MOS Transistors Chapter 3 (pp. 98-99)

- Channel types: n-channel (NMOS) and p-channel (PMOS).

Slide 2.8: Simplicity comes at a cost; however, comparing the I-V curves produced by the model to those of the actual device (BSIM4 SPICE models), a large discrepancy can be observed for intermediate values of $V_{DS}$ (around $V_{DS}=1$). When using the model for the derivation of propagation delays (performance of a CMOS gate), accuracy in this section of the overall operation region is not that crucial. What is most important is that the change of current at the highest values of $V_{DS}$ and $V_{GS}$ are predicted correctly — as these predominantly determine the charge and discharge states of the output capacitor. Hence, the propagation delay error is only a couple of percent, which is only a small penalty for a major reduction in model complexity.

Thresholds and Sub-Threshold Current

- Drain current vs. gate-source voltage:
  - $V_{DS} = 1.2V$
  - $I_{DS} = I_{TH} = I_{sub}$

For long-channel NMOS,

- $I_{DS} = 0$ for $V_{GS} < V_{TH}$ (threshold voltage)
- $I_{DS} = I_{DSS} = V_{GS} - V_{TH}$ for $V_{GS} > V_{TH}$

\[ I_{DS} = I_{DSS} = V_{GS} - V_{TH} \]

\[ V_{TH} = \sqrt{2} \cdot N_{A} \cdot E_{F} / C_{ox}, \quad C_{ox} = \frac{E_{ox}}{V_{th}} \]
For long channel nMOSFETs,

**Linear region**
\[ I_{DS} = \frac{\mu A_{ox} W}{L} \left( V_{GS} - V_{TH} \right) \left( V_{DS} - V_{TH} \right) \]

**Saturation region**
\[ I_{DS} = \frac{\mu A_{ox} W}{L} \left( V_{GS} - V_{TH} \right) \left( V_{DS} - V_{TH} \right) \]

\[ V_{DS} = V_{DS} \] when
\[ \frac{V_{GS} - V_{TH}}{1 - \lambda} = V_{DS} \] where \( \lambda \) is empirical parameter

With channel length modulation

\[ I_{DS} \] linear \( \frac{\mu A_{ox} W}{L} \left( V_{GS} - V_{TH} \right) \left( V_{DS} - V_{TH} \right) \)

\[ I_{DS \text{Sat}} = \frac{\mu A_{ox} W}{L} \left( V_{GS} - V_{TH} \right) \left( 1 + \frac{V_{DS} - V_{TH}}{V_{TH}} \right) \]
Similarly for PMOS:

\[ I_{DS\,\text{linear}} = \frac{K \cdot W}{2} \left( \frac{V_{GS} - V_T}{2V_{th}} \right)^2 \]  
\[ I_{DS\,\text{sat}} = \frac{K \cdot W}{2} \left( \frac{V_{GS} - V_T}{2V_{th}} \right)^2 \left( 1 + \frac{V_{GS} - V_T}{2V_{th}} \right) \]

which is same as

\[ \frac{K \cdot W}{2} \left[ \frac{1}{2} \left( \frac{V_{GS} - V_T}{2V_{th}} \right)^2 \right] \]

\[ I_{P} = \frac{V_{GS}}{2} \left( V_{GS} - V_T \right)^2 \]

I-V Equations for Short Channel MOSFETs

\[ n_{\text{LOF}} = \frac{n_{\infty}}{1 + \left( \frac{V_{GS} - V_T}{V \beta} \right)^n} \]  
\[ \beta = \text{empirical coefficient} \]
\[ n_{\infty} = \text{low-field electron mobility} \]

For NMOS:

\[ I_{DS\,\text{linear}} = \frac{K \cdot W}{2} \left( \frac{V_{GS} - V_T}{2V_{th}} \right)^2 \]

\[ I_{DS\,\text{sat}} = \frac{K \cdot W}{2} \left( \frac{V_{GS} - V_T}{2V_{th}} \right)^2 \left( 1 + \frac{V_{GS} - V_T}{2V_{th}} \right) \]

where \( V_{th} = \text{channel electric field} \)

Threshold Voltage of Small Geometry Devices

\[ V_T = V_{to} + k_1 \left( V_{GS} - V_T \right) + k_2 \left( V_{GS} - V_T \right)^2 \]

\[ V_T = V_{to} + k_1 \left( V_{GS} - V_T \right) + k_2 \left( V_{GS} - V_T \right)^2 \]

Similarity for PMOS:

\[ (3.57) \quad \& \quad (3.58) \]

Sub-threshold Current

- Sub-threshold behavior can be modeled physically

\[ I_{th} = I_{th0} \left( 1 - \frac{V_{GS}}{V_{th}} \right)^{a} \]

where \( a \) is the slope factor (typically around 1.6) and \( I_{th0} = \frac{W}{2L} \frac{K}{q} \)

- Very often expressed in base 10

\[ I_{th} = I_{th0} \left( \frac{1}{10} \right)^{a} \left( \frac{V_{GS}}{V_{th}} \right)^{a} \]

where \( a = \left( \frac{1}{10} \right)^{a} \), the sub-threshold slope, ranging between 80 mV and 110 mV
**Alpha Power Law Model**

- Alternate approach, useful for hand analysis of propagation delay
  
  \[ I_{on} = \frac{W}{2L}C_{ox}(V_{GD} - V_{TH}) \]

- Parameter \( \alpha \) is between 1 and 2.

- In 65-180 nm CMOS technology, \( \alpha \approx 1.2-1.3 \)

  - This is not a physical model
  - Simply empirical:
    - \( C_{ox} \) can vary
    - \( V_{GD} \) can vary
    - \( V_{TH} \) can vary

**Device Innovation Innovations**

- Strained silicon
- Silicon-on-insulator
- Dual-gate devices
- Very high mobility devices
- MEMS — transistors

**References**

- Sakuma, J., et al. (1992)

---

**Gate-Leakage Mechanisms**

- Direct tunneling
- Indirect tunneling

**Slide 2.26**

Gate leakage finds its source in two different mechanisms: Fowler-Nordheim (FN) tunneling, and direct-oxide tunneling. FN tunneling is an effect that has been effectively used in the design of nonvolatile memories, and is already quite substantial for oxide thicknesses larger than 10 nm. Its onset requires high electric-field strengths, though. With reducing oxide thicknesses, tunneling starts to occur at far lower field strengths. The dominant effect under these conditions

---

**Slide 2.45**

FinFETs — An Entirely New Device Architecture

UC Berkeley, 1999

- Suppressed short-channel effects
- Higher on-current for reduced leakage
- Undoped channel — No random dopant fluctuations

---

**Slide 2.49**

Silicon-on-insulator (SOI)

- Improved speed (15-20%)
- Improved leakage (10-20%)
- Higher on-current

**Slide 2.49**

Silicon-on-insulator (SOI)

- Improved speed (15-20%)
- Improved leakage (10-20%)
- Higher on-current

---

**Slide 2.26**

Silicon-on-insulator (SOI)

- Improved speed (15-20%)
- Improved leakage (10-20%)
- Higher on-current

---

**Slide 2.45**

FinFETs — An Entirely New Device Architecture

UC Berkeley, 1999

- Suppressed short-channel effects
- Higher on-current for reduced leakage
- Undoped channel — No random dopant fluctuations

[Ref. X. Wang, 2000]

which forms the body of the device. The dimensions of the fins determine the effective channel length of the transistor. The devices have shown the potential to scale the channel length to values that are hard, if not impossible, to accomplish in traditional planar devices. In fact, operational transistors with channel lengths down to 7 nm have been demonstrated.

In addition to a suppression of deep submicron effects, a crucial advantage of the device is again increased control of the gate arrays (almost completely around the channel).