

PHYSICS OF INTERFACES

C. A. Mead

California Institute of Technology

Pasadena, California 91109

It has long been known that when a metal is placed in contact with a semiconductor a rectifying contact often results. This rectification is a result of an energy barrier between the metal and the semiconductor. In order to form a nonrectifying or ohmic contact, two general approaches can be applied: either (1) the barrier energy can be reduced to a low enough value that the thermally excited current over the barrier is large enough for the application involved or (2) the semiconductor can be doped to a high carrier density to allow quantum mechanical tunneling to take place. The physical principles of these processes are discussed in this article.

Although many interfaces are of interest in electrochemistry, the particular interface of interest in the study of the ohmic contact is that between a metal and a semiconductor. In 1874 K. F. Braun⁽¹⁾ discovered that such an interface carried current more easily in one direction than another. Braun properly identified the current as electronic in nature and also properly identified the origin of the rectification as the interface itself. Figure 1 shows a current-voltage characteristic taken by Braun⁽²⁾ in 1877 for a contact between a metal wire and a lead sulfide crystal. The direction of the rectification was such that current flowed easily when the metal was positive with respect to the semiconductor; very little current flowed if the metal was negative. Since it was understood at the time that the current was carried by negative charges in the metal and that there were many more charge carriers in the metal than in the semiconductor, this was a very puzzling result. One would think at first glance that if there were any rectification at all the current would flow more easily from the material with larger carrier concentration into the material with lower concentration. This situation is illustrated in Figure 2 where electrons in the metal have been pulled away from the interface in the forward direction (the direction of easy current flow), and are pushed up against the interface in the reverse direction (the direction where very little current flows). Many theories were proposed giving a direction of rectification opposite to that observed experimentally. The question, of course, is not why current flows in the forward direction; there are charged carriers in both the metal and the semiconductor and it is clear that, in the absence of some other phenomenon, current should

flow. The real puzzling result is that very little current flows in the reverse direction. Many electrons in the metal are crowded against the interface and somehow cannot get across into the semiconductor.

The metal-semiconductor interface is not the only interface in which such a phenomenon is known. A vacuum is commonly known to be a good insulator. This means that when two metal plates are placed in a vacuum and a voltage is applied between them, very little current flows. The reason, of course, is not that there are no electrons in the metal, nor is it that if the electrons were free they could not cross the vacuum from one metal plate to the other. The reason is that the electrons experience great difficulty in leaving the metal and going into the vacuum. There is an energy barrier between the highest energy electron in the metal and the lowest energy in the vacuum where an electron can exist. This energy difference is illustrated in Figure 3 and is called the work function of the metal. A similar energy barrier is developed when a metal and a semiconductor are brought into contact. Electrons in the metal are distributed over a range of energies, the highest energy at which appreciable electron density exists being the Fermi level. By contrast, electrons in the semiconductor are all crowded into a very narrow energy range near the edge of the conduction band. When the metal and the semiconductor are brought into contact, in most cases the conduction band of the semiconductor is at a higher energy than the Fermi level of the metal. Electrons then flow momentarily from the conduction band of the semiconductor into the Fermi sea in the metal until (at the distances far from the interface) the two energies become nearly equal, as shown in Figure 4. The energy difference at the interface between the Fermi level of the metal and the conduction band edge is known as the barrier energy of the particular metal semiconductor interface. At present, how to calculate the barrier energy of any given metal semiconductor interface from first principles is not known. However, considerable experimental information is known about a wide variety of such interfaces. Basically, barrier energies can be grouped into two broad classes: (1) On the covalent semiconductors, such as silicon, germanium, and gallium arsenide, the barrier energies do not depend very strongly upon which metal is placed on the semiconductor surface and are thus largely a property of the semiconductor itself. (2) By contrast, on the more ionic semiconductors such as cadmium sulfide, zinc sulfide, and zinc oxide, barrier energies are a function of both the work function (or electronegativity) of the metal and of the particular semiconductor.⁽³⁾ The situation is illustrated in Figure 5 where the barrier energies for different metals on a typical covalent material, gallium arsenide, and a typical ionic material, zinc sulfide, are shown as a function of the electronegativity of the metal. Over the range of metals involved here, the electronegativity given in electron volts is equivalent to the work function with a shift in zero, and is used because a more well-defined scale is available. The straight line through the zinc sulfide points has a slope of unity, whereas in the gallium arsenide case the slope is approximately 0.1. As can be seen, the barriers on the ionic materials can be characterized by giving the intercept of a plot such as Figure 5 whereas (to a first approximation) the barriers on the covalent materials are nearly independent of the

metal and can be specified by knowing the barrier energy of any one particular metal. Figure 6 shows the barrier energies for gold on the various covalent semiconductors as a function of their bandgap.⁽³⁾ It is interesting to note that, except for one or two materials, the barriers can be given quite well by the expression

$$\varphi_B \approx \frac{2E_g}{3} \quad (1)$$

(the straight line in Figure 6).

We will now discuss the nature of the current flow in a typical metal semiconductor interface. It was mentioned earlier that electrons in the semiconductor are confined to a very narrow energy range near the conduction band edge while the highest energy electrons in the metal are at the Fermi level. This is not strictly true, of course, because there are thermally excited electrons at higher energies, although their number decreases rapidly with energy according to the well-known Boltzman distribution. The situation is illustrated schematically in Figure 7a for the case where no bias is applied between the metal and the semiconductor. When a negative voltage is applied to the semiconductor as shown in Figure 7b, the electron distribution in the semiconductor is raised in relation to that in the metal and there are now more electrons in the semiconductor with energies higher than the peak of the barrier which flow by diffusion into the metal. However, with the opposite bias applied, as shown in Figure 7c, the energy distribution in the semiconductor is lowered in relation to that of the metal. A net current flow from the metal into the semiconductor now results, because of those few electrons with energies higher than the barrier energy. However, this current is very small and does not increase appreciably with reverse bias. This situation leads to the familiar diode current-voltage characteristic shown in Figure 8. If the logarithm of the current is plotted as a function of the applied voltage in the forward direction, a straight line is obtained as shown in Figure 9. If the current at a given forward bias voltage is measured as a function of temperature, an Arrhenius plot can be constructed as shown in Figure 10. From either of these two plots the metal-semiconductor barrier energy can be deduced and the thermal nature of the current established.

If this were the entire story, the only possible way of making an ohmic contact to a semiconductor with a metal would be to decrease the barrier energy so that the thermal current which flowed in the reverse direction is large enough for the particular device application. Referring again to Figure 9, it can be seen that, for currents less than or of the order of the intercept current, the characteristic is ohmic up to voltages of 1 or 2 kT. By making the barrier sufficiently small, this current can be made quite large (i.e., a barrier of 0.25 eV can support currents in either direction of up to 10^4 amps/cm² and still remain ohmic). However, as we have seen with many of the covalent materials, the barrier energy on n-type material, for example, is approximately two-thirds of the energy gap of the semiconductor and does not change appreciably with the metal used. Hence the approach of making the barrier energy arbitrarily small cannot be used for ohmic contacts to these materials. With some highly ionic materials, such as zinc sulfide, another

difficulty is often encountered which is shown in Figure 5. That is, metals do not exist with electronegativity sufficiently small to make the resulting barrier energies low enough to form ohmic contacts for device purposes. Thus, if one considers only the thermal currents which can flow in the metal-semiconductor system, the number of semiconductors to which ohmic contact can be made becomes very small.

Thermal current, however, is not the only current which can flow in a metal semiconductor system. As the carrier concentration in the semiconductor is increased, the width of the depletion layer is decreased. At very high carrier concentrations, the depletion layer becomes sufficiently small that quantum mechanical tunneling can take place. This tunneling results from the fact that the electron probability distribution in the forbidden region is damped exponentially with distance and hence can penetrate a barrier if it is sufficiently thin. Referring again to Figures 7b and 7c, it can be seen that in both the forward and reverse directions the distance through which the electrons must tunnel decreases as the applied bias is increased. If the depletion layer is sufficiently thin, current in either direction increases markedly with voltage. The type of current-voltage characteristic resulting from tunneling of this sort is shown schematically in Figure 11. The solid line (curve A) is a typical curve for a device with low carrier concentration. In the forward direction the thermal current increases exponentially with voltage, as we have seen earlier. In the reverse direction only a very small saturation current flows until sufficient electric field is developed for avalanche breakdown to take place. However, if the carrier concentration in the semiconductor is increased, tunneling currents begin to flow in both the forward and reverse directions. In the forward direction tunneling current increases exponentially and is often difficult to distinguish from thermal current. In the reverse direction tunneling current increases exponentially but less steeply than in the forward direction. For carrier concentrations in the intermediate range (of the order of 10^{18} cm⁻³ for typical semiconductors such as gallium arsenide), the situation is illustrated by the dashed lines (curve B) in Figure 11. As the carrier concentration is increased still further to 10^{19} cm⁻³, the currents conducted in both the forward and reverse directions again increase exponentially but at very much lower voltages and there is a broad crossover range where the current is essentially ohmic. This is shown by the dot-dash line (curve C) in Figure 11. This then is the general approach used in making ohmic contacts to the traditional semiconductors: Place a metal contact on a region of very high carrier concentration and allow tunneling to take place through the barrier.

For the above reason it is important to understand in detail the nature of the tunneling process and how it is characterized in a given semiconductor. The exponential damping of the electron probability distribution in the semiconductor forbidden gap, to which we referred earlier, is quantitatively characterized by an exponential damping constant, q , which is shown in Figure 12 as a function of electron energy. Near the conduction and valence band edges, q is small and the electron penetrates deeply into the semiconductor. However, near the center of the forbidden gap, q is relatively large and the electron distribution damps very quickly.

Although the actual characteristic, q , versus energy curves has been measured for only a small number of semiconductors, the relationship to be used in any given semiconductor can often be estimated from the equation:

$$E^2 \approx \left(\frac{E_g}{2}\right)^2 + \frac{\hbar^2 q^2 E_g}{2m^*} \quad (2)$$

where E is the electron energy measured from the center of the forbidden gap, E_g is the gap energy, and m^* is the effective mass of the majority carrier in the semiconductor.

In Figure 11 current-voltage characteristics of a typical metal semiconductor contact were plotted for various semiconductor carrier concentrations. If we plot the logarithm of current versus the applied voltage for the forward direction instead, curves such as those shown in Figure 13 are obtained. For low carrier concentrations the current increases by a factor of e of every KT of applied voltage. As the carrier concentration is increased, current due to tunneling becomes important and a nearly exponential characteristic results with smaller slope and large intercept on the current axis. At still higher carrier concentrations, appreciable currents flow in the neighborhood of zero bias. These currents are of interest when the contact is used as an ohmic contact. However, a great deal of information can be obtained about the density of electrons just under the contact by studying the forward characteristics. To a first approximation the current is given by the expression

$$I \propto \exp\left[2\left(\frac{\epsilon m^*}{\hbar^2 N}\right)^{1/2} V\right] \quad (3)$$

where N is the net donor (or acceptor) concentration and V the applied voltage. Thus the forward characteristics in the tunneling range are very sensitive to the carrier concentration; if one knows the effective mass of the carrier involved, one can estimate from the slope the actual carrier concentration and from the intercept the general magnitude of current which can be carried by the contact as an ohmic contact without unreasonable voltage drop.

It should be emphasized again that the origin of these current-voltage characteristics is the exponential attenuation of the electron probability distribution in the forbidden gap of the semiconductor. In other words, the current which flows decreases in a generally exponential way with the width of the depletion layer through which the electron must tunnel, in accordance with the attenuation constant given in Figure 12. For this reason, one would expect the ohmic contact resistance to depend exponentially upon the depletion layer width at zero bias and hence upon the square root of the carrier density in the semiconductor. Figure 14 illustrates some measurements⁽⁴⁾ of ohmic contact resistance as a function of (carrier density)^{-1/2}; it can be seen that contact resistance varies exponentially with this quantity over a wide range, as is expected from this mechanism.

In summary, it can be said that there are two general classes of semiconductors with respect to their contact behavior. First are those to which ohmic contact can be made when the barrier energy is sufficiently small that the thermal currents which flow are sufficiently large for the device application involved. This is the case, for example, in the cadmium sulfide photocell. The second class is materials where this cannot be done; in general, ohmic contacts can be made by making the carrier density in the materials sufficiently large that quantum mechanical tunneling can take place. Quantitative studies of the characteristics of a particular contact can be very helpful in establishing the carrier density of the materials just under the contact and in characterizing over what range of current the contact is useful.

REFERENCES

1. K. F. Braun, Ann. Phys. Pogg. 153, 556 (1874).
2. K. F. Braun, Ann. Phys. Wied. 1, 95 (1877).
3. C. A. Mead, Solid-State Electronics, 9, 1023 (1966), and references cited therein.
4. M. Millea, M. McColl, and C. A. Mead, Phys. Rev., to be published.

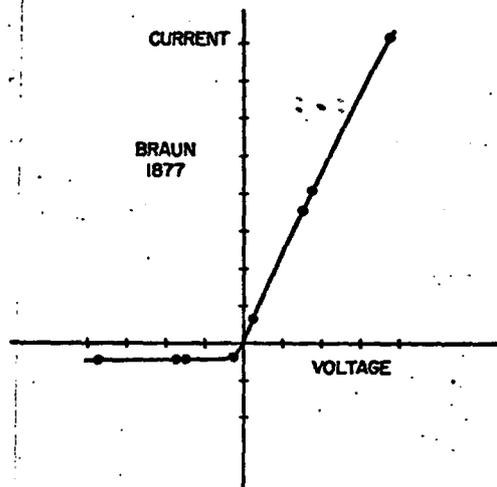


Figure 1. Current-voltage characteristic of metal wire contact to lead sulfide crystal taken from K. F. Braun (1877)

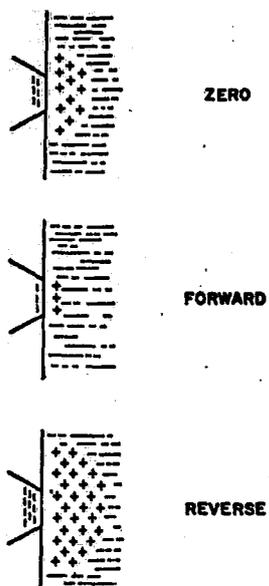


Figure 2. Schematic of situation at metal-semiconductor interface indicating the difficulty of injecting electrons from the metal into the semiconductor

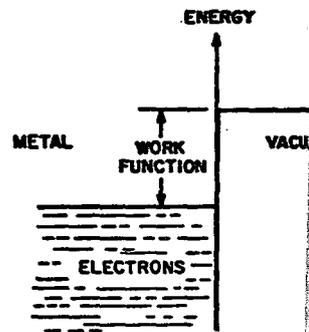


Figure 3. Diagram of electron energy versus distance through metal-vacuum interface showing the work function barrier

Figure 4. Energy diagram of a metal-semiconductor contact (a) at the moment of contact when electrons are dribbling into the metal and (b) after equilibrium is reached [from C. A. Mead in *Solid-State Electronics*, Vol. 9 (New York, Pergamon Press, 1966) p. 1026]

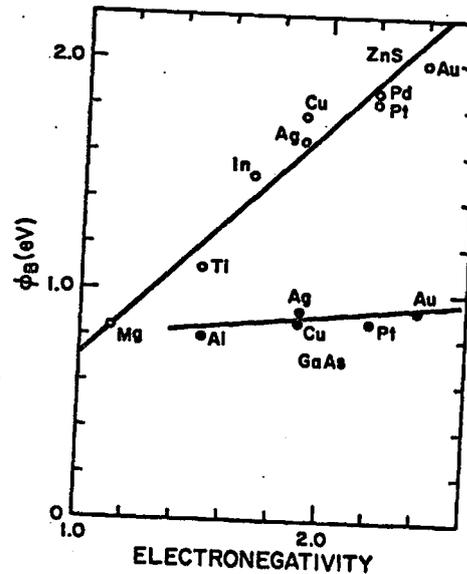
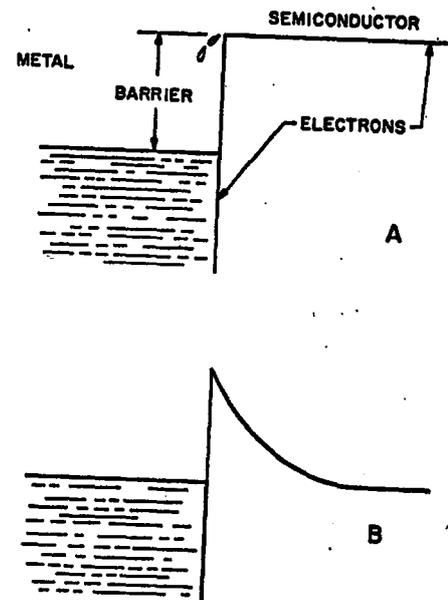


Figure 5. Measured barrier energies for several metals on an ionic semiconductor (zinc sulfide) and a covalent semiconductor (gallium arsenide) [from C. A. Mead in *Solid-State Electronics*, Vol. 9 (New York, Pergamon Press, 1966) p. 1128]

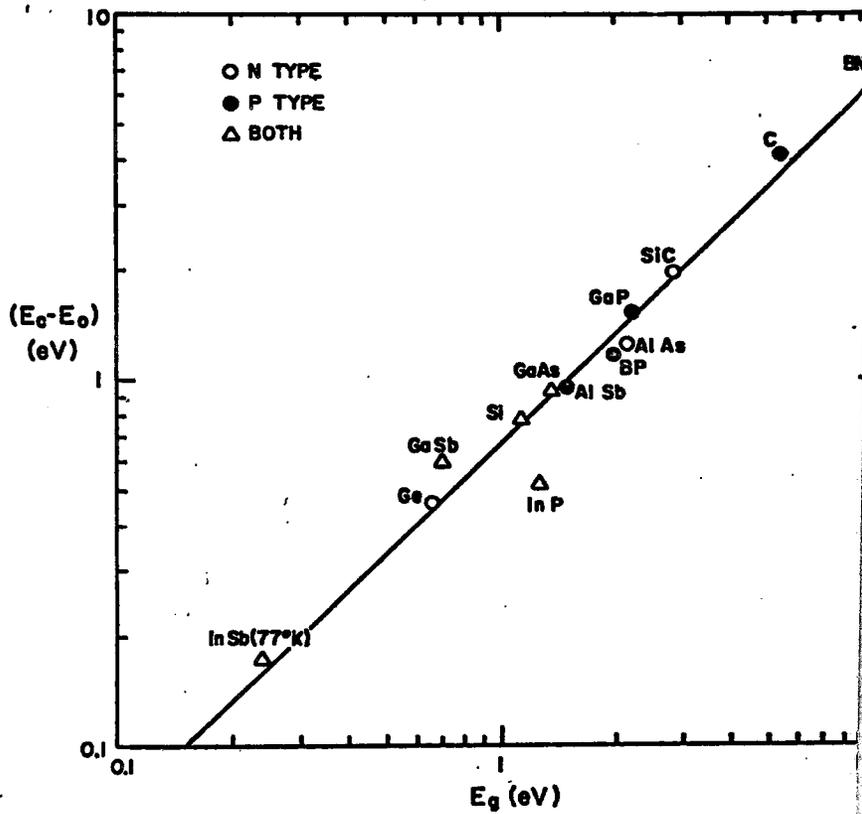


Figure 6. Barrier energies for gold on several n-type covalent semiconductors; straight line is where the barrier energy is $2/3$ of the band gap [from C. A. Mead in *Solid-State Electronics*, Vol. 9 (New York, Pergamon Press, 1966) p. 1029]

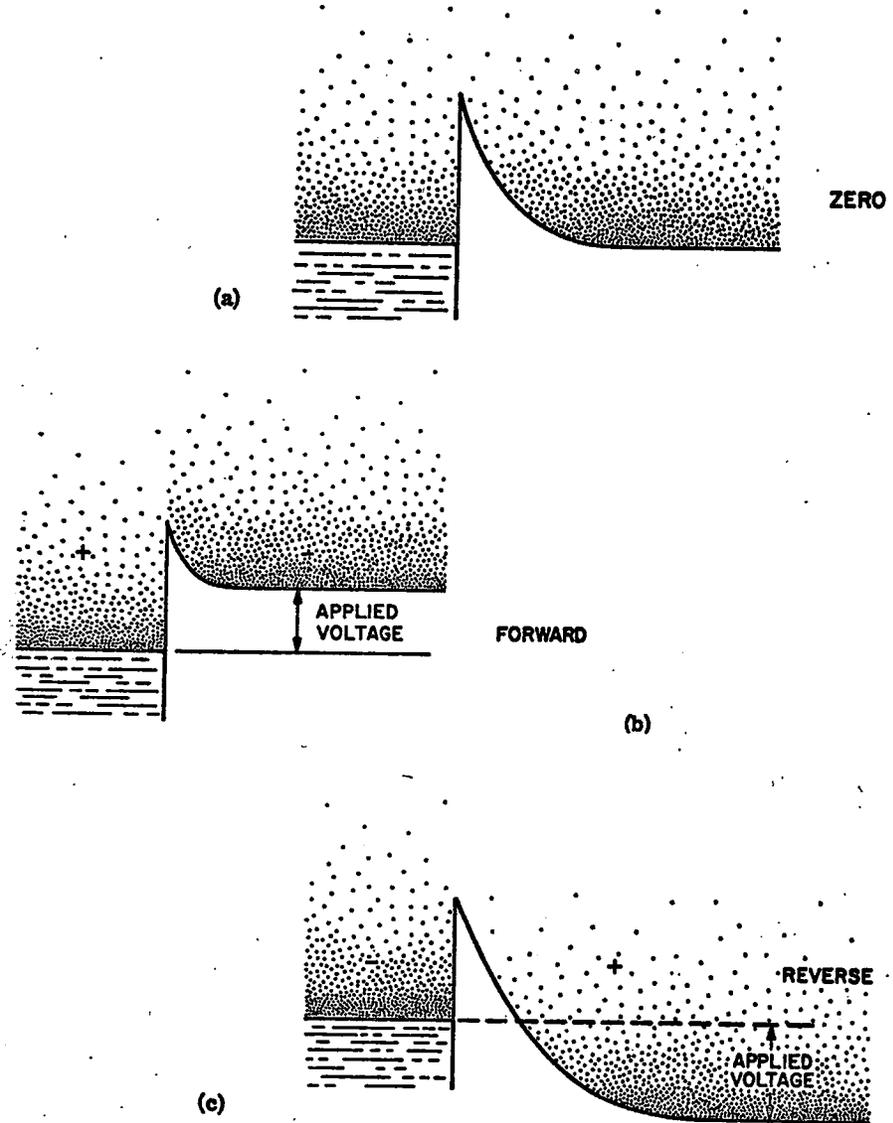


Figure 7. Energy diagrams showing schematically the thermal electron energy distribution in the case of (a) zero bias, (b) forward bias, and (c) reverse bias

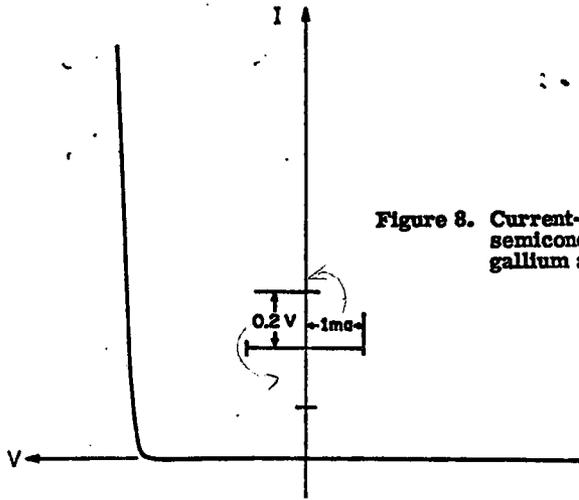


Figure 8. Current-voltage characteristic of typical semiconductor contact; this sample is gallium arsenide $\phi \approx 0.9$ eV

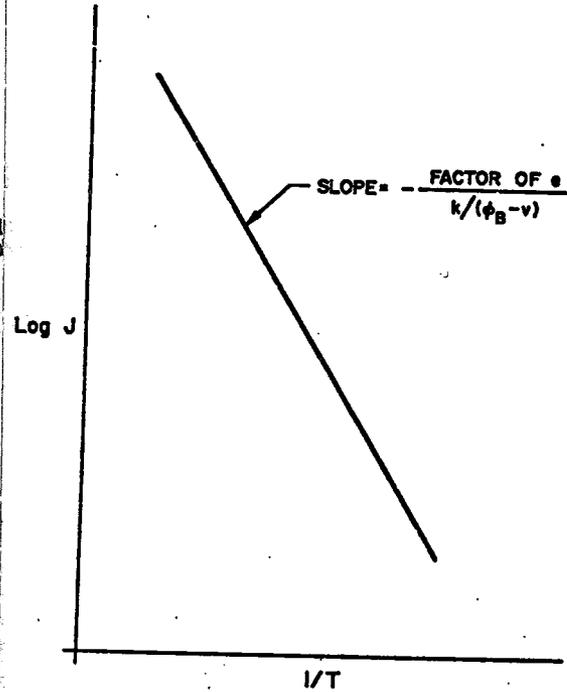


Figure 10. Schematic Arrhenius plot for forward-biased metal-semiconductor contact [from C. A. Mead in *Solid-State Electronics*, Vol. 9 (New York, Pergamon Press, 1966) p. 1025]

Figure 9. Schematic of Figure 5; I versus V plot for metal-semiconductor contact [from C. A. Mead in *Solid-State Electronics*, Vol. 9 (New York, Pergamon Press, 1966) p. 1025]

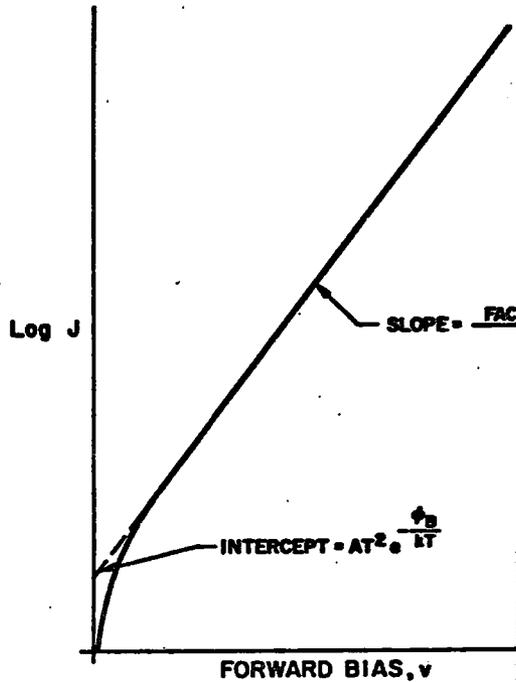
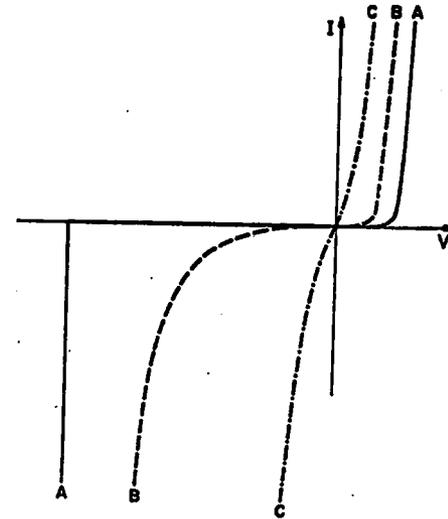


Figure 11. Schematic current-voltage curves for a metal contact on semiconductor samples of progressively higher carrier density (a) low carrier concentration, (b) approximately 10^{18} carrier/cm³ (c) approximately 10^{19} carrier/cm³



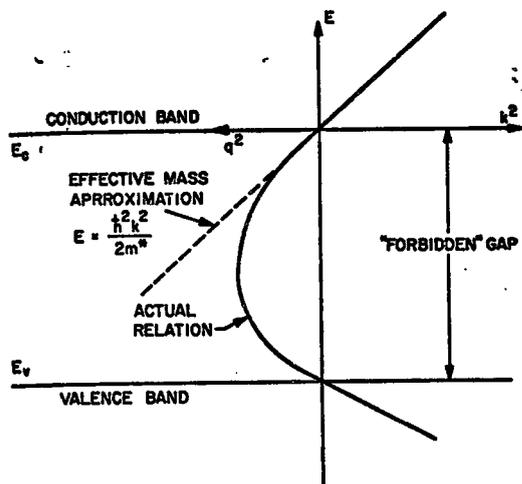


Figure 12. Propagation constant of electrons in a semiconductor showing the attenuation constant α of the forbidden gap as a function of electron energy

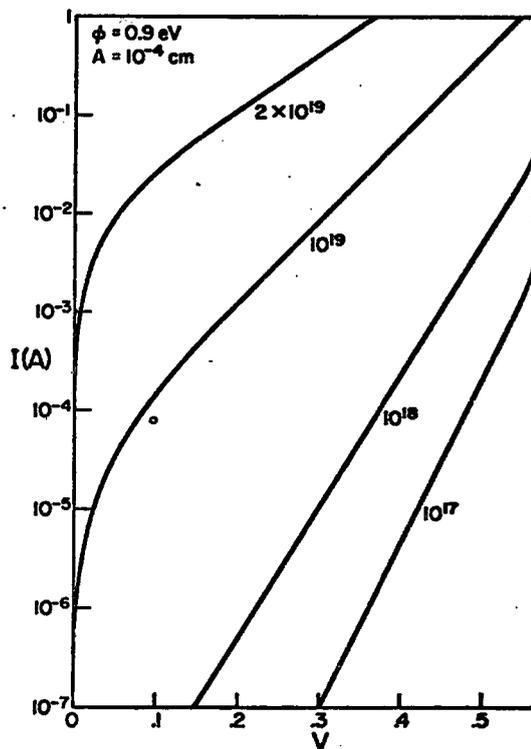


Figure 13. Schematic log current versus voltage plot for similar progression of samples as shown in Figure 11

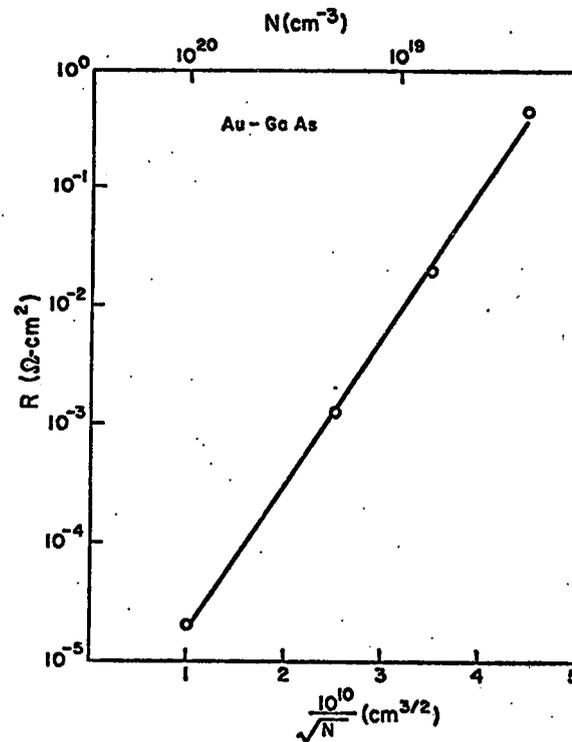


Figure 14. Contact resistance of gold on gallium arsenide contacts as a function of gallium arsenide carrier concentration (from Reference 4)