Chapter 6  MOSFET in the On-state

The MOSFET (MOS Field-Effect Transistor) is the building block of Gb memory chips, GHz microprocessors, analog, and RF circuits.

Match the following MOSFET characteristics with their applications:

- small size
- high speed
- low power
- high gain
6.1 Introduction to the MOSFET

Basic MOSFET structure and IV characteristics

What is desirable: large $I_{on}$, small $I_{off}$
6.1 Introduction to the MOSFET

Two ways of representing a MOSFET:

- Circuit Symbol
- Simple Switch
Invention of the Field-Effect Transistor

Jan. 28, 1930.
J. E. LILJENFELD

1,745,175

METHOD AND APPARATUS FOR CONTROLLING ELECTRIC CURRENTS

Filed Oct. 8, 1926

Semiconductor Devices for Integrated Circuits (C. Hu)
In 1935, a British patent was issued to Oskar Heil. A working MOSFET was not demonstrated until 1955. Using today’s terminology, what are 1, 2, and 6?
Gate oxides as thin as 1.2 nm can be manufactured reproducibly. Large tunneling current through the oxide limits oxide-thickness reduction.
6.2 Complementary MOSFETs

**NFET**

When \( V_g = V_{dd} \), the NFET is on and the PFET is off.

When \( V_g = 0 \), the PFET is on and the NFET is off.

**PFET**

When \( V_g = V_{dd} \), the NFET is on and the PFET is off.

When \( V_g = 0 \), the PFET is on and the NFET is off.
A CMOS inverter is made of a PFET *pull-up device* and a NFET *pull-down device*. $V_{out} = ?$ if $V_{in} = 0$ V.
CMOS (Complementary MOS) Inverter

- NFET and PFET can be fabricated on the same chip.

- basic layout of a CMOS inverter

Vin Vout

V_{dd} 0V

N-well P+

P-substrate N+N+ P+
P+ N+ N+ P+ P+

Contact

PFET NFET

V_{in} V_{out}

0V
6.3 Surface Mobilities of Electrons and Holes

How to measure the surface mobility:

\[ I_{ds} = W \cdot Q_{inv} \cdot v = WQ_{inv} \mu_{ns} \varepsilon = WQ_{inv} \mu_{ns} V_{ds} / L \]

\[ = WC_{oxe} (V_g - V_t) \mu_{ns} V_{ds} / L \]
Surface mobility is a function of the average of the fields at the bottom and the top of the inversion charge layer, $\mu_b$ and $\mu_t$.

From Gauss’s Law,

$$\mu_b = -\frac{Q_{dep}}{\varepsilon_s}$$

$$V_t = V_{fb} + \phi_{st} - \frac{Q_{dep}}{C_{oxe}}$$

Therefore,

$$\mu_b = \frac{C_{oxe}}{\varepsilon_s} (V_t - V_{fb} + \phi_{st})$$

$$\mu_t = -\frac{Q_{dep} + Q_{inv}}{\varepsilon_s}$$

$$= \mu_b - \frac{Q_{inv}}{\varepsilon_s} = \mu_b + \frac{C_{oxe}}{\varepsilon_s} (V_{gs} - V_t)$$

$$= \frac{C_{oxe}}{\varepsilon_s} (V_{gs} - V_{fb} + \phi_{st})$$

$$\therefore \frac{1}{2} (\mu_b + \mu_t) = \frac{C_{oxe}}{2\varepsilon_s} (V_{gs} + V_t - 2V_{fb} - 2\phi_{st})$$

$$\approx \frac{C_{oxe}}{2\varepsilon_s} (V_{gs} + V_t + 0.2 \text{ V})$$

$$= \frac{V_{gs} + V_t + 0.2 \text{ V}}{6T_{oxe}}$$
Mobility is a function of $V_{gs}$, $V_t$, and $T_{oxe}$.

What suppresses the surface mobility:
- phonon scattering
- coulombic scattering
- surface roughness scattering
**EXAMPLE:** What is the surface mobility at $V_{gs} = 1 \text{ V}$ in an N-channel MOSFET with $V_t = 0.3 \text{ V}$ and $T_{ox_e} = 2 \text{ nm}$?

**Solution:**

\[
\frac{(V_{gs} + V_t + 0.2)}{6T_{ox_e}} = \frac{1.5 \text{ V}}{12 \times 10^{-7} \text{ cm}} = 1.25 \text{ MV/cm}
\]

1 MV is a megavolt ($10^6 \text{ V}$). From the mobility figure, $\mu_{ns} = 190 \text{ cm}^2/\text{Vs}$, which is several times smaller than the bulk mobility.
6.4 MOSFET $V_t$ and the Body Effect

How to Measure the $V_t$ of a MOSFET

$V_t$ is measured by extrapolating the $I_{ds}$ versus $V_{gs}$ curve to $I_{ds} = 0$.

$$I_{dsat} = \frac{W}{L} C_{ox} (V_{gs} - V_t) \mu_n s V_{ds} \propto V_{gs} - V_t$$
6.4 MOSFET $V_t$ and the Body Effect

$$C_{dep} = \frac{\varepsilon_s}{W_{d_{max}}}$$

- Two capacitors $\Rightarrow$ two charge components

$$Q_{inv} = -C_{oxe}(V_{gs} - V_t) + C_{dep}V_{sb} = -C_{oxe}(V_{gs} - (V_t + \frac{C_{dep}}{C_{oxe}}V_{sb}))$$

- Redefine $V_t$ as

$$V_t(V_{sb}) = V_{t0} + \frac{C_{dep}}{C_{oxe}}V_{sb} = V_{t0} + \alpha V_{sb}$$
6.4 **MOSFET $V_t$ and the Body Effect**

- **body effect:**
  
  $V_t$ is a function of $V_{sb}$

- **body effect coefficient:**
  
  $\alpha = \frac{C_{dep}}{C_{oxe}} = \frac{3T_{oxe}}{W_{dmax}}$

*When the source-body junction is reverse-biased, the NFET $V_t$ increases and the PFET $V_t$ becomes more negative.*

Is the body effect a good thing? How can it be reduced?
Retrograde Body Doping Profiles

- $W_{dep}$ does not vary with $V_{sb}$.
- Retrograde doping is popular because it reduces off-state leakage.
Uniform Body Doping

When the source/body junction is reverse-biased, there are two quasi-Fermi levels \( (E_{fn} \text{ and } E_{fp}) \) which are separated by \( qV_{sb} \). An NMOSFET reaches threshold of inversion when \( E_c \) is close to \( E_{fn} \), not \( E_{fp} \). This requires the band-bending to be \( 2\phi_B + V_{sb} \), not \( 2\phi_B \).

\[
V_t = V_{t0} + \frac{\sqrt{qN_a 2\varepsilon_s}}{C_{oxe}} (\sqrt{2\phi_B + V_{sb}} - \sqrt{2\phi_B})
\]

\[
\equiv V_{t0} + \gamma (\sqrt{2\phi_B + V_{sb}} - \sqrt{2\phi_B})
\]

\( \gamma \) is the body-effect parameter.
6.5 $Q_{\text{inv}}$ in MOSFET

- Channel voltage
  
  \[ V_c = V_s \text{ at } x = 0 \text{ and } V_c = V_d \text{ at } x = L. \]

- \( Q_{\text{inv}} = -C_{\text{oxe}}(V_{gs} - V_{cs} - V_{t0} - \alpha (V_{sb} + V_{cs})) \]
  
  \[ = -C_{\text{oxe}}(V_{gs} - V_{cs} - (V_{t0} + \alpha V_{sb}) - \alpha V_{cs}) \]
  
  \[ = -C_{\text{oxe}}(V_{gs} - mV_{cs} - V_{t}) \]

- \( m \equiv 1 + \alpha = 1 + 3T_{\text{oxe}}/W_{\text{dmax}} \)

\( m \) is called the \textit{body-effect factor} or \textit{bulk-charge factor}
6.6 Basic MOSFET IV Model

\[ I_{ds} = WQ_{inv} = WQ_{inv} \mu_{ns} \]
\[ = W C_{oxe} (V_{gs} - m V_{cs} - V_t) \mu_{ns} dV_{cs} / dx \]
\[ \int_{0}^{L} I_{ds} dx = W C_{oxe} \mu_{ns} \int_{0}^{V_{ds}} (V_{gs} - m V_{cs} - V_t) dV_{cs} \]
\[ I_{ds} L = W C_{oxe} \mu_{ns} (V_{gs} - V_t - m V_{ds} / 2) V_{ds} \]

\[ I_{ds} = \frac{W}{L} C_{oxe} \mu_s (V_{gs} - V_t - \frac{m}{2} V_{ds}) V_{ds} \]
\[ \frac{dI_{ds}}{dV_{ds}} = 0 = \frac{W}{L} C_{oxe} \mu_n s (V_{gs} - V_t - mV_{ds}) \]

\[ V_{dsat} = \frac{V_{gs} - V_t}{m} \]

\( V_{dsat} : \text{Drain Saturation Voltage} \)

L = 10 \mu m, W = 10 \mu m

\( T_{oxe} = 4 \text{nm}, V_t = 0.3 \text{V} \)
\[ I = \mu_n Q_{inv} \frac{dV_{cs}}{dx} \]

\[ Q_{inv} = C_{ox} (V_g - m V_{cs} - V_t) \]

\[ V_{ds} = V_{dsat} \]

\[ V_{ds} > V_{dsat} \]

\[ V_{ds} = V_{dsat} \]

\[ V_{ds} \]

\[ E_c \]

source

drain

\[ E_c \]

source

drain

\[ V_{ds} - V_{dsat} \]
Saturation Current and Transconductance

- linear region, saturation region

\[ I_{dsat} = \frac{W}{2mL} C_{oxe} \mu_n s (V_{gs} - V_t)^2 \]

- transconductance: \( g_m = \frac{dI_{ds}}{dV_{gs}} \)

\[ g_{msat} = \frac{W}{mL} C_{oxe} \mu_n s (V_{gs} - V_t) \]

- \( L = 10\mu m, \ W = 10\mu m \)
- \( T_{oxe} = 4nm, \ V_t = 0.3V \)
6.7.1 CMOS Inverter Voltage Transfer Curve
– Regeneration of Digital Signal

Vin
PFET
NFET
S
D
D
S
Vout

Idd

Vin = 2V
Vin = 1.5V
Vin = 1V
Vin = 0.5V
Vin = 0V

Vout

Vin

Vdd

Idd (mA)

Vds (V)

0.2
0.1
0.5V
1V
1.5V
2V

0V

Vout (V)

0 0.5 1.0 1.5 2.0

0 0.5 1.0 1.5 2.0

0 0.5 1.0 1.5 2.0

0 0.5 1.0 1.5 2.0

Semiconductor Devices for Integrated Circuits (C. Hu)
6.7.2 CMOS Inverter Delay

\[ V_{d1} \]

\[ V_{1} \quad V_{2} \quad V_{3} \]

\[ C \]

\[ 2 \tau_{d} \]

\[ \tau_{d} : \text{propagation delay} \]

\[ V_{dd} \]

\[ V_{1} \quad V_{2} \quad V_{3} \]
6.7.2 CMOS Inverter Delay

\[ \tau_d \equiv \frac{1}{2} (\text{pull-down delay} + \text{pull-up delay}) \]

\[ \text{pull-up delay} \approx \frac{CV_{dd}}{2I_{dsatP}} \]

\[ \text{pull-down delay} \approx \frac{CV_{dd}}{2I_{dsatN}} \]

\[ \tau_d = \frac{CV_{dd}}{4} \left( \frac{1}{I_{dsatN}} + \frac{1}{I_{dsatP}} \right) \]

\[ R_N \text{ and } R_P = \frac{V_{dd}}{2I_{on}} = \frac{V_{dd}}{2I_{dsat}(|V_g| = V_{dd})} \]

How can the speed of an inverter circuit be improved?
6.7.3 CMOS Power Consumption

\[ P_{\text{dynamic}} = V_{dd} \times \text{average current} = CV_{dd}^2 f \]

\[ P_{\text{static}} = V_{dd} I_{\text{off}} \]

\[ P_{\text{direct-path}} \approx V_{dd} I_{\text{sat}} \frac{t_r + t_f}{2} f = 0.2CV_{dd}^2 f \]

\[ = 0.2P_{\text{dynamic}} \]

Total power consumption

\[ P = 1.2CV_{dd}^2 f + V_{dd} I_{\text{off}} \]
Logic Gates

This two-input NAND gate and many other logic gates are extensions of the inverter.
6.8 Velocity Saturation

\[
\nu = \frac{\mu_s \mu}{1 + \frac{\mu}{\mu_{sat}}}
\]

\[
\mu \ll \mu_{sat} : \nu = \mu_s \mu
\]

\[
\mu \gg \mu_{sat} : \nu = \mu_s \mu_{sat}
\]

- velocity saturation has large and deleterious effect on the \( I_{on} \) of MOSFETS
6.9 MOSFET IV Model with Velocity Saturation

\[ I_{ds} = WQ_{inv} v \]

\[ I_{ds} = W C_{oxe} (V_{gs} - m V_{cs} - V_t) \frac{\mu_{ns} dV_{cs}}{dx} \frac{dV_{cs}}{dx} / \zeta_{sat} \]

\[ \int_{0}^{L} I_{ds} \, dx = \int_{0}^{V_{ds}} [W C_{oxe} \mu_{ns} (V_{gs} - m V_{cs} - V_t) - I_{ds} / \zeta_{sat}] dV_{cs} \]

\[ I_{ds} L = W C_{oxe} \mu_{ns} (V_{gs} - V_t - \frac{m}{2} V_{ds}) V_{ds} - I_{ds} V_{ds} / \zeta_{sat} \]
6.9 MOSFET IV Model with Velocity Saturation

\[ I_{ds} = \frac{W}{L} C_{oxe} \mu_{ns} (V_{gs} - V_t - \frac{m}{2} V_{ds}) V_{ds} \]

\[ 1 + \frac{V_{ds}}{\varepsilon_{sat} L} \]

\[ I_{ds} = \frac{\text{long-channel } I_{ds}}{1 + V_{ds} / \varepsilon_{sat} L} \]
6.9 **MOSFET IV Model with Velocity Saturation**

Solving \( \frac{dI_{ds}}{dV_{ds}} = 0 \),

\[
V_{dsat} = \frac{2(V_{gs} - V_t) / m}{1 + \sqrt{1 + 2(V_{gs} - V_t) / m \varepsilon_{sat} L}}
\]

A simpler and more accurate \( V_{dsat} \) is:

\[
\frac{1}{V_{dsat}} = \frac{m}{V_{gs} - V_t} + \frac{1}{\varepsilon_{sat} L}
\]

\( \varepsilon_{sat} \equiv \frac{2\nu_{sat}}{\mu_s} \)
EXAMPLE: Drain Saturation Voltage

Question: At $V_{gs} = 1.8 \text{ V}$, what is the $V_{dsat}$ of an NFET with $T_{oxe} = 3 \text{ nm}$, $V_t = 0.25 \text{ V}$, and $W_{dmax} = 45 \text{ nm}$ for (a) $L = 10 \mu\text{m}$, (b) $L = 1 \mu\text{m}$, (c) $L = 0.1 \mu\text{m}$, and (d) $L = 0.05 \mu\text{m}$?

Solution: From $V_{gs}$, $V_t$, and $T_{oxe}$, $\mu_{ns}$ is $200 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$.

\[
\varepsilon_{sat} = \frac{2v_{sat}}{\mu_{es}} = 8 \times 10^4 \text{ V/cm}
\]

\[
m = 1 + 3\frac{T_{oxe}}{W_{dmax}} = 1.2
\]

\[
V_{dsat} = \left(\frac{m}{V_{gs} - V_t} + \frac{1}{\varepsilon_{sat}L}\right)^{-1}
\]
EXAMPLE: Drain Saturation Voltage

\[ V_{dsat} = \left( \frac{m}{V_{gs} - V_t} + \frac{1}{\varepsilon_{sat} L} \right)^{-1} \]

(a) \( L = 10 \, \mu m \), \( V_{dsat} = \left( \frac{1}{1.3V} + \frac{1}{80V} \right)^{-1} = 1.3 \, V \)

(b) \( L = 1 \, \mu m \), \( V_{dsat} = \left( \frac{1}{1.3V} + \frac{1}{8V} \right)^{-1} = 1.1 \, V \)

(c) \( L = 0.1 \, \mu m \), \( V_{dsat} = \left( \frac{1}{1.3V} + \frac{1}{.8V} \right)^{-1} = 0.5 \, V \)

(d) \( L = 0.05 \, \mu m \), \( V_{dsat} = \left( \frac{1}{1.3V} + \frac{1}{.4V} \right)^{-1} = 0.3 \, V \)
**$I_{dsat}$ with Velocity Saturation**

Substituting $V_{dsat}$ for $V_{ds}$ in $I_{ds}$ equation gives:

$$I_{dsat} = \frac{W}{2mL} C_{oxe} \mu_s \frac{(V_{gs} - V_t)^2}{1 + \frac{V_{gs} - V_t}{m \varepsilon_{sat} L}} = \text{long-channel } I_{dsat}$$

$$= \frac{V_{gs} - V_t}{1 + \frac{V_{gs} - V_t}{m \varepsilon_{sat} L}}$$

**Very short channel case:** $\varepsilon_{sat} L \ll V_{gs} - V_t$

$$I_{dsat} = \frac{W}{2} C_{oxe} \mu_s \varepsilon_{sat} (V_{gs} - V_t)$$

$$= W V_{sat} C_{oxe} (V_{gs} - V_t - \varepsilon_{sat} L)$$

- $I_{dsat}$ is proportional to $V_{gs} - V_t$ rather than $(V_{gs} - V_t)^2$, not as sensitive to $L$ as $1/L$. 

Semiconductor Devices for Integrated Circuits (C. Hu)  Slide 6-35
What is the main difference between the $V_{gs}$ dependences of the long- and short-channel length IV curves?
The PMOS IV is qualitatively similar to the NMOS IV, but the current is about half as large. How can we design a CMOS inverter so that its voltage transfer curve is symmetric?
6.10 **Parasitic Source-Drain Resistance**

- If $I_{dsat0} \propto V_g - V_t$, $I_{dsat} = \frac{I_{dsat0}}{1 + \frac{I_{dsat0} R_s}{(V_{gs} - V_t)}}$

- $I_{dsat}$ is reduced by about 15% in a 0.1\(\mu\)m MOSFET.

- $V_{dsat} = V_{dsat0} + I_{dsat} (R_s + R_d)$
Definitions of Channel Length

\[ L \equiv L_{\text{drawn}} - \Delta L \]
6.11 Extraction of the Series Resistance and the Effective Channel Length

\[ I_{ds} = \frac{WC_{oxx} \mu_s V_{ds}}{L_{\text{drawn}} - \Delta L} (V_{gs} - V_t) \]

\[ V_{ds} = \frac{I_{ds} (L_{\text{drawn}} - \Delta L)}{WC_{oxx} (V_{gs} - V_t) \mu_s} \]

Include series resistance, \( R_{ds} \equiv R_d + R_s \),

\[ \frac{V_{ds}}{I_{ds}} = R_{ds} + \frac{L_{\text{drawn}} - \Delta L}{WC_{oxx} (V_{gs} - V_t) \mu_s} \]
6.12 Source Injection Velocity Limit

- Carrier velocity is limited by the thermal velocity when they first enter the channel from the source.

- \[ I_{dsat} = WBv_{thx}Q_{inv} = WBv_{thx}C_{oxe}(V_{gs} - V_{t}) \]
6.13 Chapter Summary

- **body effect**
  \[ V_t (V_{sb}) = V_{t0} + \alpha V_{sb} \]
  for steep retrograde body doping
  \[ \alpha = \frac{3T_{oxe}}{W_{dmax}} \]

- **basic \( I_{ds} \) model**
  \[ I_{ds} = \frac{W}{L} C_{oxe} \mu_s (V_{gs} - V_t - \frac{m}{2} V_{ds}) V_{ds} \]
  \[ m = 1 + \frac{3T_{oxe}}{W_{dmax}} \approx 1.2 \]
  - Small \( \alpha \) and \( m \) are desirable. Therefore, small \( T_{oxe} \) is good.
  - Ch.7 shows that large \( W_{dmax} \) is not acceptable.
  - CMOS circuit speed is determined by \( CV_{dd}/I_{dsat} \), and its power by \( CV_{dd}^2f + V_{dd}I_{off} \).
6.13 Chapter Summary

IV characteristics can be divided into a linear region and a saturation region. $I_{ds}$ saturates at:

$$V_{dsat} = \frac{V_{gs} - V_t}{m}$$

$$I_{dsat} = \frac{W}{2mL} C_{ox} \mu_s (V_{gs} - V_t)^2$$

transconductance:

$$g_{msat} = \frac{W}{mL} C_{ox} \mu_s (V_{gs} - V_t)$$

Considering velocity saturation,

$$V_{dsat} = \left( \frac{m}{V_{gs} - V_t} + \frac{1}{\varepsilon_{sat} L} \right)^{-1}$$

$$I_{dsat} = \frac{\text{long-channel } I_{dsat}}{1 + \frac{V_{gs} - V_t}{m \varepsilon_{sat} L}}$$