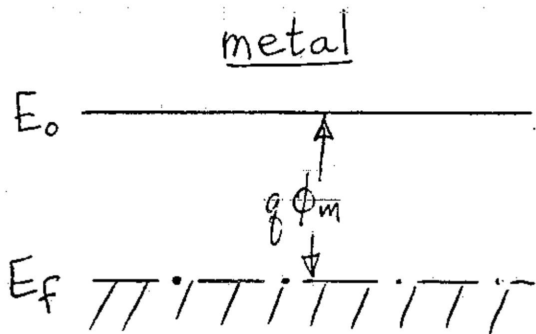
Metal-Semiconductors Contacts

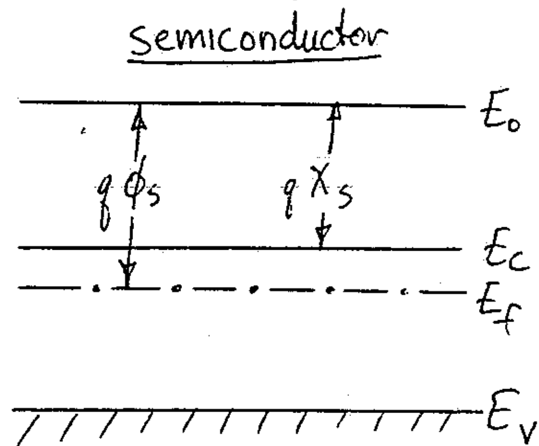
Applications :

1. Ohmic contacts to n-type and p-type regions of semiconductor devices.

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

Basic Properties - Band Diagrams





$E_0 \equiv$ reference energy that an electron just "free" of the material would have (in a vacuum)

$\phi_s \equiv$ work function

$\equiv (E_0 - E_f)/q$, therefore ϕ_s depends on doping level in semiconductor

$\equiv$ energy required to bring an electron from the Fermi level to the vacuum level

$\chi \equiv$ electron affinity

$\equiv (E_0 - E_c)/q$, therefore χ is a property of the material and does not depend on doping

$\equiv$ energy required to bring an electron from the conduction band to the vacuum level

$\phi_{ms} \equiv \phi_m - \phi_s =$ metal semiconductor work function difference.

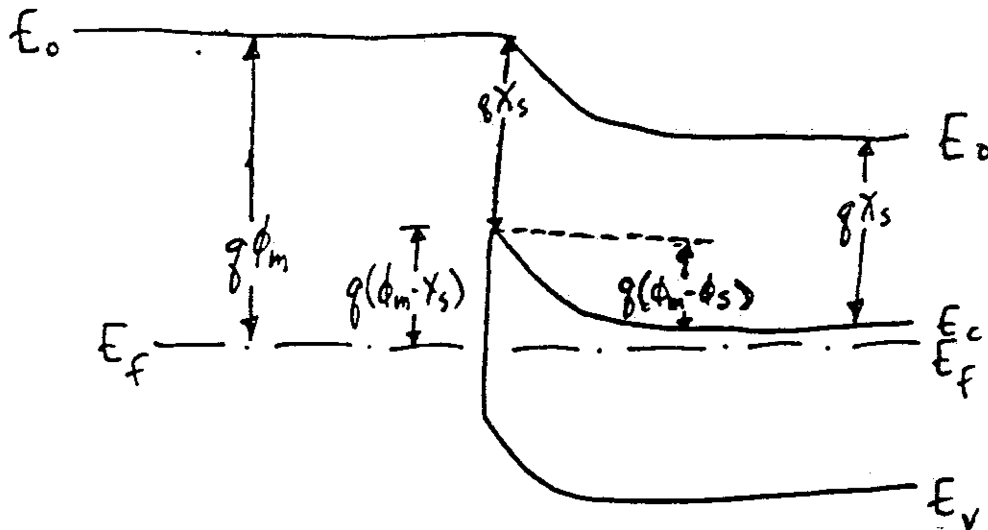
In general $\phi_{ms} \neq 0$. In other words, the average energy of an electron added to the metal is not the same as the average energy of an electron added to the semiconductor. ϕ_{ms} is just the difference in Fermi level between the two materials.

If $\phi_m > \phi_s$, the total energy of the metal/semiconductor system could be reduced by moving electrons from the semiconductor to the metal. When the two materials are brought into contact, therefore, electrons flow from the semiconductor to the metal in order to establish equilibrium and make $\mathbf{E_f} = \mathbf{constant}$.

In addition,

- E_0 must be continuous since the energy in the vacuum cannot change abruptly
- χ_s and ϕ_m are constants since they are properties of the materials.

Using these guidelines, the band diagram can be drawn for a metal in contact with a n-type semiconductor where $\phi_m > \phi_s$.



Several observations can be made about this system.

1. Electrons in the metal must pass over a potential barrier of height $q\phi_B = q(\phi_m - \chi_s)$ in order to get into the semiconductor. We will define

$$q\phi_B \equiv q\phi_m - q\chi_s \equiv \text{barrier height} \tag{1}$$

2. Electrons in the semiconductor must pass over a potential barrier of height $q\phi_i = q(\phi_m - \phi_s)$ in order to get into the semiconductor. We will define

$$q\phi_i \equiv q\phi_m - q\phi_s = q\phi_B - (E_c - E_f)_{\text{bulk}} \tag{2}$$

3. The semiconductor will be depleted of electrons near the interface. E_f is further away from E_c. Ionized donors remain so the space charge in that region will result in a built-in electric field near the interface.

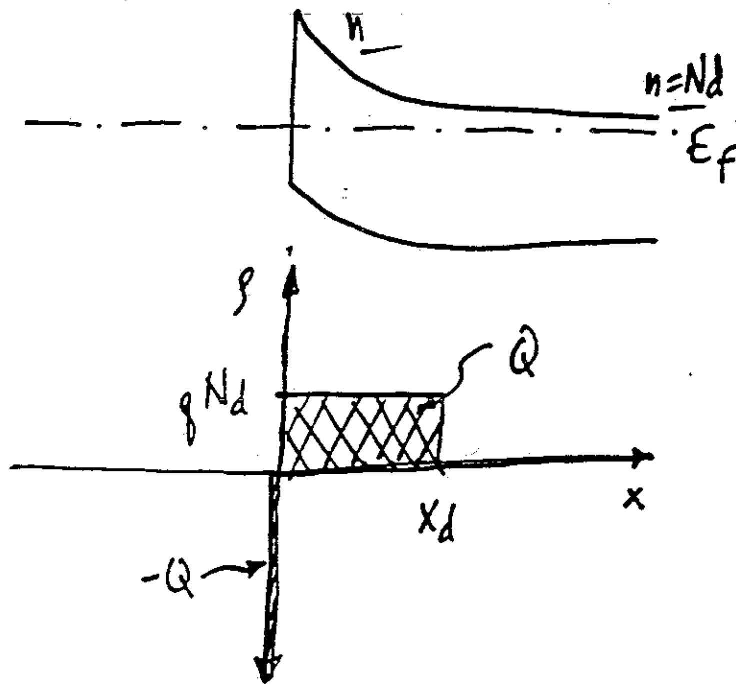
4. The built-in electric field in the semiconductor causes electrons to move away from the interface. No electric field exists in the metal since we assume it is a perfect conductor

5. Electrons which are near the interface are at a higher average energy than those in the bulk due to the electric field and therefore ϕ_i.

In order to further analyze the system, we make a set of assumptions known as the Depletion Approximation:

1. Neglect minority carrier concentration.
2. Assume that the majority carrier concentration is neglible in "depletion region" from $x = 0$ to $x = x_d$.
3. Assume that for $x > x_d$ the majority carrier concentration equals the doping concentration.

In other words, since carrier concentration varies exponentially with the difference between the Fermi level and conduction band, we will assume an abrupt boundary between the region where $n \cong N_d$ and where $n \ll N_d$.



$$Q = qN_dx_dA = \text{total charge in depletion region} \tag{3}$$

A sheet of charge is assumed to lie at the metal surface since it is considered to be a perfect conductor.

The voltage across the depletion region can be calculated using Poisson's equation:

$$\frac{d^2V}{dx^2} = \frac{q}{K\epsilon_0}\left[(n-p)-(N_d-N_a)\right] \tag{4}$$

Applying the Depletion Approximation,

$$\frac{d^2V}{dx^2} = \begin{cases} -\frac{q}{K\epsilon_0}N_d, & 0 < x < x_d \\ \frac{q}{K\epsilon_0}N_d x_d \delta(x), & x = 0 \\ 0, & \text{elsewhere} \end{cases} \tag{5}$$

We can integrate once to get the electric field ($\mathcal{E} = -dV/dx$, $\mathcal{E}(\infty) = \mathcal{E}(-\infty) = 0$).

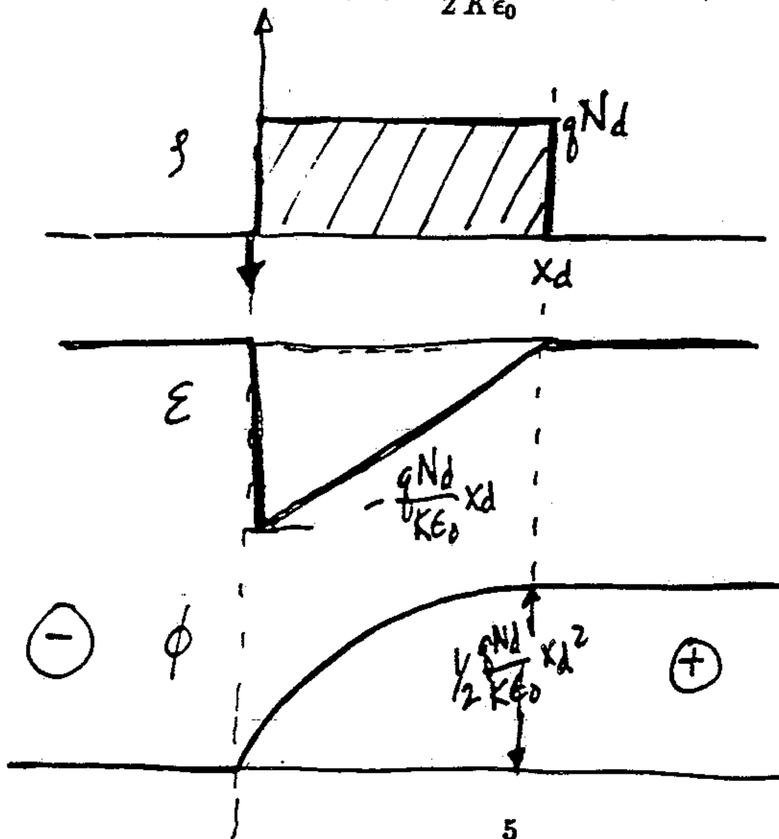
$$\mathcal{E}(x) = \frac{qN_d}{K\epsilon_0}(x - x_d), \quad 0 < x < x_d \tag{6}$$

$$\mathcal{E}_{\max} = -\frac{qN_d}{K\epsilon_0}x_d \tag{7}$$

Integrating again,

$$V(x) = \frac{1}{2}\frac{qN_d}{K\epsilon_0}\left[{x_d}^2 - (x - x_d)^2\right], \quad 0 < x < x_d \tag{8}$$

$$V(x_d) = \frac{1}{2}\frac{qN_d}{K\epsilon_0}{x_d}^2 = \phi_{ms} = \phi_i \tag{9}$$



The built-in voltage ϕ_i is just the voltage across the junction.

$$\phi_i = \phi_m - \phi_s = \frac{1}{2}\frac{qN_d}{K\epsilon_0}{x_d}^2 \tag{10}$$

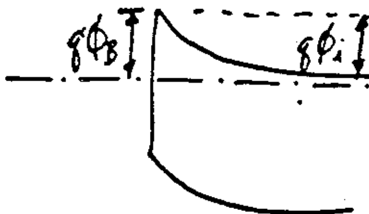
The width of the depletion layer and the total charge in depletion layer can therefore be expressed in terms of ϕ_i.

$$x_d = \sqrt{\frac{2K\epsilon_0\phi_i}{qN_d}} \tag{11}$$

$$Q = qAN_dx_d = A\sqrt{2qK\epsilon_0N_d\phi_i} \tag{12}$$

We will now apply an external voltage to the junction. Since the resistance of the metal and semiconductor bulk are much smaller than the depletion region ($n, p \cong 0$ in depletion region), almost all of the applied voltage will be dropped across the depletion region.

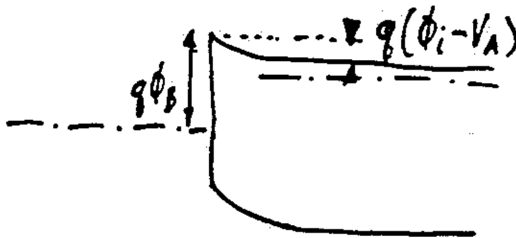
Equilibrium:



$q\phi_i = q\phi_m - q\phi_s =$ barrier for electrons S→M

$q\phi_B = q\phi_m - q\chi_s =$ barrier for electrons M→S

Forward Bias:



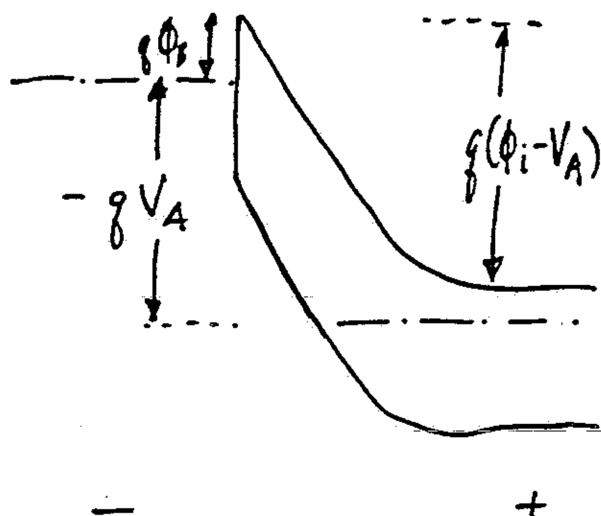
- Barrier to electron flow S→M reduced to $q(\phi_i - V_A)$
- Barrier to electron flow M→S unchanged since no voltage is dropped across metal.
- Number of electrons with energy greater than a barrier energy is expontially related to the barrier height.

$$I(\text{M}\rightarrow\text{S}) \propto \exp\left(-\frac{q\phi_B}{kT}\right) \quad \text{(unchanged)}$$

$$I(\text{S}\rightarrow\text{M}) \propto \exp\left(-\frac{q(\phi_i - V_A)}{kT}\right) \propto \exp\frac{qV_A}{kT}$$

$$I_{\text{net}} \propto \left(\exp\frac{qV_A}{kT} - 1\right)$$

Current increases exponentially with applied voltage.



Reverse Bias:

- Barrier to electron flow S→M increased to to $\phi_i + V_R$, current S→M becomes very small
- Barrier to electron flow M→S unchanged, current M→S unchanged

Note that the Fermi level is not drawn in depletion region, since with $V_A \neq 0$ we have nonequilibrium conditions. Outside the depletion region, carrier concentrations are nearly unchanged so very near to equilibrium.

Since the voltage supported by the depletion region increases with applied reverse bias, the depletion region will expand. In forward bias, the depletion region shrinks since less voltage is dropped across it.

$$x_d = \sqrt{\frac{2K\epsilon_0}{qN_d}(\phi_i - V_A)} \tag{13}$$

A rectifying metal-semiconductor contact which operates as we have described is often called a Schottky contact or Schottky diode.

A forward bias is defined to be a voltage that reduces the barrier for the majority carrier. Therefore,

- In reverse bias, ϕ, V_A add
- In forward bias, ϕ, V_A subtract

Capacitance

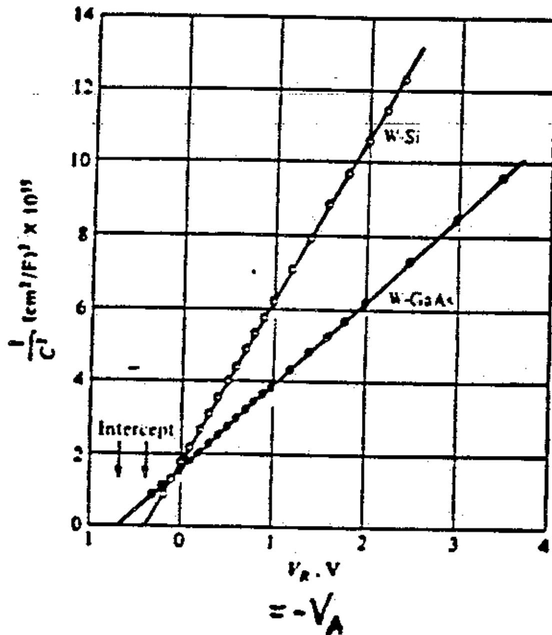
The charge in the depletion region is

$$Q_s = qAN_d x_d = A\sqrt{2qK\epsilon_0 N_d(\phi_i - V_A)}$$

Capacitance is just the change of charge with voltage.

$$C \equiv \frac{dQ_s}{dV_A} = A\sqrt{\frac{qK\epsilon_0 N_d}{2(\phi_i - V_A)}} \tag{14}$$

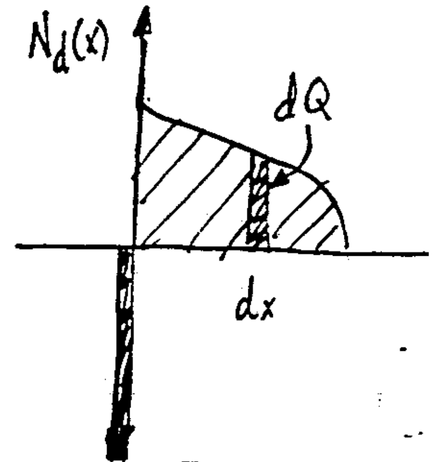
We can plot $1/C^2$ vs. V_A:



intercept $\Rightarrow \phi_i$

slope $\Rightarrow N_d$

If $N_d = f(x)$, then measurement of the incremental capacitance can be used to calculate $N_d(x)$.



$$Q_s = qA\int_0^{x_d} N(x)\,dx$$

$$dQ_s = qN(x_d)\,dx = -C\,dV_a$$

$$N(x_d) = -\frac{C}{q}\left(\frac{dx}{dV_a}\right)$$

but $C = \frac{\epsilon_s}{x_d}$, so

$$\frac{dx_d}{dV_a} = \left(\frac{dx_d}{dC}\right)\left(\frac{dC}{dV_a}\right) = -\frac{\epsilon_s}{C^2}\left(\frac{dC}{dV_a}\right)$$

$$N(x_d) = \frac{C^3}{q\epsilon_s\,(dC/dV_a)} \Rightarrow \frac{d(1/C^2)}{\ldots} = \ldots 2\,dC \ldots, \text{ so } N(x_d) = \ldots$$

Schottky Barrier Lowering

Barrier for electrons is actually altered slightly by the applied voltage due to image charges in the metal.

Assume the metal is a planar conducting sheet and there is a constant field in the region near the interface ($x \ll x_d$).

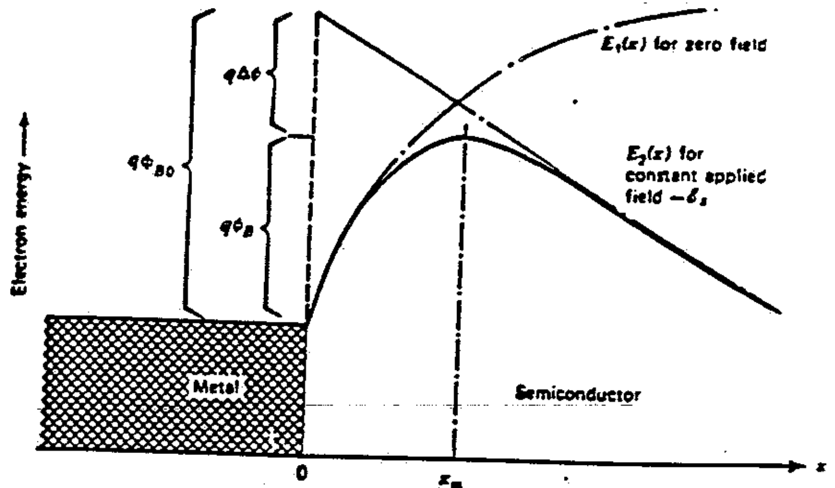
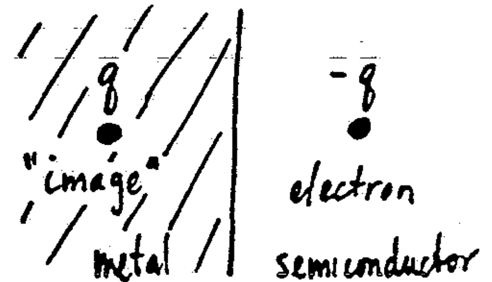


Figure 2.9 Classical energy diagram for a free electron near a plane metal surface at the equilibrium $[E_1(x)]$, and with an applied field $-\mathcal{E}_s[E_2(x)]$

The potential energy of an electron near a conducting plane is given by

$$E_1 = -\frac{q^2}{16\pi\epsilon_s x}$$



With the addition of an electric field,

$$E_2 = -\frac{q^2}{16\pi\epsilon_s x} - q\mathcal{E}x$$

We can solve this equation to find where the potential is maximum and the amount of barrier lowering.

$$q\Delta\phi = \sqrt{\frac{q^3\mathcal{E}}{4\pi\epsilon_s}}$$

The barrier is reduced $\Delta\phi$ with the barrier lowering increased for $\mathcal{E}$ large which occurs for large reverse biases.

$$\phi_B = \phi_{B0} - \Delta\phi$$

The current from the metal can be expressed as

$$J(M \to S) = J_0 \exp\left(\frac{q\Delta\phi}{kT}\right)$$

and

$$\mathcal{E} = \sqrt{\frac{2qN_d}{\epsilon_s}(\phi_i - V_A)}$$

so

$$J \propto \exp\left(K\sqrt[4]{\phi_i - V_A}\right)$$

Ohmic Contacts

Ohmic contacts do not limit the flow of majority carriers between the two materials. Current is limited by the bulk regions rather than the junction.

There are two basic ways in which ohmic contacts can be formed:

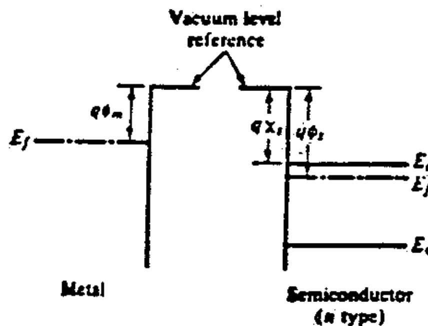
1. Choose a system with $\phi_m < \phi_s$ with an n-type semiconductor or $\phi_m > \phi_s$ with a p-type semiconductor.

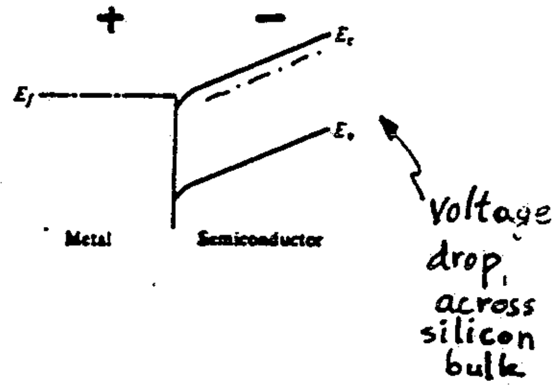
2. Dope the semiconductor heavily enough to allow tunnelling.

First consider a contact to an n-type semiconductor with $\phi_m < \phi_s$:

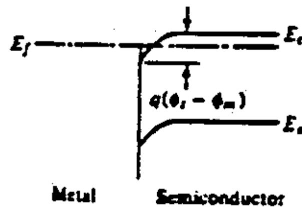
To reach equilibrium, electrons flow M→S. Electrons in the semiconductor accumulate near the junction.

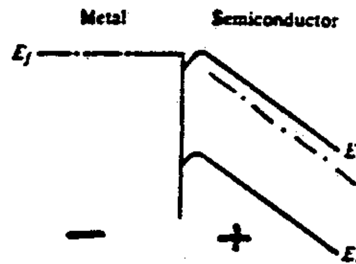
Before contact:



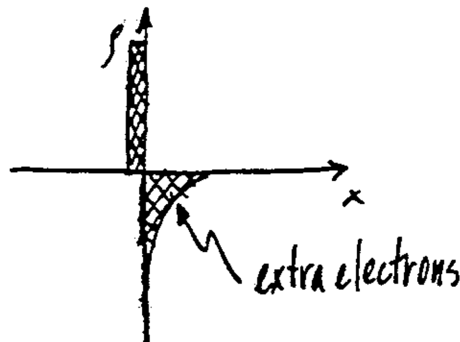


After contact:



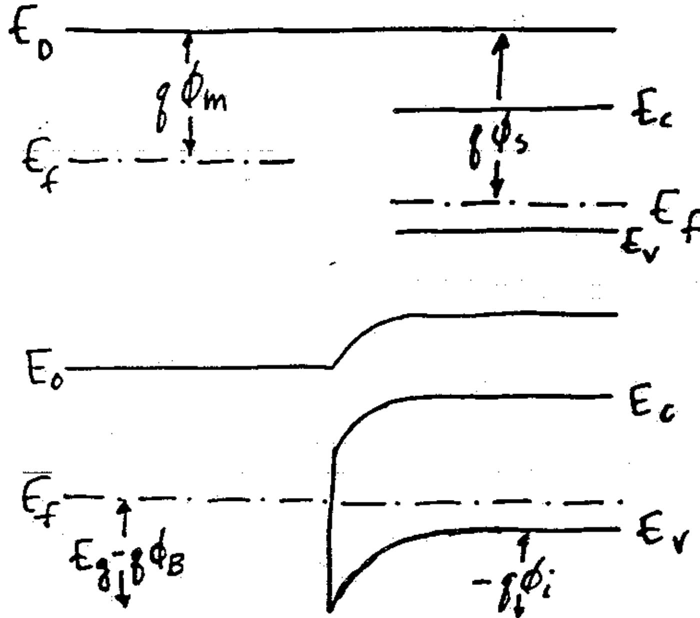


Charge density:

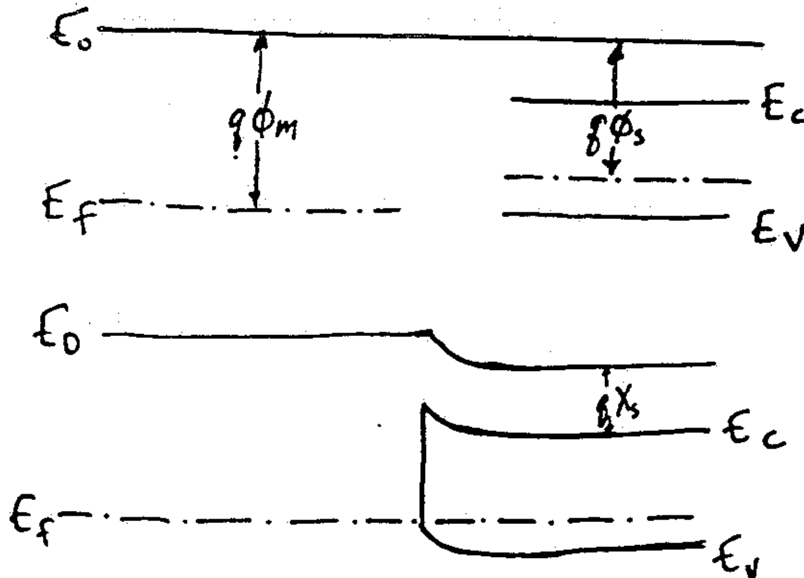


Consider also a metal contact to a p-type semiconductor.

If $\phi_m < \phi_s$, the energy of the system can be reduced by holes flowing from the semiconductor to the metal, causing a depletion region. The contact is rectifying.



If $\phi_m > \phi_s$, the energy of the system can be reduced by holes flowing from the metal to the semiconductor, causing accumulation near the junction. The contact is ohmic.



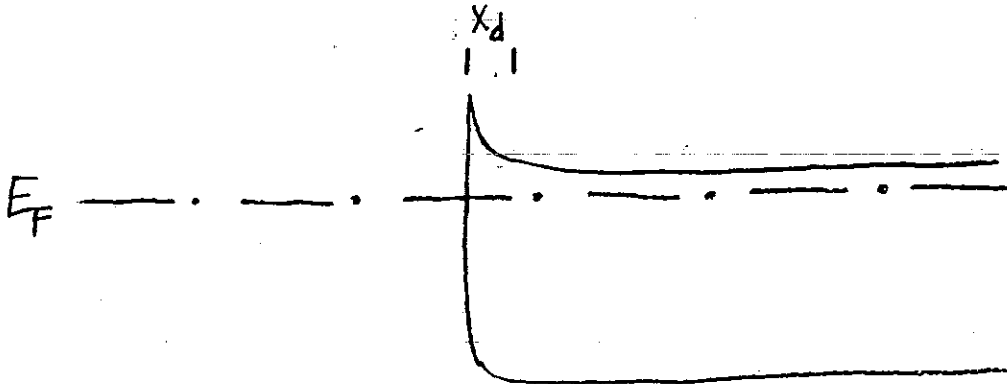
If majority carriers are enhanced near the surface, the contact will be ohmic.

Suppose N_d or N_a is very large. Then

$$x_d = \sqrt{\frac{2K\epsilon_0\phi_i}{qN_d}} \tag{15}$$

becomes very small.

When $x_d \leq$ 25-50 Å, electrons can tunnel through the barrier.



The tunnelling process can occur readily in both directions so the contact has little resistance and thus behaves as an ohmic contact.

We can roughly calculate the required doping:

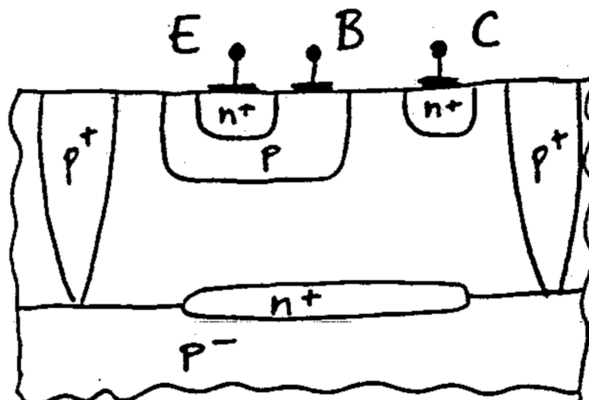
$$N_d > \frac{2K\epsilon_0\phi_i}{q{x_d}^2}$$

If we assume that $x_d = 25$ Å is required for efficient tunnelling and that $\phi_i \cong 0.3$ V, then

$$N_d > 6.2 \times 10^{19}\text{cm}^{-3}$$

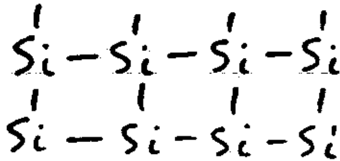
This is easily achieved and is in fact the method by which most ohmic contacts are made, particularly to n-type material, because it is independent of surface states. To contact a lightly-doped region, heavily-doped regions are added which are then contacted.

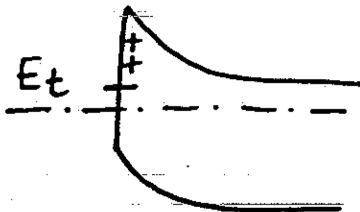
In bipolar transistors for example, the emitter diffusion is used to contact the lightly-doped collector region.



Surface States

Up to this point we have considered ideal contacts. In practice, the dangling bonds at the interface create a large density of trap levels which dominate the properties of the device. The charge state of those interface states (also generation/recombination sites) can affect the position of the Fermi level at the interface.





Assume donor-like states exist at the interface. If the Fermi level is below the donor energy level, then the levels will be empty and therefore ionized positively. These states effectively increase the donor concentration near the interface thus:

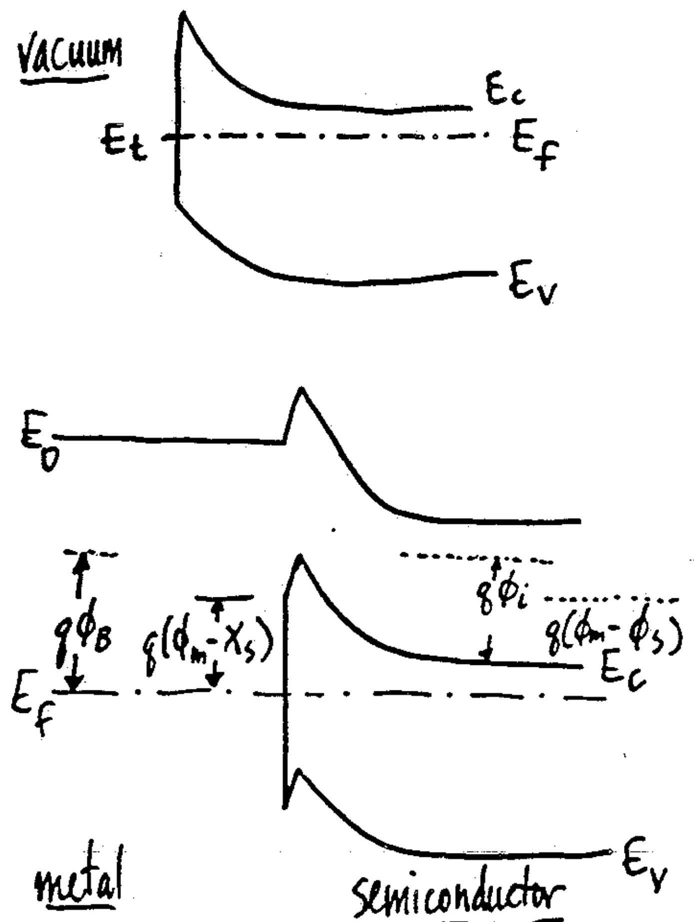
- x_d is reduced
- ϕ_i is reduced

If the number of interface states is very large, the interface traps can provide more than enough electrons to establish equilibrium. Thus the Fermi level cannot move below the trap energy and is effectively pinned near the trap level. Small changes in the Fermi level cause large changes in the electron and ion concentrations.

Similarly, if the interface traps also have acceptor levels, the Fermi level cannot rise far above the acceptor level without greatly altering the carrier concentrations.

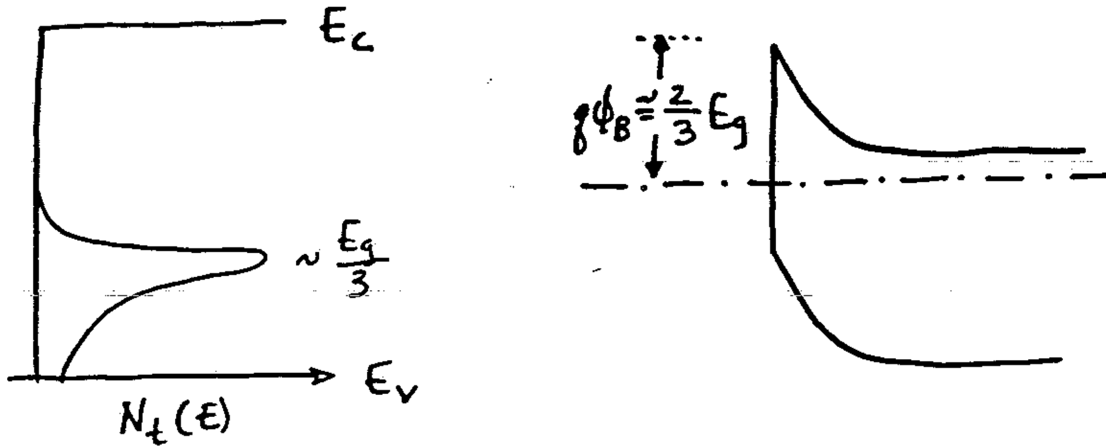
Therefore, just as in the case of a large concentration of bulk traps, the Fermi level is pinned halfway between the donor and acceptor level and $\phi_i \neq \phi_m - \phi_s$. The built in potential is determined not by the bulk material properties, but rather by the interface properties which depend critically on the fabrication techniques.

The potential difference between $\phi_m - \phi_s$ and the energy at which the Fermi level is pinned by the surface states is dropped across a very thin interface region which electrons can tunnel through easily.



The interface states do not have discrete levels, but rather have a distribution or band of energies. The peak interface state density often occurs near

$$E_t \cong E_v + \frac{1}{3}E_g.$$



Therefore, if the surface state density is large,

$$q\phi_B \cong \frac{2}{3}E_g \tag{16}$$

$$q\phi_i \cong \frac{2}{3}E_g - (E_c - E_f)_{\text{bulk}} \tag{17}$$

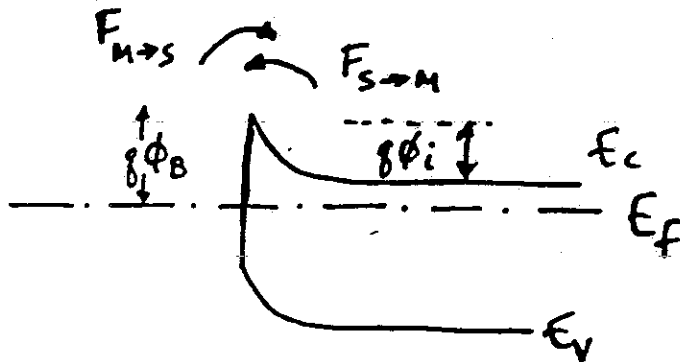
Because N_t depends on the fabrication processes and cannot be accurately predicted, ϕ_B and ϕ_i are usually measured experimentally.

Below are values for ϕ_B for metal to an n-type semiconductor. Note that the barrier heights follow the trend of $q(\phi_m - \chi_s)$ but are all near .667E_g or 0.75 eV for silicon.

Metal	ϕ_m, eV	Si (χ = 4.05)	Ge (χ = 4.13)	GaAs (χ = 4.07)	GaP (χ = 4.0)
Al	4.2	0.5–0.77	0.48	0.80	1.05
Au	4.7	0.81	0.45	0.90	1.28
Cu	4.4	0.69–0.79	0.48	0.82	1.20
Pt	5.4	0.9	—	0.86	1.45

Current-Voltage Characteristics

Under equilibrium conditions, the flow electrons from metal to semiconductor just equals that from the semiconductor to the metal. Thus $I = 0$.



An applied bias changes the barrier height for electrons from the semiconductor to the metal ($\phi_i - V_A$), thus changing the flux from the semiconductor to the metal ($F_{S\to M}$). The flux from the metal ($F_{M\to S}$) does not change since the metal being a conductor cannot sustain an electric field (ϕ_B = constant).

$$F_{S\to M} = \underbrace{D_n \frac{dn}{dx}}_{\text{diffusion}} + \underbrace{n\mu_n \mathcal{E}}_{\text{drift}}$$

By applying an external potential, the internal electric field $\mathcal{E}$ is altered. Under forward bias, the drift term is reduced, resulting in a net electron flux (or current) due to diffusion.

The number of electrons with sufficient energy to surmount the barriers (ϕ_i or ϕ_B) are given by Fermi-Dirac statistics (or the Maxwell-Boltzmann approximation if applicable).

$$n = N_c \exp\left(-\frac{E_c - E_f}{kT}\right) = \text{number of electrons in c-band}$$

We are interested in the number of electrons with sufficient energy to surmount the barrier to the metal ($E > E_c + \phi_i$). The number of electrons with sufficient energy is

$$n_s = N_c \exp\left(-\frac{E_c - E_f + q\phi_i}{kT}\right) = N_d \exp\left(-\frac{q\phi_i}{kT}\right)$$

which could have been written directly from M.B. statistics.

Since the number of electrons in the metal is much large than N_d, ϕ_B must be large than ϕ_i for the fluxes to be equal.

Therefore, in equilibrium,

$$I_{M\to S} = I_{S\to M} = KN_d \exp -\frac{q\phi_i}{kT}$$

Under bias, $\phi_i \Rightarrow \phi_i - V_A$ for electrons going from $S \to M$.

$$n_s = N_d \exp -\frac{q(\phi_i - V_A)}{kT}$$

ϕ_B is unchanged so the current from $M \to S$ is unchanged and the net current is

$$I_{\text{net}} = I_{S\to M} - I_{M\to S} = KN_d \exp -\frac{q(\phi_i - V_A)}{kT} - KN_d \exp -\frac{q\phi_i}{kT}$$

Thus,

$$I = I_0 \left(\exp \frac{qV_A}{kT} - 1\right) \tag{18}$$

where

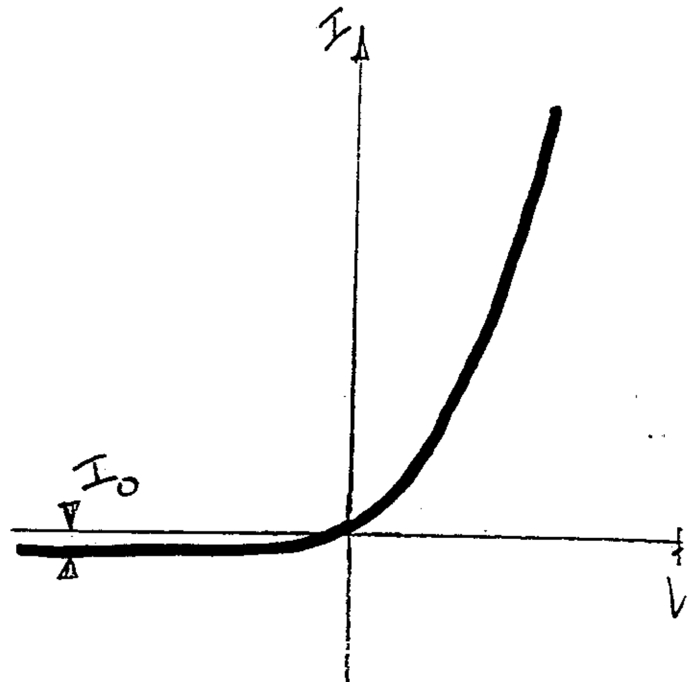
$$I_0 = KN_d \exp -\frac{q\phi_i}{kT} \tag{19}$$

This is just the ideal diode equation and I_0 is the reverse saturation current.

Forward Bias: Current increases exponentially with applied voltage.

Reverse Bias: $I \cong -I_0$, approximately constant.

Note that this result is only valid for $V_A < \phi_i$. For larger applied voltages, the depletion region disappears and the current is limited by the series resistance of the bulk semiconductor, not the junction as we assumed in this analysis.



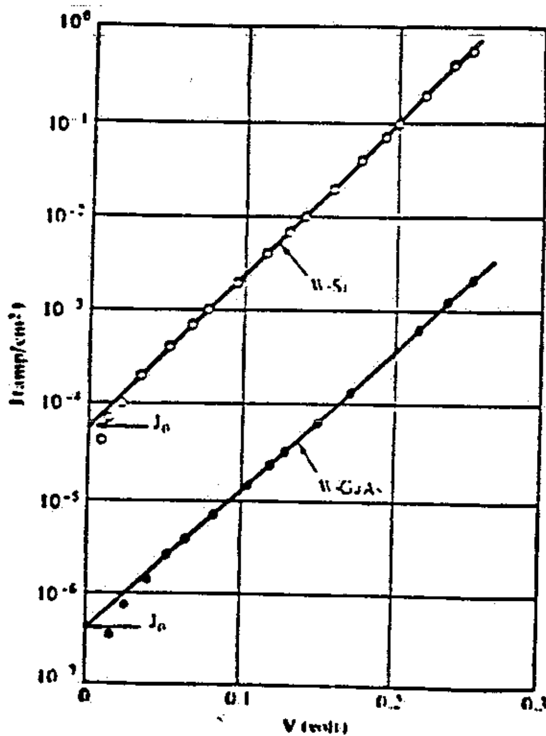
A more careful analysis of junction current (see text) reveals that I_0 is not independent of voltage. The exact relationship depends on the semiconductor doping profile. For example, with constant substrate doping,

$$I_0 = \frac{Aq^2 D_n N_c}{kT}\left(\frac{2q(\phi_i - V_A)N_d}{K_s \epsilon_0}\right)^{1/2} \exp\left(\frac{-q\phi_B}{kT}\right)$$

This is often incorporated into the diode equation by using a nonideality factor n (usually $1.0 < n < 1.2$) and keeping I_0 independent of V_A.

$$I = I_0 \left(\exp \frac{qV_A}{nkT} - 1\right) \tag{20}$$

The log of the current is often plotted versus V_A.



- I_0 can be determined by extrapolating to $V_A = 0$.
- ϕ_i and ϕ_B can then be determined through knowledge I_0 variations with temperature or by a more complete analysis of I_0 (see text).
- n can be obtained from the slope.

  Here, $n = 1.02$ for the silicon diode and $n = 1.04$ for the GaAs diode.

Maximum Voltage Capability

As we increase the reverse bias voltage on a Schottky diode,

$$x_d = \sqrt{\frac{2K\epsilon_0}{qN_d}(\phi_i + V_R)}$$

increases.

Therefore,

$$\mathcal{E} = -\frac{qN_d}{2K\epsilon_0}x_d$$

also increases.

Eventually, the electric field becomes large enough to cause avalanche breakdown and the diode conducts large amounts of currents in the reverse direction. We will consider diode breakdown in greater detail when we discuss pn diodes.

In the two or three dimensional structures of IC devices, larger electric fields which develop near edges can reduce the breakdown voltage, often drastically.

This problem can be reduced by extending the metal over the adjacent oxide or adding p^+ guard rings.