MOS Transistor Theory

- So far, we have viewed a MOS transistor as an ideal switch (digital operation)
  - Reality: less than ideal

Introduction

- So far, we have treated transistors as ideal switches
- An ON transistor passes a finite amount of current
  - Depends on terminal voltages
  - Derive current-voltage (I-V) relationships
- Transistor gate, source, drain all have capacitance
  - \( I = C \frac{\Delta V}{\Delta t} \to \Delta t = \frac{C}{I} \Delta V \)
  - Capacitance and current determine speed
- Also explore what a “degraded level” really means
MOS Transistor Theory

- Study conducting channel between source and drain
  - Modulated by voltage applied to the gate (voltage-controlled device)
  - nMOS transistor: majority carriers are electrons (greater mobility), p-substrate doped (positively doped)
  - pMOS transistor: majority carriers are holes (less mobility), n-substrate (negatively doped)

Terminal Voltages

- Mode of operation depends on $V_g$, $V_d$, $V_s$
  - $V_{gs} = V_g - V_s$
  - $V_{gd} = V_g - V_d$
  - $V_{ds} = V_d - V_s = V_{gs} - V_{gd}$
- Source and drain are symmetric diffusion terminals
  - By convention, source is terminal at lower voltage
  - Hence $V_{ds} \geq 0$
- nMOS body is grounded. First assume source is 0 too.
- Three regions of operation
  - Cutoff
  - Linear
  - Saturation
**Gate Biasing**

- $V_{gs} = 0$: no current flows from source to drain (insulated by two reverse biased pn junctions)
- $V_{gs} > 0$: electric field created across substrate
- Electrons accumulate under gate: region changes from p-type to n-type
- Conduction path between source and drain

---

**nMOS Device Behavior**

- $V_{gs} \ll V_t$: Accumulation mode
  - Enhancement-mode transistor: Conducts when gate bias $V_{gs} > V_t$
  - Depletion-mode transistor: Conducts when gate bias is zero

- $V_{gs} = V_t$: Depletion mode
- $V_{gs} > V_t$: Inversion mode

- $V_{gs} << V_t$: Polysilicon gate
- $V_{gs} = V_t$: Oxide insulator
- $V_{gs} > V_t$: Inversion Region (n-type)
nMOS Cutoff

- No channel
- $I_{ds} = 0$

nMOS Linear

- Channel forms
- Current flows from d to s
  - $e^-$ from s to d
- $I_{ds}$ increases with $V_{ds}$
- Similar to linear resistor
nMOS Saturation

- Channel pinches off
- $I_{ds}$ independent of $V_{ds}$
- We say current saturates
- Similar to current source

$I-V$ Characteristics

- In linear region, $I_{ds}$ depends on
  - How much charge is in the channel?
  - How fast is the charge moving?
Channel Charge

• MOS structure looks like parallel plate capacitor while operating in inversion
  – Gate – oxide – channel
• $Q_{\text{channel}} = CV$
  
  ![MOS Structure Diagram]

- Gate – oxide – channel
- $Q_{\text{channel}} = C(V_{gs} - V_{th})$
- $C = \frac{W \cdot L}{t_{ox}}$
Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel
- \( Q_{\text{channel}} = CV \)
- \( C = C_g = \varepsilon_{\text{ox}} WL/t_{\text{ox}} = C_{\text{ox}} WL \)
- \( V = V_{g-c} - V_t = (V_{g-s} - V_{d-s}/2) - V_t \)

\( C_{\text{ox}} = \varepsilon_{\text{ox}} / t_{\text{ox}} \)
Carrier velocity

- Charge is carried by e-
- Carrier velocity $v$ proportional to lateral E-field between source and drain
- $v = \mu E$

Carrier velocity

- Charge is carried by e-.
- Carrier velocity $v$ proportional to lateral E-field between source and drain.
  - $v = \mu E$, $\mu$ called mobility.
- $E = \frac{V_{ds}}{L}$
- Time for carrier to cross channel:
  - $t = \frac{L}{v}$
nMOS Linear I-V

- Now we know
  - How much charge $Q_{\text{channel}}$ is in the channel
  - How much time $t$ each carrier takes to cross

\[ I_{ds} = \frac{Q_{\text{channel}}}{t} \]
nMOS Linear I-V

- Now we know
  - How much charge $Q_{\text{channel}}$ is in the channel
  - How much time $t$ each carrier takes to cross

$$I_{ds} = \frac{Q_{\text{channel}}}{t}$$

$$= \mu C_{ox} \frac{W}{L} \left( V_{gs} - V_t \right) - \frac{V_{ds}}{2} V_{ds}$$

$$= \beta \left( V_{gs} - V_t - \frac{V_{ds}}{2} \right) V_{ds}$$

$\beta = \mu C_{ox} \frac{W}{L}$

nMOS Saturation I-V

- If $V_{gd} < V_t$, channel pinches off near drain
  - When $V_{ds} > V_{dss} = V_{gs} - V_t$

- Now drain voltage no longer increases current

$$I_{ds} =$$
nMOS Saturation I-V

- If $V_{gd} < V_t$, channel pinches off near drain
  - When $V_{ds} > V_{dsat} = V_{gs} - V_t$
- Now drain voltage no longer increases current
  \[ I_{ds} = \beta \left( V_{gs} - V_t - \frac{V_{dsat}}{2} \right) V_{dsat} \]
nMOS I-V Summary

- *Shockley* 1st order transistor models

\[ I_{ds} = \begin{cases} 
0 & V_{gs} < V_t \quad \text{cutoff} \\
\beta \left( V_{gs} - V_t - \frac{V_{ds}}{2} \right) V_{ds} & V_{ds} < V_{dsat} \quad \text{linear} \\
\frac{\beta}{2} (V_{gs} - V_t)^2 & V_{ds} > V_{dsat} \quad \text{saturation}
\end{cases} \]

Current-Voltage Relations

**Linear Region:** \( V_{DS} \leq V_{GS} - V_T \)

\[ I_D = \frac{\kappa'}{L} \left( V_{GS} - V_T \right) \left( V_{DS} - \frac{V_{DS}^2}{2} \right) \]

with

\[ \kappa' = \mu_n c_{ox} = \frac{\mu_n k_{FS}}{t_{ox}} \quad \text{Process Transconductance Parameter} \]

**Saturation Mode:** \( V_{DS} \geq V_{GS} - V_T \)

\[ I_D = \frac{\kappa'}{2L} \left( V_{GS} - V_T \right)^2 (1 + \lambda' V_{DS}) \]
Current-Voltage Relations

\[ k'_n: \text{transconductance of transistor} \]
\[ \frac{W}{L}: \text{width-to-length ratio} \]

• As W increases, more carriers available to conduct current

• As L increases, \( V_{ds} \) diminishes in effect (more voltage drop). Takes longer to push carriers across the transistor, reducing current flow

Example

• For a 0.6 \( \mu \)m process
  – From AMI Semiconductor
  – \( t_{ox} = 100 \) Å
  – \( \mu = 350 \text{ cm}^2/\text{V} \cdot \text{s} \)
  – \( V_t = 0.7 \text{ V} \)
• Plot \( I_{ds} \) vs. \( V_{ds} \)
  – \( V_{gs} = 0, 1, 2, 3, 4, 5 \)
  – Use \( W/L = 4/2 \lambda \)

\[ \beta = \mu C_m \frac{W}{L} = (350) \left( \frac{3.9 \times 8.85 \times 10^{-14}}{100 \times 10^{-9}} \right) \left( \frac{W}{L} \right) = 120 \frac{W}{L} \mu A/V^2 \]
pMOS I-V

- All dopings and voltages are inverted for pMOS
- Mobility $\mu_p$ is determined by holes
  - Typically 2-3x lower than that of electrons $\mu_n$
  - 120 cm$^2$/V*s in AMI 0.6 μm process
- Thus pMOS must be wider to provide same current
  - In this class, assume $\mu_n / \mu_p = 2$ to 3

Capacitance

- Any two conductors separated by an insulator have capacitance
- Gate to channel capacitor is very important
  - Creates channel charge necessary for operation
- Source and drain have capacitance to body
  - Across reverse-biased diodes
  - Called diffusion capacitance because it is associated with source/drain diffusion
Gate Capacitance

- Approximate channel as connected to source
- \( C_{gs} = \varepsilon_{ox} WL/t_{ox} = C_{ox} WL = C_{\text{permicron}} W \)
- \( C_{\text{permicron}} \) is typically about 2 fF/µm

![Gate Capacitance Diagram]

The Gate Capacitance

![The Gate Capacitance Diagram]
Diffusion Capacitance

- $C_{sb}$, $C_{db}$
- Undesirable, called *parasitic* capacitance
- Capacitance depends on area and perimeter
  - Use small diffusion nodes
  - Comparable to $C_g$ for contacted diff
  - $\frac{1}{2} C_g$ for uncontacted
  - Varies with process

\[
C_{\text{diff}} = C_{\text{bottom}} + C_{\text{sw}} = C_j \times \text{AREA} + C_{jw} \times \text{PERIMETER} = C_j L_s W + C_{jw} (2L_s + W)
\]
Parasitic Resistances

\[ R_S = \left( \frac{L_S}{W} \right) R_\sigma + R_C \]

\[ R_D = \left( \frac{L_D}{W} \right) R_\sigma + R_C \]

- \( R_C \): contact resistance
- \( R_\sigma \): sheet resistance per square of drain-source diffusion

Body Effect

- Many MOS devices on a common substrate
  - Substrate voltage of all devices are normally equal
- But several devices may be connected in series
  - Increase in source-to-substrate voltage as we proceed vertically along the chain

- Net effect: slight increase in threshold voltage \( V_{t_2} > V_{t_1} \)
Pass Transistors

- We have assumed source is grounded
- What if source > 0?
  - e.g. pass transistor passing $V_{DD}$
  
\[
\begin{align*}
V_{DD} & \\
\end{align*}
\]

- $V_g = V_{DD}$
  - If $V_s > V_{DD} - V_t$, $V_{gs} < V_t$
  - Hence transistor would turn itself off

- nMOS pass transistors pull no higher than $V_{DD} - V_{th}$
  - Called a degraded “1”
  - Approach degraded value slowly (low $I_{ds}$)

- pMOS pass transistors pull no lower than $V_{tp}$

\[
\begin{align*}
V_{DD} & \\
\end{align*}
\]
Pass Transistor Ckts

\[ V_s = V_{DD} - V_{tn} \]

\[ V_s = |V_{tp}| \]

\[ V_s = V_{DD} - 2V_{tn} \]
Effective Resistance

- Shockley models have limited value
  - Not accurate enough for modern transistors
  - Too complicated for much hand analysis
- Simplification: treat transistor as resistor
  - Replace $I_{ds}(V_{ds}, V_{gs})$ with effective resistance $R$
    - $I_{ds} = V_{ds}/R$
  - $R$ averaged across switching of digital gate
- Too inaccurate to predict current at any given time
  - But good enough to predict RC delay

RC Delay Model

- Use equivalent circuits for MOS transistors
  - Ideal switch + capacitance and ON resistance
  - Unit nMOS has resistance $R$, capacitance $C$
  - Unit pMOS has resistance $2R$, capacitance $C$
- Capacitance inversely proportional to width
- Resistance inversely proportional to width
## RC Values

- **Capacitance**
  - $C = C_g = C_s = C_d = 2 \text{ fF/\mu m of gate width}$
  - Values similar across many processes
- **Resistance**
  - $R \approx 6 \text{ K\Omega in 0.6\mu m process}$
  - Improves with shorter channel lengths
- **Unit transistors**
  - May refer to minimum contacted device ($4/2 \lambda$)
  - Or maybe 1 \text{ \mu m wide device}
  - Doesn’t matter as long as you are consistent

## Activity

1) If the width of a transistor increases, the current will
   - **increase**  
   - **decrease**  
   - **not change**

2) If the length of a transistor increases, the current will
   - **increase**  
   - **decrease**  
   - **not change**

3) If the supply voltage of a chip increases, the maximum transistor current will
   - **increase**  
   - **decrease**  
   - **not change**

4) If the width of a transistor increases, its gate capacitance will
   - **increase**  
   - **decrease**  
   - **not change**

5) If the length of a transistor increases, its gate capacitance will
   - **increase**  
   - **decrease**  
   - **not change**

6) If the supply voltage of a chip increases, the gate capacitance of each transistor will
   - **increase**  
   - **decrease**  
   - **not change**
DC Response

- DC Response: $V_{\text{out}}$ vs. $V_{\text{in}}$ for a gate
- Ex: Inverter
  - When $V_{\text{in}} = 0$ $\Rightarrow$ $V_{\text{out}} = V_{\text{DD}}$
  - When $V_{\text{in}} = V_{\text{DD}}$ $\Rightarrow$ $V_{\text{out}} = 0$
  - In between, $V_{\text{out}}$ depends on transistor size and current
  - By KCL, must settle such that $I_{\text{dsn}} = |I_{\text{dop}}|$
  - We could solve equations
  - But graphical solution gives more insight

Transistor Operation

- Current depends on region of transistor behavior
- For what $V_{\text{in}}$ and $V_{\text{out}}$ are nMOS and pMOS in
  - Cutoff?
  - Linear?
  - Saturation?
### nMOS Operation

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Linear</th>
<th>Saturated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ V_{gsn} &lt;$</td>
<td>$ V_{gsn} &gt; $</td>
<td>$ V_{gsn} &gt; $</td>
</tr>
<tr>
<td>$ V_{dsn} &lt;$</td>
<td>$ V_{dsn} &gt; $</td>
<td></td>
</tr>
</tbody>
</table>

\[ V_{DPD} \]
\[ V_{in} \rightarrow I_{ds} V_{out} \]
\[ I_{ds} \]

### nMOS Operation

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Linear</th>
<th>Saturated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ V_{gsn} &lt; V_{tn} $</td>
<td>$ V_{gsn} &gt; V_{tn} $</td>
<td>$ V_{gsn} &gt; V_{tn} $</td>
</tr>
<tr>
<td>$ V_{dsn} &lt; V_{gsn} - V_{tn} $</td>
<td>$ V_{dsn} &gt; V_{gsn} - V_{tn} $</td>
<td>$ V_{dsn} &gt; V_{gsn} - V_{tn} $</td>
</tr>
</tbody>
</table>

\[ V_{DPD} \]
\[ V_{in} \rightarrow I_{ds} V_{out} \]
\[ I_{ds} \]
### nMOS Operation

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Linear</th>
<th>Saturated</th>
</tr>
</thead>
<tbody>
<tr>
<td>( V_{gsn} &lt; V_{tn} )</td>
<td>( V_{gsn} &gt; V_{tn} )</td>
<td>( V_{gsn} &gt; V_{tn} )</td>
</tr>
<tr>
<td>( V_{dsn} &lt; V_{gsn} - V_{tn} )</td>
<td>( V_{dsn} &lt; V_{gsn} - V_{tn} )</td>
<td>( V_{dsn} &gt; V_{gsn} - V_{tn} )</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
V_{gsn} &= V_{in} \\
V_{dsn} &= V_{out}
\end{align*}
\]
### pMOS Operation

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Linear</th>
<th>Saturated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{gsp} &gt;$</td>
<td>$V_{gsp} &lt;$</td>
<td>$V_{gsp} &lt;$</td>
</tr>
<tr>
<td>$V_{dsp} &gt;$</td>
<td>$V_{dsp} &lt;$</td>
<td>$V_{dsp} &lt;$</td>
</tr>
</tbody>
</table>

![Diagram of pMOS Operation](image)

### pMOS Operation

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Linear</th>
<th>Saturated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{gsp} &gt; V_{tp}$</td>
<td>$V_{gsp} &lt; V_{tp}$</td>
<td>$V_{gsp} &lt; V_{tp}$</td>
</tr>
<tr>
<td>$V_{dsp} &gt; V_{gsp} - V_{tp}$</td>
<td>$V_{dsp} &lt; V_{gsp} - V_{tp}$</td>
<td>$V_{dsp} &lt; V_{gsp} - V_{tp}$</td>
</tr>
</tbody>
</table>

![Diagram of pMOS Operation](image)
### pMOS Operation

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Linear</th>
<th>Saturated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{gsp} &gt; V_{tp}$</td>
<td>$V_{gsp} &lt; V_{tp}$</td>
<td>$V_{gsp} &lt; V_{tp}$</td>
</tr>
<tr>
<td>$V_{dsp} &gt; V_{gsp} - V_{tp}$</td>
<td>$V_{dsp} &lt; V_{gsp} - V_{tp}$</td>
<td>$V_{dsp} &lt; V_{gsp} - V_{tp}$</td>
</tr>
</tbody>
</table>

- $V_{gsp} = V_{in} - V_{DD}$
- $V_{tp} < 0$
- $V_{DSP} = V_{out} - V_{DD}$

**Diagram:**

- $V_{DD}$
- $V_{in}$
- $V_{out}$
- $I_{dsn}$
- $I_{dsp}$
- $V_{out} < V_{in} - V_{tp}$
- $V_{gsp} = V_{in} - V_{DD}$
- $V_{dsp} = V_{out} - V_{DD}$
- $V_{tp} < 0$
I-V Characteristics

- Make pMOS is wider than nMOS such that \( \beta_n = \beta_p \)

DC Transfer Curve

- Transcribe points onto \( V_{\text{in}} \) vs. \( V_{\text{out}} \) plot
Operating Regions

- Revisit transistor operating regions

<table>
<thead>
<tr>
<th>Region</th>
<th>nMOS</th>
<th>pMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Cutoff</td>
<td>Linear</td>
</tr>
<tr>
<td>B</td>
<td>Saturation</td>
<td>Linear</td>
</tr>
<tr>
<td>C</td>
<td>Saturation</td>
<td>Saturation</td>
</tr>
<tr>
<td>D</td>
<td>Linear</td>
<td>Saturation</td>
</tr>
<tr>
<td>E</td>
<td>Linear</td>
<td>Cutoff</td>
</tr>
</tbody>
</table>

V_{in} \quad V_{DS} \quad V_{out} \quad V_{DD}
Beta Ratio

- If $\beta_p / \beta_n \neq 1$, switching point will move from $V_{DD}/2$
- Called *skewed* gate
- Other gates: collapse into equivalent inverter

\[ \frac{\beta_p}{\beta_n} \neq 1 \]

![Beta Ratio Diagram](image)

Noise Margins

- How much noise can a gate input see before it does not recognize the input?

![Noise Margins Diagram](image)
Logic Levels

- To maximize noise margins, select logic levels at
  - unity gain point of DC transfer characteristic

\[ \beta_p/\beta_n > 1 \]

\[ V_{DD} \]

\[ V_{in} \]

\[ V_{out} \]

\[ V_{OH} \]

\[ V_{OL} \]

\[ V_{IL} \]

\[ V_{IH} \]

\[ t_{ns} \]

\[ S = -1 \]

\[ \beta_p/\beta_n > 1 \]

\[ V_{DD} - |V_{tp}| \]