Overview

- **Reading**
  B. Razavi Chapter 2.

- **Introduction**
  In studying the design of integrated circuits, one of two extreme approaches can be taken:
  - Begin with quantum mechanics and understand solid-state, semiconductor device physics, device modeling, and finally the design of circuits.
  - Treat each semiconductor device as a black box whose behavior is described in terms of its terminal voltages and currents and design circuits with little attention to the internal operation of the device.
General considerations

- MOSFET as a switch
  - If the gate voltage, $V_G$, is "high", the transistor "connects" the source and the drain together.
  - If the gate voltage, $V_G$, is "low", the transistor "isolates" the source and the drain.

- MOSFET structure

MOS device

- Simple NMOS device
- Simple PMOS device

- Cross section of CMOS n-well technology

- MOS symbols
Operation of MOSFET

- A MOSFET driven by a gate voltage
- Formation of depletion region

- Onset of inversion
- Formation of inversion layer

Threshold voltage

- In semiconductor physics, the $V_{TH}$ of an NFET is defined as the gate voltage for which the interface is "as much $n$-type as the substrate is $p$-type".
- The threshold voltage can be provided that

$$V_{TH} = \phi_{MS} + 2\phi_F \frac{Q_{dep}}{C_{ox}}$$

- $\phi_{MS}$ is the difference between the work functions of the polysilicon gate and the silicon substrate.
- $\phi_F = (kT / q)\ln(N_{sub} / n_i)$, $q$ is electron charge, $N_{sub}$ is the doping concentration of the substrate.
- $Q_{dep}$ is the charge in the depletion region.
- $C_{ox}$ is the gate oxide capacitance per unit area.
I/V characteristics – Triode region

Condition: $V_{DS} \leq V_{GS} - V_{TH}$

- Equal source and drain voltage
- Unequal source and drain voltage

$Q_d = W C_{ox} (V_{GS} - V_{TH}), \quad C_{ox} : F / m^2$

$Q_d(x) = W C_{ox} [V_{GS} - V(x) - V_{TH}]$

- Derivation

$\int_{V_{DS}}^{V_{GS}} I_D(x) \cdot dx = \int_{V_{DS}}^{V_{GS}} W C_{ox} \mu_n [V_{GS} - V(x) - V_{TH}] \cdot dx$

where $V(x) = \frac{V_{DS}}{L} x$

$\Rightarrow I_D = \mu_n C_{ox} \frac{W}{L} [V_{GS} - V_{TH}] \left[V_{DS} - \frac{1}{2} V_{DS}^2\right]$ and $I_{D,max} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2$

Drain current versus drain-source voltage

Linear operation in deep triode region

With the condition $V_{GS} \ll 2(V_{GS} - V_{TH})$

$R_{on} = \frac{1}{\mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})}$
I/V characteristics – Saturation region

Condition: \( V_{DS} \leq V_{GS} - V_{TH} \)

- Saturated MOSFETs operating as current sources

Pinch-off behavior

\[
I_D = \frac{1}{2} \mu_s C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2
\]

Transconductance

- For amplification purposes, the transconductance of the device is calculated

\[
g_m = \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{GS}=\text{constant}} = \mu_s C_{ox} \frac{W}{L} (V_{GS} - V_{TH})
\]

\[
= \sqrt{2} \mu_s C_{ox} \frac{W}{L} I_D = \frac{2I_D}{V_{GS} - V_{TH}}
\]

- MOS transconductance as a function
Conceptual visualization of saturation and triode regions

Second-order effects

- Body effect
  
  $V_{TH} = V_{TH0} + \gamma(\sqrt{2q\phi_f} - V_{TH0} - \sqrt{2\phi_f})$

  where $\gamma = \sqrt{2q\phi_f} N_{sub}/C_{ox}$

  denotes the body effect coefficient.
Example: Source follower with (a) no body effect and (b) body effect

Ignoring body effect:
As \( V_{in} \) varies, \( V_{out} \) closely follows the input because the drain current remains equal to \( I_1 \). It can be written by
\[
I_1 = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{in} - V_{out} - V_{TH})^2
\]

Channel-length modulation

\( L' \) is a function of \( V_{DS} \):
Writing \( L' = L - \Delta L \), i.e., \( 1/L' \approx (1 + \Delta L / L) / L \), and assuming \( \Delta L / L = \lambda V_{DS} \).

\[
I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS})
\]

\( \lambda \) : the channel-length modulation coefficient
Subthreshold conduction, $V_{GS} < V_{TH}$

$$I_D = I_o \exp \left( \frac{V_{GS}}{\zeta V_T} \right)$$

where $\zeta > 1$ is a nonideality factor and $V_T = kT/q$.

- If $W$ increases while $I_D$ remains constant, then $V_{GS} \rightarrow V_{TH}$ and the device enters the subthreshold region. As a result, the transconductance is calculated to $g_m = I_D / (\zeta V_T)$, revealing that MOSFETs are inferior to bipolar transistors.

- The exponential dependence of $I_D$ upon $V_{GS}$ in subthreshold operation may suggest the use of MOS devices in this regime so as to achieve a higher gain. However, since such conditions are met by only a large device width or low drain current, the speed of subthreshold circuits is severely limited.

MOS devices

- Bird’s eye of a MOS device

- Vertical views of a MOS device
MOS device capacitances

- Oxide capacitance between the gate and the channel, \( C_1 = WLC_{ox} \).
- Depletion capacitance between the channel and the substrate, \( C_2 = WL\sqrt{q\epsilon_{ox}N_{sub} / (2\pi\phi_F)} \).
- Capacitance due to the overlap of the gate poly with the source and drain areas, \( C_3 \) and \( C_4 \). The overlap capacitance per unit width is denoted by \( C_{ov} \).
- Junction capacitance between the source/drain areas and the substrate.
  - bottom-plate capacitance associated with the bottom of the junction, \( C_j = C_{p} / [1 + V_{p}/(2\phi_F)]^m \).
  - sidewall capacitance due to the perimeter of the junction, \( C_{jSW} \) (note \( C_j : \text{F/m}^2, C_{jSW} : \text{F/m} \)).

\[
C_{jSW} = \frac{W}{2}EC_j + 2\left(\frac{W}{2} + E\right)C_{jSW}
\]

Example: Calculate the source and drain junction capacitances

Folded structure:

\[
C_{jSW} = \frac{W}{2}EC_j + 2\left(\frac{W}{2} + E\right)C_{jSW}
\]

The geometry of the folded structure in Fig. (b) exhibits substantially less drain junction capacitance than that in Fig. (a) while providing the same \( W/L \).
Variation of $C_{GS}$ and $C_{GD}$ versus $V_{GS}$

- Example:

For $V_X \approx 0$, $M_1$ is in the triode region, $C_{2V} = C_{EF} = (1/2)WLC_{ox} + WC_{ov}$ and $C_{FB}$ is maximum.

MOS small-signal model

- Basic MOS small-signal model

$$g_m = \frac{\partial I_D}{\partial V_{GS}}$$

- Channel-length modulation represented by a depend current source

- Channel-length modulation represented by a resistor

$$r_a = \frac{\partial V_{gs}}{\partial I_D} = \frac{1}{I_D} \frac{\partial I_D}{\partial V_{gs}} = \frac{1}{I_D} \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 \lambda = \frac{1}{I_D}$$
Body effect represented by a dependent current source

In the saturation region, \( g_{mb} \) can be expressed as

\[
g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH}) \left( -\frac{\partial V_{TH}}{\partial V_{BS}} \right)
\]

We also have

\[
\frac{\partial V_{TH}}{\partial V_{BS}} = -\frac{\partial V_{TH}}{\partial V_{SB}} = \frac{2}{2} (2\phi_F + V_{SB})^{1/2}
\]

Thus

\[
g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = g_m \cdot \frac{\gamma}{2\sqrt{2\phi_F + V_{SB}}} = \eta g_m
\]

where \( \eta = \frac{g_{mb}}{g_m} \).

Example: Sketch \( g_m \) and \( g_{mb} \) of \( M \) as a function of the bias current \( I_1 \).

Since \( g_m = \sqrt{2\mu_n C_{ox} (W/L)} I_D \), we have \( g_m \propto \sqrt{I_1} \).

The dependence of \( g_{mb} \) upon \( I_1 \) is less straightforward.

As \( I_1 \) increases, \( V_X \) decreases and so does \( V_{SB} \).
- Complete MOS small-signal model

- Reduction of gate resistance by folding

---

**MOS SPICE models**

<table>
<thead>
<tr>
<th>Table 2.1</th>
<th>Level 1 SPICE Models for NMOS and PMOS Devices.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NMOS Model</strong></td>
<td></td>
</tr>
<tr>
<td>LEVEL = 1</td>
<td>VTO = 0.7</td>
</tr>
<tr>
<td>NSUB = 9e+14</td>
<td>LD = 0.08e–6</td>
</tr>
<tr>
<td>TOX = 9e–9</td>
<td>PB = 0.9</td>
</tr>
<tr>
<td>MJ = 0.45</td>
<td>MJSW = 0.2</td>
</tr>
<tr>
<td><strong>PMOS Model</strong></td>
<td></td>
</tr>
<tr>
<td>LEVEL = 1</td>
<td>VTO = –0.8</td>
</tr>
<tr>
<td>NSUB = 5e+14</td>
<td>LD = 0.09e–6</td>
</tr>
<tr>
<td>TOX = 9e–9</td>
<td>PB = 0.9</td>
</tr>
<tr>
<td>MJ = 0.5</td>
<td>MJSW = 0.3</td>
</tr>
</tbody>
</table>
The parameters of Spice models are defined as below:

- VTO: threshold voltage with zero $V_{th}$ (unit: V)
- GAMMA: body effect coefficient (unit: $V^{1/2}$)
- PHI: $2\Phi_F$ (unit: V)
- TOX: gate oxide thickness (unit: m)
- NSUB: substrate doping (unit: cm$^{-3}$)
- LD: source/drain side diffusion (unit: m)
- UO: channel mobility (unit: cm$^2$/V/s)
- LAMBDA: channel-length modulation coefficient (unit: $V^{-1}$)
- CJ: source/drain bottom-plate junction capacitance per unit area (unit: F/m$^2$)
- CJSW: source/drain sidewall junction capacitance per unit length (unit: F/m)
- PB: source/drain junction built-in potential (unit: V)
- MJ: exponent in CJ equation (unitless)
- MJSW: exponent in CJSW equation (unitless)
- CGDO: gate-drain overlap capacitance per unit width (unit: F/m)
- CGSO: gate-source overlap capacitance per unit width (unit: F/m)
- JS: source/drain leakage current per unit area (unit: A/m$^2$)