Section 12: Intro to Devices

Extensive reading materials on reserve, including

Robert F. Pierret, Semiconductor Device Fundamentals

EE143 – Ali Javey

Bond Model of Electrons and Holes

- Silicon crystal in a two-dimensional representation.

- When an electron breaks loose and becomes a conduction electron, a hole is also created.
Semiconductors, Insulators, and Conductors

- Totally filled bands and totally empty bands do not allow current flow. (Just as there is no motion of liquid in a totally filled or totally empty bottle.)
- Metal conduction band is half-filled.
- Semiconductors have lower $E_G$’s than insulators and can be doped

Dopants in Silicon

- As, a Group V element, introduces conduction electrons and creates N-type silicon, and is called a donor.
- B, a Group III element, introduces holes and creates P-type silicon, and is called an acceptor.
Types of charges in semiconductors

Mobile Charge Carriers
they contribute to current flow
with electric field is applied.

Immobile Charges
they DO NOT
contribute to current flow
with electric field is applied.
However, they affect the
local electric field

Fermi Function—The Probability of an Energy State Being Occupied by an Electron

\[
f(E) = \frac{1}{1 + e^{(E-E_f)/kT}}
\]

\(E_f\) is called the Fermi energy or the Fermi level.

Boltzmann approximation:

\[
f(E) \approx e^{-(E-E_f)/kT} \quad E - E_f >> kT
\]

\[
f(E) \approx 1 - e^{-(E_f-E)/kT} \quad E - E_f << -kT
\]
**Electron and Hole Concentrations**

\[
n = N_c e^{(E_F - E_C)/kT}
\]

\[
p = N_v e^{(E_V - E_F)/kT}
\]

*\(N_c\) is called the effective density of states.*

*\(N_v\) is called the effective density of states of the valence band.*

Remember: the closer \(E_f\) moves up to \(E_c\), the larger \(n\) is; the closer \(E_f\) moves down to \(E_v\), the larger \(p\) is.

For Si, \(N_c = 2.8 \times 10^{19} \text{ cm}^{-3}\) and \(N_v = 1.04 \times 10^{19} \text{ cm}^{-3}\).

**Shifting the Fermi Level**

[Image of energy band diagrams showing occupancy factors and carrier distributions for different Fermi levels]
General Effects of Doping on n and p

I. \( N_d - N_a >> n_i \) (i.e., N-type)
\[
\begin{align*}
    n &= N_d - N_a \\
    p &= \frac{n_i^2}{n}
\end{align*}
\]

If \( N_d >> N_a \), \( n = N_d \) and \( p = \frac{n_i^2}{N_d} \)

II. \( N_a - N_d >> n_i \) (i.e., P-type)
\[
\begin{align*}
    n &= \frac{n_i^2}{p} \\
    p &= N_a - N_d
\end{align*}
\]

If \( N_a >> N_d \), \( p = N_a \) and \( n = \frac{n_i^2}{N_a} \)

---

Carrier Drift

- When an electric field is applied to a semiconductor, mobile carriers will be accelerated by the electrostatic force. This force superimposes on the random thermal motion of carriers:

\[ E = 0 \]

E.g. Electrons \textit{drift} in the direction opposite to the \( E \)-field
\( \rightarrow \) Current flows

Average drift velocity = \( |\mathbf{v}| = \mu E \)

Carrier mobility
Carrier Mobility

- Mobile carriers are always in random thermal motion. If no electric field is applied, the average current in any direction is zero.

- Mobility is reduced by
  1) collisions with the vibrating atoms “phonon scattering”
  2) deflection by ionized impurity atoms “Coulombic scattering”

Total Mobility

\[ \frac{1}{\tau} = \frac{1}{\tau_{\text{phonon}}} + \frac{1}{\tau_{\text{impurity}}} \]

\[ \frac{1}{\mu} = \frac{1}{\mu_{\text{phonon}}} + \frac{1}{\mu_{\text{impurity}}} \]
**Conductivity and Resistivity**

\[ J_{p,\text{drift}} = q_p v = q_p \mu_p E \]

\[ J_{n,\text{drift}} = -q_n v = q_n \mu_n E \]

\[ J_{\text{drift}} = J_{n,\text{drift}} + J_{p,\text{drift}} = \sigma E = (q_n \mu_n + q_p \mu_p) E \]

\[ \therefore \text{ conductivity of a semiconductor is } \sigma = q_n \mu_n + q_p \mu_p \]

Resistivity, \( \rho = 1/\sigma \)

---

**Relationship between Resistivity and Dopant Density**

![Graph showing the relationship between resistivity and dopant density for N-type and P-type semiconductors.](image)

\[ \rho = 1/\sigma \]
Sheet Resistance

\[ R = \rho \frac{L}{Wt} = R_s \frac{L}{W} \]

\( R_s \) is the resistance when \( W = L \) (in ohms/square)

\[ R_s \equiv \frac{\rho}{t} \quad \text{if } \rho \text{ is independent of depth } x \]

- \( R_s \) value for a given conductive layer (e.g. doped Si, metals) in IC or MEMS technology is used
  - for design and layout of resistors
  - for estimating values of parasitic resistance in a device or circuit

Diffusion Current

Particles diffuse from higher concentration to lower concentration locations.
**Diffusion Current**

\[ J_{n,\text{diffusion}} = qD_n \frac{dn}{dx} \quad J_{p,\text{diffusion}} = -qD_p \frac{dp}{dx} \]

\( D \) is called the diffusion constant. Signs explained:

- Electron flow
- Hole flow

**Generation/Recombination Processes**

- (a) Band-to-band recombination
- (b) R–G center recombination
- (c) Auger recombination
- (d) Band-to-band generation
- (e) R–G center generation
- (f) Carrier generation via impact ionization

Recombination continues until excess carriers = 0.
Time constant of decay is called recombination lifetime
A PN junction is present in almost every semiconductor device.

Energy Band Diagram and Depletion Layer

\[ n \approx 0 \text{ and } p \approx 0 \text{ in the depletion layer} \]

\[ \phi_{bi} = \frac{kT}{q} \ln \frac{N_d N_a}{n_i^2} \]
**Qualitative Electrostatics**

- Band diagram
- Built-in potential
- From $\varepsilon = -\frac{dV}{dx}$

**Effect of Bias on Electrostatics**

- Equilibrium ($V_A = 0$)
- Forward bias ($V_A > 0$)
- Reverse bias ($V_A < 0$)
Current Flow - Qualitative

(a) Equilibrium ($V_A = 0$)

(b) Forward bias ($V_A > 0$)

(c) Reverse bias ($V_A < 0$)

(d) Exponential and Constant

PN Diode IV Characteristics

$I = I_0 (e^{qV/kT} - 1)$

$I_0 = A q n_i^2 \left( \frac{D_p}{L_p N_d} + \frac{D_n}{L_n N_a} \right)$

$I_r = I_0 + A q n_i W_{dep} \frac{1}{\tau_{dep}}$
Solar Cells

![Diagram of solar cells](image)

**Equation (4.9.4)**

$$q(0)$$

**Equation (4.12.1)**

Maximum power-output

$p-i-n$ Photodiodes

- Only electron-hole pairs generated in depletion region (or near depletion region) contribute to current
- Only light absorbed in depletion region contributes to generation
  - Stretch depletion region
  - Can also operate near avalanche to amplify signal
**Light Emitting Diodes (LEDs)**

- LEDs are typically made of compound semiconductors
  - Why not Si

<table>
<thead>
<tr>
<th>Semiconductor</th>
<th>Color</th>
<th>Peak λ(μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs&lt;sub&gt;0.4&lt;/sub&gt;P&lt;sub&gt;0.6&lt;/sub&gt;</td>
<td>Red</td>
<td>0.650</td>
</tr>
<tr>
<td>GaAs&lt;sub&gt;0.35&lt;/sub&gt;P&lt;sub&gt;0.65&lt;/sub&gt;:N</td>
<td>Orange-Red</td>
<td>0.630</td>
</tr>
<tr>
<td>GaAs&lt;sub&gt;0.14&lt;/sub&gt;P&lt;sub&gt;0.86&lt;/sub&gt;:N</td>
<td>Yellow</td>
<td>0.585</td>
</tr>
<tr>
<td>GaP:N</td>
<td>Green</td>
<td>0.565</td>
</tr>
<tr>
<td>GaP:Zn-O</td>
<td>Red</td>
<td>0.700</td>
</tr>
<tr>
<td>AlGaAs</td>
<td>Red</td>
<td>0.650</td>
</tr>
<tr>
<td>AlInGaP</td>
<td>Orange</td>
<td>0.620</td>
</tr>
<tr>
<td>AlInGaP</td>
<td>Yellow</td>
<td>0.585</td>
</tr>
<tr>
<td>SiC</td>
<td>Blue</td>
<td>0.470</td>
</tr>
<tr>
<td>GaN</td>
<td>Blue</td>
<td>0.450</td>
</tr>
</tbody>
</table>

**MOS Capacitors**

**MOS: Metal-Oxide-Semiconductor**

-MOS capacitor

-MOS transistor
Ideal MOS Capacitor

- Oxide has zero charge, and no current can pass through it.
- No charge centers are present in the oxide or at the oxide-semiconductor interface.
- Semiconductor is uniformly doped
- \( \Phi_M = \Phi_S = \chi + (E_C - E_F)_{FB} \)

At Equilibrium:

\[ E_F \quad \Phi_M \quad E_0 \quad \Phi_S \quad E_C \quad E_V \]

Figure 16.4
**Ideal MOS Capacitor Under Bias**

- Let us ground the semiconductor and start applying different voltages, $V_G$, to the gate.

- $V_G$ can be positive, negative or zero with respect to the semiconductor.

- $E_{F,\text{metal}} - E_{F,\text{semiconductor}} = -qV_G$

- Since oxide has no charge (it’s an insulator with no available carriers or dopants), $\frac{dE_{\text{oxide}}}{dx} = \frac{\rho}{\varepsilon} = 0$; meaning that the $E$-field inside the oxide is constant.

**Inversion condition**

If we continue to increase the positive gate voltage, the bands at the semiconductor bends more strongly. At sufficiently high voltage, $E_i$ can be below $E_F$ indicating large concentration of electrons in the conduction band.

We say the material near the surface is “inverted”. The “inverted” layer is not gotten by chemical doping, but by applying $E$-field. Where did we get the electrons from?

When $E_i(\text{surface}) - E_i(\text{bulk}) = 2[E_F - E_i(\text{bulk})]$, the condition is start of “inversion”, and the voltage $V_G$ applied to gate is called $V_T$ (threshold voltage). For $V_G > V_T$, the Si surface is inverted.
Ideal MOS Capacitor – *n*-type Si
Electrostatic potential, $\phi(x)$

Define a new term, $\phi(x)$ taken to be the potential inside the semiconductor at a given point $x$. [The symbol $\phi$ instead of $V$ used in MOS work to avoid confusion with externally applied voltage, $V$]

$$\phi(x) = \frac{1}{q}[E_i(\text{bulk}) - E_i(x)]$$  \hspace{1cm} \text{Potential at any point } x

$$\phi_S = \frac{1}{q}[E_i(\text{bulk}) - E_i(\text{surface})]$$  \hspace{1cm} \text{Surface potential}

$$\phi_F = \frac{1}{q}[E_i(\text{bulk}) - E_F]$$  \hspace{1cm} |\phi_F| \text{ related to doping concentration}

$\phi_F > 0$ means p-type  \hspace{1cm} $\phi_F < 0$ means n-type
Electrostatic potential

φ_S is positive if the bands bend ……?

φ_S = 2φ_F at the depletion-inversion transition point (threshold voltage)

Charge Density - Accumulation

p-type silicon accumulation condition  V_G < 0

The accumulation charges in the semiconductor are ……, and appear close to the surface and fall-off rapidly as x increases. One can assume that the free carrier concentration at the oxide-semiconductor interface is a δ-function.

Charge on metal = −Q_M
Charge on semiconductor = − (charge on metal)

|Q_Accumulation| = |Q_M|
**Charge Density - Depletion**

**p-type Si, depletion condition**

The depletion charges in Si are immobile ions - results in depletion layer similar to that in pn junction or Schottky diode.

\[ |qN_A\cdot W| = |Q_M| \]

(-) \quad (+)

If surface potential is \( \phi_s \), then the depletion layer width \( W \) will be

\[ W = \frac{2\varepsilon_{Si}}{qN_A} \phi_s \]

Does this equation look familiar?

**Charge Density - Inversion**

**p-type Si, strong inversion**

Once inversion charges appear, they remain close to the surface since they are ……… Any additional voltage to the gate results in extra \( Q_M \) in gate and get compensated by extra inversion electrons in semiconductor.

So, the depletion width does not change during inversion. Electrons appear as \( \delta \)-function near the surface. Maximum depletion layer width \( W = W_T \).
**MOS C-V characteristics**

- The measured MOS capacitance (called gate capacitance) varies with the applied gate voltage
  - A very powerful diagnostic tool for identifying threshold voltage, oxide thickness, substrate doping concentration, and flat band voltage.
  - It also tells you how close to an ideal MOSC your structure is.

- Measurement of C-V characteristics
  - Apply any dc bias, and superimpose a small (15 mV) ac signal (typically 1 kHz – 1 MHz)

---

**C-V: under accumulation**

Consider p-type Si under accumulation.

\[ V_G < 0. \]

Looks similar to parallel plate capacitor.

\[ C_G = C_{ox} \]

where \( C_{ox} = \frac{\varepsilon_{ox} A}{x_{ox}} \)

\( C_G \) is constant as a function of \( V_G \)
**C-V: under depletion**

Depletion condition: $V_G > 0$

$C_G$ is $C_{ox}$ in series with $C_s$ where $C_s$ can be defined as “semiconductor capacitance”

$$C_{ox} = \varepsilon_{ox} A / x_{ox}$$

$$C_s = \varepsilon_{Si} A / W$$

$$C_G = C_{ox} C_s / (C_{ox} + C_S)$$

$W = \sqrt{\frac{2\varepsilon_{Si} \phi_s}{qN_A}}$

where $\phi_s$ is surface potential

$C_G$ decreases with increasing $V_G$

**C-V: under inversion (high frequency)**

$V_G = V_T$ and $V_G > V_T$

Inversion condition $\phi_s = 2 \phi_F$

$$W = W_T = \sqrt{\frac{2\varepsilon_{Si} 2\phi_F}{qN_A}}$$

At high frequency, inversion electrons are not able to respond to ac voltage.

$$C_{ox} = \varepsilon_{ox} A / x_{ox}$$

$$C_s = \varepsilon_{Si} A / W_T$$

$$C_G (\omega \rightarrow \infty) = C_{ox} C_s / (C_{ox} + C_S)$$

$C_G$ will be constant for $V_G \geq V_T$
**C-V: under inversion (low frequency)**

At low frequency, the inversion electrons will be able to respond to the ac voltage. So, the gate capacitance will be equal to the “oxide capacitance” (similar to a parallel plate capacitance).

\[
C_G (\omega \to 0) = C_{ox} = \varepsilon_{ox} A / x_{ox}
\]

\[
C_G \text{ increases for } V_G \geq V_t \text{ until it reaches } C_{ox}
\]
In 1935, a British patent was issued to Oskar Heil.
A working MOSFET was not demonstrated until 1955.
Today’s MOSFET Technology

Gate oxides as thin as 1.2 nm can be manufactured reproducibly. Large tunneling current through the oxide limits oxide-thickness reduction.

Modern Field Effect Transistor (FET)

- An electric field is applied normal to the surface of the semiconductor (by applying a voltage to an overlying electrode), to modulate the conductance of the semiconductor
  → Modulate drift current flowing between 2 contacts ("source" and "drain") by varying the voltage on the "gate" electrode
Introduction to the MOSFET

Basic MOSFET structure and IV characteristics

![MOSFET Diagram]

What is desirable: large $I_{on}$, small $I_{off}$

Introduction to the MOSFET

Two ways of representing a MOSFET:

- Circuit Symbol
- Simple Switch
Complementary MOSFETs (CMOS)

**N-channel vs. P-channel**

- For current to flow, $V_{GS} > V_{Th}$
- For current to flow, $V_{GS} < V_{Th}$

**Complementary MOSFETs (CMOS)**

NFET

- $V_g = V_{dd}$
- $V_{ds} > 0$
- When $V_g = V_{dd}$, the NFET is on and the PFET is off.

PFET

- $V_g = 0$
- $V_{ds} < V_{dd}$
- When $V_g = 0$, the PFET is on and the NFET is off.
**Qualitative discussion: n-MOSFET**

- \( V_G > V_T; \ V_{DS} \approx 0 \)
  - \( I_D \) increases with \( V_{DS} \)

- \( V_G > V_T; \ V_{DS} \) small, > 0
  - \( I_D \) increases with \( V_{DS} \), but rate of increase decreases.

- \( V_G > V_T; \ V_{DS} \approx \text{pinch-off} \)
  - \( I_D \) reaches a saturation value, \( I_{D, \text{sat}} \)
  - The \( V_{DS} \) value is called \( V_{DS, \text{sat}} \)

\( V_G > V_T; \ V_{DS} > V_{DS, \text{sat}} \)
- \( I_D \) does not increase further, saturation region.

**Threshold voltage for NMOS and PMOS**

When \( V_G = V_T, \ \phi_s = 2 \ \phi_F \); we get expression for \( V_T \).

\[
V_T = 2 \phi_F + x_{ox} \frac{\varepsilon_{Si}}{\varepsilon_{ox}} \sqrt{\frac{2 q N_A}{\varepsilon_{Si}}} 2 \phi_F
\]

- Ideal n-channel
- (p-silicon) device
- both terms positive

\[
V_T = 2 \phi_F + \left( -x_{ox} \frac{\varepsilon_{Si}}{\varepsilon_{ox}} \sqrt{\frac{2 q N_D}{\varepsilon_{Si}}} \left| 2 \phi_F \right| \right)
\]

- Ideal p-channel
- (n-silicon) device
- both terms negative

\[ \varepsilon_{Si} / \varepsilon_{ox} = 11.9 / 3.9 \approx 3 \]

\[ \phi_F = \frac{1}{q} \left[ E_i(\text{bulk}) - E_F \right] \]

\( \phi_F > 0 \) means p-type

\( \phi_F < 0 \) means n-type
How to Measure the $V_T$ of a MOSFET

$V_T$ is measured by extrapolating the $I_{ds}$ versus $V_{gs}$ (at low $V_{ds}$) curve to $I_{ds} = 0$.

$$I_{dsat} = \frac{W}{L} C_{oxe} (V_{gs} - V_T) \mu_{ns} V_{ds} \propto V_{gs} - V_T$$

Quantitative $I_D$-$V_{DS}$ Relationships

“Square Law”

$$I_D = \frac{Z \mu_n}{L} C_{ox} \left[ (V_G - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right] \quad 0 < V_{DS} < V_{DS,\text{sat}} ; \quad V_G > V_T$$

$I_D$ will increase as $V_{DS}$ is increased, but when $V_G - V_{DS} = V_T$, pinch-off occurs, and current saturates when $V_{DS}$ is increased further. This value of $V_{DS}$ is called $V_{DS,\text{sat}}$, i.e., $V_{DS,\text{sat}} = V_G - V_T$ and the current when $V_{DS} = V_{DS,\text{sat}}$ is called $I_{DS,\text{sat}}$.

$$I_{D,\text{sat}} = \frac{Z \mu C_{ox}}{2L} (V_G - V_T)^2 \quad V_D > V_{DS,\text{sat}} ; \quad V_G > V_T$$

Here, $C_{ox}$ is the oxide capacitance per unit area, $C_{ox} = \varepsilon_{ox} / x_{ox}$.
$I_D - V_{DS}$ characteristics expected from a long channel ($\Delta L \ll L$) MOSFET (n-channel), for various values of $V_G$.

$V_G > V_T$  \quad $V_G < V_T$

N-channel MOSFET

Layout (Top View)

4 lithography steps are required:
1. active area
2. gate electrode
3. contacts
4. metal interconnects
Simple NMOS Process Flow

1) Thermal oxidation
   (~10 nm “pad oxide”)

2) Silicon-nitride (Si$_3$N$_4$) deposition by CVD
   (~40nm)

3) Active-area definition
   (lithography & etch)

4) Boron ion implantation
   (“channel stop” implant)

5) Thermal oxidation to grow oxide in “field regions”

6) Si$_3$N$_4$ & pad oxide removal

7) Thermal oxidation
   (“gate oxide”)

8) Poly-Si deposition by CVD

9) Poly-Si gate-electrode patterning (litho. & etch)

10) P or As ion implantation to form n+ source and drain regions

Cross-Section View

Top View of Masks

EE143 – Ali Javey
Simple NMOS Process Flow

11) SiO₂ CVD

12) Contact definition (litho. & etch)

13) Al deposition by sputtering

14) Al patterning by litho. & etch to form interconnects

EE143 – Ali Javey