6.1 INTRODUCTION
6.2 BASIC DEVICE CHARACTERISTICS
6.3 NONUNIFORM DOPING AND BURIED-CHANNEL DEVICE
6.4 DEVICE SCALING AND SHORT-CHANNEL EFFECTS
MOSFETs

- Speed
- Performance
- Density

Fig. 1 Minimum gate dimension in commercial integrated circuit as a function of the year of production.
MOSFETs

FET – Field Effect
PET – Potential Effect

Fig. 2 Distinction between (a) field-effect transistor (FET) and (b) potential-effect transistor (PET).

Fig. 3 Family tree of field-effect transistors (FETs).

All differ on how the gate capacitor is formed.
Why FET

• Good for applications in analog switching, high-input-impedance amplifiers, and microwave amplifiers

• Higher input impedance than bipolar transistors

• Thermally stable
  Negative temperature coefficient at high current levels preventing thermal runaway

• High switching speed
  No minority-carrier storage as a result of no forward-biased p-n junctions

• Linear devices
  Intermodulation and cross-modulation products are smaller than those of bipolar transistors
FET Variants

N-channel MOSFET
Formed by electrons and more conductive with $+V_g$

P-channel MOSFET
Formed by Holes and more conductive with $-V_g$

$V_g = 0$, Mosfet is
- **Enhancement Mode = Normally off**
  Low channel conductance. Must apply $V_g$ to form a conductive channel

- **Depletion Mode = Normally on**
  High channel conductance. Must apply $V_g$ to turn device off
FET Variant

Surface Inversion ~ 5nm thick

Buried channels
Free from surface effects, scattering
Better carrier mobility
Minus - Lower and variable conductance from gate-channel distance
FET Structure

Source and drain – ion implantation
SiO₂ gate dielectric – Thermal oxidation
Gate electrode - heavily doped polysilicon or polysilicon + silicide

Basic Parameters:
- Channel Length, L
- Channel Width, Z
- Insulator thickness, d
- Junction depth, rj
- substrate doping, Na

* Field oxide to distinguish it from the gate oxide or a trench filled with insulator to electrically isolate it from adjacent devices
**MOSFET characteristic**

**Inversion Charge**

Equilibrium conditions

1) In fig b, Flat band
   \[ V_g = V_d = V_{bs} = 0 \]

2) In fig c, Gate bias
   \[ V_d = V_{bs} = 0, \; V_g \uparrow 0 \]

Quasi-fermi conditions

\[ V_g, \; V_d \downarrow 0 \]

for inversion to take place at the drain

\[ \Psi_s \; \text{is} \; \sim \; \text{twice} \; \Psi_b \] (weak inversion)

\[ \Psi_b = \text{fermi level from intrinsic fermi level} \]
MOSFETs

MOSFET characteristic

Inversion Charge

Equilibrium case
Surface depletion region reaches a maximum width $W_{Dm}$, at inversion

Quasi-fermi case
Depletion-layer width is deeper than $W_{Dm}$ and is a function of the drain bias $V_d$,

For strong inversion to occur

$$\psi_s(\text{inv}) \approx V_D + 2\psi_B$$
Current-Voltage Characteristics

Basic MOSFET characteristics under the following idealized conditions

(1) It has an ideal gate structure with no interface traps nor mobile oxide charge
(2) Only drift current will be considered;
(3) doping in the channel is uniform
(4) little or no reverse leakage current
(5) the transverse field (x-direction) in the channel is >> than the longitudinal field (y-direction).

Inversion charge

ignore condition (1), then \( V_g \), is replaced by \( V_g - V_{FB} \)

\[
|Q_n(y)| = [V_G - V_{FB} - \Delta \psi(y) - 2\psi_B]C_{ox} - \sqrt{2\varepsilon_s q N_A[\Delta \psi(y) + 2\psi_B]}
\]
For a given $V_g$, the $I_d$ first increases linearly with $V_d$ (the linear region), then gradually levels off (the nonlinear region), and finally approaching a saturated value (the saturation region).
MOSFETs

**MOSFET characteristic**

**Constant Mobility**

**Condition 1**

If an applied $+V_g > *V_T$ (large enough to cause an inversion at the semiconductor surface.)

Then apply a small drain $V_d$ (current will flow from the source to the drain through the conducting channel)

**Effect**: The channel acts as a resistor, $I_d \propto V_d$. This is the linear region.

*$V_T$ is the gate bias beyond flat-band just starting to induce an inversion charge sheet
MOSFET characteristic

Constant Mobility

**Condition 2**

**Effect**: The inversion charge at the drain end $Q_{n,L} = 0$

Pinch off point occurs there because the relative voltage between the gate and the semiconductor is reduced

Practically, $Q_n \approx 0$ because of high field and high carrier velocity

Increase $V_D$ (the charge near the drain end is reduced by the channel potential $\Psi_i$)
**Constant Mobility**

**Condition 3**

**Effect:** the pinch-off point starts to move toward the source, but the pinch-off point voltage is the same

change in L increases $I_d$ only when the shortened amount is a substantial fraction of the channel length

Increase $V_d$ more… ($V_d > V_{sat}$)
MOSFETs

MOSFET characteristic

Velocity-Field Relationship

Reduce Channel length $L$, the internal longitudinal field $E_y$ in the channel also increases as a result.

Recall:

Mobility $u$ is defined as $v/E$.

Linear graph = Constant mobility = long channel from low field $>> E_y$, $v_s$ saturates

$v_s$ at room temp for Si = $1E7$ cm/s

In btw constant mobility and saturated $v_s$

**Best fit:**

$n = 2$ (electrons), $n = 1$ (holes)
- Field-Dependent Mobility: Two-Piece Linear Approximation
- Field-Dependent Mobility: Empirical Formula
- Velocity Saturation
- Ballistic Transport

In ultra-short channel lengths whose dimensions are on the order of or shorter than the mean free path, channel carriers do not suffer from scattering. They can gain energy from the field.

They acquire a velocity \( \gg v_s \)

At room temperature, \( v_s \) almost equal the thermal velocity \( v_{th} \) for silicon
- **Threshold Voltage**

$V_T$ is the gate bias beyond flat-band just starting to induce an inversion charge sheet and is given by the sum of voltages across the semiconductor ($2\psi_B$) and the gate material.

Its shift from $V_{FB}$ = zero as a result of:
1) Fixed oxide charge, $Q_f$
2) the work-function difference $\Phi_{ms}$ between the gate material and the semiconductor.

\[
V_T = V_{FB} + 2\psi_B + \frac{\sqrt{2\varepsilon_s q N_A (2\psi_B)}}{C_{ox}}
\]

By applying a –ve substrate bias (negative for n-channel or p-substrate),

\[
V_T = V_{FB} + 2\psi_B + \frac{\sqrt{2\varepsilon_s q N_A (2\psi_B - V_{BS})}}{C_{ox}}
\]

Practically, the difference between equations (1) and (2) should be low as possible by:
- low substrate doping
- thin oxide thickness
MOSFETs

MOSFET characteristic

Measuring Threshold Voltage
use the linear region by applying a small $V_d$
(a) Plot $I_D$ vs $V_G$

Swing $S$ (inverse of sub-threshold slope)
Quantifies how sharply the transistor is turned off by the $V_g$
(b) defined as $\Delta V_g$ required to induce a $\Delta I_d$ one order of magnitude

Fig. 15 Transfer characteristics ($I_D$ vs. $V_G$) in the linear region ($V_D \ll V_G$). (a) $I_D$ in linear scale to deduce $V_T$. Deviation from linearity at higher $V_G$ is due to lower mobility. (b) $I_D$ in logarithmic scale to show subthreshold swing.
Sub-threshold Region

1) When the gate bias below the threshold
2) the semiconductor surface is in weak inversion or depletion
It tells how sharply the current drops with gate bias

Importance
For low-voltage, low-power applications, such as when the MOSFET is used as a switch in digital logic and memory applications

For a sharp sub-threshold slope (small Swing), it becomes advisable to have
- low channel doping
- thin oxide thickness
- low interface-trap density
- low-temperature operation
Temperature Dependence

1) Mobility  (2) Threshold voltage  (3) sub-threshold characteristics

The effective mobility in inversion layer is such that gives rise to higher current and trans-conductance at lower temperature.

**Fig. 18** Threshold-voltage shift \( (dV_T/dT) \) of a Si-SiO\(_2\) system at room temperature vs. substrate doping, with oxide thickness \( d \) as a parameter.
Temperature Dependence

The MOSFET characteristics improve, as temperature decreases, especially in the subthreshold region.

Low temperature operation offers
Subthreshold swing $S$,
Higher mobility,… thus,
Higher current and
Trans-conductance, lower power consumption,
Lower junction leakage current, And
Lower metal-line resistance

Demerit
Special coolant to operate

Fig. 19 Subthreshold characteristics for a long-channel MOSFET ($L = 9 \, \mu m$) with temperature as a parameter. (After Ref. 36.)
Non Uniform Doping and Buried Channel Device

In non uniform doping profiles, threshold can be analyzed using

\[ V_{FB} + 2 \psi_B + \frac{q}{\varepsilon_{ox}} \int_0^{W_{Dm}} N(x) \, dx \]

**Fig. 20** Nonuniform channel doping profiles. (a) High-low profile. (b) Low-high (retrograde) profile. (c) – (d) Their approximations using step profiles.
Non Uniform Doping

High-Low Profile

$x_s \rightarrow$ step depth

$= \text{sum of projected range and standard deviation of original implant}$

As $x_s$ get wider or equal to $W_{Dm}$,
The surface region can be considered a uniform doped region with a higher concentration

If $W_{Dm} > x_s$

$$V_T = V_{FB} + 2\psi_B + \frac{qN_B W_{Dm} + q\Delta N x_s}{C_{ox}}$$

$$W_{Dm} = \sqrt{\frac{2\varepsilon_s}{qN_B} \left(2\psi_B - \frac{q\Delta N x_s^2}{2\varepsilon_s}\right)}$$

*from the eqtns, a higher surface concentration $N_B$ decreases $W_{Dm}$ but increases $V_T$
Non Uniform Doping
High-Low Profile

For a typical case, the threshold voltage depends on the implanted dose $D_I$ and the centroid of the dose $x_c$

$$D_I = \int_0^{W_{Dm}} \Delta N(x) \, dx,$$

$$x_c = \frac{1}{D_I} \int_0^{W_{Dm}} x \Delta N(x) \, dx$$

$$V_T = V_{FB} + 2 \psi_B + \frac{1}{C_{ox} N_B} \sqrt{2q \varepsilon_s N_B \left( 2 \psi_B - \frac{q x_c D_I}{\varepsilon_s} \right) + \frac{q D_I}{C_{ox}}}$$

$$W_{Dm} = \sqrt{\frac{2 \varepsilon_s}{q N_B} \left( 2 \psi_B - \frac{q D_I x_c}{\varepsilon_s} \right)}.$$
Non Uniform Doping
Low-High Profile → Retrograde profile

Subtraction of $\Delta N$ from background doping

$$V_T = V_{FB} + 2\psi_B + \frac{qN_B W_{Dm} - q\Delta N x_s}{C_{ox}}$$

$$= V_{FB} + 2\psi_B + \frac{1}{C_{ox} N} \left[ 2q\varepsilon_s N_B \left( 2\psi_B + \frac{q\Delta N x_s^2}{2\varepsilon_s} \right) \right] - \frac{q\Delta N x_s}{C_{ox}}$$

$$W_{Dm} = \frac{2\varepsilon_s}{qN_B \left[ 2\psi_B + \frac{q\Delta N x_s^2}{2\varepsilon_s} \right]}.$$

*lower surface concentration increases $W_{Dm}$ but decreases $V_T$
Buried Channel Device

Caused by
- the surface doping of the opposite type of the substrate
- part of the surface doped layer is not fully depleted

Vg can resize the opening
i.e. >>Vg = larger channel, >> -Vg = smaller channel

Fig. 22 (a) Schematic of a buried-channel MOSFET under bias. (b) Its doping profile and depletion regions

Fig b) The net channel thickness is reduced from $x_s$ by the amounts of surface depletion $W_{Ds}$ and the bottom p-n junction depletion $W_{Dn}$
**Buried Channel Device**

- **Fig 23.** Energy-band diagrams of a buried-channel MOSFET, for the bias conditions of (a) flat band (\(V_g = V_{FB}^*\)), (b) surface depletion, and (c) threshold (\(V_g = V_T\))

\[ V_{FB}^* \text{ below refers to the condition that the surface } n\text{-layer has flat band, as opposed to the p-substrate} \]

\[ V_{FB}^* = V_{FB} + \psi_{bi} \]

\[ V_T < V_G < V_{FB}^* \]

\[ Q = Q_B = (x_s - W_{Ds} - W_{Dn}) N_D \]

- Channel Charge, \(Q_B\) = Bulk Charge

\[ V_{FB}^* < V_G \]

\[ Q = Q_B + Q_I \]

\[ = (x_s - W_{Dn}) N_D + C_{ox}(V_G - V_{FB}^*) \]

Assuming long-channel constant-mobility model, Substitute \(Q\) into the eqtn to get \(I_D\)

\[ I_D = \frac{Z}{L} \int_{0}^{L} |Q_n(y)| \nu(y) dy \]
DEVICE SCALING AND SHORT-CHANNEL EFFECTS

As the MOSFET dimensions shrink, they need to be designed properly to preserve the long-channel behavior as much as possible.

As the channel length decreases, the depletion widths of the source and drain become comparable to the channel length and punch-through between the drain and source will eventually occur.

Even with best practice, keeping long-channel behavior becomes impossible. **Short-channel effects** cause
- 2-d potential distribution
- high channel electric field
**DEVICE SCALING AND SHORT-CHANNEL EFFECTS**

**Fig 25.** Physical parameters for MOSFET scaling. Scaling factors for constant-field are indicated.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Scaling factor: Constant-$\kappa$</th>
<th>Scaling factor: Actual</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L$</td>
<td>$1/\kappa$</td>
<td>$/ 1/\kappa$</td>
<td>/</td>
</tr>
<tr>
<td>$\mathcal{C}$</td>
<td>$1$</td>
<td>$&gt; 1$</td>
<td>/</td>
</tr>
<tr>
<td>$d$</td>
<td>$1/\kappa$</td>
<td>$&gt; 1/\kappa$</td>
<td>Tunneling, defects</td>
</tr>
<tr>
<td>$r_J$</td>
<td>$1/\kappa$</td>
<td>$&gt; 1/\kappa$</td>
<td>Resistance</td>
</tr>
<tr>
<td>$V_T$</td>
<td>$1/\kappa$</td>
<td>$\gg 1/\kappa$</td>
<td>Off current</td>
</tr>
<tr>
<td>$V_D$</td>
<td>$1/\kappa$</td>
<td>$\gg 1/\kappa$</td>
<td>System, $V_T$</td>
</tr>
<tr>
<td>$N_A$</td>
<td>$\kappa$</td>
<td>$&lt; \kappa$</td>
<td>Junction breakdown</td>
</tr>
</tbody>
</table>

**Table 2** MOSFET Scaling
DEVICE SCALING AND SHORT-CHANNEL EFFECTS

Drain-Induced Barrier Lowering (DIBL)

when the source and drain depletion regions are a substantial fraction of the channel length, short-channel effects start to occur

When \( W_s + W_D \sim L \Rightarrow \) punch-through
\( \Rightarrow \) high leakage current between source and drain

Fig 28. Energy-band diagram at the semiconductor surface from source to drain, for (a) long channel and (b) short-channel MOSFETs, showing the DIBL effect in the latter. Dashed lines \( V_D = 0 \). Solid lines \( V_D > 0 \).
DEVICE SCALING AND SHORT-CHANNEL EFFECTS
Multiplication and Oxide Reliability

As channel carriers (electrons) go through the high-field region, they acquire extra energy from the field without losing it to the lattice. These energetic carriers are called **hot carriers**.

Carriers going to the gate or source will produce very nasty effects.
MOSFETs

DEVICE SCALING AND SHORT-CHANNEL EFFECTS
Multiplication and Oxide Reliability

$I_G$ is due to hot carriers over the barrier and it differs from carriers tunneling through the barrier.

It peaks at $V_G \sim V_D$
Their impact creates problems
=> Create oxide charges and Interface traps
=> VT shifts, higher value
=> Degrading trans conductance, $g_m$
=> Reduced Channel mobility
=> Larger Sub threshold swing $S$

Fig 31. Drain current, substrate current, and gate current vs. gate voltage of a MOSFET. L/W = 03um/30um