eqs. 12 and 13. In the absence of any space-charge effects in the interfacial layer, an exactly equal and opposite charge, Q_M (C/cm^2), develops on the metal surface. For thin interfacial layers such space-charge effects are negligible and Q_M can be written as

$$Q_M = -(Q_{ss} + Q_{sc}). \tag{14}$$

The potential Δ across the interfacial layer can be obtained by applying Gauss' law to the surface charge on the metal and semiconductor:

$$\Delta = -\frac{\delta Q_M}{\varepsilon_i} \tag{15}$$

where ε_i is the permittivity of the interfacial layer and δ its thickness. Another relation for Δ can be obtained by inspection of the energy-band diagram of Fig. 5:

$$\Delta = \phi_m - (\chi + \phi_{Bn0}). \tag{16}$$

This relation results from the fact that the Fermi level must be constant throughout this system at thermal equilibrium.

If Δ is eliminated from Eqs. 15 and 16, and Eq. 14 is used to substitute for Q_M, we obtain

$$\phi_m - \chi - \phi_{Bn0} = \sqrt{\frac{2q\varepsilon_s N_D \delta^2}{\varepsilon_i^2}\left(\phi_{Bn0} - \phi_n - \frac{kT}{q}\right)} - \frac{qD_{it}\delta}{\varepsilon_i}(E_g - q\phi_0 - q\phi_{Bn0}). \tag{17}$$

Equation 17 can now be solved for ϕ_{Bn0}. We introduce the quantities

$$c_1 \equiv \frac{2q\varepsilon_s N_D \delta^2}{\varepsilon_i^2}, \tag{18}$$

$$c_2 \equiv \frac{\varepsilon_i}{\varepsilon_i + q^2 \delta D_{it}} \tag{19}$$

which contain all the interfacial properties. Equation 18 can be used to calculate c_1 if values of δ and ε_i are estimated. For vacuum-cleaved or well-cleaned semiconductor substrates the interfacial layer will have a thickness of atomic dimensions (i.e., 4 or 5 Å). The permittivity of such a thin layer can be well approximated by the free-space value and since this approximation represents a lower limit for ε_i, it leads to an over-estimation of c_2. For $\varepsilon_s \approx 10\varepsilon_0$, $\varepsilon_i = \varepsilon_0$, and $N_D < 10^{18}$ cm^{-3}, c_1 is small, of the order of 0.01 V and the square-root term in Eq. 17 is estimated to be less than 0.1 V. Neglecting this square-root term, Eq. 17 reduces to

$$\phi_{Bn0} = c_2(\phi_m - \chi) + (1 - c_2)\left(\frac{E_g}{q} - \phi_0\right) \equiv c_2 \phi_m + c_3 . \tag{20}$$

With known c_2 and c_3 from experiments of varying ϕ_m, the interfacial properties are given by

$$\phi_0 = \frac{E_g}{q} - \frac{c_2 \chi + c_3}{1 - c_2}, \tag{21}$$

$$D_{it} = \frac{(1 - c_2)\varepsilon_i}{c_2 \delta q^2} . \tag{22}$$

Using the previous assumptions for δ and ε_i, we obtain $D_{it} \approx 1.1 \times 10^{13}(1 - c_2)/c_2$ states/cm^2-eV.

There are two limiting cases which can be obtained directly from Eq. 20:

1. When $D_{it} \to \infty$, then $c_2 \to 0$ and

$$q\phi_{Bn0} = E_g - q\phi_0 . \tag{23}$$

In this case the Fermi level at the interface is *pinned* by the surface states at the value $q\phi_0$ above the valence band. The barrier height is independent of the metal work function and is determined entirely by the surface properties of the semiconductor.

2. When $D_{it} \to 0$, then $c_2 \to 1$ and

$$q\phi_{Bn0} = q(\phi_m - \chi) . \tag{24}$$

This equation for the barrier height of an ideal Schottky barrier where surface-state effects are neglected, is identical to Eq. 1.

The experimental results of the metal-*n*-silicon system are shown in Fig. 6a. A least-square straight-line fit to the data yields

$$q\phi_{Bn0} = 0.27 q\phi_m - 0.52 . \tag{25}$$

Comparing this expression with Eq. 20 ($c_2 = 0.27$, $c_3 = -0.52$) and using Eqs. 21 and 22, we obtain $q\phi_0 = 0.33$ eV, and $D_{it} = 4 \times 10^{13}$ states/cm^2-eV. Similar results are obtained for GaAs, GaP, and CdS, which are shown in Fig. 6b and listed in Table 1.

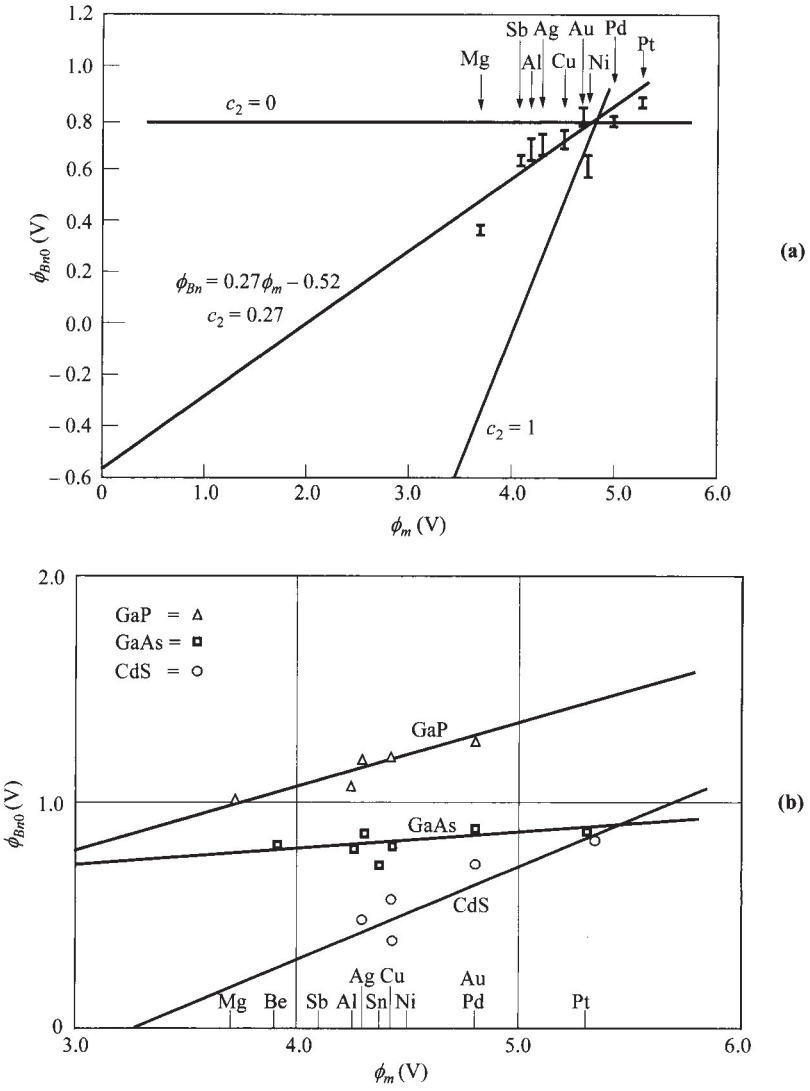


Fig. 6 Experimental barrier heights for different metals on *n*-type (a) silicon and (b) GaAs, GaP, and CdS. (After Ref. 14.)

It should be pointed out that in spite of nonideal factors such as interface states, the relationship of Eq. 3, that the sum of barrier heights on *n*- and *p*-type substrates equals the energy gap of the semiconductor, is still generally valid.

We note that the values of $q\phi_0$ for Si, GaAs, and GaP are very close to one-third of the bandgap. Similar results are obtained for other semiconductors.[16] This fact indicates that most covalent semiconductor surfaces have a high peak density of

Table 1 Summary of Barrier Height Data and Calculations of Interface Properties for Si, GaAs, GaP, and CdS (After Ref. 14)

Semi-conductor	c_2	c_3 (V)	χ (V)	D_{it} (10^{13} /eV-cm^2)	$q\phi_0$ (eV)	$q\phi_0/E_g$
Si	0.27±0.05	−0.52±0.22	4.05	2.7±0.7	0.30±0.36	0.27
GaAs	0.07±0.05	0.51±0.24	4.07	12.5±10.0	0.53±0.33	0.38
GaP	0.27±0.03	0.02±0.13	4.0	2.7±0.4	0.66±0.2	0.294
CdS	0.38±0.16	−1.17±0.77	4.8	1.6±1.1	1.5±1.5	0.6

surface states or defects near the neutral level and that the neutral level is about one-third of the bandgap from the valence-band edge. The theoretical calculation by Pugh[17] for ⟨111⟩ diamond indeed gives a narrow band of surface states slightly below the center of the forbidden gap. It is thus expected that a similar situation may exist for other semiconductors.

For III-V compounds, extensive measurements using photoemission spectroscopy indicate that the Schottky-barrier formation is due mainly to defects generated near the interface by deposition of the metal.[18] It has been shown that on a few compound semiconductors such as GaAs, GaSb, and InP, the surface Fermi-level positions obtained from a number of metals are pinned at an energy level quite independent of the metal.[19] This pinning of surface Fermi level can explain the fact that for most III-V compounds, the barrier height is essentially independent of metal work function.

For ionic semiconductors such as CdS and ZnS, the barrier height generally depends strongly on the metal and a correlation has been found between interface behavior and the electronegativity. The electronegativity X_M is defined as the power of an atom in a molecule to attract electrons to itself. Figure 7 shows Pauling's electronegativity scale. Note that the periodicity is similar to that for the work function (Fig. 2).

Figure 8a shows a plot of the barrier height versus the electronegativity of metals deposited on Si, GaSe, and SiO_2. From the plot we define the slope as an index of interface behavior:

$$S \equiv \frac{d\phi_{Bn0}}{dX_M}. \tag{26}$$

Note the comparison of S to c_2 ($= d\phi_{Bn0}/d\phi_m$). We can also plot the index S as a function of the electronegativity difference (ionicity ΔX) of the semiconductors, shown in Fig. 8b. The electronegativity difference is defined as the difference in the Pauling electronegativities between the cation and the anion of the semiconductor. Note a sharp transition from the covalent semiconductors (such as GaAs with $\Delta X = 0.4$) to ionic semiconductors (such as AlN with $\Delta X = 1.5$). For semiconductors with $\Delta X < 1$, the index S is small, indicating that the barrier height is only weakly dependent on metal electronegativity (or the work function). On the other hand, for $\Delta X > 1$, the

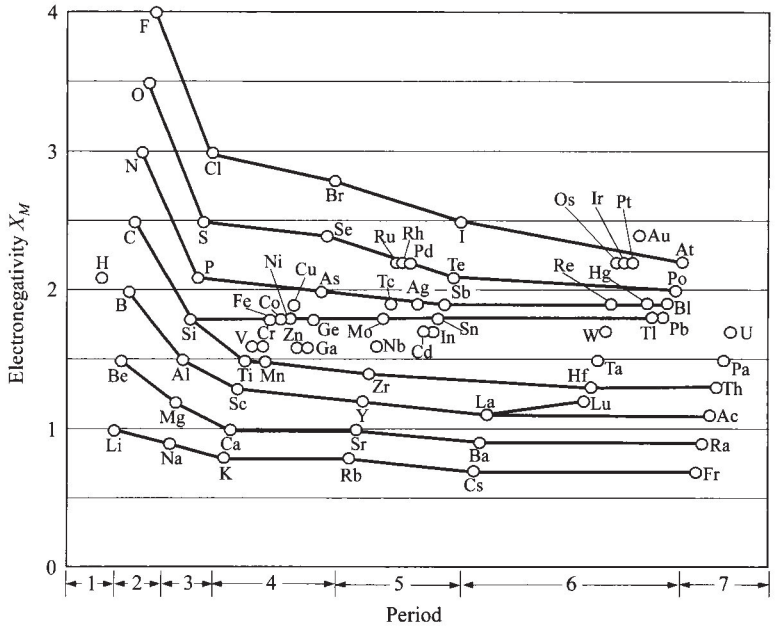


Fig. 7 Pauling's electronegativity scale. Note the trend of increasing electronegativity within each group. (After Ref. 20.)

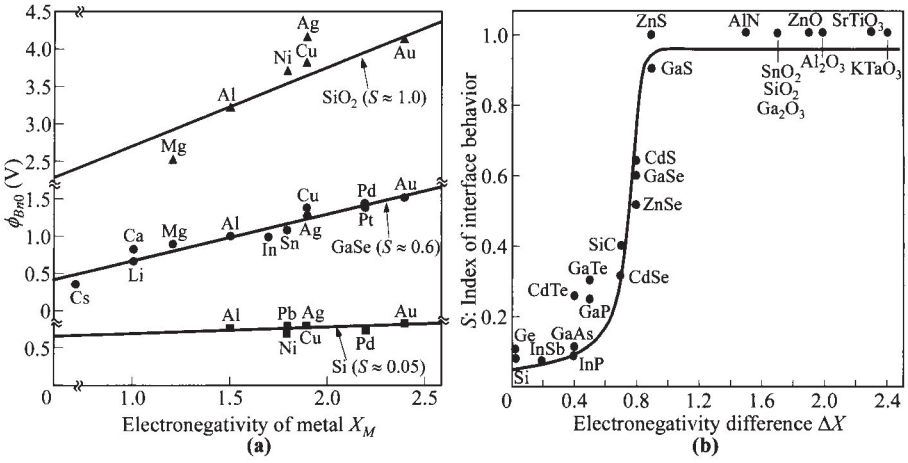


Fig. 8 (a) Barrier height versus electronegativity of metals deposited on Si, GaSe, and SiO_2. (b) Index of interface behavior S as a function of the electronegativity difference of the semiconductors. (After Ref. 21.)

index S approaches 1, and the barrier height is strongly dependent on the metal electronegativity (or the work function).

For technological applications in silicon integrated circuits, an important class of Schottky barrier contacts has been developed in which a chemical reaction between the metal and the underlying silicon is induced to form silicides.[22] The formation of metal silicides by solid-solid metallurgical reaction provides more reliable and reproducible Schottky barriers, because the interface chemical reactions are well defined and can be maintained under good control. It is thought that since the silicide interfacial properties depends on the eutectic temperature, there should be a correlation between the barrier height and the eutectic temperature. Figure 9 shows such an empirical fit for the barrier heights on n-type silicon of transition-metal silicides plotted against the eutectic temperature of the silicides. Similar correlation had been observed when barrier heights are plotted against the heat of formation of silicides.[24]

3.2.4 Image-Force Lowering

The image-force lowering, also known as the Schottky effect or Schottky-barrier lowering, is the image-force-induced lowering of the barrier energy for charge carrier emission, in the presence of an electric field. Consider a metal-vacuum system first. The minimum energy necessary for an electron to escape into vacuum from an initial energy at the Fermi level is the work function $q\phi_m$ as shown in Fig. 10. When an electron is at a distance x from the metal, a positive charge will be induced on the metal surface. The force of attraction between the electron and the induced positive charge is equivalent to the force that would exist between the electron and an equal positive

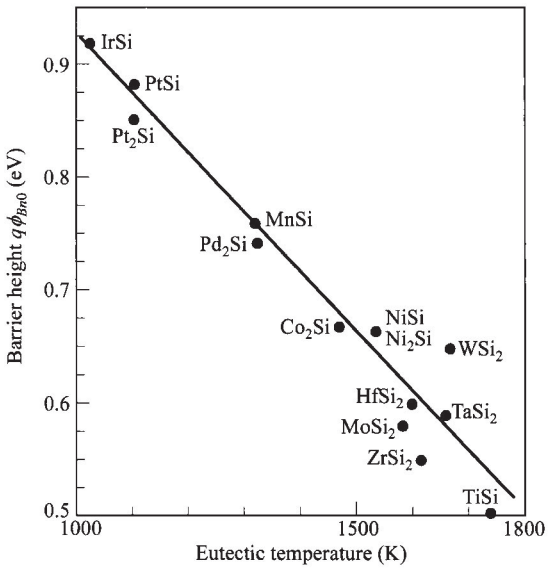


Fig. 9 Correlation of barrier height of transition-metal silicides with their eutectic temperature. (After Ref. 23.)

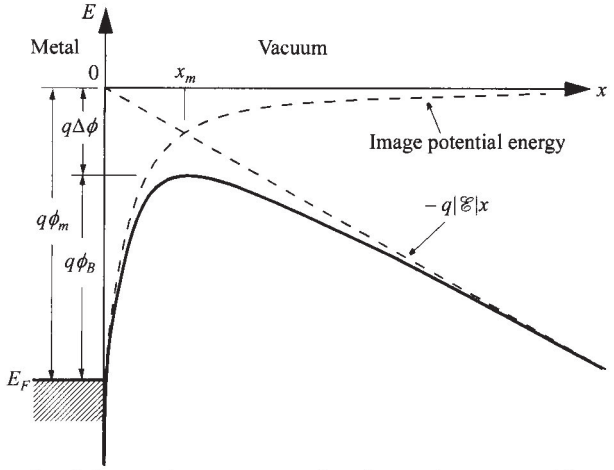


Fig. 10 Energy-band diagram between a metal surface and a vacuum. The metal work function is $q\phi_m$. The effective barrier is lowered when an electric field is applied to the surface. The lowering is due to the combined effects of the field and the image force.

charge located at $-x$. This positive charge is referred to as the image charge. The attractive force toward the metal, called the image force, is given by

$$F = \frac{-q^2}{4\pi\varepsilon_0(2x)^2} = \frac{-q^2}{16\pi\varepsilon_0 x^2} \tag{27}$$

where ε_0 is the permittivity of free space. The work done to an electron in the course of its transfer from infinity to the point x is given by

$$E(x) = \int_{\infty}^{x} F dx = \frac{-q^2}{16\pi\varepsilon_0 x}. \tag{28}$$

This energy corresponds to the potential energy of an electron placed at a distance x from the metal surface, shown in Fig. 10, and is measured downwards from the x-axis. When an external field $\mathscr{E}$ is applied (in this example in the $-x$ direction), the total potential energy PE as a function of distance is given by the sum

$$PE(x) = -\frac{q^2}{16\pi\varepsilon_0 x} - q|\mathscr{E}|x. \tag{29}$$

This equation has a maximum value. The image-force lowering $\Delta\phi$ and the location of the lowering x_m (as shown in Fig. 10), are given by the condition $d(PE)/dx = 0$, or

$$x_m = \sqrt{\frac{q}{16\pi\varepsilon_0|\mathscr{E}|}} \tag{30}$$

$$\Delta\phi = \sqrt{\frac{q|\mathscr{E}|}{4\pi\varepsilon_0}} = 2|\mathscr{E}|x_m. \tag{31}$$

From Eqs. 30 and 31 we obtain $\Delta\phi = 0.12$ V and $x_m = 6$ nm for $\mathscr{E} = 10^5$ V/cm; and $\Delta\phi = 1.2$ V and $x_m = 1$ nm for $\mathscr{E} = 10^7$ V/cm. Thus at high fields the Schottky barrier is considerably lowered, and the effective metal work function for thermionic emission ($q\phi_B$) is reduced.

These results can be applied to metal-semiconductor systems. However, the field should be replaced by the appropriate field at the interface, and the free-space permittivity ε_0 should be replaced by an appropriate permittivity ε_s characterizing the semiconductor medium, that is,

$$\Delta\phi = \sqrt{\frac{q\mathscr{E}_m}{4\pi\varepsilon_s}}. \tag{32}$$

Note that inside a device such as metal-semiconductor contact, the field is not zero even without bias due to the built-in potential. Because of the larger values of ε_s in a metal-semiconductor system, the barrier lowering is smaller than that in a corresponding metal-vacuum system. For example, for $\varepsilon_s = 12\varepsilon_0$, $\Delta\phi$ as obtained from Eq. 32 is only 0.035 V for $\mathscr{E} = 10^5$ V/cm and even smaller for smaller fields. Also a typical value for x_m is calculated to be less than 5 nm. Although the barrier lowering is small, it does have a profound effect on current transport processes in metal-semiconductor systems. These are considered in Section 3.3.

In a practical Schottky-barrier diode, the electric field is not constant with distance, and the maximum value at the surface based on the depletion approximation can be used,

$$\mathscr{E}_m = \sqrt{\frac{2qN|\psi_s|}{\varepsilon_s}}, \tag{33}$$

where the surface potential ψ_s (on n-type substrate) is

$$|\psi_s| = \phi_{Bn0} - \phi_n + V_R. \tag{34}$$

Substituting $\mathscr{E}_m$ into Eq. 32 gives

$$\Delta\phi = \sqrt{\frac{q\mathscr{E}_m}{4\pi\varepsilon_s}} = \left[\frac{q^3N|\psi_s|}{8\pi^2\varepsilon_s^3}\right]^{1/4}. \tag{35}$$

Figure 11 shows the energy diagram incorporating the Schottky effect for a metal on n-type semiconductor under different biasing conditions. Note that for forward bias (V > 0), the field and the image force are smaller and the barrier height $q\phi_{Bn0} - q\Delta\phi_F$ is slightly larger than the barrier height at zero bias of

$$q\phi_{Bn} = q\phi_{Bn0} - q\Delta\phi. \tag{36}$$

For reverse bias ($V_R > 0$), the barrier height $q\phi_{Bn0} - q\Delta\phi_R$ is slightly smaller. In effect, the barrier height becomes bias dependant.

The value ε_s may also be different from the semiconductor static permittivity. If during the emission process, the electron transit time from the metal-semiconductor interface to the barrier maximum x_m is shorter than the dielectric relaxation time, the semiconductor medium does not have enough time to be polarized, and smaller per-

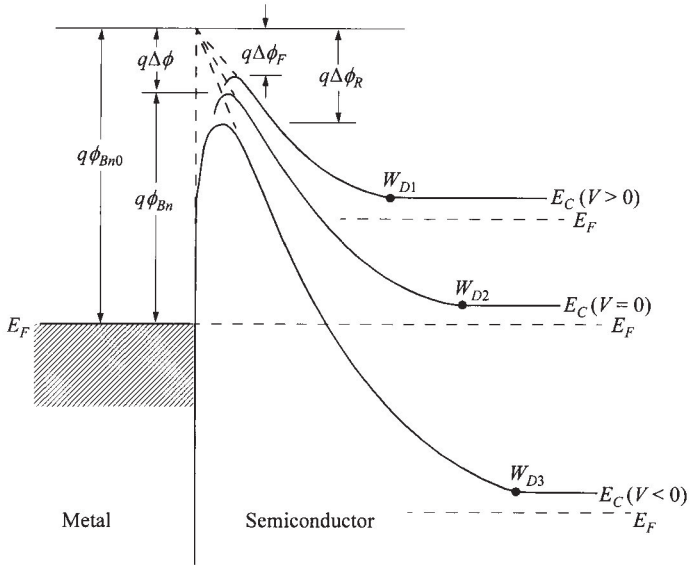


Fig. 11 Energy-band diagram incorporating the Schottky effect for a metal n-type semiconductor contact under different biasing conditions. The intrinsic barrier height is $q\phi_{Bn0}$. The barrier height at thermal equilibrium is $q\phi_{Bn}$. The barrier lowerings under forward and reverse bias are $\Delta\phi_F$ and $\Delta\phi_R$ respectively. (After Ref. 10.)

mittivity than the static value is expected. It will be shown, however, that for Si the appropriate permittivities are about the same as their corresponding static values.

The dielectric constant ($K_s = \varepsilon_s/\varepsilon_0$) in gold-silicon barriers has been obtained from photoelectric measurements, which will be discussed in Section 3.4.4. The experimental results are shown in Fig. 12, where the measured barrier lowering is plotted as a function of the square root of the maximum electric field.[25] From Eq. 35 the image-force dielectric constant is determined to be 12 ± 0.5. For $\varepsilon_s/\varepsilon_0 = 12$, the distance x_m varies between 1 and 5 nm for the field range shown in Fig. 12. Assuming a carrier velocity of the order of 10^7 cm/s, the transit time for these distances should be between $1\text{–}5\times10^{-14}$ s. The image-force dielectric constant should thus be comparable to the value of approximately 12 for electromagnetic radiation of roughly these periods (wavelengths between 3 and 15 μm).[26] The dielectric constant of bulk silicon is essentially constant (11.7) from dc to $\lambda = 1$ μm, therefore the lattice has time to polarize while the electron is traversing the depletion layer. The photoelectric measurements and data deduced from the optical constants are in excellent agreement. For Ge and GaAs, the dependence of the optical dielectric constant on wavelength is similar to that of Si. The image-force permittivities of these semiconductors in the foregoing field range are thus expected to be approximately the same as the corresponding static bulk values.

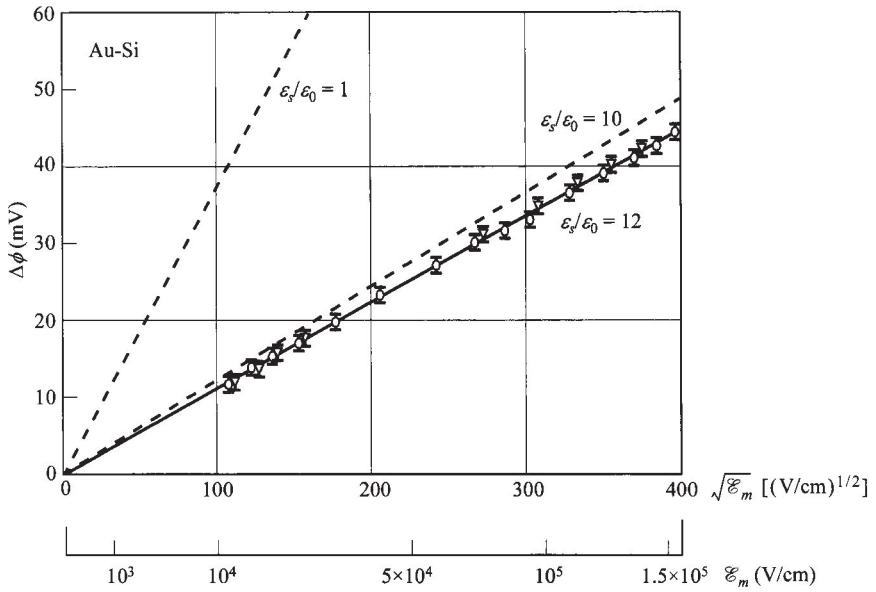


Fig. 12 Measurement of barrier lowering as a function of the electric field in a Au-Si diode. (After Ref. 25.)

3.2.5 Barrier-Height Adjustment

For an ideal Schottky barrier, the barrier height is determined primarily by the characters of the metal and the metal-semiconductor interface properties and is nearly independent of the doping. Usual Schottky barriers on a given semiconductor (e.g., *n*- or *p*-type Si) therefore give a finite number of choices for barrier height. However, by introducing a thin layer ($\approx$ 10 nm or less) of controllable number of dopants on a semiconductor surface (e.g., by ion implantation), the effective barrier height for a given metal-semiconductor contact can be varied.[27–29] This approach is particularly useful in order to select a metal having the most desirable metallurgical properties required for reliable device operation and at the same time to be able to adjust the effective barrier height between this metal and the semiconductor in a controlled manner.

Figure 13a shows the idealized controlled barrier contacts with a thin n^+-layer or a thin p^+-layer on an *n*-type substrate for barrier reduction or barrier increase, respectively. Consider the reduction of barrier first. The field distribution in Fig. 13b is given by

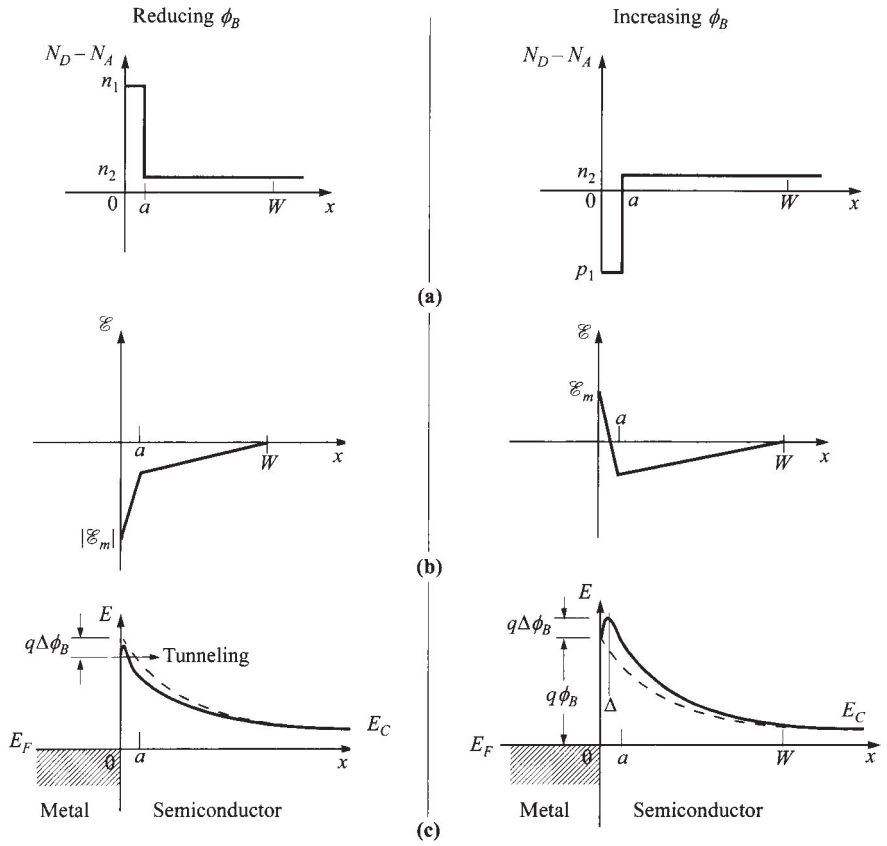


Fig. 13 Idealized controlled barrier contacts with a thin n^+-layer or a thin p^+-layer on an n-type substrate for barrier reduction (left) or barrier increase (right), respectively. Dashed lines indicate original barrier with uniform doping.

$$\begin{aligned} \mathscr{E} &= -|\mathscr{E}_m| + \frac{qn_1x}{\varepsilon_s} \qquad \text{for} \quad 0 < x < a \\ &= -\frac{qn_2}{\varepsilon_s}(W-x) \qquad \text{for} \quad a < x < W \end{aligned} \tag{37}$$

where $\mathscr{E}_m$ is the maximum electric field at the metal-semiconductor interface, and is given by

$$|\mathscr{E}_m| = \frac{q}{\varepsilon_s}[n_1a + n_2(W-a)]. \tag{38}$$

The image-force lowering due to $\mathscr{E}_m$ is given by Eq. 35. For Si and GaAs Schottky barriers with n_2 of the order of 10^{16} cm^{-3} or less, the zero-bias value of $n_2(W-a)$ is

about 10^{11} cm^{-2}. Therefore, if $n_1 a$ is made sufficiently larger than 10^{11} cm^{-2}, Eqs. 38 and 35 can be reduced to

$$|\mathscr{E}_m| \approx \frac{q n_1 a}{\varepsilon_s}, \tag{39}$$

$$\Delta\phi \approx \frac{q}{\varepsilon_s}\sqrt{\frac{n_1 a}{4\pi}}. \tag{40}$$

For $n_1 a = 10^{12}$ and 10^{13} cm^{-2}, the corresponding lowerings are 0.045 and 0.14 V, respectively.

Although the image-force lowering contributes to the barrier reduction, generally the tunneling effect is more significant. For $n_1 a = 10^{13}$ cm^{-2}, the maximum field from Eq. 39 is 1.6×10^6 V/cm, which is the zero-bias field of a Au-Si Schottky diode with a doping of 10^{19} cm^{-3}. The increased saturation current density due to tunneling for such a diode is about 10^{-3} A/cm^2, corresponding to an effective barrier height of 0.6 V (see discussion later on current vs. barrier height), a reduction of 0.2 V from the 0.8 V barrier of the original Au-Si diode. The calculated effective barrier height as a function of $\mathscr{E}_m$ is shown in Fig. 14 for Si and GaAs barriers. By increasing the maximum field from 10^5 V/cm to 10^6 V/cm, one generally can reduce the effective barrier by 0.2 V in Si and over 0.3 V in GaAs.

For a given application, the parameters n_1 and a should be properly chosen so that in the forward direction the larger Schottky-barrier lowering and the added tunneling current will not substantially degrade the ideality factor η. And in the reverse direction, they will not cause large leakage current in the required bias range.

If opposite doping is formed in the thin semiconductor layer at the interface, the effective barrier can be increased. As indicated in Fig. 13a, if the n^+-region is replaced

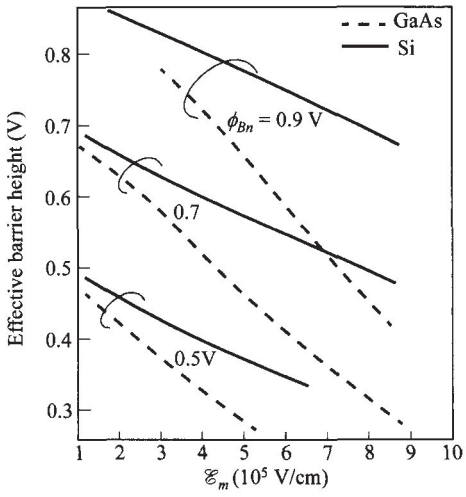


Fig. 14 Calculated reduced effective barrier height from tunneling for Si and GaAs metal-semiconductor contacts. (After Ref. 30.)

by p^+-region, it can be shown that the energy-band profile will be $q\phi_B$ at $x = 0$ and reach a maximum at $x = \Delta$, where

$$\Delta = \frac{1}{p_1}[ap_1 - (W-a)n_2]. \tag{41}$$

The effective barrier height occurs at $x = \Delta$ and is given by

$$\phi_B' = \phi_B + \mathscr{E}_m\Delta - \frac{qp_1\Delta^2}{2\varepsilon_s}. \tag{42}$$

Equation 42 approaches $(\phi_B + qp_1a^2/2\varepsilon_s)$ if $p_1 \gg n_2$ and $ap_1 \gg Wn_2$. Therefore, as the product ap_1 increases, the effective barrier height will increase accordingly.

Figure 15 shows the measured results of Ni-Si diodes with shallow antimony implantation on the surface. As the implant dose increases, the effective barrier height decreases for n-type substrates and increases for p-type substrates.

3.3 CURRENT TRANSPORT PROCESSES

The current transport in metal-semiconductor contacts is due mainly to majority carriers, in contrast to p-n junctions where the minority carriers are responsible. Figure 16 shows five basic transport processes under forward bias (the inverse processes occur under reverse bias).[8] These five processes are (1) emission of electrons from the semiconductor over the potential barrier into the metal [the dominant process for Schottky diodes with moderately doped semiconductors (e.g., Si with $N_D \leq 10^{17}$ cm^{-3}) operated at moderate temperatures (e.g., 300 K)], (2) quantum-mechanical tunneling of electrons through the barrier (important for heavily doped semiconductors and responsible for most ohmic contacts), (3) recombination in the space-charge region [identical to the recombination process in a p-n junction (refer to

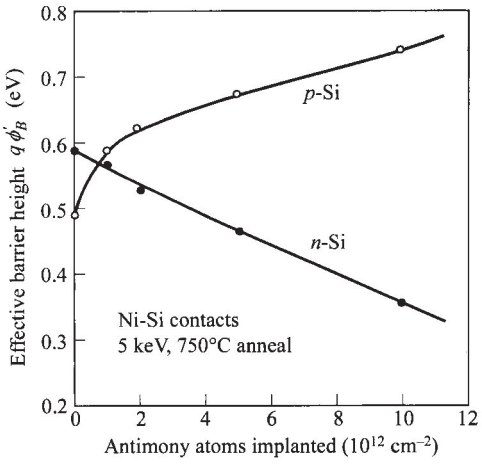


Fig. 15 Effective barrier height for holes in p-type substrates and for electrons in n-type substrates as a function of the implanted antimony dose. (After Ref. 30.)

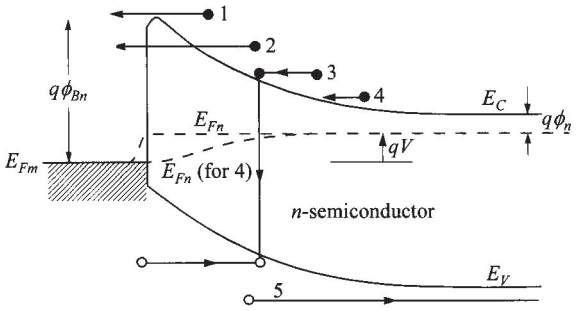


Fig. 16 Five basic transport processes under forward bias. (1) Thermionic emission. (2) Tunneling. (3) Recombination. (4) Diffusion of electrons. (5) Diffusion of holes.

chapter 2)], (4) diffusion of electrons in the depletion region, and (5) holes injected from the metal that diffuse into the semiconductor (equivalent to recombination in the neutral region). In addition, we may have edge leakage current due to a high electric field at the metal-contact periphery or interface current due to traps at the metal-semiconductor interface. Various methods have been used to improve the interface quality, and many device structures have been proposed to reduce or eliminate the edge leakage current (see Section 3.5).

For common high-mobility semiconductors (e.g., Si and GaAs) the transport can be adequately described by this thermionic-emission theory. We shall also consider the diffusion theory applicable to low-mobility semiconductors and a generalized thermionic-emission-diffusion theory that is a synthesis of the preceding two theories.

Schottky diode behavior is to some extent electrically similar to a one-sided abrupt p-n junction, and yet the Schottky diode can be operated as a majority-carrier device with inherent fast response. Thus, the terminal functions of a p-n junction diode can general be performed by a Schottky diode with one exception as a charge-storage diode. This is because the charge-storage time in a majority-carrier device is extremely small. Another difference is the larger current density in a Schottky diode due to the smaller built-in potential as well as the nature of thermionic emission compared to diffusion. This results in a much smaller forward voltage drop. By the same token, the disadvantage is the larger reverse current in the Schottky diode and a lower breakdown voltage.

3.3.1 Thermionic-Emission Theory

The thermionic-emission theory by Bethe[6] is derived from the assumptions that (1) the barrier height $q\phi_{Bn}$ is much larger than kT, (2) thermal equilibrium is established at the plane that determines emission, and (3) the existence of a net current flow does not affect this equilibrium so that one can superimpose two current fluxes—one from metal to semiconductor, the other from semiconductor to metal, each with a different quasi Fermi level. If thermionic emission is the limiting mechanism, then E_{Fn} is flat

throughout the deletion region (Fig. 16). Because of these assumptions, the shape of the barrier profile is immaterial and the current flow depends solely on the barrier height. The current density from the semiconductor to the metal $J_{s\to m}$ is then given by the concentration of electrons with energies sufficient to overcome the potential barrier and traversing in the x-direction:

$$J_{s\to m} = \int_{E_{Fn}+q\phi_{Bn}}^{\infty} qv_x dn \tag{43}$$

where $E_{Fn} + q\phi_{Bn}$ is the minimum energy required for thermionic emission into the metal, and v_x is the carrier velocity in the direction of transport. The electron density in an incremental energy range is given by

$$\begin{aligned} dn &= N(E)F(E)dE \\ &\approx \frac{4\pi(2m^*)^{3/2}}{h^3}\sqrt{E-E_C}\exp\left(-\frac{E-E_C+q\phi_n}{kT}\right)dE \end{aligned} \tag{44}$$

where $N(E)$ and $F(E)$ are the density of states and the distribution function, respectively.

If we postulate that all the energy of electrons in the conduction band is kinetic energy, then

$$E - E_C = \frac{1}{2}m^*v^2 \tag{45}$$

$$dE = m^*v dv \tag{46}$$

$$\sqrt{E-E_C} = v\sqrt{\frac{m^*}{2}}. \tag{47}$$

Substituting Eqs. 45–47 into Eq. 44 gives

$$dn \approx 2\left(\frac{m^*}{h}\right)^3 \exp\left(-\frac{q\phi_n}{kT}\right)\exp\left(-\frac{m^*v^2}{2kT}\right)(4\pi v^2 dv)\ . \tag{48}$$

Equation 48 gives the number of electrons per unit volume that have velocities between v and $v + dv$, distributed over all directions. If the velocity is resolved into its components along the axes with the x-axis parallel to the transport direction, we have

$$v^2 = v_x^2 + v_y^2 + v_z^2. \tag{49}$$

With the transformation $4\pi v^2 dv = dv_x dv_y dv_z$, we obtain from Eqs. 43, 48, and 49

$$\begin{aligned} J_{s\to m} &= 2q\left(\frac{m^*}{h}\right)^3 \exp\left(-\frac{q\phi_n}{kT}\right)\int_{v_{0x}}^{\infty} v_x \exp\left(-\frac{m^*v_x^2}{2kT}\right)dv_x \\ &\qquad \int_{-\infty}^{\infty}\exp\left(-\frac{m^*v_y^2}{2kT}\right)dv_y \int_{-\infty}^{\infty}\exp\left(-\frac{m^*v_z^2}{2kT}\right)dv_z \\ &= \left(\frac{4\pi q m^* k^2}{h^3}\right)T^2\exp\left(-\frac{q\phi_n}{kT}\right)\exp\left(-\frac{m^*v_{0x}^2}{2kT}\right)\ . \end{aligned} \tag{50}$$

The velocity v_{0x} is the minimum velocity required in the x-direction to surmount the barrier and is given by

$$\frac{1}{2}m^*v_{0x}^2 = q(\psi_{bi} - V). \tag{51}$$

Substituting Eq. 51 into Eq. 50 yields

$$\begin{aligned} J_{s\to m} &= \left(\frac{4\pi q m^* k^2}{h^3}\right) T^2 \exp\left(-\frac{q\phi_{Bn}}{kT}\right) \exp\left(\frac{qV}{kT}\right) \\ &= A^* T^2 \exp\left(-\frac{q\phi_{Bn}}{kT}\right) \exp\left(\frac{qV}{kT}\right) , \end{aligned} \tag{52}$$

and

$$A^* = \frac{4\pi q m^* k^2}{h^3} \tag{53}$$

is the effective Richardson constant for thermionic emission, neglecting the effects of optical-phonon scattering and quantum mechanical reflection (see Section 3.3.3). For free electrons ($m^* = m_0$) the Richardson constant A is 120 A/cm^2-K^2. Note that when the image-force lowering is considered, the barrier height ϕ_{Bn} in Eq. 52 is reduced by $\Delta\phi$.

For semiconductors with isotropic effective mass in the lowest minimum of the conduction band such as n-type GaAs, A^*/A simply is equal to m^*/m_0. For multiple-valley semiconductors the appropriate Richardson constant associated with a single energy minimum is given by[31]

$$\frac{A_1^*}{A} = \frac{1}{m_0}\sqrt{l_1^2 m_y^* m_z^* + l_2^2 m_z^* m_x^* + l_3^2 m_x^* m_y^*} \tag{54}$$

where l_1, l_2, and l_3 are the direction cosines of the normal to the emitting plane relative to the principal axes of the ellipsoid, and m_x^*, m_y^*, and m_z^* are the components of the effective mass tensor.

For Si the conduction band minima occur in the ⟨100⟩-directions and $m_l^* = 0.98m_0$, $m_t^* = 0.19m_0$. The minimum value of A^* occurs for the ⟨100⟩-directions:

$$\left(\frac{A^*}{A}\right)_{n\text{-Si}\langle 100\rangle} = \frac{2m_t^*}{m_0} + \frac{4\sqrt{m_l^* m_t^*}}{m_0} = 2.1. \tag{55}$$

In the ⟨111⟩-directions all minima contribute equally to the current, yielding the maximum A^*:

$$\left(\frac{A^*}{A}\right)_{n\text{-Si}\langle 111\rangle} = \frac{6}{m_0}\sqrt{\frac{(m_t^*)^2 + 2m_l^* m_t^*}{3}} = 2.2. \tag{56}$$

For holes in Si and GaAs the two energy maxima at $\boldsymbol{k} = 0$ give rise to approximately isotropic current flow from both the light and heavy holes. Adding the currents due to these carriers, we obtain

$$\left(\frac{A^*}{A}\right)_{p\text{-type}} = \frac{m_{lh}^* + m_{hh}^*}{m_0}. \tag{57}$$

Table 2 gives a summary of the values of A^*/A for Si and GaAs.

Since the barrier height for electrons moving from the metal into the semiconductor remains the same under bias, the current flowing into the semiconductor is thus unaffected by the applied voltage. It must therefore be equal to the current flowing from the semiconductor into the metal when thermal equilibrium prevails (i.e., when $V = 0$). This corresponding current density is obtained from Eq. 52 by setting $V = 0$,

$$J_{m\to s} = -A^*T^2\exp\left(-\frac{q\phi_{Bn}}{kT}\right). \tag{58}$$

The total current density is given by the sum of Eqs. 52 and 58.

$$\begin{aligned} J_n &= \left[A^*T^2\exp\left(-\frac{q\phi_{Bn}}{kT}\right)\right]\left[\exp\left(\frac{qV}{kT}\right) - 1\right] \\ &= J_{TE}\left[\exp\left(\frac{qV}{kT}\right) - 1\right] \end{aligned} \tag{59}$$

where

$$J_{TE} \equiv A^*T^2\exp\left(-\frac{q\phi_{Bn}}{kT}\right). \tag{60}$$

Equation 59 is similar to the transport equation for p-n junctions. However, the expressions for the saturation current densities are quite different.

An alternative approach to derive the thermionic-emission current is the following.[8] Without decomposing the velocity components, only electrons with energy above the barrier will contribute to the forward current. This number of electrons above the barrier is given by

$$n = N_C\exp\left[\frac{-q(\phi_{Bn} - V)}{kT}\right]. \tag{61}$$

It is known that for a Maxwellian distribution of velocities, the current from random motion of carriers across a plane is given by

$$J = nq\frac{v_{ave}}{4} \tag{62}$$

where v_{ave} is the average thermal velocity,

Table 2 Values of A^*/A (After Ref. 31)

Semiconductor	Si	GaAs
p-type	0.66	0.62
n-type ⟨100⟩	2.1	0.063 (low field) 0.55 (high field)
n-type ⟨111⟩	2.2	,, ,,

$$v_{ave} = \sqrt{\frac{8kT}{\pi m^*}} \quad . \tag{63}$$

Substitution of Eqs. 61 and 63 into Eq. 62 gives

$$J = \frac{4(kT)^2 q \pi m^*}{h^3} \exp\left[\frac{-q(\phi_{Bn} - V)}{kT}\right] \tag{64}$$

which is identical to Eq. 52.

3.3.2 Diffusion Theory

The diffusion theory by Schottky[4] is derived from the assumptions that (1) the barrier height is much larger than kT, (2) the effect of electron collisions within the depletion region, i.e., diffusion, is included, (3) the carrier concentrations at $x = 0$ and $x = W_D$ are unaffected by the current flow (i.e., they have their equilibrium values), and (4) the impurity concentration of the semiconductor is nondegenerate.

Since the current in the depletion region depends on the local field and the concentration gradient, we must use the current density equation:

$$\begin{aligned} J_x = J_n &= q\left(n\mu_n \mathscr{E} + D_n \frac{dn}{dx}\right) \\ &= qD_n\left(\frac{n}{kT}\frac{dE_C}{dx} + \frac{dn}{dx}\right). \end{aligned} \tag{65}$$

Under the steady-state condition, the current density is independent of x, and Eq. 65 can be integrated using $\exp[E_C(x)/kT]$ as an integrating factor. We then have

$$J_n \int_0^{W_D} \exp\left[\frac{E_C(x)}{kT}\right] dx = qD_n \left\{ n(x) \exp\left[\frac{E_C(x)}{kT}\right] \right\} \Bigg|_0^{W_D} \tag{66}$$

and the boundary conditions using $E_{Fm} = 0$ as the reference (see Fig. 16 but ignore image force for diffusion):

$$E_C(0) = q\phi_{Bn}, \tag{67}$$

$$E_C(W_D) = q(\phi_n + V), \tag{68}$$

$$n(0) = N_C \exp\left[-\frac{E_C(0) - E_{Fn}(0)}{kT}\right] = N_C \exp\left(-\frac{q\phi_{Bn}}{kT}\right), \tag{69}$$

$$n(W_D) = N_D = N_C \exp\left(-\frac{q\phi_n}{kT}\right). \tag{70}$$

Substituting Eqs. 67–70 into Eq. 66 yields

$$J_n = qN_C D_n \left[\exp\left(\frac{qV}{kT}\right) - 1\right] \Big/ \int_0^{W_D} \exp\left[\frac{E_C(x)}{kT}\right] dx. \tag{71}$$

For Schottky barriers, neglecting image-force effect, the potential distribution is given by Eq. 6. Substituting this expression for $E_C(x)$ into Eq. 71 and expressing W_D in terms of $\psi_{bi} + V$ leads to

$$J_n \approx \frac{q^2 D_n N_C}{kT}\sqrt{\frac{2qN_D(\psi_{bi}-V)}{\varepsilon_s}}\exp\left(-\frac{q\phi_{Bn}}{kT}\right)\left[\exp\left(\frac{qV}{kT}\right)-1\right]$$
$$\approx q\mu_n N_C \mathscr{E}_m \exp\left(-\frac{q\phi_{Bn}}{kT}\right)\left[\exp\left(\frac{qV}{kT}\right)-1\right] = J_D\left[\exp\left(\frac{qV}{kT}\right)-1\right]. \quad (72)$$

The current density expressions of the diffusion and thermionic-emission theories, Eqs. 59 and 72, are basically very similar. However, the saturation current density for the diffusion theory J_D is dependent on the bias and is less sensitive to temperature compared to the saturation current density of the thermionic-emission theory J_{TH}.

3.3.3 Thermionic-Emission-Diffusion Theory

A synthesis of the thermionic-emission and diffusion approaches described above has been proposed by Crowell and Sze.[32] This approach is derived from the boundary condition of a thermionic recombination velocity v_R near the metal-semiconductor interface.

Since the diffusion of carriers is strongly affected by the potential configuration in the region through which the diffusion occurs, we consider the electron potential energy [or $E_C(x)$] versus distance incorporating the Schottky lowering effect as shown in Fig. 17. We consider the case where the barrier height is large enough that the charge density between the metal surface and $x = W_D$ is essentially that of the ionized donors (i.e., depletion approximation). As drawn, the applied voltage V between the

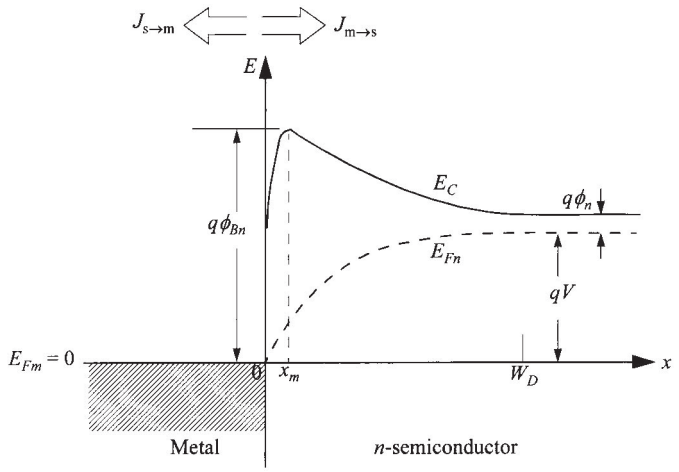


Fig. 17 Energy-band diagram incorporating the Schottky effect to show the derivations of thermionic-emission-diffusion theory and tunneling current.

metal and the semiconductor bulk would give rise to a flow of electrons toward the metal. The electron quasi-Fermi level E_{Fn} in the barrier is also shown schematically as a function of distance. Throughout the region between x_m and W_D,

$$J = n\mu_n \frac{dE_{Fn}}{dx} \tag{73}$$

where the electron density at any point x is given by

$$n = N_C \exp\left(-\frac{E_C - E_{Fn}}{kT}\right). \tag{74}$$

We will assume that the region between x_m and W_D is isothermal and that the electron temperature T is equal to the lattice temperature.

If the portion of the barrier between x_m and the interface ($x = 0$) acts as a sink for electrons, we can describe the current flow in terms of an effective recombination velocity v_R at the potential energy maximum x_m:

$$J = q(n_m - n_0)v_R \tag{75}$$

where n_m is the electron density at x_m when the current is flowing,

$$n_m = N_C \exp\left[\frac{E_{Fn}(x_m) - E_C(x_m)}{kT}\right] = N_C \exp\left[\frac{E_{Fn}(x_m) - q\phi_{Bn}}{kT}\right]. \tag{76}$$

n_0 is a quasi-equilibrium electron density at x_m, the density that would occur if it were possible to reach equilibrium without altering the magnitude or position of the potential energy maximum, i.e., $E_{Fn}(x_m) = E_{Fm}$

$$n_0 = N_C \exp\left(-\frac{q\phi_{Bn}}{kT}\right). \tag{77}$$

Another boundary condition, taking $E_{Fm} = 0$ as reference, is

$$E_{Fn}(W_D) = qV. \tag{78}$$

If n is eliminated from Eqs. 73 and 74 and the resulting expression for E_{Fn} is integrated between x_m and W_D,

$$\exp\left[\frac{E_{Fn}(x_m)}{kT}\right] - \exp\left(\frac{qV}{kT}\right) = \frac{-J}{\mu_n N_C kT}\int_{x_m}^{W_D} \exp\left(\frac{E_C}{kT}\right)dx \ . \tag{79}$$

Then from Eqs. 75 and 79, $E_{Fn}(x_m)$ can be solved as

$$\exp\left[\frac{E_{Fn}(x_m)}{kT}\right] = \frac{v_D \exp(qV/kT) + v_R}{v_D + v_R} \tag{80}$$

where

$$v_D \equiv D_n \exp\left(\frac{q\phi_{Bn}}{kT}\right) \Big/ \int_{x_m}^{W_D} \exp\left[\frac{E_C}{kT}\right]dx \tag{81}$$

is an effective diffusion velocity associated with the transport of electrons from the edge of the depletion layer W_D to the potential energy maximum x_m. Substituting Eq. 80 into Eq. 75 gives the end result of the thermionic-emission-diffusion theory

$$J_{TED} = \frac{qN_C v_R}{1 + (v_R/v_D)} \exp\left(-\frac{q\phi_{Bn}}{kT}\right)\left[\exp\left(\frac{qV}{kT}\right) - 1\right]. \tag{82}$$

In this equation, the relative values of v_R and v_D determines the relative contribution of thermionic emission versus diffusion. The parameter v_D can be evaluated as the Dawson's integral and can be approximated by $v_D \approx \mu_n \mathscr{E}_m$ in this case of depletion region.[8] If the electron distribution is Maxwellian for $x \geq x_m$, and if no electrons return from the metal other than those associated with the current density $qn_0 v_R$, the semiconductor acts as a thermionic emitter. Then v_R is the thermal velocity given by

$$\begin{aligned} v_R &= \int_0^\infty v_x \exp\left(\frac{-m^* v_x^2}{2kT}\right) dv_x \Big/ \int_{-\infty}^\infty \exp\left(\frac{-m^* v_x^2}{2kT}\right) dv_x \\ &= \sqrt{\frac{kT}{2m^*\pi}} = \frac{A^* T^2}{qN_C} \end{aligned} \tag{83}$$

where A^* is the effective Richardson constant, as shown in Table 2. At 300 K, v_R is 5.2×10^6 and 1.0×10^7 cm/s for $\langle 111 \rangle$ n-type Si and n-type GaAs respectively. It can be seen that if $v_D \gg v_R$, the pre-exponential term in Eq. 82 is dominated by v_R and the thermionic-emission theory applies ($J_{TED} = J_{TE}$). If, however, $v_D \ll v_R$, the diffusion process is the limiting factor ($J_{TED} = J_D$).

In summary, Eq. 82 gives a result that is a synthesis of Schottky's diffusion theory and Bethe's thermionic-emission theory, and it predicts currents in essential agreement with the thermionic-emission theory if $\mu\mathscr{E}(x_m) > v_R$. The latter criterion is more rigorous than Bethe's condition $\mathscr{E}(x_m) > kT/q\lambda$, where λ is the carrier mean free path.

In the preceding section a recombination velocity v_R associated with thermionic emission was introduced as a boundary condition to describe the collecting action of the metal in a Schottky barrier. In many cases an appreciable probability exists that an electron which crosses the potential energy maximum will be back-scattered by electron optical-phonon scattering.[33,34] As a first approximation the probability of electron emission over the potential maximum can be given by $f_p = \exp(-x_m/\lambda)$. In addition, the electron energy distribution can be further distorted from a Maxwellian distribution because of quantum-mechanical reflection of electrons by the Schottky barrier, and also because of tunneling of electrons through the barrier.[35,36] The ratio f_Q of the total current flow, considering the quantum-mechanical tunneling and reflection, to the current flow neglecting these effects depends strongly on the electric field and the electron energy measured from the potential maximum.

The complete expression of the J-V characteristics taking into account f_p and f_Q is thus

$$J = A^{**} T^2 \exp\left(-\frac{q\phi_{Bn}}{kT}\right)\left[\exp\left(\frac{qV}{kT}\right) - 1\right] \tag{84}$$

where

$$A^{**} = \frac{f_p f_Q A^*}{1 + (f_p f_Q \upsilon_R / \upsilon_D)}. \tag{85}$$

The impacts of these effects are reflected in the reduced effective Richardson constant from A^* to A^{**}, by as much as 50%. Figure 18 shows the calculated room-temperature values of A^{**} for metal-Si systems with an impurity concentration of 10^{16} cm^{-3}. We note that for electrons (n-type Si), A^{**} in the field range 10^4 to 2×10^5 V/cm remains essentially at a constant value of about 110 A/cm^2-K^2. For holes (p-type Si), A^{**} in this field range also remains essentially constant but at a considerably lower value ($\approx$ 30 A/cm^2-K^2). For n-type GaAs, A^{**} has been calculated to be 4.4 A/cm^2-K^2.

We conclude from the foregoing discussions that at room temperature in the electric field range of 10^4 to about 10^5 V/cm, the current transport mechanism in most Si and GaAs Schottky-barrier diodes is mainly due to thermionic emission of majority carriers. The spatial dependence of the electron Fermi level E_{Fn} near the metal-semiconductor interface has been studied by substituting Eqs. 6 and 74 into Eq. 73 and evaluating the difference, $E_{Fn}(W_D) - E_{Fn}(0)$. The E_{Fn} as shown in Fig. 16 is essentially flat throughout the depletion region.[38] The difference $E_{Fn}(W_D) - E_{Fn}(0)$ for a Au-Si diode with $N_D = 1.2\times10^{15}$ cm^{-3}, is only 8 meV for a forward bias of 0.2 V at 300 K. At higher doping levels the difference is even smaller. These results further confirm that for high-mobility semiconductors with moderate dopings, the thermionic-emission theory is applicable.

3.3.4 Tunneling Current

For more heavily doped semiconductors and/or for operation at low temperatures, the tunneling current may become more significant. In the extreme of an ohmic contact,

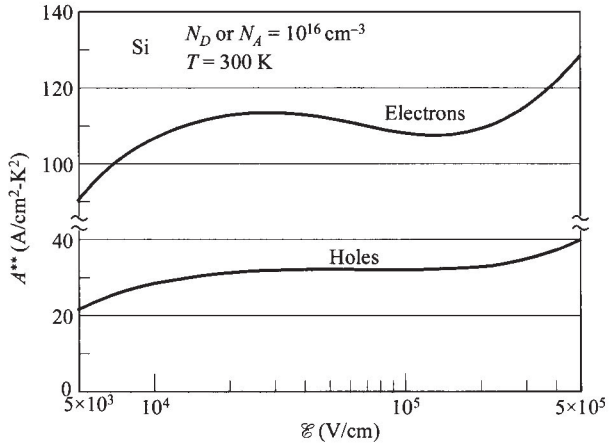


Fig. 18 Calculated effective Richardson constant A^{**} versus electric field for metal-silicon barriers. (After Ref. 37.)

which is a metal contact on degenerate semiconductor, the tunneling current is the dominant transport process. We will concentrate on ohmic contacts in the last section of this chapter.

The tunneling current from semiconductor to metal $J_{s\to m}$ is proportional to the quantum transmission coefficient (tunneling probability) multiplied by the occupation probability in the semiconductor and the unoccupied probability in the metal, that is,[36]

$$J_{s\to m} = \frac{A^{**}T^2}{kT}\int_{E_{Fm}}^{q\phi_{Bn}} F_s T(E)(1-F_m)dE. \tag{86}$$

F_s and F_m are the Fermi-Dirac distribution functions for the semiconductor and the metal respectively, and $T(E)$ is the tunneling probability which depends on the width of the barrier at a particular energy. A similar expression can be given for the current $J_{m\to s}$ which traverses in the opposite direction. In that case F_s and F_m would be interchanged in using the same equation. The net current density is the algebraic sum of the two components. Further analytical expression for the above equation is difficult, and the results can be obtained by numerical evaluation by computer.

Theoretical and experimental values of typical current-voltage characteristics for Au-Si barriers are shown in Fig. 19. We note that the total current density, which consists of both thermionic emission and tunneling, can be conveniently expressed as

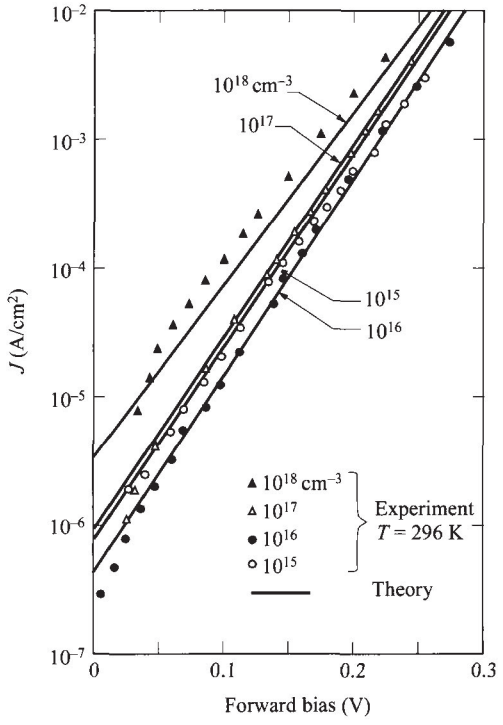


Fig. 19 Theoretical and experimental current-voltage characteristics for Au-Si Schottky barriers. Increased current is due to tunneling. (After Ref. 36.)

$$J = J_0\left[\exp\left(\frac{qV}{\eta kT}\right) - 1\right] \tag{87}$$

where J_0 is the saturation current density obtained by extrapolating the current density from the log-linear plot to $V = 0$, and η is the ideality factor, related to the slope. With little or no tunneling current or depletion-layer recombination, J_0 is determined by that of thermionic emission and η is close to unity. For higher doping and/or lower temperature, tunneling starts to occur and both J_0 and η increase.

The saturation current density J_0 and η are plotted in Fig. 20 for Au-Si diodes as a function of doping concentration, with temperature as a parameter. Note that J_0 is essentially a constant for low dopings but begins to increase rapidly when $N_D > 10^{17}$ cm^{-3}. The ideality factor η is very close to unity at low dopings and high temperatures. However, it can depart substantially from unity when the doping is increased or the temperature is lowered.

Figure 21 shows the ratio of the tunneling current to the thermionic current of a Au-Si barrier diode. Note that for $N_D \leq 10^{17}$ cm^{-3} and $T \geq 300$ K, the ratio is much less than unity and the tunneling component can be neglected. However, for higher dopings and lower temperatures, the ratio can become much larger than unity, indicating that the tunneling current becomes dominant.

Alternatively, the tunneling current can be expressed analytically and will give more physical insight. This formulation, based on the work of Padovani and Stratton,[39] is also used to derive the ohmic contact resistance. Referring to the energy-band diagrams in Fig. 22, we can roughly categorize the components into three types:

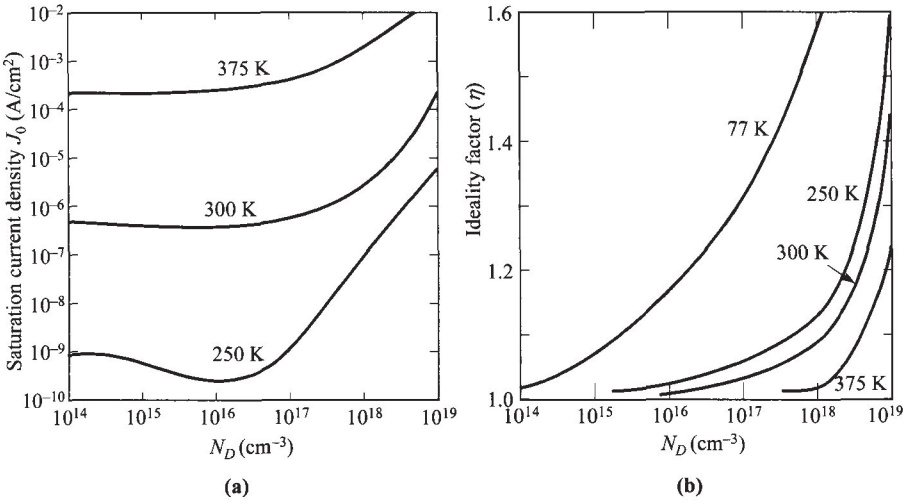


Fig. 20 (a) Saturation current density versus doping concentration for Au-Si Schottky barriers at three temperatures. (b) Ideality factor η versus doping concentration at different temperatures. (After Ref. 36.)

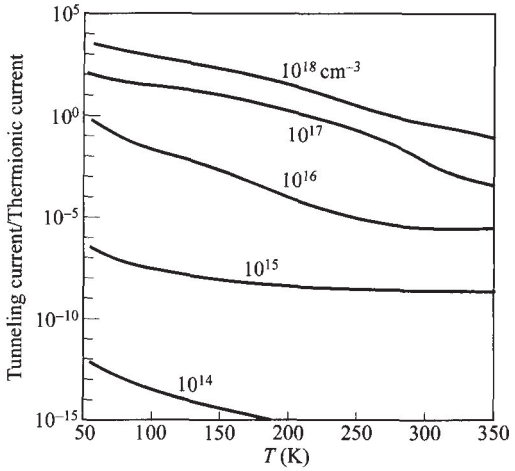


Fig. 21 Ratio of tunneling-current component to the thermionic-current component of a Au-Si barrier. The tunneling current will dominate at higher dopings and lower temperatures. (After Ref. 36.)

(1) thermionic emission (TE) over the barrier, (2) field emission (FE) near the Fermi level, and (3) thermionic-field emission (TFE) at an energy between TE and FE. While FE is a pure tunneling process, TFE is tunneling of thermally excited carriers which see a thinner barrier than FE. The relative contributions of these components depend on both temperature and doping level. A rough criterion can be set by comparing the thermal energy kT to E_{00} which is defined as

$$E_{00} \equiv \frac{q\hbar}{2}\sqrt{\frac{N}{m^* \varepsilon_s}}. \tag{88}$$

When $kT \gg E_{00}$, TE dominates and the original Schottky-barrier behavior prevails without tunneling. When $kT \ll E_{00}$, FE (or tunneling) dominates. When $kT \approx E_{00}$, TFE is the main mechanism which is a combination of TE and FE.

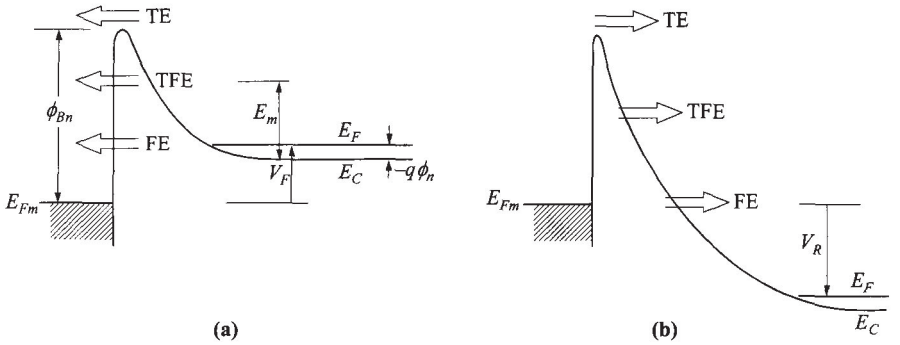


Fig. 22 Energy-band diagrams showing qualitatively tunneling currents in a Schottky diode (on *n*-type degenerate semiconductor) under (a) forward bias and (b) reverse bias. TE = thermionic emission. TFE = thermionic-field emission. FE = field emission.

Under forward bias, the current due to FE can be expressed as[39]

$$J_{FE} = \frac{A^{**}T\pi \exp[-q(\phi_{Bn}-V_F)/E_{00}]}{c_1 k \sin(\pi c_1 kT)}[1 - \exp(-c_1 q V_F)]$$
$$\approx \frac{A^{**}T\pi \exp[-q(\phi_{Bn}-V_F)/E_{00}]}{c_1 k \sin(\pi c_1 kT)} \tag{89}$$

where

$$c_1 \equiv \frac{1}{2E_{00}} \log\left[\frac{4(\phi_{Bn} - V_F)}{-\phi_n}\right]. \tag{90}$$

(ϕ_n is negative for degenerate semiconductors.) Notice the much weaker temperature dependence here (absent in the exponential term) compared to TE which is a characteristic of tunneling. The current due to TFE is given by

$$J_{TFE} = \frac{A^{**}T\sqrt{\pi E_{00} q(\phi_{Bn} - \phi_n - V_F)}}{k \cosh(E_{00}/kT)} \exp\left[\frac{-q\phi_n}{kT} - \frac{q(\phi_{Bn} - \phi_n)}{E_0}\right] \exp\left(\frac{qV_F}{E_0}\right), \tag{91}$$

$$E_0 \equiv E_{00} \coth\left(\frac{E_{00}}{kT}\right). \tag{92}$$

This TFE peaks roughly at an energy

$$E_m = \frac{q(\phi_{Bn} - \phi_n - V_F)}{\cosh^2(E_{00}/kT)} \tag{93}$$

where E_m is measured from E_C of the neutral region.

Under reverse bias, the tunneling current can be much larger because a large voltage is possible. The currents due to FE and TFE are given by

$$J_{FE} = A^{**}\left(\frac{E_{00}}{k}\right)^2 \left(\frac{\phi_{Bn} + V_R}{\phi_{Bn}}\right) \exp\left(-\frac{2q\phi_{Bn}^{3/2}}{3E_{00}\sqrt{\phi_{Bn} + V_R}}\right), \tag{94}$$

$$J_{TFE} = \frac{A^{**}T}{k}\sqrt{\pi E_{00} q\left[V_R + \frac{\phi_{Bn}}{\cosh^2(E_{00}/kT)}\right]} \exp\left(\frac{-q\phi_{Bn}}{E_0}\right) \exp\left(\frac{qV_R}{\varepsilon'}\right), \tag{95}$$

where

$$\varepsilon' = \frac{E_{00}}{(E_{00}/kT) - \tanh(E_{00}/kT)}. \tag{96}$$

These analytical expressions, although complicated, can be easily evaluated if all the parameters are known. These equations are also used to derive the ohmic contact resistance in the last section of this chapter.

3.3.5 Minority-Carrier Injection

The Schottky-barrier diode is mainly a majority-carrier device. The minority-carrier injection ratio γ, which is the ratio of minority-carrier current to total current, is small because the minority-carrier diffusion is much smaller than the majority-carrier ther-

mionic-emission current. However, at sufficiently large forward bias, the drift component of the minority carriers cannot be ignored anymore and the increased drift component will increase the overall injection efficiency. Both drift and diffusion of holes lead to the total current of

$$J_p = q\mu_p p_n \mathscr{E} - qD_p \frac{dp_n}{dx}. \tag{97}$$

The increased field is set up by the large majority-carrier thermionic-emission current,

$$J_n = q\mu_n N_D \mathscr{E}. \tag{98}$$

We consider the energy-band diagram shown in Fig. 23 where x_1 is the boundary of the depletion layer, and x_2 marks the interface between the n-type epitaxial layer and the n^+-substrate. From the junction theory discussed in Chapter 2, the minority-carrier density at x_1 is

$$p_n(x_1) = p_{no}\exp\left(\frac{qV}{kT}\right) = \frac{n_i^2}{N_D}\exp\left(\frac{qV}{kT}\right). \tag{99}$$

This quantity $p_n(x_1)$ can also be expressed as a function of the forward current density, obtained from Eqs. 84 and 99:

$$p_n(x_1) \approx \frac{n_i^2}{N_D}\frac{J_n}{J_{n0}}, \tag{100}$$

where J_{n0} (saturation current density) and J_n are representations of the thermionic-emission current (Eq. 84) in the following form:

$$J_n = J_{n0}\exp\left[\left(\frac{qV}{kT}\right) - 1\right]. \tag{101}$$

The other boundary condition for $p_n(x_2)$ is also necessary to calculate the diffusion current. We use the term transport velocity S_p (or surface recombination velocity) for the minority carriers to relate the current and concentration by

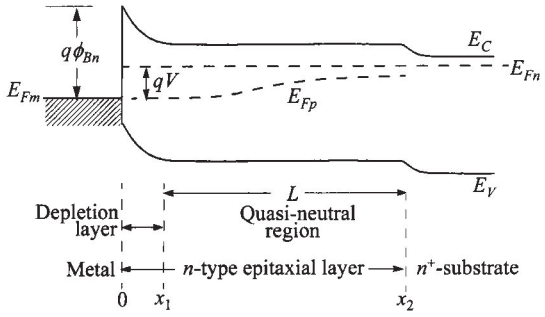


Fig. 23 Energy-band diagram of an epitaxial Schottky barrier under forward bias.

$$J_p(x_2) = qS_p[p_n(x_2) - p_{no}]. \tag{102}$$

We first consider the case with $S_p = \infty$ or equivalently $p_n(x_2) = p_{no}$. Under this boundary condition, the diffusion component has a standard form as in a p-n junction. From Eqs. 97, 98, and 100 we obtain the total hole current as (for $L \ll L_p$)

$$\begin{aligned} J_p &= q\mu_p p_n \mathscr{E} + \frac{qD_p n_i^2}{N_D L}\exp\left[\left(\frac{qV}{kT}\right) - 1\right] \\ &= \frac{\mu_p n_i^2 J_n^2}{\mu_n N_D^2 J_{n0}} + \frac{qD_p n_i^2}{N_D L}\exp\left[\left(\frac{qV}{kT}\right) - 1\right]. \end{aligned} \tag{103}$$

The injection ratio is given by

$$\gamma \equiv \frac{J_p}{J_p + J_n} \approx \frac{J_p}{J_n} \approx \frac{\mu_p n_i^2 J_n}{\mu_n N_D^2 J_{n0}} + \frac{qD_p n_i^2}{N_D L J_{n0}}. \tag{104}$$

For Au-Si diodes, the injection ratio has been measured to be very small, of the order of 10^{-5}, in agreement with the above equation.[40] Notice that γ has two terms. The second term is due to diffusion and is bias independent. This is the injection ratio for low-level bias,

$$\gamma_0 = \frac{qD_p n_i^2}{N_D L J_{n0}}. \tag{105}$$

The first term is due to the drift process, and is bias (or current) dependent. It can surpass the diffusion component at high current.

It is evident that to reduce the minority-carrier injection ratio (to reduce the charge storage time to be discussed below) one must use a metal-semiconductor system with large N_D (corresponding to low resistivity material), large J_{n0} (corresponding to small barrier height), and small n_i (corresponding to large bandgap). Furthermore, high-level bias is to be avoided. As an example, a gold-n-silicon diode with $N_D = 10^{15}$ cm^{-3} and $J_{n0} = 5\times10^{-7}$ A/cm^2 would give a low-bias injection γ_0 of $\approx 5\times10^{-4}$. But it would be expected to have an injection ratio of about 5% at a current density of 350 A/cm^2.

The above assumes that $p_n(x_2) = p_{no}$. Notice that at x_2, there is a barrier for holes that causes the holes to build up. These intermediate cases have been considered by Scharfetter using S_p as a parameter.[41] The computed results are shown in Fig. 24a, where the normalization factors are given by γ_0 and

$$J_{00} \equiv \frac{qD_n N_D}{L}. \tag{106}$$

J_{00} is the majority-carrier current at which the hole drift and diffusion components become equal, obtained by equating the two terms in Eq. 103.

Another quantity associated with the injection ratio is the minority-carrier storage time τ_s, which is defined as the minority carrier stored in the quasi-neutral region per unit current density:

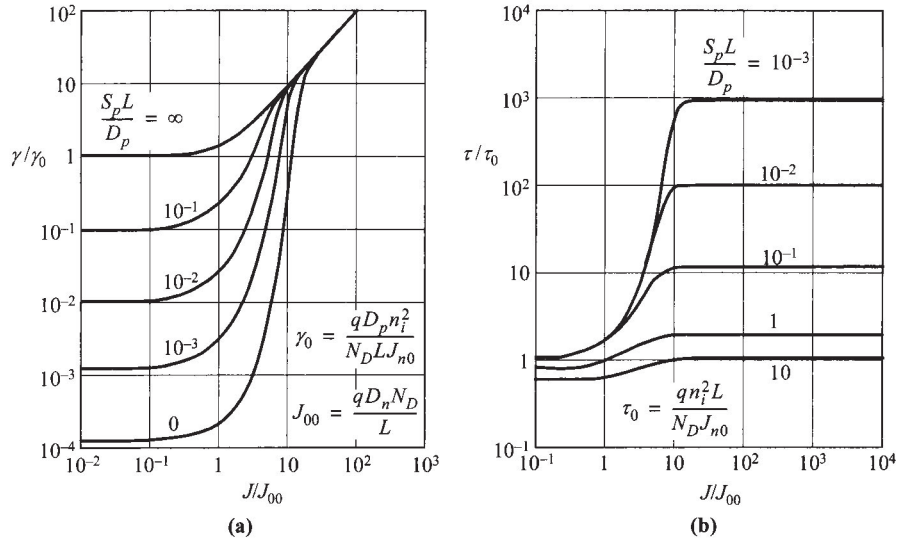


Fig. 24 (a) Normalized minority-carrier injection ratio versus normalized current density. (b) Normalized minority-carrier storage time versus normalized current density. $L/L_p = 10^{-2}$. (After Ref. 41.)

$$\tau_s \equiv \int_{x_1}^{x_2} qp(x)dx/J \tag{107}$$

For the low-current limit, depending on $p_n(x_2)$ or S_p, τ_s is given approximately by (for $L \ll L_p$)

$$\tau_s \approx \frac{qn_i^2L}{N_DJ_{n0}}, \tag{108}$$

and is independent of current. For high-current biases, $p_n(x_2)$ can become much higher, even to the extent that it is larger than in the rest of the quasi-neutral region L, i.e. a profile that increases with distance. The general results, using S_p again as a parameter, for τ_s versus the current density are shown in Fig. 24b. It can be seen that for finite S_p ($S_p \neq \infty$), τ_s can increase by orders of magnitude. Also, a high doping level is critical to reduce storage time in all cases.

3.3.6 MIS Tunnel Diode

In the *metal-insulator-semiconductor* (MIS) tunnel diode, a thin interfacial layer such as an oxide is intentionally (or sometimes unintentionally) introduced before metal deposition.[42,43] This interfacial-layer thickness lies in the range of 1–3 nm. This device differs from the MIS capacitor (to be considered in Chapter 4) in having appreciable current and under bias the semiconductor is not in equilibrium, i.e., the

quasi Fermi levels for electrons E_{Fn} and holes E_{Fp} split. The major differences of this structure compared to a conventional metal-semiconductor contact are: (1) reduced current because of the added interfacial layer, (2) lower barrier height (some potential is developed across the interfacial layer), and (3) higher ideality factor η. The energy-band diagram is similar to Fig. 5.

The current equation can be written as[42]

$$J = A^*T^2 \exp(-\sqrt{\zeta}\,\delta) \exp\left(\frac{-q\phi_B}{kT}\right)\left[\exp\left(\frac{qV}{\eta kT}\right) - 1\right]. \tag{109}$$

The derivation for this equation can be found in Section 8.3.2. For the same barrier, the current is suppressed by the tunneling probability $\exp(-\sqrt{\zeta}\,\delta)$. Here ζ (in eV) and δ (in Å) are the effective barrier and thickness of the interfacial layer. (A constant of $[2(2m^*/\hbar^2)]^{1/2}$ which has the value of 1.01 $eV^{-1/2}Å^{-1}$ is omitted.) This added tunneling probability can be considered as a modification to the effective Richardson constant, as discussed before. The ideality factor is increased to[42]

$$\eta = 1 + \left(\frac{\delta}{\varepsilon_i}\right)\frac{(\varepsilon_s/W_D) + qD_{its}}{1 + (\delta/\varepsilon_i)qD_{itm}} \tag{110}$$

where D_{its} and D_{itm} are interface traps in equilibrium with the semiconductor and metal, respectively. In general, when the oxide thickness is less than 3 nm, the interface traps are in equilibrium with the metal, whereas for thicker oxides, these traps tend to be in equilibrium with the semiconductor.

The interfacial layer reduces the majority-carrier thermionic-emission current without affecting the minority-carrier current, which is from diffusion, and raises the minority injection efficiency. This phenomenon is exploited in improving the injection efficiency of an electroluminescent diode and the open-circuit voltage of the Schottky-barrier solar cell.

3.4 MEASUREMENT OF BARRIER HEIGHT

Basically, four methods are used to measure the barrier height of a metal-semiconductor contact: the (1) current-voltage, (2) activation-energy, (3) capacitance-voltage, and (4) photoelectric methods.

3.4.1 Current-Voltage Measurement

For moderately doped semiconductors, the *I-V* characteristics in the forward direction with $V > 3kT/q$ is given by Eq. 84:

$$J = A^{**}T^2 \exp\left(-\frac{q\phi_{B0}}{kT}\right)\exp\left[\frac{q(\Delta\phi + V)}{kT}\right]. \tag{111}$$

Since both A^{**} and $\Delta\phi$ (image-force lowering) are weak functions of the applied voltage, the forward *J-V* characteristic (for $V > 3kT/q$) is represented by $J = J_0\exp(qV/\eta kT)$, as given previously in Eq. 87, where η is the ideality factor:

$$\eta \equiv \frac{q}{kTd(\ln J)}\frac{dV}{} = \left[1 + \frac{d\Delta\phi}{dV} + \frac{kTd(\ln A^{**})}{q}\frac{}{dV}\right]^{-1}. \tag{112}$$

Typical examples are shown in Fig. 25, where $\eta = 1.02$ for the W-Si diode and $\eta = 1.04$ for the W-GaAs diode. The extrapolated value of current density at zero voltage is the saturation current J_0, and the barrier height can be obtained from the equation

$$\phi_{Bn} = \frac{kT}{q}\ln\left(\frac{A^{**}T^2}{J_0}\right). \tag{113}$$

The value of ϕ_{Bn} is not very sensitive to the choice of A^{**}, since at room temperature, a 100% increase in A^{**} will cause an increase of only 0.018 V in ϕ_{Bn}. The theoretical relationship between J_0 and ϕ_B (ϕ_{Bn} or ϕ_{Bp}) at room temperature is plotted in Fig. 26 for $A^{**} = 120$ A/cm^2-K^2. For other values of A^{**}, parallel lines can be drawn on this plot to obtain the proper relationship.

In the reverse direction, the dominant voltage dependence is due mainly to the Schottky-barrier lowering, or

$$J_R \approx J_0 \qquad (\text{for } V_R > 3kT/q)$$
$$\approx A^{**}T^2\exp\left[-\frac{q(\phi_{B0} - \sqrt{q\mathscr{E}_m/4\pi\varepsilon_s})}{kT}\right] \tag{114}$$

where

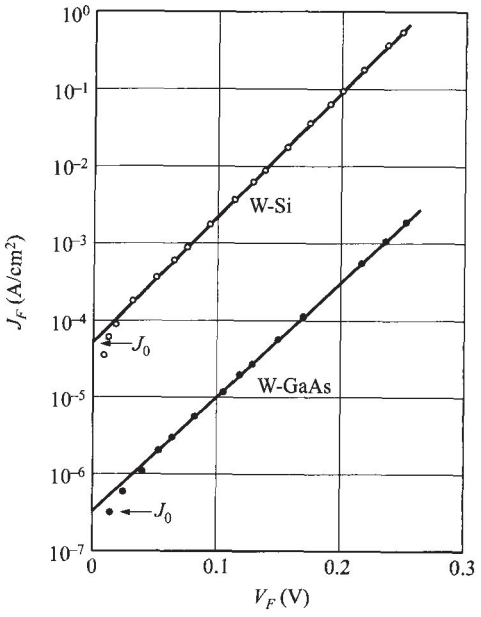


Fig. 25 Forward current density versus applied voltage of W-Si and W-GaAs diodes. (After Ref. 44.)

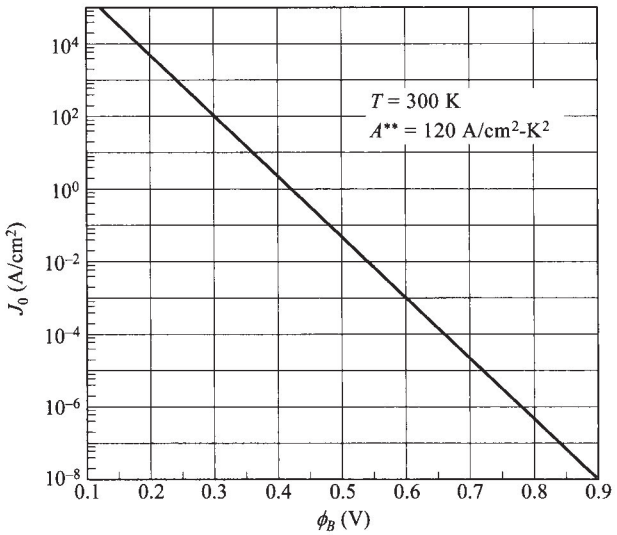


Fig. 26 Theoretical saturation current density at 300 K versus barrier height for an effective Richardson constant of 120 A/cm^2-K^2.

$$\mathscr{E}_m = \sqrt{\frac{2qN_D}{\varepsilon_s}\left(V_R + \psi_{bi} - \frac{kT}{q}\right)}. \tag{115}$$

If the barrier height $q\phi_{Bn}$ is sufficiently smaller than the bandgap so that the depletion-layer generation-recombination current is small in comparison with the Schottky emission current, then the reverse current will increase gradually with the reverse bias as given by Eq. 114, due mainly to image-force lowering.

For most of the practical Schottky diodes, however, the dominant reverse current component is the edge leakage current, which is caused by the sharp edge around the periphery of the metal plate. This sharp-edge effect is similar to the junction-curvature effect (with $r_j \to 0$) as discussed in Chapter 2. To eliminate this effect, metal-semiconductor diodes have been fabricated with a diffused guard ring (these structures will be discussed later). The guard ring is a deep p-type diffusion, and the doping profile is tailored to give the p-n junction a higher breakdown voltage than that of the metal-semiconductor contact. Because of the elimination of the sharp-edge effect, near-ideal forward and reverse I-V characteristics have been obtained. Figure 27 shows a comparison between experimental measurement from a PtSi-Si diode with guard ring, and theoretical calculation based on Eq. 114. The agreement is excellent. The sharp increase of current near 30 V is due to avalanche breakdown and is expected for the diode with a donor concentration of 2.5×10^{16} cm^{-3}.

The efficacy of guard ring structures in preventing premature breakdown and surface leakage can be ascertained by studying the reverse leakage current as a function of diode diameter at constant reverse bias. For this purpose, arrays of Schottky

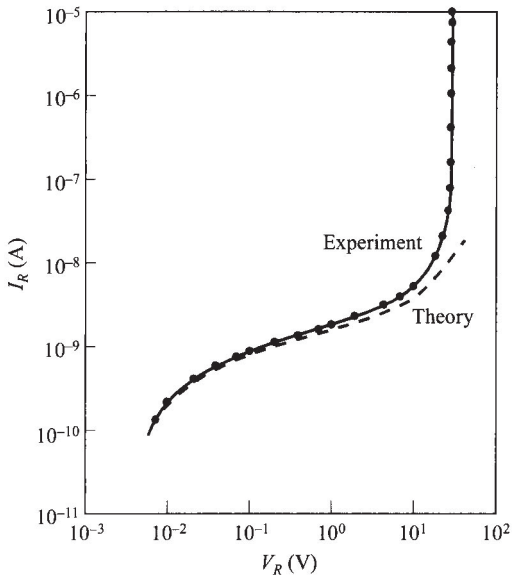


Fig. 27 Comparison of measurement with the theoretical prediction of reverse-bias current from Eq. 114 for a PtSi-Si diode. (After Ref. 45.)

diodes with different diameters can be formed on the semiconductor. The reverse leakage current can be measured and plotted as a function of diode diameter.[46] If the experimental data have a slope equal to two, the leakage currents are proportional to the device area. If, on the other hand, the leakage currents are dominated by edge effects, the data would be expected to lie along a straight line with a slope equal to unity.

For some Schottky diodes, the reverse current has an additional component. This component arises from the fact that if the metal-semiconductor interface is free from intervening layers of oxide and other contaminants, the electrons in the metal have wave functions that penetrate into the semiconductor energy gap. This is a quantum-mechanical effect that results in a static dipole layer at the metal-semiconductor interface. The dipole layer causes the intrinsic barrier height to vary slightly with the field, so $d\phi_{B0}/d\mathscr{E}_m \neq 0$. To a first approximation the static lowering can be expressed as

$$\Delta\phi_{\text{static}} \approx \alpha\mathscr{E}_m \tag{116}$$

or $\alpha \equiv d\phi_{B0}/d\mathscr{E}_m$. Figure 28 shows good agreement between the theory and measurements of the reverse current in a RhSi-Si diode, based on an empirical value of $\alpha = 1.7$ nm.

3.4.2 Activation-Energy Measurement

The principal advantage of Schottky-barrier determination by means of an activation energy measurement is that no assumption of electrically active area is required. This feature is particularly important in the investigation of novel or unusual metal-semiconductor interfaces because often the true value of the contacting area is not known. In the case of poorly cleaned or incompletely reacted surfaces, the electrically active

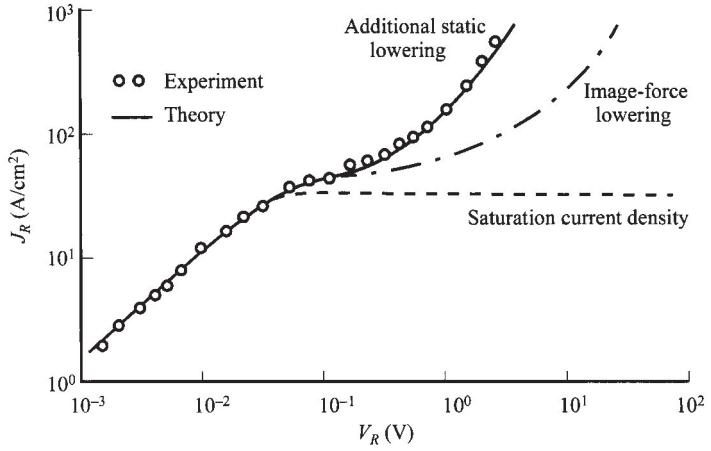


Fig. 28 Theory and experimental results of reverse characteristics for a RhSi-Si diode. (After Ref. 37.)

area may be only a small fraction of the geometric area. On the other hand, a strong metallurgical reaction could result in rough nonplanar metal-semiconductor interface with an electrically active area that is larger than the apparent geometric area.

If Eq. 84 is multiplied by A, the electrically active area, we obtain

$$\ln\left(\frac{I_F}{T^2}\right) = \ln(AA^{**}) - \frac{q(\phi_{Bn} - V_F)}{kT} \tag{117}$$

where $q(\phi_{Bn} - V_F)$ is considered the activation energy. Over a limited range of temperature around room temperature, the value of A^{**} and ϕ_{Bn} are essentially temperature independent. Thus for a fixed forward bias V_F, the slope of a plot of $\ln(I_F/T^2)$ versus $1/T$ yields the barrier height ϕ_{Bn}, and the ordinate intercept at $1/T = 0$ yields the product of the electrically active area A and the effective Richardson constant A^{**}.

To illustrate the importance of the activation-energy method in the investigation of interfacial metallurgical reactions, Fig. 29 shows the activation-energy plots of the saturation current in Al-n-Si contacts of different barrier heights, formed simply by annealing at various temperatures.[47] The slopes of these plots indicate a nearly linear increase of effective Schottky barrier height from 0.71 to 0.81 V for annealing temperatures between 450°C and 650°C. These observations were also confirmed with I-V and C-V measurements. Also supposedly when the Al-Si eutectic temperature ($\approx$ 580°C) is reached, the true metallurgical nature of the metal-semiconductor interface must be considerably modified. Determination of the ordinate intercepts from the plots shown in Fig. 29 indicates that the electrically active area increases by a factor of two, when the annealing temperature exceeds the Al-Si eutectic temperature.

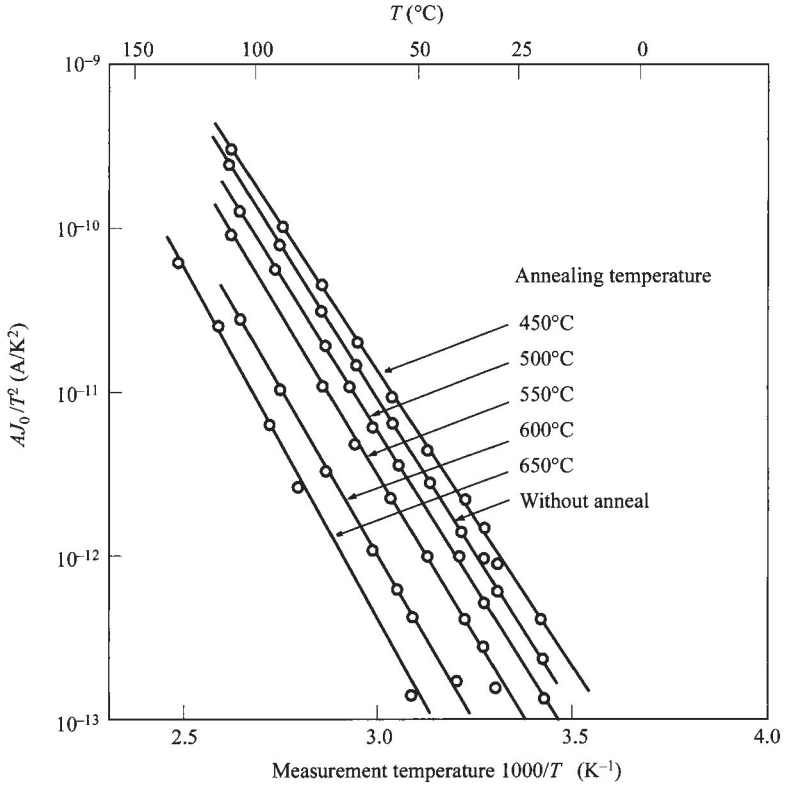


Fig. 29 Activation energy plots for determination of barrier height. (After Ref. 47.)

3.4.3 Capacitance-Voltage Measurement

The barrier height can also be determined by the capacitance measurement. When a small ac voltage is superimposed upon a dc bias, incremental charges of one sign are induced on the metal surface and charges of the opposite sign in the semiconductor. The relationship between C (depletion-layer capacitance per unit area) and V is given by Eq. 10. Figure 30 shows some typical results where $1/C^2$ is plotted against the applied voltage. The intercept on the voltage axis gives the built-in potential ψ_{bi} from which the barrier height can be determined:[44,48]

$$\phi_{Bn} = \psi_{bi} + \phi_n + \frac{kT}{q} - \Delta\phi. \tag{118}$$

From the slope the carrier density can also be determined (Eq. 11) and it can be used to calculate ϕ_n.

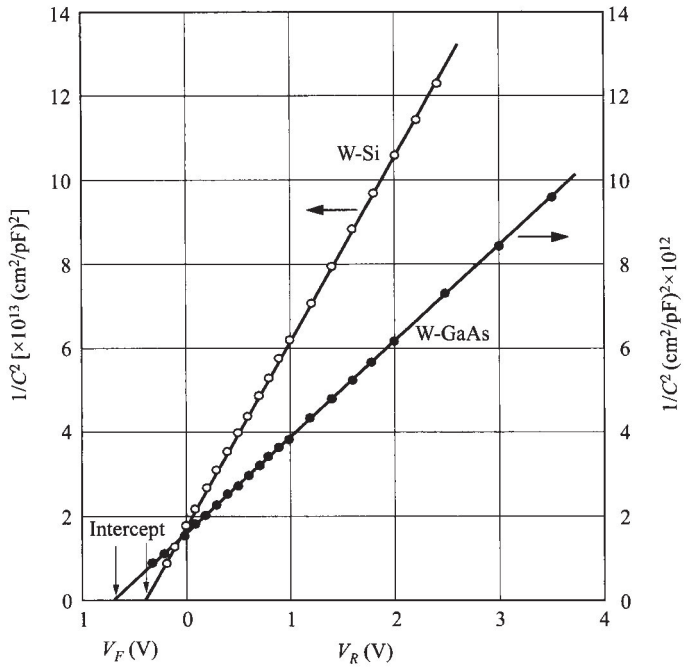


Fig. 30 $1/C^2$ versus applied voltage for W-Si and W-GaAs diodes. (After Ref. 44.)

To obtain the barrier height of semiconductor which contain both shallow-level and deep-level impurities (Fig. 4), we need to measure the C-V curves at two different temperatures at multiple frequencies.[49]

3.4.4 Photoelectric Measurement

The photoelectric measurement is an accurate and direct method of determining the barrier height.[50] When a monochromatic light is incident upon a metal surface, photocurrent may be generated. The basic setup is shown in Fig. 31. In a Schottky-barrier diode, two kinds of carrier excitation can occur that contribute to photocurrent; excitation over the barrier (process-1) and band-to-band excitation (process-2). In measuring the barrier height, only process-1 is useful and the most useful wavelengths should be in the range of $q\phi_{Bn} < h\nu < E_g$. Furthermore, the most critical light absorption region is at the metal-semiconductor interface. For front illumination, the metal film should be thin so light can penetrate to that interface. There is no such restriction in using back illumination since light is transparent in the semiconductor if $h\nu < E_g$, and the highest light intensity would be at the metal-semiconductor interface. Note that photocurrent can be collected without bias.

The photocurrent per absorbed photon (photoresponse R) as a function of the photon energy $h\nu$, is given by the Fowler theory:[51]

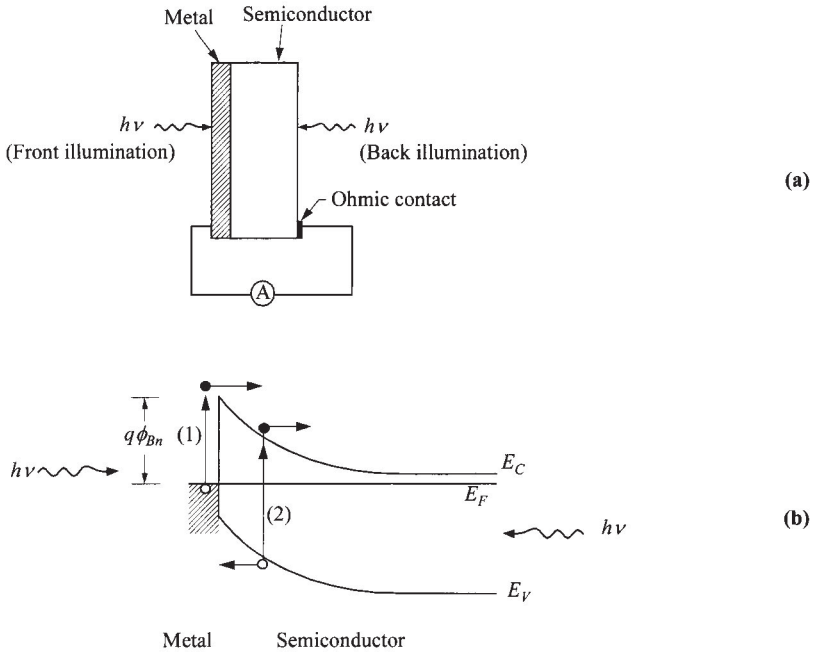


Fig. 31 (a) Schematic setup for photoelectric measurement. (b) Energy-band diagram for photoexcitation processes.

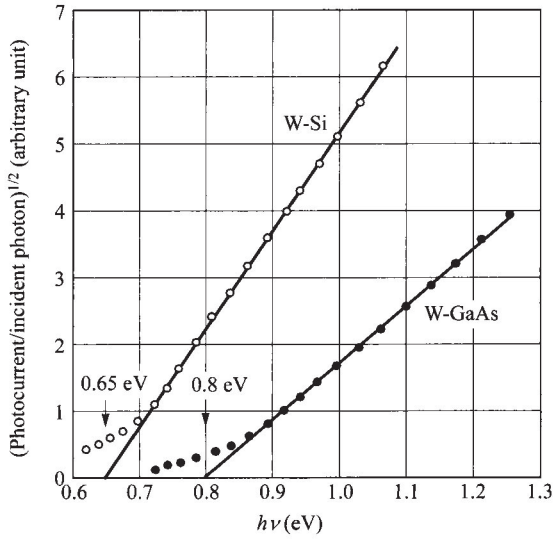


Fig. 32 Square root of the photoresponse versus photon energy for W-Si and W-GaAs diodes. The extrapolated values are the corresponding barrier height $q\phi_{Bn}$. (After Ref. 44.)

$$R \propto \frac{T^2}{\sqrt{E_s - h\nu}}\left\{\frac{x^2}{2} + \frac{\pi^2}{6} - \left[\exp(-x) - \frac{\exp(-2x)}{4} + \frac{\exp(-3x)}{9} - \ldots\right]\right\} \quad \text{for } x \geq 0 \tag{119}$$

where E_s is the sum of $h\nu_0$ (= barrier height $q\phi_{Bn}$) and the Fermi energy measured from the bottom of the metal conduction band, and $x \equiv h(\nu - \nu_0)/kT$. Under the condition of $E_s \gg h\nu$ and $x > 3$, Eq. 119 reduces to

$$R \propto (h\nu - h\nu_0)^2. \tag{120}$$

When the square root of the photoresponse is plotted as a function of photon energy, a straight line should be obtained, and the extrapolated value on the energy axis should give directly the barrier height. Figure 32 shows the photoresponse of W-Si and W-GaAs diodes, and the barrier heights of 0.65 and 0.80 eV are obtained respectively.

The photoelectric measurement can be used to study other device and material parameters. It has been used to determine the image-force dielectric constant of Au-Si diodes.[25] By measuring the shift of the photothreshold under different reverse biases, one can determine the image-force lowering $\Delta\phi$. From a plot of $\Delta\phi$ versus $\sqrt{\mathscr{E}_m}$, the dielectric constant ($\varepsilon_s/\varepsilon_0$) can be determined, as shown previously in Fig. 12. Photoelectric measurement has been used to study the temperature dependence of the barrier height.[52] The photothreshold is measured as a function of the temperature of Au-Si diodes. The shift of photothreshold correlates reasonably well with the temperature dependence of the silicon bandgap. This result implies that the Fermi level at the Au-Si interface is pinned in relation to the valence-band edge and this is in agreement with our discussion in Section 3.2.3.

3.4.5 Measured Barrier Heights

The *I-V*, *C-V*, activation-energy and photoelectric methods have been used to measure the barrier heights. For intimate contacts with a clean interface, these methods generally yield consistent barrier heights within ± 0.02 V. A large discrepancy between different methods may result from such causes as contamination in the interface, an intervening insulating layer, edge leakage current, or deep impurity levels.

The measured Schottky barrier heights for some elemental and compound semiconductors are listed in Table 3. The barrier heights are representative values for metal-semiconductor contacts made by deposition of high-purity metals in a good vacuum system onto cleaved or chemically cleaned semiconductor surfaces. As expected, silicon and GaAs metal-semiconductor contacts are most extensively studied. Among the metals, gold, aluminum, and platinum are most commonly used. The barrier heights of metal silicides on *n*-type silicon and some of their properties are listed in Table 4.

It should be pointed out that the barrier height is generally sensitive to pre-deposition surface preparation and post-deposition heat treatments.[63] Figure 33 shows the barrier heights on *n*-type Si and GaAs measured at room temperature after annealing at various temperatures. When an Al-Si diode is annealed above 450°C, the barrier height begins to increase,[47] presumably due to diffusion of Si in Al (also see Fig. 29).

Table 3 Measured Schottky-Barrier Heights ϕ_{Bn} (V) at 300 K on *n*-type Semiconductors. Each Entry Represents the Highest Value Reported for that System. Barrier Heights on *p*-type Can Be Estimated by $\phi_{Bp} + \phi_{Bn} \approx E_g/q$ (After Refs. 8, 53–59)

	Si	GaAs	Ge	AlAs	SiC	GaP	GaSb	InP	ZnS	ZnSe	ZnO	CdS	CdSe	CdTe	PbO
E_g	1.12	1.42	0.66	2.16	3.0	2.24	0.67	1.29	3.6	2.82	3.2	2.43	1.7	1.6	
Ag	0.83	1.03	0.54			1.2	0.45	0.54	1.81	1.21		0.56	0.43	0.8	0.95
Al	0.81	0.93	0.48		1.3	1.06	0.6	0.5	0.8	0.75	0.68			0.76	
Au	0.83	1.05	0.59	1.2	1.4	1.3	0.61	0.52	2.2	1.51	0.65	0.78	0.7	0.86	
Bi		0.9					0.2			1.14				0.78	
Ca	0.4	0.56													
Co	0.81	0.86	0.5		1.4										
Cr	0.60	0.82			1.2	1.18		0.45							
Cu	0.8	1.08	0.5		1.3	1.2	0.47	0.42	1.75	1.1	0.45	0.5	0.33	0.82	
Fe	0.98	0.84	0.42							1.11				0.78	
Hf	0.58	0.82				1.84									
In		0.83	0.64				0.6		1.5	0.91	0.3			0.69	0.93
Ir	0.77	0.91	0.42												
Mg	0.6	0.66				1.04	0.3		0.82	0.49					
Mo	0.69	1.04			1.3	1.13									
Ni	0.74	0.91	0.49		1.4	1.27		0.32				0.45		0.83	0.96
Os	0.7		0.4									0.53			
Pb	0.79	0.91	0.38							1.15		0.59		0.68	0.95
Pd	0.8	0.93			1.2		0.6	0.41	1.87		0.68	0.62		0.86	
Pt	0.9	0.98		1.0	1.7	1.45			1.84	1.4	0.75	1.1	0.37	0.89	
Rh	0.72	0.90	0.4												
Ru	0.76	0.87	0.38												
Sb		0.86					0.42			1.34				0.76	
Sn		0.82						0.35							
Ta		0.85							1.1		0.3				
Ti	0.6	0.84			1.1	1.12						0.84			
W	0.66	0.8	0.48												

Table 4 Barrier Height of Metal Silicide on *n*-type Si. For Each System, the Barrier-Height Entry Represents the Highest Reported Value (After Refs. 8, 22, 56, 60–62)

Metal Silicide	ϕ_{Bn} (V)	Structure	Forming Temperature (°C)	Melting Temperature (°C)
CoSi	0.68	Cubic	400	1460
$CoSi_2$	0.64	Cubic	450	1326
$CrSi_2$	0.57	Hexagonal	450	1475
$DySi_2$	0.37			
$ErSi_2$	0.39			
$GdSi_2$	0.37			
HfSi	0.53	Orthorhombic	550	2200
$HoSi_2$	0.37			
IrSi	0.93		300	
Ir_2Si_3	0.85			
$IrSi_3$	0.94			
MnSi	0.76	Cubic	400	1275
$Mn_{11}Si_{19}$	0.72	Tetragonal	800[a]	1145
$MoSi_2$	0.69	Tetragonal	1000[a]	1980
Ni_2Si	0.75	Orthorhombic	200	1318
NiSi	0.75	Orthorhombic	400	992
$NiSi_2$	0.66	Cubic	800[a]	993
Pd_2Si	0.75	Hexagonal	200	1330
PtSi	0.87	Orthorhombic	300	1229
Pt_2Si	0.78			
RhSi	0.74	Cubic	300	
$TaSi_2$	0.59	Hexagonal	750[a]	2200
$TiSi_2$	0.60	Orthorhombic	650	1540
VSi_2	0.65			
WSi_2	0.86	Tetragonal	650	2150
YSi_2	0.39			
$ZrSi_2$	0.55	Orthorhombic	600	1520

[a] Can be ≤ 700°C under clean interface condition.

Also, for metals that form silicides on silicon, the barrier height changes abruptly when the eutectic temperature is reached. The barrier height of a Pt-Si diode is 0.9 V. After annealing at 300°C or higher temperatures, PtSi is formed at the interface and ϕ_{Bn} decreases to 0.85 V.[64] For Pt-GaAs contact the barrier height increases from 0.84 V to 0.87 V when $PtAs_2$ is formed at the interface.[65] For a W-Si diode the barrier height remains constant until the annealing temperature is above 1000°C, when WSi_2 is formed.[66]

So far in all the Schottky diodes discussed above, the metal layers are deposited so they are polycrystalline or amorphous in structure. For certain silicide contacts on silicon, it has been demonstrated that single-crystalline form can be grown epitaxially

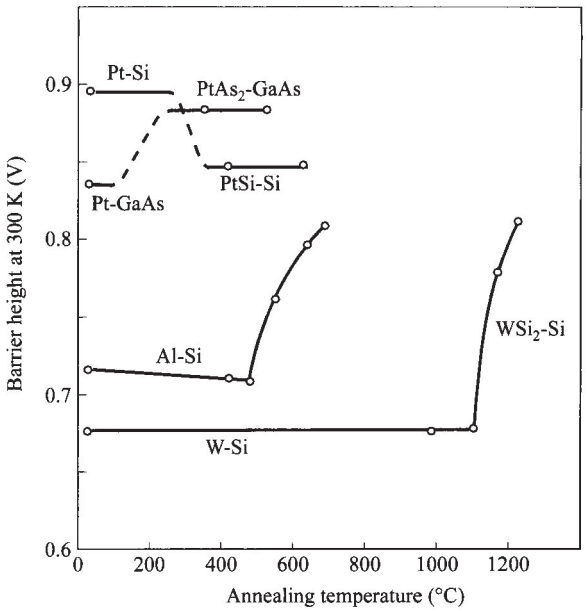


Fig. 33 Barrier heights on *n*-type Si and GaAs measured at room temperature after annealing at various temperatures.

from the underlying single-crystalline silicon.[67] These epitaxial silicides include $NiSi_2$, $CoSi_2$, $CrSi_2$, Pd_2Si, $ErSi_{2-x}$, $TbSi_{2-x}$, YSi_{2-x}, and $FeSi_2$. The epitaxial silicides have the properties of high uniformity and thermal stability. They provide a unique opportunity to study the fundamental relationship of barrier height to the microscopic interfacial configuration. It has been demonstrated that even on the same orientation of Si surface, different types (A and B) and interface structures (6-, 7-, or 8-folded) can be formed and that they give a difference in barrier height of as much as 0.14 eV. With this insight, the range of barrier heights observed on the same metal-semiconductor system can be rationalized, due to the statistical spacial distribution of the interfacial structure.

3.5 DEVICE STRUCTURES

The earliest device structure is the point-contact rectifier using a small metal wire with a sharp point making contact with a semiconductor. The contact may be just a simple mechanical contact or it may be formed by an electrical discharge processes that may result in a small alloyed *p-n* junction.

A point-contact rectifier usually has poor forward and reverse *I-V* characteristics compared to a planar Schottky diode. Its characteristics are also difficult to predict