OUTLINE
Using pn-diodes to isolate transistors in an IC
The metal-oxide-field-effect transistor (MOSFET)
  Structure of the MOSFET
  The MOSFET as a controlled resistance
  Pinch-off and current saturation in the MOSFET
  Channel-length modulation
  Velocity saturation in a short-channel MOSFET

Reading
  Rabaey et al.  Ch. 3.3.1-3.3.2
  Hambley   Ch. 12.1
Why are pn Junctions Important for ICs?

• The basic building block in digital ICs is the MOS transistor, whose structure contains reverse-biased diodes.
  – pn junctions are important for electrical isolation of transistors located next to each other at the surface of a Si wafer.
  – The junction capacitance of these diodes can limit the performance (operating speed) of digital circuits
Device Isolation using pn Junctions

No current flows if voltages are applied between n-type regions, because two pn junctions are “back-to-back”

=> n-type regions isolated in p-type substrate and vice versa
Figure 0.1 Example of a densely populated integrated circuit – the DRAM
We can build large circuits consisting of many transistors without worrying about current flow between devices. The p-n junctions isolate the transistors because there is always at least one reverse-biased p-n junction in every potential current path.
Modern Field Effect Transistor (FET)

- An electric field is applied normal to the surface of the semiconductor (by applying a voltage to an overlying “gate” electrode), to modulate the conductance of the semiconductor

→ Modulate drift current flowing between 2 contacts (“source” and “drain”) by varying the voltage on the “gate” electrode

Metal-oxide-semiconductor (MOS) FET:
A GATE electrode is placed above (electrically insulated from) the silicon surface, and is used to control the resistance between the SOURCE and DRAIN regions.

- **NMOS**: N-channel Metal Oxide Semiconductor
- \( L \) = channel length
- \( W \) = channel width

**MOSFET**

"Metal" (heavily doped poly-Si)

\[ \text{oxide insulator} \]

\[ \text{p-type silicon} \]

\[ \text{source} \]

\[ \text{drain} \]
• Without a gate-to-source voltage applied, no current can flow between the source and drain regions.

• Above a certain gate-to-source voltage (*threshold voltage* $V_T$), a conducting layer of mobile electrons is formed at the Si surface beneath the oxide. These electrons can carry current between the source and drain.
N-channel vs. P-channel MOSFETs

- For current to flow, $V_{GS} > V_T$
- Enhancement mode: $V_T > 0$
- Depletion mode: $V_T < 0$
  - Transistor is ON when $V_G=0V$

- For current to flow, $V_{GS} < V_T$
- Enhancement mode: $V_T < 0$
- Depletion mode: $V_T > 0$
  - Transistor is ON when $V_G=0V$

(“n+” denotes very heavily doped n-type material; “p+” denotes very heavily doped p-type material)
MOSFET Circuit Symbols

(a) NMOS transistor as 4-terminal device
(b) NMOS transistor as 3-terminal device
(a) PMOS transistor as 4-terminal device
(d) PMOS transistor as 3-terminal device
Figure 0.1 Schematic symbol and water model for a p-channel MOSFET
The voltage applied to the GATE terminal determines whether current can flow between the SOURCE & DRAIN terminals.

- For an n-channel MOSFET, the SOURCE is biased at a lower potential (often 0 V) than the DRAIN (Electrons flow from SOURCE to DRAIN when $V_G > V_T$)

- For a p-channel MOSFET, the SOURCE is biased at a higher potential (often the supply voltage $V_{DD}$) than the DRAIN (Holes flow from SOURCE to DRAIN when $V_G < V_T$)

The BODY terminal is usually connected to a fixed potential.

- For an n-channel MOSFET, the BODY is connected to 0 V
- For a p-channel MOSFET, the BODY is connected to $V_{DD}$
**NMOSFET $I_G$ vs. $V_{GS}$ Characteristic**

Consider the current $I_G$ (flowing into G) versus $V_{GS}$:

The gate is insulated from the semiconductor, so there is no significant steady gate current.

always zero!
The MOSFET as a Controlled Resistor

- The MOSFET behaves as a resistor when $V_{DS}$ is low:
  - Drain current $I_D$ increases linearly with $V_{DS}$
  - Resistance $R_{DS}$ between SOURCE & DRAIN depends on $V_{GS}$
    - $R_{DS}$ is lowered as $V_{GS}$ increases above $V_T$

**NMOSFET Example:**

- $I_D = 0$ if $V_{GS} < V_T$
- $V_{GS} = 1 \, \text{V} > V_T$
- $V_{GS} = 2 \, \text{V}$

Inversion charge density $Q_i(x) = -C_{ox}[V_{GS}-V_T-V(x)]$

where $C_{ox} \equiv \frac{\varepsilon_{ox}}{t_{ox}}$
Sheet Resistance Revisited

Consider a sample of n-type semiconductor:

\[ R_s = \frac{\rho}{t} = \frac{1}{\sigma t} = \frac{1}{q \mu_n n t} = \frac{1}{\mu_n Q_n} \]

where \( Q_n \) is the charge per unit area.
Next consider $I_D$ (flowing into D) versus $V_{DS}$, as $V_{GS}$ is varied:

Above threshold ($V_{GS} > V_T$): “inversion layer” of electrons appears, so conduction between S and D is possible

Below “threshold” ($V_{GS} < V_T$): no charge $\rightarrow$ no conduction
MOSFET as a Controlled Resistor (cont’d)

\[ I_D = \frac{V_{DS}}{R_{DS}} \]

\[ R_{DS} = R_s \left( \frac{L}{W} \right) = \frac{L/W}{\mu_n Q_i} = \frac{L/W}{\mu_n C_{ox} (V_{GS} - V_T - \frac{V_{DS}}{2})} \]

\[ I_D = \mu_n C_{ox} \frac{W}{L} \left( V_{GS} - V_T - \frac{V_{DS}}{2} \right) V_{DS} \]

We can make \( R_{DS} \) low by
- applying a large “gate drive” \( (V_{GS} - V_T) \)
- making \( W \) large and/or \( L \) small
Charge in an N-Channel MOSFET

\( V_{GS} < V_T: \)

- Depletion region
- No inversion layer at surface

\( V_{GS} > V_T: \)

- Inversion layer
- \( I_D = WQ_{inv} \nu \)
  \[ = WQ_{inv} \mu_n E \]
  \[ = WQ_{inv} \mu_n \left( \frac{V_{DS}}{L} \right) \]

Average electron velocity \( \nu \) is proportional to lateral electric field \( E \).
What Happens at Larger $V_{DS}$?

$V_{GS} > V_T$:

$V_{DS} = V_{GS} - V_T$

Inversion-layer is “pinched-off” at the drain end

$V_{DS} > V_{GS} - V_T$

As $V_{DS}$ increases above $V_{GS} - V_T \equiv V_{DSAT}$, the length of the “pinched-off” region $\Delta L$ increases:

- “extra” voltage ($V_{DS} - V_{Dsat}$) is dropped across the distance $\Delta L$
- the voltage dropped across the inversion-layer “resistor” remains $V_{Dsat}$

$\Rightarrow$ the drain current $I_D$ saturates

Note: Electrons are swept into the drain by the $E$-field when they enter the pinch-off region.
Summary of $I_D$ vs. $V_{DS}$

- As $V_{DS}$ increases, the inversion-layer charge density at the drain end of the channel is reduced; therefore, $I_D$ does not increase linearly with $V_{DS}$.

- When $V_{DS}$ reaches $V_{GS} - V_T$, the channel is “pinched off” at the drain end, and $I_D$ saturates (i.e. it does not increase with further increases in $V_{DS}$).

$$I_{DSAT} = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_T)^2$$
$I_D$ vs. $V_{DS}$ Characteristics

The MOSFET $I_D$-$V_{DS}$ curve consists of two regions:

1) Resistive or “Triode” Region: $0 < V_{DS} < V_{GS} - V_T$

\[
I_D = k'_n \frac{W}{L} \left[ V_{GS} - V_T - \frac{V_{DS}}{2} \right] V_{DS}
\]

where $k'_n = \mu_n C_{ox}$

process transconductance parameter

2) Saturation Region:

$V_{DS} > V_{GS} - V_T$

\[
I_{DSAT} = \frac{k'_n}{2} \frac{W}{L} (V_{GS} - V_T)^2
\]

where $k'_n = \mu_n C_{ox}$

“CUTOFF” region: $V_G < V_T$
If $L$ is small, the effect of $\Delta L$ to reduce the inversion-layer “resistor” length is significant

$\rightarrow I_D$ increases noticeably with $\Delta L$ (i.e. with $V_{DS}$)

Channel-Length Modulation

\[ I_D = I'_D(1 + \lambda V_{DS}) \]

$\lambda$ is the slope

$I'_D$ is the intercept

$\Delta L \sim L$

$\Delta L \ll L$
Velocity Saturation

At high electric fields, the average velocity of carriers is NOT proportional to the field; it saturates at \( \sim 10^7 \) cm/sec for both electrons and holes:
Current Saturation in Modern MOSFETs

• In digital ICs, we typically use transistors with the shortest possible gate-length for high-speed operation.

• In a very short-channel MOSFET, $I_D$ saturates because the carrier velocity is limited to $\sim 10^7$ cm/sec.
Consequences of Velocity Saturation

1. $I_D$ is lower than that predicted by the mobility model

2. $I_D$ increases **linearly** with $V_{GS} - V_T$ rather than quadratically in the saturation region

\[
I_{DSAT} = WC_{ox} \left[ V_{GS} - V_T - \frac{V_{DSAT}}{2} \right] v_{sat}
\]

where $V_{DSAT} = \frac{L}{\mu_n} v_{sat}$
P-Channel MOSFET $I_D$ vs. $V_{DS}$

- As compared to an n-channel MOSFET, the signs of all the voltages and the currents are reversed:

Note that the effects of velocity saturation are less pronounced than for an NMOSFET. Why is this the case?