CHAPTER IV

THEORETICAL MODELS

From the DACCT measurement, and based on the theoretical models of Tantraporn (1970) and Tantraporn (1972) which deal with the intimate contacts, it is possible to use his models to the case believed to be more common, i.e. the contacts which have thin interfacial layer between metal and semiconductor. Section 4.1 deals further with the pseudo-Richardson model, which is appropriate in explaining the ohmic behavior of the contacts. It gives the range of current and temperature that the contact should be referred as "ohmic". As already mentioned in the section 2.4, there are several definitions of the ohmic contact, but it seems that they are difficult to access by the experiments. So, a practical ohmic criterion is defined in the section 4.2. This criterion is easily access by the DACCT measurement.

To obtain more advantage from the measurement, another model in section 4.3 which is more complex is described. To ensure the correctness of the model, it calls on the basic principles, such as thermionic-field and WKB approximations. From consideration in subsection 4.3.1, ambiguous unknowns are not fixed, a priori, in the computer fitting procedure. Even the conduction area and tunneling effective mass are regarded as unknown parameters in fitting.

4.1 Restatement of Pseudo-Richardson Model

As mentioned in section 2.6 when there is a thin insulator layer between metal and insulator, hole (electron) have to tunnel through both the insulator barrier and across the Schottky barrier. These two processes are in series. So, the right hand side of eq. 3.2.2.1 should be multiplied by Ti, where Ti is the average tunneling coefficient of hole in tunneling through the insulator barrier (Card and Rhoderick, 1971). The outcome is the reduction of area A_c, which will be called hereafter the effective area A_{eff}. Clearly, the current that a contact can supply depends not only on its effective barrier but also on the effective area. So, it is appropriate to state that :

A metal-thin insulator-semiconductor contact can be represented by an effective barrier B_{eff} with an effective area A_{eff}. The equilibrium flux across the barrier is given by the pseudo-Richardson equation. As long as the current demand is less than "Richardson" current, the current is symmetric and limited by the "bulk" and there is no barrier effect.

In the same sense as subsection 3.2.2, there is the split temperature T_s that pseudo-Richardson equation can be applied, i.e.

$$I = A_{eff} A^* T_s^2 \exp(-B_{eff}/kT_s) \qquad (4.1.1)$$

all symbols are the same as subsection 3.2.2. The validity of this model will be seen in the text.

ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

4.2 The Criterion of Ohmic Contact

The meaning of "ohmic" by mean of pseudo-Richardson model was individually used by Tantrapom (1970,1972). Because it has two prominent points, easy access by experiment and independent of the "bulk" sample geometry. So, it should be more useful to regard this meaning as the ohmic criterion. That is :

" A contact is ohmic as long as the current demanded is less than Richardson current."

4.3 Basic Conduction Across a Small Barrier

As usual, one would formulate a theoretical model that is the most pertinent to his experiment, and obtain the model parameters from fitting. In our case, we are interested in ohmic contacts, involving a small barrier, and the attention should be on the properties of the barrier of this kind. However, before discussing the model in subsection 4.3.2, we will first consider some aspects on the fitting procedure in the following subsection.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4.3.1 Consideration on Fitting Theoretical Expression To Experimental Data

In this subsection the concept of error surfaces in multiparameter space is introduced. This consideration was mention in section 2 in the work of Tantraporn (1972). It is described again in details here, for our purpose.

One point, i.e. "physical reality test", of the original work is excluded here. It is one important point in view of Tantraporn (1972). According to his view, the investigator should fit his theoretical expression with the experimental data by not fixing the values of the parameters, a priori. The parameters derived from fitting are regarded as merely adjustable parameters, and no physical significance can be assigned to the model until a physical reality test is met. The test can be performed by comparing the value of the parameter found at fitting with its known value. For example, if one chooses the electronic charge as an unknown testing parameter, the numerical value of this parameter that emerges from fitting should be near 1.6×10^{-19} Coulombs to meet the test.

In practice, the sensitivity of the test depends on the sensitivity of the testing parameter. In Fig.4 one can see that on the minimum error surface at point Q, the uncertainty of parameter a_1 is large. So, it is the most sensitive for testing. On the other hand, if one chose a_3 as testing parameter, because of the small uncertainty of a_3 the test would not be sensitive. As can be seen in the text , all parameters contribute to the current in some exponential manner, except the area A_{app}. A slight change in another parameter causes a great change in A_{app}. So, A_{app} is the most sensitive for testing. Tantraporn used this fact to establish the sensitive physical reality test, namely "electrode area testing". The test is considered met when A_{app} is found comparable to the electrode area.

For metal-thin insulator-semiconductor contact, because A_{app} is already algebraically modified to be much smaller than contact area and it better be found

that small from the fitting, it can no longer be used as the testing parameter. So, this thesis will omit only one point from his consideration. We instead pay our attention to two pertinent parameters, i.e. the barrier height and tunneling effective mass. Their values do not vary too much, i.e. like a_3 in Fig.4, and are regarded as acceptable values or significant barrier parameters from fitting.

Let X represent a set of the experimental variables x_1, x_2, x_3... (x_1 could be the temperature, x_2 the voltage, x_3 the pressure, etc.). The experimental value of y_1 is obtained for a property y under the experimental condition X_1, For example, y may be the current through the sample. The experimental data therefore consist of N data points;

$$\text{At} \quad X_i \, , y = y_i \quad \text{for } i = 1, 2, ...N. \qquad (4.3.1.1)$$

Note that we do not use $y(X_i)$ which may imply a knowledge of the functional dependence, a priori.

Let there be a theoretical model to be tested. This model contains a set A of parameters a_1, a_2, a_3,...(a_1 may be the effective mass of the semiconductor, a_2 the Fermi level, etc.). The property y is then theoretically given as a function of X for a given set of A;

$$y_{t,i} = y_t(X_i;A) \, , i = 1, 2, ... N. \qquad (4.3.1.2)$$

where the subscript t denotes that the quantity is "theoretical". "The least squares" fit is obtained when ε, defined by

$$\varepsilon = \sum_{i=1}^{N} (y_i - y_{t,i})^2 \qquad (4.3.1.3)$$

is minimum with respect to A. In other words, one is to find the best choice of a_1, a_2 , ... so that ε is a minimum. The right hand side of eq. 4.3.1.3 may contain weight factors, should one have some subjective judgment on the accuracy of the individual experimental points.

For a given set l^{th} of A there is only one error number

$$\varepsilon_l = \varepsilon(A_l). \qquad (4.3.1.4)$$

For illustration purposes, we depict a case of three parameters (a_1,a_2,a_3) in Fig.4. A single point in parameters space represents a set A. At the lth point (a_1,a_2,a_3) one may assign an error number ε_l. A sufficient number of points can be computed in the A-space to form equi-error-surfaces such as those schematically shown in Fig.4. If ε is small, the equi-error-surfaces may approximate ellipsoids according to eq. 4.3.1.3 and a minimum -error point exists at, say, $(a_1,a_2,a_3)^3_{min}$ = Q, where the superscript 3 represents "in 3 dimensions".

Suppose the same theoretical model is to contain only two parameters by fixing, a priori, a_3 = constant. Then instead of the three-dimensional equi-error-surfaces one has a equi-error contour curves, also depicted in Fig.4. There also exists the minimum error point at

$$S = (a_1,a_2)^2_{min} \qquad (4.3.1.5)$$

where again the superscript denotes the number of dimensions. Clearly, the two-dimensional minimum error value is larger than the three dimensional one. The values of a_1 and a_2 at the two-dimensional minimum error point S are also different from those at Q, where $Q = (a_1, a_2 ,a_3)^3_{min}$.

This illustrates the danger of deriving the physical parameters using a set of other "known" parameters, when the latter in reality should be treated as unknowns.

On the equi-error-surface at minimum error point Q in Fig.4, called the "box". Among each side of the box, a_3 has minimum length, i.e. a_3 get minimum uncertainty. So, the value of a_3 from fitting is the most acceptable, while the value of a_1 is more uncertain.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

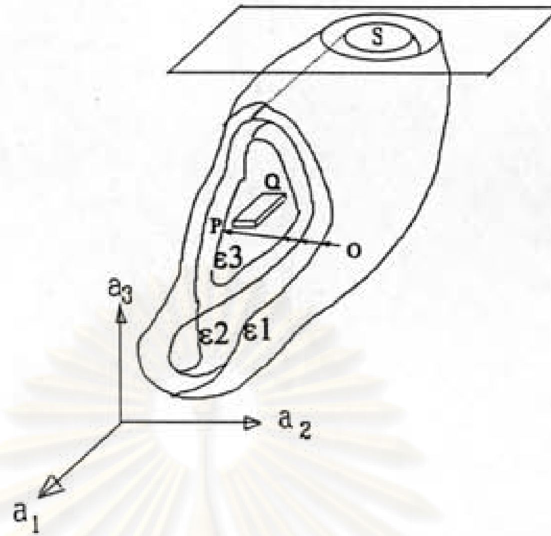

Fig.4 Sample equi-error-surfaces in a three-dimensional parameter space.

Figure 4 shows equi-error-surfaces in a three-dimensional parameter space with portions cut out to reveal the successively smaller error surfaces. If point O is selected as a beginning point, the computer program described in Appendix C will give minimum error point within a small "box" at Q (suggestively drawn in this figure to show the uncertainty of each parameters). The uncertainties of each parameter depend on the nature of the parameters themselves. The parameter a_1, may be A_{app}, carries a large uncertainty. The parameter a_3, may be m^+, does not vary too much and can be regarded as acceptable parameter. In the case of more restricted parameters, say, a_3 = constant, a two-dimensional localization of the minimum error may be found at S, illustrating the wrong values of a_1, a_2 found at S for two-parameters best fitting as compared with a_1, a_2 values found at the true minimum Q.

4.3.2 The Basic Model For a Small barrier

In line with above consideration, the theoretical model should be formulated starting as close to "first principle" as possible. Only the physical parameters that are clearly known, are to be assigned numerical values, a priori. Other parameters, whose physical existence may be questionable, or which can vary from case to case, must be kept as algebraic unknowns.

Since we are interested in p-type semiconductor, all quantities which have unit of energy are measured downward in the energy band diagram, and "barrier" in this subsection means the energy barrier profile.

This subsection base on section3 of the work of Tantraporn (1972). However, in the present work, tunneling effective mass m^+ is introduced as unknown parameter. This is because, when hole(electron) is in the forbidden gap, its effective mass (assuming the effective mass approximation still valid) would differ from that in the allowed band. So, it should be the "clouding" parameter. We assume here only one m^+ for each contact. So m^+ can reflect the influence of the heavy and light holes in tunneling. Further more, m^+ should absorb the deviation from one dimensional model as will be discussed later in the section7.5. (The author is indebted Dr. W. Tantraporn for this intersting aspect of m^+.)

Another different point, as mentioned in subsection 4.1, included in the present work is the tunneling through barrier in series with a thin layer of

insulator. However, this point is not important. It different from original work only the terminology, i.e. A_{app} is used here.

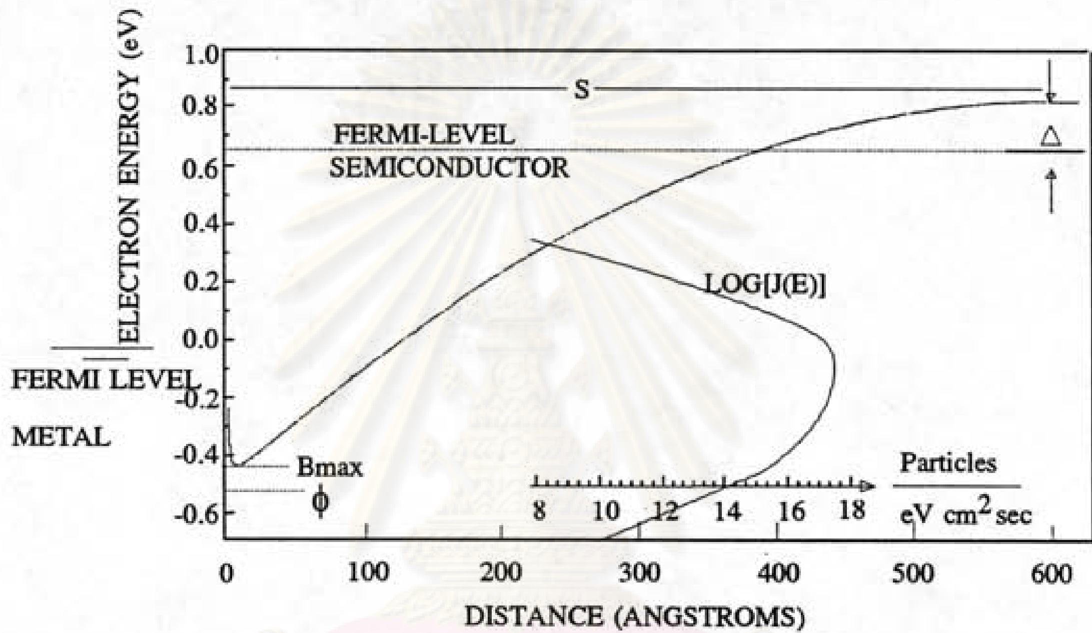

Fig.5 Definition of the various quantities used in the basic model, and quantitative plots of the barrier profile vs. distance and the current-energy distribution for the case of V = 0.624 volts, T = 232.2 K in Table 6.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Referring to Fig.5, the general expression for the hole current density J flowing from the metal one-dimensionally toward the positive x direction at a given temperature T when the semiconductor Fermi level at large x is negatively biased with respect to the metal by the voltage V is given by

$$I(V,T) = A_{cond}\, T_i\, J(V,T)$$
$$= A_{app}\, J(V,T) \qquad (4.3.2.1)$$
$$J(V,T) = \int_{-v+\Delta}^{\infty} J(E)\, dE$$

where E is the hole kinetic energy measured from the metal's Fermi level, A_{cond} the conduction area through which uniform conduction takes place, and T_i the average transmission coefficient in tunneling through the insulator barrier. F is the incident hole flux distribution and D the transmission coefficient through the Schottky barrier.

The product $A_{cond} \cdot T_i$ form the compound parameter, hereafter, called apparent area A_{app}. Setting metal Fermi level as zero, the current density is given by (see Appendix A)

$$J(V,T;E_x) = \frac{4\pi q(m_l^* + m_h^*)kT}{h^3} \int_{B\,max}^{\infty} \ln\left[\frac{1+e^{-E_x/kT}}{1+e^{-(E_x+qV)/kT}}\right] dE_x$$
$$+ \frac{4\pi q(m_l^* + m_h^*)kT}{h^3} \int_{-v+\Delta}^{B\,max} D(Ex) \ln\left[\frac{1+e^{-E_x/kT}}{1+e^{-(E_x+qV)/kT}}\right] dE_x \qquad (4.3.2.2)$$

where q, k, h are the hole charge, Boltzmann and Planck constants, respectively. m_l^* and m_h^* are the effective masses at the edge of the valence band of the light and heavy hole.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

[The numerical value of $\log(4\pi m_0 kT/h^3)$ used here is 24.94, where m_0 is electronic mass. This value is different from 26.594 of Tantraporn (1964a), and 25.94 of Tantraporn (1964b). However, this point is not serious because this numerical value is normalized by the same value of the base. So, the different result is A_{app}.]

The transmission coefficient D from WKB approximation, is given by

$$D(E) = \exp\left\{ -\frac{4\pi}{h} \int_{x_1(E)}^{x_2(E)} \sqrt{2m^+\left[B(x;V) - E\right]}\, dx \right\} \qquad (4.3.2.3)$$

Where m^+ is the tunneling effective mass, $B(x;V)$ is the energy barrier profile similar to that shown in Fig.5, and $x_1(E)$ and $x_2(E)$ are, respectively, the small and large roots of $B(x;V) - E = 0$. D is assumed unity for E greater than the maximum of the barrier.

Thus if one can write the functional dependence $B(x;V)$ then in principle one can obtain the theoretical value of I as a function of V and T from 4.3.2.1 Herein lies the heart of the study. How does one write the barrier profile $B(x;V)$ and also relate $B(x;V)$ to other physical parameters amenable to experimental study and control? This is a different approach from the conventional one in that one starts with an energy barrier profile derived from an assumed model, then from experimental-theory fitting ends up concluding whether the WKB theory is valid or not (Tantraporn, 1972). We shall also assume a model from which the

profile B(x;V) is obtained, so that I(V,T) can be calculated from 4.3.2.1 and compare with the experiment to obtain the parameters.

Consider a plausible doping profile, with an acceptor doping density as a function of the position from the insulator-semiconductor interface (x=0) given by

$$N_a(x) = N_a(b) + N_a(s) \exp(- \alpha x) \qquad (4.3.2.4)$$

Which might be the case if $N_a(b)$ is the level of the semiconductor bulk.

[As point out by Tantraporn (1972), if there are more than two species of acceptors, eq. 4.3.2.4 may be written, say, $N_a(x) = N_a(b) + N_{a1}(s) \exp(-\alpha x) + N_{a2}(s) \exp(-\beta x) + ...$ However, if within the range of interest N_{a1} is rapidly decreasing with x, then for a practical purpose $\exp(- \beta x)$,...etc. terms would be essentially constant so that $N_a(x) = \{N_a(b) + N_{a2}(s) + N_{a3}(s) +...\} + N_{a1}(s) \exp(-\beta x)$. In other words, $N_a(b)$ in eq. 4.3.2.4 needs not necessarily be the acceptor density at large x in the bulk semiconductor.]

$N_a(s) \exp(-\alpha x)$ represents a doping profile near the interface as a result of dopant diffusion from the interface or gradual departure from stoichiometry in the case of compound semiconductors.

With respect to the Fermi level, let the position of the valence band at x=0 be at ϕ (See Fig.5). That is :

$$B(0; V) = \phi \qquad (4.3.2.5)$$

[For intimate contact ϕ is independent of the applied bias, i.e. eq.4.3.2.5 above is valid (see also Section 2.6). However for the contacts of interest, the insulator layer should be so thin that the variation of ϕ can be neglected.]

The reverse applied voltage V causes the semiconductor Fermi level at x=S to be at V "below" the metal Fermi level. Let Δ be the position of the valence band "above" its Fermi level at x > S. So, for reverse bias

$$B(S; V) = -V + \Delta \qquad (4.3.2.6)$$

then the width of the depletion region S is given as the solution of

$$\phi + V - \Delta = \frac{q}{K\varepsilon_o} \int_0^S \int_x^S Na(\xi)\, d\xi\, dx \qquad (4.3.2.7)$$

Equation 4.3.2.7 implies that S is a function of ϕ , Δ, K , $N_a(b)$, $N_a(s)$, α and V; and is derived from the Poisson equation with the usual boundary condition that (dV/dx) = 0 at x = S; K is the dielectric constant that is appropriate for hole in traverse across the barrier.

The energy barrier profile for this basic model is then

±16725190

$$B(x:V) = \frac{q}{K\varepsilon_o}\int_x^s\int_\eta^s Na(\xi)\,d\xi\,d\eta - V + \Delta - \frac{q}{16\pi K\varepsilon_o x} \qquad (4.3.2.8)$$

where the effect of the image force is also included and is represented by the last term.

The theoretical value of I(V,T) based on this model then can be calculated according to eq.4.2.2.1 for each set of (V,T). The parameters included in the theoretical model are A_{app}, m^*, m^+, K, ϕ, Δ, $N_a(b)$, $N_a(s)$, and α.

The voltage V used in the barrier model above can be obtained from the DACCT measurement. In the range that the lower branch is due to the "bulk", the difference in the values of the voltage of the two polarities at a constant current is the voltage directly sustained by the barrier region, called the split voltage V_S, which is the pertinent voltage in the barrier model. For various values of the current I one may obtain a set of data I as a function of Vs and T, and can extract the above barrier's parameters from fitting procedure. This fitting procedure will be in the section 6.2.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย