METAL SEMICONDUCTOR CONTACTS

Applications: 1) Ohmic contacts to N⁺ and P⁺ regions of semiconductor devices

2) Rectifying junctions which switch faster than PN junctions since majority carrier phenomena dominate

**Basic Properties – Band Diagrams**

**Metal**

- \( E_F \) in conduction band
- lots of free \( e^- \)

**N Type Semiconductor**

- \( E_F \) in forbidden band
- Valence band \( \approx \) full
- Conduction band \( \approx \) empty

\[ E_o \Delta \text{ reference energy that an } e^- \text{ just "free" of the material would have (in a vacuum)} \]

\[ \phi \Delta \text{ work function } \Delta E_o - E_F \therefore \phi_s \text{ depends on doping level in semiconductor} \]

= energy required to bring an \( e^- \) from the Fermi level to the vacuum level

\[ \chi \Delta \text{ electron affinity } \Delta E_o - E_c \therefore \chi \text{ is a property of the material and does not depend on doping} \]

= energy required to bring an \( e^- \) from the conductor band edge to the vacuum level
\[ \phi_{\text{MS}} = \phi_M - \phi_S = \text{metal semiconductor work function difference} \]

Note that \( \phi_{\text{MS}} \neq 0 \) in general \( \therefore \) e\(^-\) are in general not at the same average energy in the metal and in the semiconductor. \( \phi_{\text{MS}} \) is the difference in Fermi levels between the two materials.

When the two materials are joined, e\(^-\) will flow from the semiconductor to the metal (if \( \phi_M > \phi_S \)) to establish equilibrium or \( E_F = \text{CONSTANT} \)

In terms of the band diagram, the result will be

Note:

1) e\(^-\) in the metal must cross a barrier whose height is \( q\phi_B = q(\phi_M - \chi_S) \) to get into the semiconductor.

\[ \therefore \phi_B \triangleq \text{barrier height} \]
2) $e^-$ in the semiconductor must cross a barrier whose height is $q\phi_i = q(\phi_M - \phi_s)$ to get into the metal.

$$\phi_i = \phi_B - (E_c - E_F)$$

3) The region near the surface will be depleted of carriers in the semiconductor. ($E_F$ moves further away from the CB). $\therefore$ There will be a built-in $\mathbf{E}$ field in the semiconductor near the surface.

To proceed, we make a number of assumptions:

a) Neglect the free hole concentrations

b) $n \equiv 0$ in a region from $X = 0$ to $X = X_d$

c) Beyond $X_d$, $n = N_D$

\[ Q = q N_D X_d A \]

$Q$ is total charge in depletion region

$Q$ is composed of ionized donor atoms

Sheet of charge assumed at metal surface (perfect conductor)
The voltage across the depletion region is governed by Poisson's equation

\[
\frac{d^2V}{dx^2} = \frac{q}{K\varepsilon_0} \left[ (n-p) - (N_D - N_A) \right]
\]

\(\equiv -\frac{q}{K\varepsilon_0} N_D\)  

Integrating twice from \(X_d\) to \(X = 0\), we obtain

\[V_{\text{max}} = \frac{1}{2} \frac{qN_D}{K\varepsilon_0} X_d^2\]  \(\text{(2)}\)

Since \(\varepsilon = -\frac{dV}{dx}\), we have directly that

\[\varepsilon_{\text{max}} = -\frac{qN_D}{K\varepsilon_0} X_d\]  \(\text{(3)}\)

Thus we have

\(\Phi_i\) is called the built in voltage

\[\Phi_i = \frac{1}{2} \frac{qN_D}{K\varepsilon_0} X_d^2 = \Phi_M - \Phi_s\]

\[X_d = \sqrt{\frac{2K\varepsilon_0}{qN_D}} \Phi_i\]

= width of depletion layer

\[Q = A\sqrt{2q\varepsilon_0 KN_D\Phi_i}\]

= charge in depletion layer
Suppose we apply an external potential to the semiconductor. Since the impedance of the depletion region is much greater than the metal or the bulk semiconductor \((n \equiv p \equiv 0\) in depletion region), most of the voltage appears across the depletion region.

**Equilibrium**

\[
\phi_i = \phi_M - \phi_s = \frac{1}{2} \frac{qN_D}{K \varepsilon_0} \chi_d
\]

\[
\phi_B = \phi_M - \chi_s = \text{barrier to } e^- \text{ flow } M \rightarrow S.
\]

**Forward Bias**

\[
e^- \text{ flow } S \rightarrow M \text{ now has a barrier of } \phi_i - V_A
\]

\[
e^- \text{ flow } M \rightarrow S \text{ has same barrier of } \phi_B \text{ (no voltage can be sustained across metal).}
\]

Since number of \(e^-\) with \(E > \phi_i - V_A\) is exponentially related to \(V_A\) (F.D. probability function) \(\therefore \) I ought to increase exponentially with \(V_A\).

**Reverse Bias**

\[
e^- \text{ flow } M \rightarrow S \text{ still has same barrier}
\]

\[
e^- \text{ flow } S \rightarrow M \text{ now has higher barrier } \therefore \text{ essentially no current flow.}
\]

Note that \(E_F\) is not drawn in the depletion region since this is a nonequilibrium condition.
Under reverse bias, the depletion region will expand since it supports essentially the entire voltage.

\[ X_d = \sqrt{\frac{2K\varepsilon_o}{qN_D}} (\phi_i - V_A) \]  \hspace{1cm} (4)

A retifying diode of the type we have been describing is often called a Schottky diode after W. Schottky who first developed its theory in 1938.

\[
\begin{align*}
\text{Signs:} & \quad \text{Reverse bias} \quad \phi_i, V_A \quad \text{add} \\
\text{Forward bias} & \quad \phi_i, V_A \quad \text{subtract}
\end{align*}
\]

The barrier that exists across a Schottky diode is actually slightly smaller than the idealized treatment given above because of "image force lowering".

An e\(^{-}\) outside a metal induces a a + image charge in the metal

\[ F = -\frac{q^2}{16\pi \varepsilon_o X^2} \]  \hspace{1cm} (5)

where \( X \) is the distance of the e\(^{-}\) from the metal.
When an external field is applied, the total potential energy as a function of $X$ is:

$$PE(X) = \frac{q^2}{16\pi \varepsilon_0} X + q\varepsilon_x$$ (6)

where the first term is obtained by integrating (5).

The result is a reduced barrier under reverse bias and a slightly increased barrier under forward bias. Typically $\Delta\phi_f \approx 25 - 50\text{mV}$ which is small, but it depends exponentially on $\phi_F$.

**Capacitance**

In the depletion region,

$$Q = A \sqrt{2qK\varepsilon_0 N_D(\phi_i - V_A)}$$

$$\therefore C = \frac{dQ}{dV} = A \left[ \frac{qK\varepsilon_0 N_D}{2(\phi_i - V_A)} \right]^{1/2}$$ (7)
Example: From the data in the figure, for W–Si:

\[
1/C^2 = 6 \times 10^{15} \text{ at } V_R = 1 \text{ volt}
\]

\[
1/C^2 = 10.6 \times 10^{15} \text{ at } V_R = 2 \text{ volts}
\]

\[
1/C^2 = \frac{2}{qK\varepsilon_0 N_D A^2} (\phi_i - V_A)
\]

slope = \[
\frac{\Delta 1/C^2}{\Delta (\phi_i - V_A)} = \frac{2}{qK\varepsilon_0 N_D A^2}
\]

\[
N_D = \frac{2}{qK\varepsilon_0 A^2} \frac{d(\phi_i - V_A)}{d(1/c^2)} \quad (A = 1)
\]

\[
= \frac{2}{(1.6 \times 10^{-19}) (11.8) (8.84 \times 10^{-14})} \frac{IV}{4.6 \times 10^{15}}
\]

\[
N_D = 2.6 \times 10^{15}/\text{cm}^3
\]
From the intercept of the curve, \( \phi_i \approx 0.4 \) volts.

The position of the Fermi level in the bulk of the N silicon is given by

\[
E_F = E_c - kT \ln \frac{N_c}{N_D} \approx 0.24 \text{ volts below } E_c
\]

Thus we have:

\[
\begin{align*}
&0.64\text{eV} \quad \uparrow 0.4\text{eV} \quad \downarrow 0.24\text{eV} \\
\text{Metal} &- \text{Semiconductor}
\end{align*}
\]

Thus with no applied bias, the barrier to \( e^- \) flow from S \( \rightarrow \) M \( \approx 0.4\text{eV} \) and from M \( \rightarrow \) S \( \approx 0.64\text{eV} \).

**Practical Considerations – Surface States**

The example described above is an ideal metal-semiconductor diode. In actual practice, surface states often dominate the observed characteristics.

\[
\begin{align*}
\text{Si} & - \text{Si} - \text{Si} - \text{Si} & \text{dangling bonds at the surface create trap levels at the interface (same as G-R centers we considered earlier)}. \\
\text{Si} & - \text{Si} - \text{Si} - \text{Si}
\end{align*}
\]

\[ \therefore E_F \] may be altered from the ideal picture we previously used.

\[
\begin{align*}
E_t & \quad \text{Suppose a donor like state exists at } E_t \text{ due to these dangling bonds}
\end{align*}
\]
If the Fermi level is below $e^-$ as shown, these states will have donated $e^-$ and will be positively charged. $\therefore$ These states end up contributing some of the $e^-$ and $+$ space charge needed for the structure to come to equilibrium

$\therefore \phi_i$ is effectively reduced

$X_d$ is effectively reduced

If the density of traps at $E_t$ is large, essentially all of the $e^-$ and fixed $+$ charge needed to reach equilibrium can be provided by these states. $\therefore E_F$ is "pinned" near $E_t$ (since small variations of $E_F$ around $E_t$ greatly change $e^-$ and $+$ concentrations).

$\therefore \phi_i \neq \phi_M - \phi_S$ and the built in potential is determined not only by material properties (work function), but also by practical technological questions.

Experimentally it is found that surface states often show a pronounced peak in density at

$E_t \approx 1/3 E_g, \quad E_t \equiv E_v + 1/3 E_g$

Thus we have:

$\phi_B \equiv \frac{2}{3} E_g \quad \phi_i \equiv \frac{2}{3} E_g - (E_c - E_F)$

(8)  (9)
In practice, because of the great difficulty in predicting $E_t$ theoretically (depends on technology), $\phi_i$ and $\phi_B$ are usually measured experimentally.

Note that the measured barrier heights do not change as much as $\phi_M$.

### Current Voltage Characteristics

Under applied bias we can change $F_{S \rightarrow M}$ by changing the barrier height $\phi_i - V_A$. We cannot change $F_{M \rightarrow S}$ since no electric field can be supported in a perfect conductor.
\[ F_{S \rightarrow M} = D_n \frac{dn}{dx} + n \mu E \]

By applying a potential externally, we change \( E \). Under forward bias, the drift term is decreased, allowing diffusion to produce a net flux or current.

The numbers of \( e^- \) with sufficient energy to surmount the respective barriers \( \phi_i \) and \( \phi_B \) are given by F.D. statistics (or the M.B. approximation)

\[
n = N_c e^{-\frac{(E_c - E_F)}{kT}} = \text{number } e^- \text{ in C.B.}
\]

But if we are interested in the number of electrons with \( E > E_c + \phi_i \), then

\[
n_s = N_c e^{-q\phi_B/kT} = \text{number of } e^- \text{ with } E > E_c + \phi_i
\]

But \( n = N_D = N_c e^{-\frac{(E_c - E_F)}{kT}} \) and \( \phi_B = \phi_i + (E_c - E_F) \)

\[
\therefore n_s = N_D e^{-\frac{q\phi_i}{kT}}
\]

= number of \( e^- \) getting over barrier \( S \rightarrow M \).

(This expression could really have been written by inspection since \( n = N_D = \text{number of } e^- \text{ in C.B. and the exponential comes directly from M.B. statistics.})

Note that the number of \( e^- \) in the metal \( \gg N_D \therefore \text{ barrier must be higher than } \phi_i \text{ (i.e. } \phi_B > \phi_i \text{) to get the same flux in both directions.}
\[ \therefore I_{S \rightarrow M} = I_{M \rightarrow S} = K N_D e^{-\frac{q \phi_i}{kT}} \]

Under bias, \( \phi_i \rightarrow \phi_i - V_A \)

\[ \therefore n_s = N_D e^{-\frac{q \phi_i - V_A}{kT}} \]

Since the current flowing \( M \rightarrow S \) is unchanged by \( V_A \), we have:

\[ \text{Net } I = I_{S \rightarrow M} - I_{M \rightarrow S} \]
\[ = K N_D e^{-\frac{(q \phi_i - V_A)}{kT}} - K N_D e^{-\frac{q \phi_i}{kT}} \]

(11)

or \[ I = I_o \left[ e^\left(\frac{q V_A}{kT}\right) - 1 \right] \] Ideal Diode Equation

(12)

where \[ I_o = K N_D e^{-\frac{q \phi_i}{kT}} \]

(13)

\( I_o = \text{reverse saturation current} \)

Forward bias: \( I \uparrow \) exponentially with \( V_A \)

Reverse bias: \( I \cong I_o \) (constant)

Note that \( V_A \) under forward bias is constrained to be \( < \phi_i \). In practice, series resistance of the semiconductor resistance limits \( I \) for \( V_A \approx \phi_i \)
More accurate analysis shows that $I_o$ is not independent of voltage but rather is given by

$$I_o = \frac{Aq^2 D_n N_c}{kT} \left[ \frac{2q(\phi_i - V_A)N_D}{K\varepsilon_0} \right]^{1/2} e^{-\frac{q\phi_B}{kT}} \quad (14)$$

This is often incorporated into the diode equation as

$$I = I_o \left[ e^{\frac{qV}{nkT}} - 1 \right] \quad (15)$$

Where $I_o$ is independent of $V_A$ and the non ideal effects are incorporated into $n$. $n$ is usually 1.0 to 1.2.

a) $I_o$ is obtained by extrapolating to $V_A = 0$

b) $\phi_i$ and $\phi_B$ can then be obtained from (14)

c) $n$ is obtained from the slope

In this example,

- $n = 1.02$ for Si diode
- $n = 1.04$ for GaAs diode

Typically,

$$X_d \approx 0.1 - 1.0\mu$$

$$I_o \approx 10^{-8} - 10^{-4} \frac{A}{cm^2} \text{ (off)}$$

$$I \approx 10^{-2} - 1 \frac{A}{cm^2} \text{ (on)}$$
**Ohmic Contacts**

An ohmic contact is defined as one in which there is an unimpeded transfer of majority carriers from one material to another i.e. the contacts do not limit the current.

There are two basic ways to achieve such a contact:

a) Choose a system where \( \phi_M < \phi_S \)

b) Dope the semiconductor heavily enough that tunneling is possible

Consider a) first:

\[
q\phi_M < \phi_S
\]
Note:

1) If $\phi_M < \phi_S$, then equilibrium is established by $e^-$ flow $M \rightarrow S$.

2) Most of the common metals do not exhibit this behavior on N type silicon. From a practical point of view, this does not happen ($\phi_M < \phi_S$) typically because surface traps $E_t$ pin the Fermi level. While it is theoretically possible, metal to N$^-$ contacts are not made this way if they are to be ohmic.

Consider a metal – P semiconductor contact.

In general, if majority carriers are enhanced near the surface, then the contact will be ohmic.

From a practical point of view considering surface states, as well as "normal" values of $\phi_M$ and $\phi_S$,

1) Metal to N$^-$ contacts are nearly always retifying in silicon

2) Metal to P$^-$ silicon contacts are usually ohmic
.: For contact to \( N^- \) device regions and for reduced resistance to contact \( P^- \) regions, it is usual to heavily dope the Si regions \( N^+ \) or \( P^+ \) so that a different phenomenon insures ohmic contacts.

Suppose \( N_D \) or \( N_A \) in the semiconductor is very large. Then

\[
X_d = \sqrt{\frac{2K\varepsilon_0}{qN_D}} \phi_i \quad \text{becomes very small}
\]

Thus we have

\[\begin{array}{c}
\text{M} \\
\downarrow \\
X_d \\
\downarrow \\
\text{S} \\
\end{array}\]

When \( X_d \approx 25 \text{ to } 50 \text{Å} \), e\(^-\) can "tunnel" through the barrier.

This process occurs in both directions \( M \rightarrow S \) and \( S \rightarrow M \) so the contact shows very little resistance and becomes ohmic.

To calculate an approximate value for the required doping we have:

\[
N_{D_{\text{min}}} \equiv \frac{2K\varepsilon_0 \phi_i}{qX_d^2} 
\]

(16)

Let \( X_d = 25 \text{Å} = 2.5 \times 10^{-7} \text{ cm} \) for tunneling to occur efficiently.

\[
\therefore N_{D_{\text{min}}} \approx \frac{(2) (11.8) (8.84 \times 10^{-14}) (0.3)}{(1.6 \times 10^{-19}) (2.5 \times 10^{-7})^2} 
\]

\[
\approx 6.2 \times 10^{19}/\text{cm}^3
\]
This is a relatively easy value to achieve in practice and this is normally how ohmic contacts are made to $N^-$ regions in integrated circuits.

**Example**

Collector contacts in bipolar devices are made by using the $N^+$ emitter diffusion to contact the $N^-$ region.

**Maximum Voltage Capability**

As we increase the reverse bias on a Schottky diode,

$$
\chi_d = \sqrt{\frac{2K\varepsilon_0}{qN_D}} (\phi_i - V_A)
$$

increases

$$
\varepsilon_{\text{MAX}} = -\frac{qN_D}{K\varepsilon_0} \chi_d
$$

increases

Eventually an electric field high enough to cause avalanche breakdown will occur and the diode will begin to conduct large currents in the reverse direction. We shall consider this process in detail when we talk about PN junctions.
Practical Realization & Applications

Schottky diodes have found widespread application in NPN structures in bipolar ICs. They are often used as fast switching diodes and as clamping diodes to prevent saturation of NPN bipolar transistors.

Depletion layer crowding at the surface in the simplest structure often drastically reduces the breakdown.

Extending the metal over the oxide or using p+ guard rings can substantially improve B.V.
Compared with PN junctions, Schottky diodes have

1) Lower forward voltage
2) Higher reverse currents i.e.
   \[ I = I_0 \left( e^{\frac{qV_A}{kT}} - 1 \right) \]
   \( \text{bigger} \quad \therefore \text{I larger for a given } V_A \)

3) No minority carrier charge storage effects \( \therefore \) faster
4) More sensitivity to technology (surface preparation critical)
5) Lower reverse BV

---

Schottky clamped transistor:
When the NPN transistor is saturated, the C-B junction is forward biased by \( \approx 0.5 \) volts. This results in excess minority carrier injection and charge storage and \( \therefore \) slow switching.

Since Schottky diodes have lower forward voltage drops than PN diodes, placing one in \( \parallel \) with the C-B diode \( \Rightarrow \) Schottky carries most of the I under forward bias; \( \therefore \) majority carriers \( \therefore \) faster.
Schottky barrier transistor or MESFET.

This device is a form of JFET in which the current flowing S → D is controlled by changing the gate voltage. This increases or decreases the depletion layer width. This is the dominant device used in GaAs today.