SMA5111 - Compound Semiconductors

Lecture 2 - Metal-Semiconductor Junctions - Outline

• Introduction
  Structure - What are we talking about?
  Behaviors: Ohmic, rectifying, neither

• Band picture in thermal equilibrium
  (Establishing the baseline)
  Ideal junction - no surface states
  Real junctions - surface states and Fermi level pinning

• Applying voltage bias (i-v and c-v)
  (Where it gets interesting, i.e. useful)
  Forward bias, current flow
    1. General comments; 2. Thermionic emission theory;
    3. Drift-diffusion theory; 4. Real junctions
  Reverse bias, image-force lowering
  Switching dynamics
    1. Step response; 2. High frequency response

• Applications
  (Benefiting from these simple structures)
  Ohmic contacts
  Doping profiling
  Shunt diodes
  FET gate (MESFETs)
  UV photodiodes
Metal-Semiconductor Junctions - the structure

The structure is very simple

This is the "junction" we're talking about

but also very interesting, important, and useful
Metal-Semiconductor Junctions - barrier basics

- The evolution of the electrostatic barrier at the interface
  Initially we assume no surface states, i.e. bulk bands right to surface

- The energy band picture in isolation
  An isolated metal and an isolated semiconductor; neither "sees" the other

The vacuum reference levels are equal.
Both materials are neutral.
Note definitions of $\phi$ (work function) and $\chi$ (electron affinity)

Note: no surface states for now; they come later
Metal-Semiconductor Junctions - barrier basics

- The metal and semiconductor shorted electrically

  The short imposes a constant Fermi level throughout

  The combination remains neutral, but the two materials become charged as electrons flow from the semiconductor to the metal until the Fermi levels are the same. The semiconductor surface is slightly depleted at large separation; the depletion increases as they approach.
Metal-Semiconductor Junctions - barrier basics

- Shorted metal and semiconductor in physical contact

  As the distance between the metal and semiconductor decreases to zero, the depletion region grows

  \[
  q\phi_b = q\phi_m - q\chi_s
  \]

  The final depletion region width is that needed to support a potential change equal to the built-in potential, \( \phi_b (= \phi_m - \chi_s) \)

  The total structure is neutral, but there is now a dipole layer between the metal and semiconductor

  To model this we use the depletion approximation
Metal-Semiconductor Junctions - barrier basics

• Depletion approximation

The charge in the metal is approximated as a sheet (impulse) charge density at the surface, and charge in the semiconductor is approximated by a fully depleted layer $X_D$ wide:

$$Q^* \delta(x) \quad \text{for} \quad x \leq 0$$

$$\rho(x) \approx \begin{cases} 
   qN_D & \quad \text{for} \quad 0 < x \leq X_D \\
   0 & \quad \text{for} \quad X_D < x
\end{cases}$$

Charge neutrality requires $Q^* = -qN_D X_D$ 
Remember we are dealing with sheet charge density, Coul/cm$^2$
• Depletion approximation (cont)

Integrating the charge divided by the dielectric constant yields the electric field

\[ E(x) = \int \frac{\rho(x)}{\varepsilon} \, dx \]

We get:

\[ E(x) \approx \begin{cases} 0 & \text{for } x \leq 0 \\ qN_D \frac{(x - X_D)}{\varepsilon} & \text{for } 0 < x \leq X_D \\ 0 & \text{for } X_D < x \end{cases} \]
Depletion approximation (cont)

Integrating the electric field yields the electrostatic potential

\[ \phi(x) = - \int E(x) \, dx \]

We get:

\[ \phi(x) \approx \begin{cases} \phi_b & \text{for } x \leq 0 \\ qN_D(x - X_D)^2 / 2\varepsilon & \text{for } 0 < x \leq x_D \\ 0 & \text{for } x_d < x \end{cases} \]

Requiring that \( \phi(x) \) be continuous at \( x = 0 \) we find that the depletion region width, \( X_D \), must be

\[ X_D \approx \left( 2\varepsilon \phi_b / qN_D \right)^{1/2} \]

The profile is now fully determined. (i.e., we're done)
Real semiconductor surfaces - surface states

- **Surface states**
  There will be additional energy states on the surface of a semiconductor because the perfectly periodic lattice ends at the surface and many bonds are not "satisfied"
  These states... can have a very high density
  have a narrow distribution of energies within bandgap

- **The energy bands in a semiconductor with surface states**
  The surface states typically are sufficiently dense that in equilibrium the Fermi level falls within them at the surface and the surface is depleted:

\[
q\phi_{sd} = fE_g - kT \ln(N_C/N_D)
\]

Note: \(0 < f < 1\); for many III-V's \(f \approx 0.6-0.7\)
Real semiconductor surfaces - surface states, cont.

- Estimating the number of surface states
  Unit cell 5.5Å by 5.5Å → 3 x 10^{14} cells/cm^2 at surface
  4 unsatisfied bonds per cell → ≈ 10^{15} states/cm^2
  If the states fall within 0.1 eV of each other → ≈ 10^{16} states/cm^2-eV
  This is very large!!

- What does this mean as a practical matter?
  Suppose \( \Phi_m - \chi_s = 0.5 \) V, and that the effective separation of the charge in the surface states and metal is 25nm. The sheet charge density induced in this situation is:
  \[ Q^* = \varepsilon \Delta V / d = 10^{-12} \times 0.5 / 2.5 \times 10^{-6} = 2 \times 10^{-6} \text{ coul/cm}^2 \]
  The corresponding state density is \( Q^*/q \approx 10^{13} \text{ cm}^{-2} \)
  
  If all the surface states are active, the Fermi level at the surface will change only 1 mV; if only 10% are active it is only 10 mV. Only if 1%, or less, are active can the surface be unpinned.

- Conclusion
  The metal work function is often not the main determinant of the potential barrier in a metal-semiconductor junction.
Metal-Semiconductor Junctions - w. surface states

- The energy band picture in isolation with surface states

  The surface of the semiconductor is depleted because of the charged surface states, independent of there being any metal nearby.

  \[ q\Phi_m \]
  \[ q\chi_s \]
  \[ fE_g \]
  \[ q\phi_{sd} = fE_g - kT \ln(N_C/N_D) \]

  Note: \( 0 < f < 1; \) for many III-V's \( f \approx 0.6-0.7 \)

C. G. Fonstad, 2/03
Metal-Semiconductor Junctions - w. surface states (cont.)

- Shorted metal and semiconductor, with surface states, in physical contact

When the density of surface states is high, as it typically is, the potential barrier that develops is dominated by the location of the surface states in the semiconductor band gap, rather than by the work function of the metal.

Otherwise, nothing is different and the same modeling holds.
Barrier heights
\[ \text{vs.} \]
metal work function

-\( \rightarrow \) the impact of surface states on metal-semiconductor barrier heights

-\( \rightarrow \) the barrier height varies much less than does the work function of the metal


Applying bias to a metal-semiconductor junction

• What happens globally
  Potential step crossing junction changes
  Depletion region width and electric field change
  Current flows across junction

• Potential step change

Assuming all the bias appears across the junction, the potential barrier changes from $\phi_b$ to $\phi_b - V_{AB}$

$\phi_b \rightarrow \phi_b - V_{AB}$

Note: Forward bias decreases the barrier
Reverse bias increases the barrier
Applying bias to a metal-semiconductor junction, cont.

- **Depletion region width and field changes**
  
  Wherever $\phi_b$ appears in the expressions for depletion region width and electric field, it is replaced by $\phi_b - V_{AB}$:

  **Depletion region width:**
  
  $$X_D \rightarrow \left[ \frac{2\varepsilon (\phi_b - V_{AB})}{qN_D} \right]^{1/2}$$

  Note: The depletion region width decreases in forward bias. Reverse bias increases the depletion region width.

  **Peak electric field:**
  
  $$E_{pk} = \left[ \frac{2\varepsilon \phi_b qN_D}{\varepsilon} \right]^{1/2} \rightarrow \left[ \frac{2\varepsilon (\phi_b - V_{AB}) qN_D}{\varepsilon} \right]^{1/2}$$

  Note: The peak electric field decreases in forward bias. Reverse bias increases the field strength.

- **Note:** potential step and depletion region changes are the same as happens in a p-n junction.
Applying bias to a metal-semiconductor junction, cont.

- **Currents**
  
  *Note:* the barrier seen by electrons in the metal does not change with bias, whereas the barrier seen by those in the semiconductor does.

  Thus the carrier flux (current) we focus on is that of majority carriers from the semiconductor flowing into the metal. Metal-semiconductor junctions are primarily *majority carrier* devices.

Minority carrier injection into the semiconductor can usually be neglected; more about this later.
Applying bias to a metal-semiconductor junction, cont.

- **Currents, cont.**

  The **net current** is the current from the semiconductor to the metal, minus the current from the metal to the semiconductor:

  \[ i_D(v_{AB}) = i_{Dm->s}(v_{AB}) - i_{Ds->m}(v_{AB}) \]

  **Semiconductor to metal, \( i_{Ds->m}(v_{AB}) \)**

  **Four factors:**

  1. \( N_{Dn} \exp[-q(\phi_b - v_{AB})/kT] \), the number of carriers that can cross the barrier, \((\phi_b - v_{AB})\)
  2. \( R \), the rate at which the carriers that can cross, get across
  3. \( A \), the cross-sectional area
  4. \(-q\), the charge per carrier

  \[ i_{Ds->m}(v_{AB}) = -qARN_{Dn} \exp[-q(\phi_b - v_{AB})/kT] \]

  **Metal to semiconductor, \( i_{Dm->s}(v_{AB}) \)**

  Not a function of voltage  
  (because barrier seen from metal doesn't change)

  Must equal \( i_{Ds->m}(v_{AB}) \) when \( v_{AB} = 0 \), i.e. \( i_{Ds->m}(0) \)

  \[ i_{Dm->s}(v_{AB}) = i_{Ds->m}(0) = -qARN_{Dn} \exp[-q\phi_b/kT] \]
Applying bias to a metal-semiconductor junction, cont.

- **Currents, cont.**

  Thus, the net current is:

  \[
  i_D(v_{AB}) = qA R N_{Dn} \exp(-q\Phi_b/kT) \left[\exp(qv_{AB}/kT) - 1\right]
  \]

  ******

  What we haven't done yet is say anything about \( R \)
  The modeling meat is in \( R \)!

- **Barrier transit rate models (models for \( R \))**

  Different models assume that different factors are limiting the flow, and they result in different dependences of \( R \) (and thus of the \( i_D \)) on the device and material parameters and temperature.

  **Thermionic emission theory** - the flow is limited by the rate at which carriers try to cross the barrier

  **Drift-diffusion theory** - the flux is limited by the rate at which carriers cross the depletion region and reach the barrier

  **Combination theories** - both of the above factors play a role and must be included in the modeling
Applying bias to a metal-semiconductor junction, cont.

- **Image force barrier lowering**

  An electron leaving a metal sees an image force pulling it back:

  \[ \phi = \frac{q^2}{16\pi d} \]

  [Diagram showing metal and vacuum with an image force barrier]

  We see that the potential step at the surface of a metal is not abrupt as we have modeled it:

  \[ q\phi_m \]

  [Graph showing potential step with reduced barrier]

  This reduces the barrier seen by the carriers. (next foil)
Applying bias to a metal-semiconductor junction, cont.

- Image force barrier lowering (cont.)

  The image force reduces the barrier:

Furthermore the barrier reduction increases with increasing reverse bias:

This means the current does not saturate in reverse bias (unlike the case in a p-n diode).
Comparison of m-s junctions and p-n junctions

Lessons from i-v modeling results:
– Comparing metal to n-Si and p⁺-Si to n-Si diodes, i.e. same n-sides

• The m-s current is higher at the same bias (m-s barrier is always lower)
\[ i_{D,m-s}(v_{AB}) > i_{D,p-n}(v_{AB}) \text{ @ same } v_{AB} \]

• There is no minority carrier injection or storage in the m-s diode
  modulation and switching can be much faster

• The reverse bias, or "off" current of an m-s diode does not truly saturate
  turn-off is not has hard, but we can still have sharp breakdown and avalanche

The first two differences play major roles in several applications of m-s diodes
What metal-semiconductor junctions are good for

Note: The key features that make m-s junctions useful are…
- majority carrier devices, negligible minority carrier injection
- relatively low barrier to forward current flow
- depletion and field extend to surface

Important Applications

• Ohmic contacts
  an essential component of any electronic device

• Determining doping profiles
  a key diagnostic technique in device fabrication/processing

• Shunt diodes
  to reduce switching transients in bipolar transistor logic

• Microwave diodes
  another use taking advantage of negligible excess carrier injection

• FET gate (MESFETs)
  the subject of Lecture 9

• Ultraviolet detectors
  to be discussed in Lecture 21