Lecture 3
The MOS Transistor
Long Channel Behavior

EE 115C: Digital Electronic Circuits
Metal-Oxide Semi-conductor Field Effect Transistor (MOSFET)

NMOS Cross section View

Body

Source

Gate

Poly Si

Drain

p+

n

n

P substrate

SiO₂
This is a 4 terminal Device

Source  Drain  Gate  Body

Both NMOS and PMOS transistors are available

What’s the difference between NMOS and PMOS?

**NMOS:** current is carried by electrons moving through an n-type channel

**PMOS:** current is carried by holes moving through an p-type channel
Analog representation shows current flow from drain to source when $V_G$ is high.

Digital representation shows the logic: when $V_G = \text{high}$, transistor is “on”
For PMOS we have a similar representation

Analog Representation: when \( V_G = \text{low} \), current flows from source to drain.

Digital Representation: when \( V_G = \text{low} \), transistor is “on”
If $V_a > V_c$ and the NMOS is on

- current flow from A to C $\Rightarrow$ the electron flow from C to A
- From Definition, “source” is the source of electron flow into the channel
- Therefore, C is the “**Source**” while A is **Drain**

**Alternatively:** for an NMOS the source is the terminal with the lower potential.
Consider $V_s = V_d = \text{GND}$ and $V_g$ increase.

- $V_g \uparrow \Rightarrow V_{gate} \uparrow \Rightarrow E\text{-field} \uparrow$

Electrons are stripped away from atoms deep in the substrate and attracted to the channel. Simultaneously, holes are pushed away from the surface and into the substrate.
Initially electrons fill the valence band of the acceptor (p-type) atom below the gate to form a depletion layer (no mobile charge available).

As $V_g$ increases further, more electrons are attracted below the gate, thus creating an "Inversion Layer" (channel) consisting of "Mobile electron" (electrons that will flow in response to an applied voltage between Drain and source).
The voltage $V_{gs} = V_g - V_s$ at which the channel is initially formed is referred to as the threshold voltage, $V_t$. 
Linear (Triode) Operation

Once the channel has been formed

The application of a potential between the source and the drain causes current to flow

The shape of the channel region changes (it becomes more tapered off)

Under the above set of conditions, the transistor is said to be in the triode mode. The inversion layer spans the channel and the current is given by:

\[
I_{DS} = \mu_n C_{ox} \frac{W}{L} \left[ (V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2} \right]
\]
Why is Inversion Layer Tapered?

Why is the inversion layer tapered as we go from source to drain???

- We can consider the channel as a series of resistance.
- Less potential drop across gate-channel at drain for $V_g \sim$ fixed
As $V_{ds} \uparrow$ we reach a point at which $V_{gs} - V_{ds} = V_t \Rightarrow$ no inversion layer near drain

The channel is said to be “pinched off”

as $V_{ds}$ increase further, the pinch-off point starts moving toward the source.

For $V_{ds} > V_{gs} - V_t$ the transistor is said to be operated in the saturation or pinch-off
Current Equation in Saturation

The current equation for a long-channel NMOS in saturation is the following:

\[ I_{DS} = \mu_n \frac{C_{ox} W}{2 L} (V_{gs} - V_t)^2 \]

\[ \mu_n = \text{electron mobility} \]

\[ C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} = \text{gate capacitance} \]

\[ t_{ox} = \text{thickness of gate–oxide} \]

\[ W = \text{Channel Width} \]

\[ L = \text{Channel Length} \]
Long Channel NMOS I-V characteristic:

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

- **Linear**
- **Saturation**

\[ V_{GS} = 2.5 \text{ V} \]
\[ V_{GS} = 2.0 \text{ V} \]
\[ V_{GS} = 1.5 \text{ V} \]
\[ V_{GS} = 1.0 \text{ V} \]
Current Determinates

For a fixed $V_{DS}$ and $V_{GS} (> V_T)$, $I_{DS}$ is a function of

- the distance between the source and drain – $L$
- the channel width – $W$
- the threshold voltage – $V_T$
- the thickness of the SiO$_2$ – $t_{ox}$
- the dielectric of the gate insulator (SiO$_2$) – $\varepsilon_{ox}$
- the carrier mobility
  - for nfets: $\mu_n = 500$ cm$^2$/V-sec
  - for pfets: $\mu_p = 180$ cm$^2$/V-sec
**P-Channel Device Enhancement Mode** \((V_t < 0)\)

Equation are the same as the NMOS if the following substitutions are made:

<table>
<thead>
<tr>
<th>NMOS</th>
<th>PMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(V_{GS})</td>
<td>(V_{SG})</td>
</tr>
<tr>
<td>(V_{DS})</td>
<td>(V_{SD})</td>
</tr>
<tr>
<td>(V_{Tn})</td>
<td>(</td>
</tr>
<tr>
<td>(\mu_n)</td>
<td>(\mu_p)</td>
</tr>
</tbody>
</table>

**Cut off**  
\(V_{SG} \leq |V_{TP}|\)

**Triode**  
\(V_{SG} > |V_{TP}|\)

**Saturation**  
\(V_{SG} \geq |V_{TP}|\)  
\(V_{GD} \leq -|V_{TP}|\)  
\(V_{GD} \geq -|V_{TP}|\)
Alternatively for the PMOS use the same equation as for the NMOS

Keep in mind that $V_{tp} < 0$

**Cut off**

- $V_{gs} > V_{tp}$

**Triode**

- $V_{gs} \leq V_{tp}$
- $V_{ds} \geq V_{gs} - V_t$

**Pinch-off**

- $V_{gs} \leq V_{tp}$
- $V_{ds} \leq V_{gs} - V_t$
Finite Output Resistance in Saturation  
(Channel-length Modulation)

Due to channel length modulation:
- The Pinch-off point moves as a function of $V_{DS}$
- $I_d$ changes as a function of $V_{DS}$

\[
\frac{V_{ds} - VA}{V_{drop}} = -\frac{1}{\lambda}
\]

$V_{dsat} = V_{gs} - V_t$

$V_{drop} = V_{ds} - V_{dsat}$
Modified Current equation in Saturation

\[ i_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{gs} - V_t)^2 \left( 1 + \lambda V_{DS} \right) \]

\[ \lambda \propto \frac{1}{L} \]

\[ \lambda \] - empirical constant referred to as the channel length modulation
The Body Effect

- For an NMOS
- The substrate is usually connected to the most negative supply
- The Substrate behaves like a second gate
  i.e. The channel width can be controlled by the substrate voltage $V_B$
- The Threshold Voltage was defined as the gate-source voltage required to induce a channel when $V_{SB}=0$. For $V_{SB}$ not equal to zero we define:

$$V_T = V_{T_0} + \gamma (\sqrt{V_{SB} - 2\phi_F} - \sqrt{-2\phi_F})$$

$V_{T_0}$ = threshold voltage with $V_{SB} = 0$

$\phi_F = \text{physical parameter (work function fermi potential)}$

$$\phi_F = \phi_T \ln\left(\frac{N_A}{n_i^2}\right) \approx -0.3V$$

$$\gamma = \frac{\sqrt{2q\varepsilon_{si}N_A}}{C_{ox}} \approx 0.4V^{1/2}$$

$\therefore V_B \uparrow \Rightarrow V_T \uparrow \Rightarrow I_D \downarrow$
MOS Capacitance in Cutoff

Lateral diffusion (overlap) capacitance

\[ C_{GSO} = C_{GDO} = C_{ox} W x_d = C_o W \]

where \[ C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \]

Gate to Channel Capacitance

\[ C_{GCB} = C_{ox} W L \]

Total Gate Capacitance

\[ C_g = C_{ox} W L + 2 C_{ox} W x_d \]
$C_{sb}$ & $C_{db}$ are reverse bias junction capacitance

Recall \[ C_j = \frac{C_{jo}}{\sqrt{1 - \frac{V_D}{\phi_o}}} \] where $C_{jo} \propto A_D$

What is $A_D$ for the source (drain)?
\[ A_D = \text{Bottom Plate Area} + \text{Side Wall Area} \]
\[ = WL_s + (2L_s + W)x_j \]

\[ \therefore C_{sb} = C_j WL_s + (2L_s + W)C_{jsw} \]

$C_{jsw} = C_{jsw}'x_j$ where $C_{jsw}$ is the perimeter Capacitance per unit length
\( C_{\text{GCB}} = 0 \) since a channel exists. However, we have a gate to channel capacitance of \( C_{\text{ox}}WL \) which is divided equally between \( C_{\text{GCS}} \) and \( C_{\text{GCD}} \)

Therefore: \( C_{\text{GCS}} = C_{\text{GCD}} = \frac{C_{\text{ox}}WL}{2} \)

Total gate capacitance \( C_g = C_{\text{ox}}WL + 2C_{\text{ox}}Wx_d \)

\( C_{\text{sb}} = C_{\text{db}} = C_jWL_s + (2L_s + W)C_{jsw} \)

\( C_{\text{overlap}} \)

\( C_{\text{gs}} \)

\( C_{\text{gc}} \)

\( C_{\text{gd}} \)

\( C_{\text{db}} \)

\( C_{\text{overlap}} \)

Inversion Layer
Saturation

\[ C_{GCB} = 0 \]

\[ C_{GCS} = \frac{2}{3} C_{ox} WL \]

Total Gate Capacitance

\[ C_g = \frac{2}{3} C_{ox} W_{\text{eff}} + 2C_{ox} W x_d \]
### Average Distribution of Channel Capacitance

<table>
<thead>
<tr>
<th>Operation Region</th>
<th>( C_{GCB} )</th>
<th>( C_{GCS} )</th>
<th>( C_{GCD} )</th>
<th>( C_{GC} )</th>
<th>( C_G )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cutoff</td>
<td>( C_{ox} \cdot WL )</td>
<td>0</td>
<td>0</td>
<td>( C_{ox} \cdot WL )</td>
<td>( C_{ox} \cdot WL + 2C_o \cdot W )</td>
</tr>
<tr>
<td>Resistive</td>
<td>0</td>
<td>( C_{ox} \cdot WL/2 )</td>
<td>( C_{ox} \cdot WL/2 )</td>
<td>( C_{ox} \cdot WL )</td>
<td>( C_{ox} \cdot WL + 2C_o \cdot W )</td>
</tr>
<tr>
<td>Saturation</td>
<td>0</td>
<td>( (2/3)C_{ox} \cdot WL )</td>
<td>0</td>
<td>( (2/3)C_{ox} \cdot WL )</td>
<td>( (2/3)C_{ox} \cdot WL + 2C_o \cdot W )</td>
</tr>
</tbody>
</table>

- Channel capacitance components are nonlinear and vary with operating voltage.
- Most important regions are cutoff and saturation since that is where the device spends most of its time.
Transistor Capacitance Values for 0.25\(\mu\phi\)

\[
\begin{align*}
C_{\text{GSO}} &= C_{\text{GDO}} = C_{\text{ox}} x_d W = C_o W = \\
C_{\text{GC}} &= C_{\text{ox}} WL = \\
\text{so } C_{\text{gate\_cap}} &= C_{\text{ox}} WL + 2C_o W = \\
C_{\text{bp}} &= C_j L_S W = \\
C_{\text{sw}} &= C_{\text{jsw}} (2L_S + W) = \\
\text{so } C_{\text{diffusion\_cap}} &= \\

\begin{array}{|c|c|c|c|c|c|c|c|}
\hline
& C_{\text{ox}} & C_o & C_j & m_j & \phi_b & C_{\text{jsw}} & m_{\text{jsw}} & \phi_{\text{bsw}} \\
& (\text{fF/\(\mu\text{m}^2\)}) & (\text{fF/\(\mu\text{m}\)}) & (\text{fF/\(\mu\text{m}^2\)}) & & (\text{V}) & (\text{fF/\(\mu\text{m}\)}) & & (\text{V}) \\
\hline
\text{NMOS} & 6 & 0.31 & 2 & 0.5 & 0.9 & 0.28 & 0.44 & 0.9 \\
\hline
\text{PMOS} & 6 & 0.27 & 1.9 & 0.48 & 0.9 & 0.22 & 0.32 & 0.9 \\
\hline
\end{array}
\end{align*}
\]
Example 1 (long channel devices)

Determinate the region of operation and the drain current $I_d$

$$k_n' = 60 \mu A/V^2 \quad V_{ton} = 0.7V \quad \lambda = 0.1V^{-1}$$

$$k_p' = 20 \mu A/V^2 \quad V_{top} = -0.8V \quad \lambda = 0.1V^{-1}$$

where both $\frac{W}{L} = 1$

NMOS: $V_{GS} = 3.3V \quad V_{DS} = 2.2V$

PMOS: $V_{GS} = -3.3V \quad V_{DS} = -2.6$
Example 2 (long channel)

NMOS  \[ 0 < V_{SB} < 4V \]

\[ V_{To} = 1V \]

Find the range of \( V_t \) resulting of \( \gamma = 0.5V^{-0.5}|2\phi| = 0.6V \)
Example 3 (long channel)

Find all label current & voltage:

\[ \mu_n C_{ox} = 20 \mu A / V^2, \quad V_t = 1V, \quad \lambda = 0, \quad \gamma = 0, \quad V_{T1} = -1V \]

\[ W = 30 \mu m, \quad L = 10 \mu m \]

Note: M1 is a depletion-type NMOS Device with \( V_t = -1V \)
Example 4

Given an NMOS transistor with \( t_{ox} = 6 \text{ nm}, L = 0.24 \ \mu \text{m}, W = 0.36 \ \mu \text{m}, \)
\( L_D = L_S = 0.625 \ \mu \text{m}, \) \( C_0 = 3 \times 10^{-10} \text{ F/m}, \) \( C_{j0} = 2 \times 10^{-3} \text{ fF/m}^2, \)
\( C_{jsw} = 2.75 \times 10^{-10} \text{ F/m}, \) and \( \phi_0 = 0.65 \text{V} \)

Determine the value of all capacitances with the transistor is in the active region with \( V_g = 3 \text{V} \) and \( V_d = 1.5 \text{V}. \)