Lecture #12

OUTLINE

- Diode analysis and applications continued
- The MOSFET
  - The MOSFET as a controlled resistor
  - Pinch-off and current saturation
  - Channel-length modulation
  - Velocity saturation in a short-channel MOSFET

Reading

- Rabaey
  - Section 3.3 (page 87-128)
- Hambley
  - Chapter 10

Summary of pn-Junction (Lec. 11-12)

- Two major currents in a pn junction
  - Diffusion current: carriers flow from where there are many to where there are few
    - Electrons diffuse from n-side to p-side and holes diffuse from p-side to n-side → both result in positive current from p to n side
  - Drift current: carriers flow due to electric field
    - Electric field exists in the depletion region only (this is the depletion approximation)
    - Electric field points from n side to p side, sweeping electrons from p-side (minority) to n-side and holes from n-side (minority also) to p-side → resulting positive current from n to p side

\[ I_D \]

\[ V_D \]
Summary of \( pn \)-Junction (Lec. 11-12)

- Under zero bias (0V), the two currents are equal \( \rightarrow \) net current \( I_D = 0 \)
- Under forward bias, the potential barrier is reduced \( \rightarrow \) drift current is reduced and diffusion current increases (positive net current)
  - Current \( I_D \) increases exponentially with increasing forward bias
  - The carriers become minority carriers once they cross the junction; as they diffuse in the quasi-neutral regions, they recombine with majority carriers (supplied by the metal contacts)
- Under reverse bias, the potential barrier is increased \( \rightarrow \) drift current dominates (negative net current)
  - But since the carriers that create drift currents are minority carriers, drift current is carrier limited and remains to be very small and saturates with large reverse bias \( V \)

![Graph](Image)

Light Emitting Diode (LED)

- LEDs are made of compound semiconductor materials
- Carriers diffuse across a forward-biased junction and recombine in the quasi-neutral regions
  \( \rightarrow \) optical emission

<table>
<thead>
<tr>
<th>Semiconductor</th>
<th>Color</th>
<th>Peak ( \lambda (\mu m) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaAs(<em>{0.6})P(</em>{0.4})</td>
<td>Red</td>
<td>0.650</td>
</tr>
<tr>
<td>GaAs(<em>{0.35})P(</em>{0.65}):N</td>
<td>Orange-Red</td>
<td>0.630</td>
</tr>
<tr>
<td>GaAs(<em>{0.14})P(</em>{0.86}):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>AlInGaP</td>
<td>Green</td>
<td>0.570</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>
Solar cell: Example of simple PN junction

- What is a solar cell?
  - Device that converts sunlight into electricity

- How does it work?
  - In simple configuration, it is a diode made of PN junction
  - Incident light is absorbed by material
  - Creates electron-hole pairs that transport through the material through:
    - Diffusion (concentration gradient)
    - Drift (due to electric field)

Optoelectronic Diodes (cont’d)

- Light incident on a pn junction generates electron-hole pairs

- The carriers that are generated in the depletion region and the minority carriers that are generated in the quasi-neutral regions that diffuse into the depletion region and are swept across the junction by the electric field

- This results in an additional component of current flowing in the diode:

\[ I_D = I_S \left( e^{qV_D/kT} - 1 \right) - I_{\text{optical}} \]

where \( I_{\text{optical}} \) is proportional to the intensity of the light
Photovoltaic (Solar) Cell

\[ I_D = I_S \left( e^{qV_D/kT} - 1 \right) - I_{optical} \]

Photodiode

- An intrinsic region is placed between the p-type and n-type regions
  - \( W_j \approx W_{i-region} \), so that most of the electron-hole pairs are generated in the depletion region
  - faster response time (~10 GHz operation)
Photodetector Circuit Using Load Line

As light shines on the photodiode, carriers are generated by absorption. These excess carriers are swept by the electric field at the junction creating drift current, which is in the same direction as the reverse bias current and hence negative current. The current is proportional to light intensity and hence can provide a direct measurement of light intensity → photodetector.

- What happens when $R_{th}$ is too large?
- Why use $V_{th}$?

Diodes in Circuits

- Use piece-wise linear model

"Practical" diode model
ideal with offset

Open segment

Short segment

$V_D = 0$

$0.6V$

$i_D = 0$

$0.6V$

$i_D$

$V_D$
Power Conversion Circuits

- Converting AC to DC
- Potential applications: Charging a battery

\[ V_I = V_m \sin(\omega t) \]

Equivalent circuit

- \( V > 0.6V \), diode = short circuit
  \[ \Rightarrow V_o = V_I - 0.6 \]
- \( V < 0.6V \), diode = open circuit
  \[ \Rightarrow V_o = 0 \]
Half-wave Rectifier Circuits

- Adding a capacitor: what does it do?

\[ V_m \sin(\omega t) \rightarrow \text{Capacitor} \rightarrow V_0 \]

Half-wave Rectifier

Current charging up capacitor
Why are pn Junctions Important for ICs?

- The basic building block in digital ICs is the MOS transistor, whose structure contains reverse-biased diodes.
  - pn junctions are important for electrical isolation of transistors located next to each other at the surface of a Si wafer.
  - The junction capacitance of these diodes can limit the performance (operating speed) of digital circuits

Device Isolation using pn Junctions

No current flows if voltages are applied between n-type regions, because two pn junctions are “back-to-back”

=> n-type regions isolated in p-type substrate and vice versa
We can build large circuits consisting of many transistors without worrying about current flow between devices. The p-n junctions isolate the transistors because there is always at least one reverse-biased p-n junction in every potential current path.

**Modern Field Effect Transistor (FET)**

- An electric field is applied normal to the surface of the semiconductor (by applying a voltage to an overlying “gate” 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

Metal-oxide-semiconductor (MOS) FET:
A GATE electrode is placed above (electrically insulated from) the silicon surface, and is used to control the resistance between the SOURCE and DRAIN regions.

Without a gate voltage applied, no current can flow between the source and drain regions.

Above a certain gate-to-source voltage (threshold voltage $V_T$), a conducting layer of mobile electrons is formed at the Si surface beneath the oxide. These electrons can carry current between the source and drain.
**N-channel vs. P-channel MOSFETs**

- For current to flow, $V_{GS} > V_T$
- Enhancement mode: $V_T > 0$
- Depletion mode: $V_T < 0$
  - Transistor is ON when $V_G = 0V$

- For current to flow, $V_{GS} < V_T$
- Enhancement mode: $V_T < 0$
- Depletion mode: $V_T > 0$
  - Transistor is ON when $V_G = 0V$

(“n+” denotes very heavily doped n-type material; “p+” denotes very heavily doped p-type material)

**MOSFET Circuit Symbols**

- (a) NMOS transistor as 4-terminal device
- (b) NMOS transistor as 3-terminal device
- (c) PMOS transistor as 4-terminal device
- (d) PMOS transistor as 3-terminal device
MOSFET Terminals

- The voltage applied to the GATE terminal determines whether current can flow between the SOURCE & DRAIN terminals.
  - For an n-channel MOSFET, the SOURCE is biased at a lower potential (often 0 V) than the DRAIN
    (Electrons flow from SOURCE to DRAIN when $V_G > V_T$)
  - For a p-channel MOSFET, the SOURCE is biased at a higher potential (often the supply voltage $V_{DD}$) than the DRAIN
    (Holes flow from SOURCE to DRAIN when $V_G < V_T$)

- The BODY terminal is usually connected to a fixed potential.
  - For an n-channel MOSFET, the BODY is connected to 0 V
  - For a p-channel MOSFET, the BODY is connected to $V_{DD}$

NMOSFET $I_G$ vs. $V_{GS}$ Characteristic
Consider the current $I_G$ (flowing into G) versus $V_{GS}$:

The gate is insulated from the semiconductor, so there is no significant (steady) gate current.
The MOSFET as a Controlled Resistor

- The MOSFET behaves as a resistor when $V_{DS}$ is low:
  - Drain current $I_D$ increases linearly with $V_{DS}$
  - Resistance $R_{DS}$ between SOURCE & DRAIN depends on $V_{GS}$
    - $R_{DS}$ is lowered as $V_{GS}$ increases above $V_T$

**NMOSFET Example:**

\[
I_{DS} = I_D \text{ = 0 if } V_{GS} < V_T
\]

\[
V_{GS} = 1 \text{ V} > V_T
\]

\[
V_{GS} = 2 \text{ V}
\]

\[
I_{DS} = 0 \text{ if } V_{GS} < V_T
\]

\[
V_{DS} = 1 \text{ V} > V_T
\]

\[
V_{GS} = 2 \text{ V}
\]

Inversion charge density $Q_i(x) = -C_{ox}[V_{GS}-V_T-V(x)]$

where $C_{ox} = \varepsilon_{ox}/t_{ox}$

---

### Sheet Resistance Revisited

Consider a sample of n-type semiconductor:

\[
R_s = \rho = \frac{1}{t} = \frac{1}{\sigma t} = \frac{1}{q\mu_n n t} = \frac{1}{\mu_n Q_n}
\]

where $Q_n$ is the charge per unit area
**NMOSFET $I_D$ vs. $V_{DS}$ Characteristics**

Next consider $I_D$ (flowing into D) versus $V_{DS}$, as $V_{GS}$ is varied:

- **Above threshold ($V_{GS} > V_T$):**
  - "inversion layer" of electrons appears, so conduction between S and D is possible

- **Below “threshold” ($V_{GS} < V_T$):**
  - no charge → no conduction

**MOSFET as a Controlled Resistor (cont’d)**

\[
I_D = \frac{V_{DS}}{R_{DS}}
\]

\[
R_{DS} = R_s \frac{L}{W} = \frac{L}{W} \frac{1}{\mu_n Q_i} = \frac{L}{W} \frac{1}{\mu_n C_{ox} (V_{GS} - V_T - \frac{V_{DS}}{2})}
\]

\[
I_D = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T - \frac{V_{DS}}{2}) V_{DS}
\]

We can make $R_{DS}$ low by
- applying a large "gate drive" ($V_{GS} - V_T$)
- making $W$ large and/or $L$ small
Charge in an N-Channel MOSFET

\( V_{GS} < V_T \):

\( V_{GS} > V_T \):

\( V_{DS} \approx 0 \)

\( V_{DS} > 0 \) (small)

Average electron velocity \( v \) is proportional to lateral electric field \( E \)

What Happens at Larger \( V_{DS} \)?

\( V_{GS} > V_T \):

\( V_{DS} = V_{GS} - V_T \)

Inversion-layer is “pinched-off” at the drain end

\( V_{DS} > V_{GS} - V_T \)

As \( V_{DS} \) increases above \( V_{GS} - V_T = V_{DSAT} \),

the length of the “pinch-off” region \( \Delta L \) increases:

- “extra” voltage \( (V_{DS} - V_{DSAT}) \) is dropped across the distance \( \Delta L \)
- the voltage dropped across the inversion-layer “resistor” remains \( V_{DSAT} \)

\( \Rightarrow \) the drain current \( I_D \) saturates

Note: Electrons are swept into the drain by the \( E \)-field when they enter the pinch-off region.
Summary of $I_D$ vs. $V_{DS}$

- As $V_{DS}$ increases, the inversion-layer charge density at the drain end of the channel is reduced; therefore, $I_D$ does not increase linearly with $V_{DS}$.
- When $V_{DS}$ reaches $V_{GS} - V_T$, the channel is “pinched off” at the drain end, and $I_D$ saturates (i.e. it does not increase with further increases in $V_{DS}$).

\[
I_D = \frac{k_n' W}{L} \left[ V_{GS} - V_T - \frac{V_{DS}}{2} \right] V_{DS}
\]
where \( k_n' = \mu_n C_{ox} \)

When $V_{DS}$ reaches $V_{GS} - V_T$, the channel is “pinched off” at the drain end, and $I_D$ saturates (i.e. it does not increase with further increases in $V_{DS}$).

\[
I_{DSAT} = \frac{k_n' W}{2 L} (V_{GS} - V_T)^2
\]
where \( k_n' = \mu_n C_{ox} \)

$I_D$ vs. $V_{DS}$ Characteristics

The MOSFET $I_D$-$V_{DS}$ curve consists of two regions:

1) Resistive or “Triode” Region: $0 < V_{DS} < V_{GS} - V_T$

2) Saturation Region:

\[ V_{DS} > V_{GS} - V_T \]

\[
I_{DSAT} = \frac{k_n' W}{2 L} (V_{GS} - V_T)^2
\]
where \( k_n' = \mu_n C_{ox} \)

“CUTOFF” region: $V_G < V_T$
Channel-Length Modulation

If $L$ is small, the effect of $\Delta L$ to reduce the inversion-layer “resistor” length is significant

$\rightarrow I_D$ increases noticeably with $\Delta L$ (i.e. with $V_{DS}$)

$I_D = I_D'(1 + \lambda V_{DS})$

$\lambda$ is the slope

$I_D'$ is the intercept

P-Channel MOSFET $I_D$ vs. $V_{DS}$

As compared to an n-channel MOSFET, the signs of all the voltages and the currents are reversed:

Note that the effects of velocity saturation are less pronounced than for an NMOSFET. Why is this the case?