CHAPTER II

METAL - SEMICONDUCTOR CONTACTS

This chapter views the metal-semiconductor contacts in some aspects pertinent to our work. In section 2.1 and 2.2 we discuss the Schottky barrier formation, and in section 2.3 the general transport mechanisms of the carriers traversing the Schottky barrier. In section 2.4, the definitions of the ohmic contact are reviewed and the ohmic criterion is proposed in section 4.2. The section 2.5 discusses theoretical contact resistivities based on various mechanisms. We will calculate the values needed for theory-experiment fitting in subsection 6.2.2. The section 2.6, the effect of an interfacial layer is discussed as the tunneling of holes through the insulator's barrier which is observed in the calcution in section 6.2 by the reduction of apparent area A_{app}.

2.1 The Schottky and Bardeen Models of Metal Semiconductor Contacts

The simplest way of modeling the electrical properties of metal-semiconductor (MS) contacts is to assume no electronic states in the semiconductor band gap. In this way, Schottky (1938) proposed that the barrier height of the MS contacts is given by

$$\phi_n = \phi_m - \chi_s \qquad (2.1.1)$$

$$\phi_p = E_g + \chi_s - \phi_m \qquad (2.1.2)$$

where

ϕ_n = Schottky barrier height of n-type semiconductor,

ϕ_p = Schottky barrier height of p-type semiconductor,

ϕ_m = metal work function,

χ_s = electron affinity of the semiconductor (the energy from the bottom of the conduction band to the vacuum level), and

E_g = energy gap of the semiconductor.

The model predicts that for an appropriate choice of the metal workfunction, the MS contact can be accumulation, neutral, or depletion contacts. For p-type semiconductor which is the type of special interest in this thesis, an accumulation (ohmic) contact should be obtained when $\phi_m \geq E_g + \chi_s$. However, in most of semiconductor, including p-CuInSe$_2$, this condition is not applicable.

While the Schottky model predicts that the barrier height should vary directly with ϕ_m, in reality the barrier height is often found to be relatively independent of ϕ_m. This led Bardeen (1947) to point out the importance of the surface states. Additionally, Bardeen supposed that there could exist a thin insulator layer between the metal and the semiconductor. Since surface states can screen the semiconductor from the electric field in the insulator which arises from the difference between metal and semiconductor work functions. So, the metal work function ϕ_m is less important to the barrier height. In the case that there exists sufficient number of surface states, the barrier height would be independent of ϕ_m. The Fermi level at the interface is then said to be "pinned", since its position could not much vary in the band gap.

2.2   Mechanism of Barrier Formation

The Bardeen model, introduced above, requires an interfacial layer for screening, the layer is often a thin oxide layer. However, it is possible to prepare an intimate MS interface which precludes oxide layer formation. For example, freshly cleaved (110) surfaces of GaAs, InP and GaSb in high vacuum. However, even without externally-caused interfacial surfaces, there are still the natural interface states (Spicer et al., 1979). The first explanation of this effect was offered by Heine (1965). He suggested that the decay of the metallic wavefunctions into the semiconductor could create interface states. So, this model is called "metal induced gap states" (MIGS).

A serious limitation of the MIGS model is that it seems improbable to form the interface state until a sufficient thickness of the metal has been deposited. The critical thickness is about 5 Angstroms. However, Spicer et al. (1979) have demonstrated that the Fermi level on the surface of a covalently bonded semiconductor is pinned by a metal coverage of only a tenth of a monolayer. Furthermore, a similar coverage of oxygen atoms pin the Fermi level at the same position in the semiconductor band gap. These results led Spicer et al. (1980) to propose that the release of the heat of condensation by metal, or oxygen, atoms arriving on the surface of the semiconductor creates native point defects in the semiconductor. These defects then pin the Fermi level at the MS interface. So, the model is called the "unified defect model".

All models mention above do not consider the possibility of diffusion and chemical process at the interface. It has become clear that these processes often occur. In many semiconductors, the presence of a metal layer promotes dissociation of the semiconductor, even when there is no direct chemical reaction between metal and semiconductor. For example, deposition a gold film on a silicon crystal, results in a weakening of the bonds close to the interface. There is some mixing of gold and silicon of unknown stoichiometry (Braicovich et al., 1979). Similarly, there exists metal-semiconductor alloying in Au contacts to GaAs, GaSb and InP, resulting from the out-diffusion of one or both of the semiconductor components into the Au film. Some in-diffusion of Au may also occur (Hiraki et al., 1971; Sinha and Poate, 1973; Chye et al., 1978; Komatsu et al., 1981).

More complex interactions are found if there is a chemical reaction between the metal and one component of the semiconductor, for example, the contact of Al onto GaAs. The semiconductor dissociates under the contact and Ga diffuses away from the interface and replace by Al with the formation of AlAs. The importance of chemical effects at the interface was first considered by Phillips (1974), Andrew and Phillips (1975), and Brillson (1978). Brillson (1979) proposed that the electrical properties of MS contacts can be predicted from whether a chemical reaction occurs and whether the semiconductor is ionic and covalent bonded.

Different from Brillson model, Freeouf and Woodall (1981), with the spirit of Schottky, proposed that the properties of contacts are controlled primarily by the workfunctions of the pseudo-alloys formed by the mixing of metal and

semiconductor at the interface. This model is called the effective workfunction model.

However, it is believed that there is no simple model that can completely explain the formation of the barrier. More importantly perhaps, one should not bother except to accept that the symmetry properties differ between that of a point on the surface and one in the bulk, leading to some degree of surface states.

2.3   Current Transport Mechanisms Through the Schottky Barrier

Generally, the mechanisms of hole (electron) transport across the barrier are (Rhoderick, 1978) described as :-

a) emission of holes over the top of the barrier; this is called thermionic emission (TE),

b) quantum mechanical tunneling through the barrier; this is called the field emission (FE),

c) recombination in the space charge region, and

d) recombination in the neutral region.

The last two, as well as transport by minority carriers, are considered negligible for an energy barrier which is small compared to the band gap.
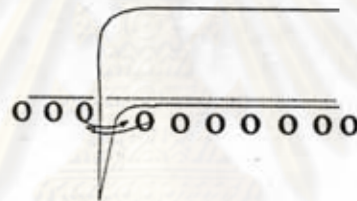
The thermionic emission, field emission and the combination of these two mechanisms, i.e. the thermionic-field emission, are shown below.
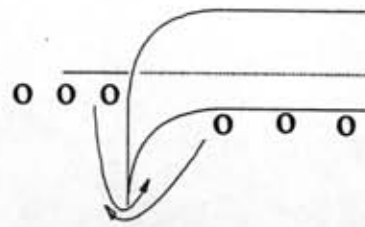
Low N_a ; Thermionic emission.
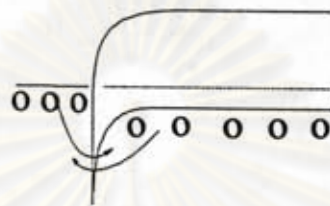
Intermediate N_a ; Thermionic-field emission.

High N_a ; Field emission.

Fig.1 Conduction mechanisms for metal/p-semiconductor contacts with increasing semiconductor doping concentration.

The ratio of TE /FE can be obtained by the expression kT/E_{00}, where k is the Boltzmann constant and T the absolute temperature. The energy E_{00} is given by (Padovani and Strattron, 1966)

$$E_{00} = (qh/4\pi)(N_a/K\varepsilon_0 M_t)^{1/2} \qquad 2.3.1$$

where

N_a =   acceptor concentration,

K   =   dielectric constant of the semiconductor, and

M_t =   appropriate effective mass.

The other symbols have their usual meaning.

For lightly doped semiconductors ($kT/E_{00} \gg 1$; $N_a < 10^{17}$ cm^{-3}), the current is the result of TE. For heavily doped semiconductors ($kT/E_{00} \ll 1$; $N_a > 10^{19}$ cm^{-3}), the barrier is thin enough at or near the bottom of the valence band where there are copious number of holes. So that, FE dominates. In the intermediate range ($10^{17} < N_a < 10^{19}$ cm^{-3}), TFE results. The material used in this study has the doping concentration in the order of 10^{17} cm^{-3}. So, it is expected that the pertinent mechanism is TFE.

2.4 Definitions of Ohmic Contact

There are many definition of ohmic contact, including the following (Look, 1989):

a) a perfect source and sink of both carrier types and having no tendency to inject or collect either minority carriers,

b) a source of carriers with an internal resistance R_c which is totally negligible compared with the semiconductor resistance,

c) a source of carriers with a non-negligible R_c, but one which obeys Ohm's law for current densities of interest.

Definition (a) has only academic significance since no actual contacts ever fulfill this condition. More practically used are definition (b) or (c). However the ohmic property in definition (b) depend on the relative dimensions. Definition (c) is not acceptable in the case of contact resistance R_c is so high that sufficient current can not be passed, even R_c is ohmic.

2.5 Contact Resistivity

When metal-semiconductor contact is used as the ohmic contact, it is qualified by the contact resistivity ρ_c, which is defined by (Schroder and Meier, 1984)

$$\rho_c = \left(\partial J/\partial V\right)^{-1}_{V=0} \tag{2.5.1}$$

where J is the current density, and V is the applied voltage, the units of ρ_c is Ω-cm^2. Note that in the range that the contact is qualified by ρ_c, the contact resistance (in Ω) is independent of the polarities of bias. In practice, it holds in the small-current range.

The contact resistivity for TE is given by (Sze, 1981)

$$\rho_c(TE) = \rho_1 \exp(q \, \phi_{bp} /kT) \tag{2.5.2}$$

where $\quad \rho_1 = k / (qA*T)$,

$A* = 4\pi\, qk^2\, m* / h^3$ is the Richardson constant, here m* is the density of state effective mass,

ϕ_{bp} = barrier height including the image force lowering.

For TFE , ρ_c is given by (Yu, 1970)

$$\rho_c(TFE) \quad = \quad \rho_1\, C_1 \exp(q\, \phi_{bp} / E_0) \qquad (2.5.3)$$

For FE , it is (Yu, 1970)

$$\rho_c(FE) \quad = \quad \rho_1\, C_2 \exp(q\, \phi_{bp} / E_{00}) \qquad (2.5.4)$$

Where C_1 and C_2 are functions of Na, T and ϕ_{bp}. E_0 is related to E_{00} by

$$E_0 \quad = \quad E_{00} \coth (E_{00} / kT) \qquad (2.5.5)$$

Several of the above parameters vary with temperature and doping concentration, leading to ρ_c varying in different ways. The $\rho_c(TE)$ is very sensitive to temperature variations, while it is nearly independent of doping. In contrast, $\rho_c(FE)$ is only weakly temperature dependent, the dependence is contained in C_2, while strongly dependent on doping through E_{00}.

It should be noted that eqs. 2.5.2, 2.5.3 and 2.5.4 resulted from linearization (truncation) of the Taylor series, and as such are of limited validity for obvious reasons.

2.6 The Effect of an Interfacial Layer

If the Schottky diodes are not manufactured by cleaving the semiconductor in an ultra-high vacuum, a thin oxide layer between metal and the semiconductor is formed. This interfacial layer may be considered nonconducting, even though it may be so thin that it does not possess the band-structure characteristic of a thick oxide. To see the effect of an insulating layer, let us look at the usual Schottky diode current-voltage relationship (Rhoderick, 1978)

$$J = A^* \, T^2 \, \exp(- q \, \phi_{bp} \, /kT) \, [\, \exp(qV \, /nkT) - 1] \qquad (2.6.1)$$

where   J   =   current density,

n   =   nonideality factor.

the other symbols are the same   as that of subsection 2.5. The insulator layer affects the Schottky diode in three ways (Rhoderick, 1978) :

a)   the zero bias barrier height is lower than it would be in ideal diode. This is due to some part of potential drop in the insulator layer,

b)   holes have to tunnel through the thicker barrier. So, the current is reduced, and reflect by the reduction of A^*,

c)   the barrier height (not including image force lowering) is dependent of applied bias. This change the shape of the current-voltage characteristics in a manner which can manifest itself in the non ideality factor n.