CHAPTER II

METAL-SEMICONDUCTOR CONTACTS

The earliest systematic investigation on metal-semiconductor rectifying systems is generally attributed to Braun who in 1874 noted the dependence of the total resistance on the polarity of the applied voltage. The point-contact rectifier in various forms found practical applications beginning in 1904. In 1931 Wilson formulated the transport theory of semiconductors based on the band theory of solids. This theory was then applied to the 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. The model arising from this consideration is known as the Schottky barrier. In 1938 Mott also devised an appropriate theoretical model for swept-out metal-semiconductor contacts that is known as the Mott barrier. The basic theory and the historical development of rectifying metal-semiconductor contacts were summarized by Henisch in 1957(2).

2.1 The Schottky Model

Barriers at metal-semiconductor contacts, if they follow the simple Schottky model, are determined by the difference in the work function, ϕ_m, of the metal and the electron affinity, χ_s, or work function, ϕ_s, of the semiconductor. The energy diagrams that are expected are shown in Figs.1 and 2 for metal contacts on n-and p-type semiconductors. In Fig. 1(b), for instance, the barrier to the movement of electrons from the n-semiconductor into the metal is $(\phi_m - \phi_s)$, and the barrier to the reverse flow of electrons from the metal to the semiconductor is $(\phi_m - \chi_s)$. If the junction is forward biased with an external voltage source, V_a, so that the semiconductor is negative with respect to the metal, the forward barrier becomes $q(V_D - V_a)$. Where V_D is the diffusion voltage. However, the reverse barrier $(\phi_m - \chi_s)$ remains, in a first-order model, relatively unaffected by applied voltage or by the doping level of the semiconductor. Therefore $(\phi_m - \chi_s)$ will be defined as the barrier of the metal-semiconductor pair.

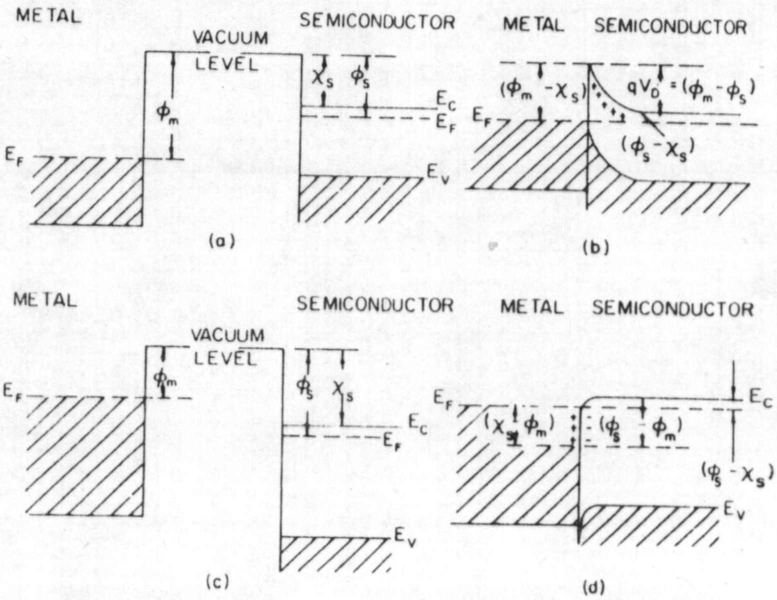

Fig. 1 Energy level diagrams of metal contacts to n-type semiconductors. (a) and (b) with $\phi_m > \phi_s$, (c) and (d) with $\phi_m < \phi_s$. Contact (b) acts as a rectifier, since a barrier $(\phi_m - \phi_s)$ exists in the conduction band of the semiconductor. The forward bias direction is for the semiconductor negative with respect to the metal and electrons flow from the simiconductor into the metal. Contact (d) is ohmic since virtually no barrier exists in the conduction band.(3)

Comparison of Fig. 1(b) and (d) shows that the junction is rectifying at "low" T for an n-type semiconductor if $\phi_m > \phi_s$, and ohmic if $\phi_m < \phi_s$. For a p-type semiconductor (Fig. 2) the converse is expected to be true. The experimental evidence for most semiconductor, whether n- or p-type, however, cannot be explained by this simple model which neglects the effects of interface states.

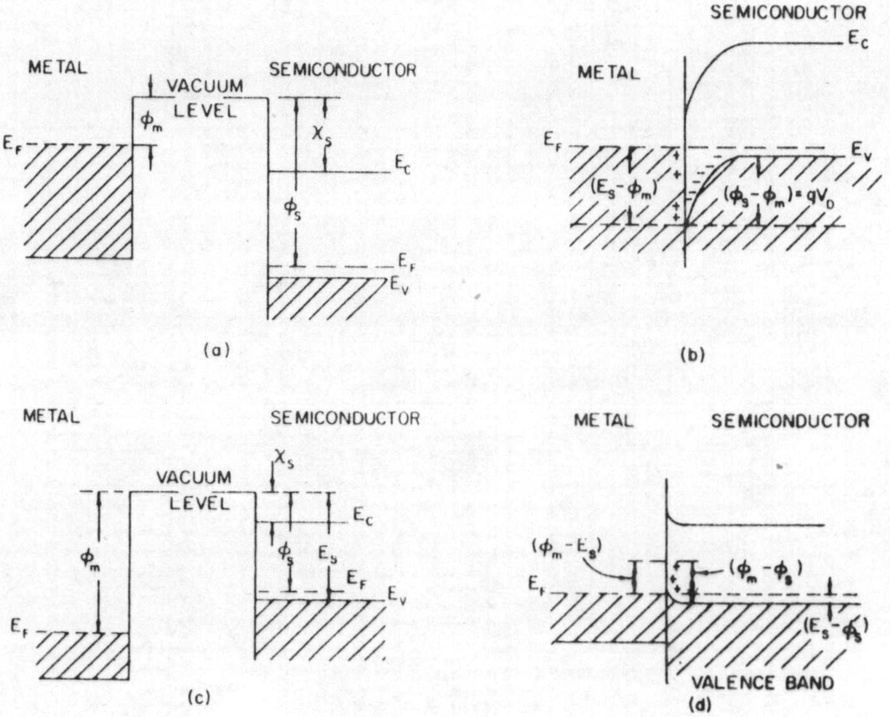

Fig. 2 Energy level diagrams of metal contacts to p-type semiconductors. (a) and (b) are for $\phi_m < \phi_s$. Contact (b) therefore acts as a rectifier since a barrier $(\phi_s - \phi_m)$ exists to the flow of holes. This barrier to hole flow is lowered if the semiconductor is made positive with respect to the metal. E_s is $\chi_s + E_g$. In(c) and (d) since $\phi_m > \phi_s$ there is no barrier to current flow and the contact is ohmic.$^{(3)}$

2.2 Contact Resistance and Contact Resistivity (14,15)

One of the obvious measures that have been taken to obtain higher densities in monolithic circuits has been to decrease the area provided for the metal-semiconductor interconnection contacts on the circuit chips. This has put a severe burden on the technologist, who has to develop proper production processes for small area contacts of low resistance and high reliability.

2.2.1 Definition of Contact Resistance

The access to a semiconductor region via a metal contact usually exhibits higher resistance than expected from an ideal contact. The additional resistance may be imagined as a series resistor connected to the ideal contact and it will be refered to as "contact resistance"

The theoretical treatment of semiconductor devices has been facilitated by this mental image. First, it attributes the consideration of the device behavior under simplifying condition of ideal contacts and after adding the experimentally or theoretically derived contact resistances that provide the corrections for the actual contacts. However, the still rather vague notion of contact resistance as a series resistance describing the deviation of an actual contact from its ideal condition can be strengthened by a precise definition contact resistance.

Practical contacts are nonideal for two basic reasons

1. In a thin surface layer of the semiconductor beneath the contact metal the charge carrier density differs from that of the semiconductor bulk (accumulation or depletion layer) due to differences in work function of metal and semiconductor or due to surface states of the semiconductor

2. Layers of foreign matter sometimes impede a rigorous contact of metal and semiconductor.

In contrast, an ideal contact does not exhibit any interface layers. Hence, its potential may be identified with the potential Ø' in the contact plane of the semiconductor after extending the latter by its mirror image on that contact plane (Fig.3).

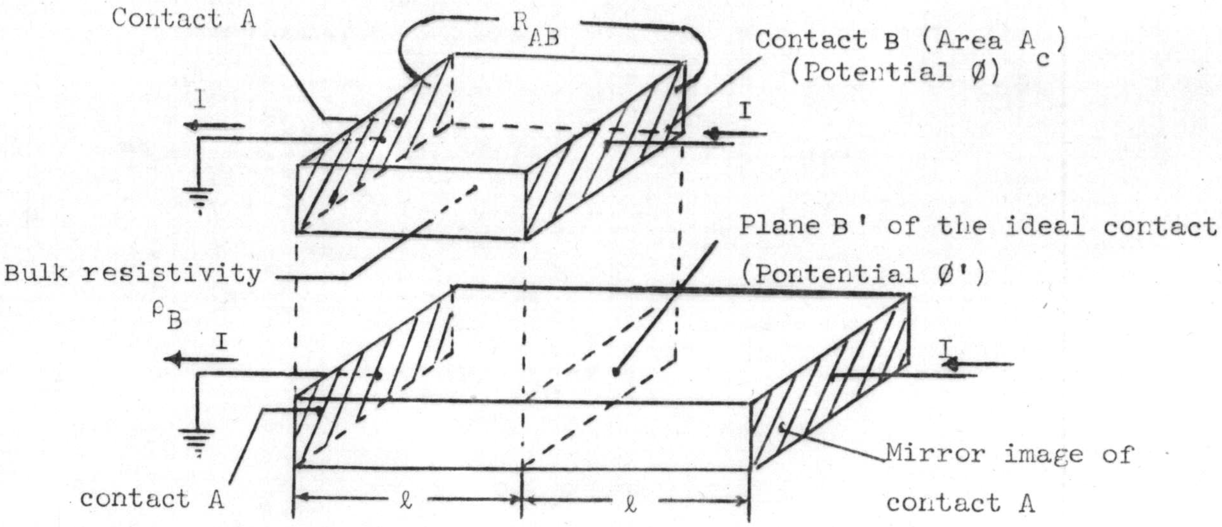

Fig.3 Derivation of the potential Ø' of the ideal contact (B') pertaining to the real contact (B) by mirror-imaging.[4]

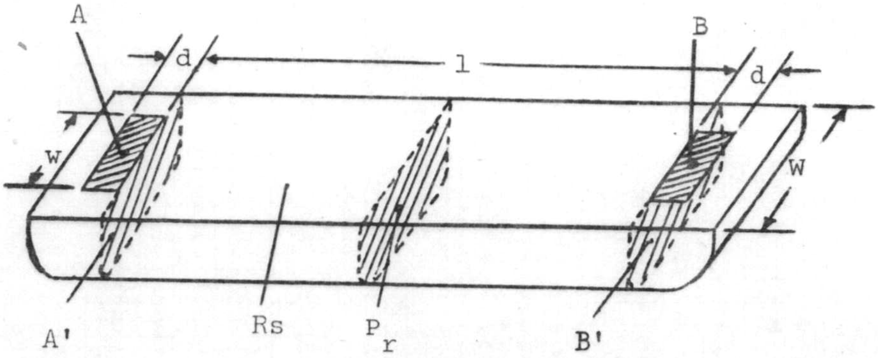

Fig.4 Planar resistor with real contacts A and B and the most practicable locations A' and B' for the corresponding ideal contacts.(4)

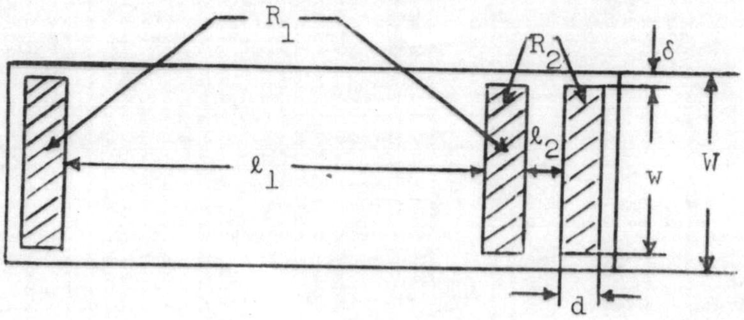

Fig.5 Planar resistor having three contacts for contact resistance determination.(4)

To compare the access via the actual contact with that via its ideal counterpart, a common reference plane P_r within the semiconductor must be identified. This plane must coincide with a plane of constant potential and must carry the total contact current. The resistance between P_r and the actual contact is R_a, that between P_r and the ideal contact R_i. The contact resistance is then defined by the difference of these two resistances

$$R_c = R_a - R_i \tag{2.1}$$

Of course, for this comparison of R_a with R_i, the reference plane P_r must (practically) remain a plane of constant potential when the actual contact is replaced by its ideal counterpart. Hence, the mental image underlying the contact resistance concept fails where such a reference plane cannot be found. In such cases the contacts must be considered an integral part of the whole semiconductor device.

With the sample shown in Fig. 3, however, any plane parallel to the contacts may be taken for P_r provided that the contacts are uniform. Although no assumptions have been made in the definition of contact resistance regarding the position of the ideal contact, for the sample in Fig. 3, it is most convenient to let ideal and actual contact coincide. Taking, e.g., the grounded contact A as the reference plane P_r, the contact resistance of contact B may be written according to Eq.2.1 as the difference of the potential $\emptyset$ of the actual contact and the potential $\emptyset'$ of its ideal counterpart, divided by the current I

$$R_c = \frac{\emptyset - \emptyset'}{I} \qquad (2.2)$$

With other structures it might prove advantageous to choose a plane for the ideal contact that does not coincide with that of the actual contact. This is the case with the contact on a diffused resistor shown in Fig.4. By using the planes A' and B' for the ideal contacts, one facilitates the calculation of R_i in Eq.(2.1). Thus, with the reference plane P_r in the center of the resistor one obtains for equal contacts

$$R_c = \frac{1}{2}\left(R_{AB} - R_s \cdot \frac{l}{W}\right) \qquad (2.3)$$

where R_s is the sheet resistance of the resistor layer. Solved for R_{AB}

$$R_{AB} = R_s \cdot \frac{\ell}{W} + 2R_c \qquad (2.4\,)$$

it corresponds to the widely used formulation in terms of an "end correction" (equivalent number of squares k of the resistor layer)

$$R_{AB} = R_s\left(\frac{\ell}{W} + 2k\right) \qquad (2.5)$$

Equation (2. 4) leads to a practicable way of measuring the so defined contact resistance of a planar resistor. Such a resistor having three equal contacts A, B, C arranged in different distances ι_1 and ι_2 (Fig. 5) exhibits different resistances between A and B (R_1) and B and C (R_2). The contact resistance R_c can be eliminated by applying Eq. (2. 4) to R_1 and R_2, which results in

$$R_c = \frac{R_2 \cdot \ell_1 - R_1 \cdot \ell_2}{2(\ell_1 - \ell_2)} \qquad (2.6)$$

It might appear peculiar that in this paper contact resistance has been defined in such a way as to allow a free choice of the position of the ideal contact, thus permitting the arbitrary introduction of semiconductor bulk portions into the contact resistance. Clearly, this would be a disadvantage if there were reasons to assume that coinciding ideal and actual contact planes would generally lead to a contact resistance being free of any semiconductor bulk influence and thus describing solely the metal-semiconductor interface. However, the replacement of the actual by the ideal contact alters the boundary conditions for the electric field in the semiconductor and hence changes the current distribution in the vicinity of the contact, unless it is fixed by the geometry of the sample as in Fig. 3 . Obviously, change of current distribution will introduce different bulk influence into R_a and R_i Eq. (2.1) and thus introduce bulk resistance into the contact resistance. Generally attempts to split the total contact resistance into interface and bulk contribution in the form of concentrated series resistors must be rather fruitless according to the foregoing discussion. Actually, it suffices to make this separation only for infinitesimal sections of the contact. This leads to the term "contact resistivity" solely describing the metal-semiconductor interface independent of contact geometry(5)

2.2.2. Contact Resistivity

The expression for the contact resistivity ρ_c can be derived formally from the contact resistance definition. Consider a tube of current streamlines in the semiconductor device carrying a portion ΔI of the total current. In the plane of the contact the tube has a cross section of ΔA_c of the contact may now be considered as adevice with a contact, for which a contact resistance R_c can be defned according to Eq.(2.1). For this purpose, the plane of the ideal contact coincides with that of the actual one. Furthermore, it is assumed that the current flow within the interface portion of the tube is directed vertically to the contact plane. Then contact resistivity is defined as

$$\rho_c \text{ ohm-cm}^2 = \lim_{\Delta A_c \to 0} (R_c . \Delta A_c), \qquad (2.7)$$

where an infinitesimal tube of streamlines are taken.

Substituting Eq. (2.2) into Eq. (2.7), one get

$$\rho_c = \lim_{\Delta A_c \to 0} \left[(\phi - \phi') \frac{\Delta A_c}{\Delta I} \right] \tag{2.8}$$

$$= \frac{1}{j_c} \lim_{\Delta A_c \to 0} (\phi - \phi') \tag{2.9}$$

where j_c stands for current density at the contact.

For a homogeneous contact having uniform current density as in Fig. 3, Eq. (2.7) simplifies to

$$\rho_c = R_c . A_c \quad (j_c = \text{const.}) \tag{2.10}$$

This equation has been used to determine ρ_c via contact resistance and contact area A_c. However, it must be kept in mind, that it is valid only for contacts having uniform current distribution dictated by the sample geometry.

The assumption of vertical current flow in the interface layer is, of course, fulfilled with the sample in Fig. 3. It is also practically fulfilled even with adverse sample geometries, if the resistivity of the interface layer is much larger than that of the undisturbed semiconductor bulk. Most practical contacts indeed exhibit very thin depletion layers of very high resistivities.(6) For these, without noticeable error, the thin depletion layer may even be considered as lying outside of the semiconductor like a possible layer of foreign matter. The potential difference $\phi - \phi'$ in Eq. (2.9) then becomes the total voltage drop v_c across these outside layers

$$\rho_c \approx \frac{v_c}{j_c} \tag{2.11}$$

This is the form(7) most readily used in practical applications of ρ_c. Of course, a definite value ρ_c can only be given for contacts which are sufficiently ohmic within the range of current density under study.