Introduction to
CMOS VLSI
Design

MOSFETs:
The Long Channel, Ideal, or
Shockley Model

Peter Kogge
University of Notre Dame
Fall 2015, 2018

Based on material from
Prof. Jay Brockman, Joseph Nahas: University of Notre Dame
Prof. David Harris, Harvey Mudd College
http://www.cmosvlsi.com/coursematerials.html

Outline

- Lecture A
  - IEEE Notation and IV curves
  - MOS Gate
  - Water Model
  - nMOS Ideal Long Channel I-V Model
  - Supplementary Material – More Careful Computation

- Lecture B
  - Reading the I-V Curves
  - Sample Technologies
  - Load Lines and an NMOS Inverter
  - A CMOS Inverter

- Lecture C
  - DC Transfer Curves for an Inverter
  - Ideal vs Real
  - Real-World Effects
IEEE Notation
And IV Curves

IEEE Standard Device Notation

- $V_{ab}$ = voltage between terminals with “a” positive and “b” negative.
  - By definition: $V_{ab} = -V_{ba}$
- $V_a$ = voltage at terminal “a” relative to some standard terminal
- $I_a$ is current into terminal “a”.
  - Electrons are flowing out.
- Denoting dependence on time:
  - Upper case, V or I, denote time independent (DC) values;
  - Lower case, v or i, denote time dependent values.
**Introduction**

- So far, transistors = ideal switches
- Reality: ON transistor passes finite current
  - Depends on terminal voltages
  - Derived from current-voltage (I-V) relationships
- Transistor gate, source, drain also have capacitance
  - \( I = C \left( \Delta V/\Delta t \right) \rightarrow \Delta t = (C/I) \Delta V \)
  - Capacitance and current determine speed
- \( V_g, V_s, V_d \): voltages as measured from body

\[ V_g, \quad V_s, \quad V_d \]

**IV Curves: Simple 2-Terminal Devices**

What is curve for I as a function of V?
**IV Curves: Switched Resistance**

Switch has 3 Positions: A, B, C, D

What is curve for I as a function of V and switch position?

Position B: Slope = 1/4R

**IV Curves: Potentiometer**

Fully Counter Clockwise: Slope = 1/R

What is curve for I as a function of V and rotation?
IV Curves: Diode

Assume resistance $R$ a function of $V$:
- $R_1$ (large) for $V < V_T$
- $R_2$ (small) for $V > V_T$

IV Curves: Voltage-Sensitive Resistor

Assume resistance $R$ a function of $V$:
- $R_1$ (i.e. constant) for $V < V_{sat}$
- $V_0(R_1/V_{sat})$ for $V > V_{sat}$
Assume resistance $R$ a function of $V$:
- $R_1$ (i.e. constant) for $V < V_{\text{sat}}$
- $V_1(R_1/V_{\text{sat}})$ for $V > V_{\text{sat}}$

AND $R_1$ is a function of $V_{gs}$
- Larger $V_{gs}$ reduces $R_1$

**MOS Gate**

and Effect on Channel
MOS Gate Capacitor

- Gate and body form MOS capacitor
- Operating modes

Accumulation mode:
Holes attracted below gate
MOS Gate Capacitor

- Gate and body form MOS capacitor
- Operating modes

Accumulation

Depletion Mode:
- Holes repelled from under gate

Inversion Mode:
- Holes further repelled
- Free electrons attracted

$V_T$ is the gate voltage at the THRESHOLD between Depletion and Inversion

VT is the gate voltage at the THRESHOLD between Depletion and Inversion
**Channel Current**

- What happens if right side more positive relative to left?
- Electrons in Inversion Region leaves channel to right
- Replacement electrons enter from left
- Current moves from right to left

**Terminal Voltages**

- Mode of operation depends on $V_g$, $V_d$, $V_s$
  - $V_{gs} = V_g - V_s$
  - $V_{gd} = V_g - V_d$
  - $V_{ds} = V_d - V_s = V_{gs} - V_{gd}$
- Source and drain are symmetric diffusion terminals
  - By convention, source is terminal at lower voltage (nmos)
  - Hence $V_{ds} \geq 0$
- Also assume nMOS body is grounded.
Regions of Operation

- Assume
  - nMOS body is grounded.
  - First assume source is 0 V too.
- Three regions of operation
  - Cutoff
  - Linear
  - Saturation
- Each region has different relationship between currents and voltages
- Based largely on $V_{gs}$ vs $V_T$

$V_{gs} < V_T$: nMOS Cutoff

- No “channel”
- $I_d = I_{ds}$ (source & body same)
- $I_{ds} = I_d = 0$

Note: source and drain do have “free” electrons
\[ V_{gs} > V_T \text{ and } V_{ds} < (V_{gs} - V_T) : \text{nMOS Linear} \]

- Channel forms
  - "Inversion" of charges
- \( V_{ds} > 0 \), but small
- Current flows from drain to source
  - E from source to drain
- \( I_{ds} \) increases with \( V_{ds} \)
  - As long as \( V_{ds} < (V_{gs} - V_T) \)
- Similar to linear resistor

\[ V_{gs} > V_T & V_{ds} < (V_{gs} - V_T) : \text{nMOS Linear} \]

\[ V_{gs} > V_T \]

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

\[ V_{ds} = 0 \]

\[ V_{ds} > V_{gs} - V_T \]

\[ 0 < V_{ds} < V_{gs} - V_T \]

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

\[ g \]

\[ s \]

\[ d \]

\[ V_{ds} \]

\[ I_{ds} \]

\[ V_{gs} > V_T \]

\[ V_{gd} < V_T \]

\[ V_{ds} > V_{gs} - V_T \]

\[ 0 < V_{ds} < V_{gs} - V_T \]

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

\[ g \]

\[ s \]

\[ d \]

\[ V_{ds} \]

\[ I_{dsat} \]

\[ V_{gs} > V_T \]

\[ V_{gs} < V_i \]

\[ V_{gs} = V_{ds} - V_i \]

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

\[ g \]

\[ s \]

\[ d \]

\[ V_{ds} \]

\[ I_{dsat} \]

\[ MOSFETs-A \]

\[ CMOS \text{ VLSI Design} \]

\[ Slide 21 \]

\[ nMOS \text{ Saturation} \]

- When \( V_{ds} \) becomes sufficiently large, channel "pinches off"
  - \( V_{ds} > V_{gs} - V_T \)
- \( I_d \) becomes independent of \( V_{ds} \)
- We say "current saturates"
- Similar to current source
- Denoted \( I_{dsat} \)
Summary Modes

- **Cutoff**: $V_{gs} < V_T$
  - No current flow across channel
- **Linear**: $V_{gs} > V_T$ and $V_{ds} > 0$ but small
  - Current approximately linear with $V_{ds}$
- **Saturation**: $V_{gs} > V_T$ and $V_{ds} >> 0$
  - Current independent of $V_{ds}$

Water Model
Water Model (C. Sequin)

- Source/drain each have deep container of fluid
  - Applying positive voltage lowers top of container
- Gate has plunger
  - Starts at height $V_T$ above surface
  - Positive voltage on gate lowers plunger

Regions of Operation

- $V_{GS} < V_T$: **cutoff**
- $V_{GS} > V_T$: **no current**
- $V_{DS} < V_{GS} - V_T$: **current linear with $V_{DS}$**
Regions of Operation (cont)

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

“pinch-off” saturated cutoff

Idealized MOS I-V Characteristics: Different Lines for different $V_{gs}$ values

- Linear $V_{gs} > V_T$
- Saturation $V_{gs} > V_T$
- Cutoff $V_{gs} < V_T$
“Long Channel”
MOS Mathematical Model

MOS Device Notation

- $V_{ds}$ is the voltage of the drain relative to the source.
  - By definition: $V_{ds} = -V_{sd}$
- $V_{gs}$ is the voltage of the gate relative to the source.
- $I_{ds} = I_d$ is the current into the drain terminal.

![MOSFET Diagrams](image-url)
MOS Device Notation

- **NMOS Typically:**
  - $V_{ds} \geq 0$
  - $V_{gs} \geq 0$
  - $I_d \geq 0$

- **PMOS Typically:**
  - $V_{ds} \leq 0$
  - $V_{gs} \leq 0$
  - $I_d \leq 0$

MOS I-V Characteristics

- **In Linear region, $I_{ds}$ depends on:**
  - How much charge is in the channel
  - How fast is the charge moving

- $I_{ds} = \frac{Q_{\text{channel}}}{t}$
  - $t = \text{time for charge to transit channel}$
Key Dimensions

Note: We often define $\lambda$ as $\lambda = L/2$, when $L$ is smallest possible
Or $L = 2\lambda$

Calculating MOS I-V Relations

- $I_{ds} = \frac{\text{Charge in channel}}{\text{Transit time}}$
- $\tau = \text{transit times}$
- $Q = \text{charge in transit}$
- $\mu = \text{electron mobility}$
- $C_g = \text{gate capacitance}$
- $E = \text{electric field}$
- $\varepsilon = \text{permittivity of gate dielectric}$
What’s the Charge in the Channel?

- MOS gate structure looks like parallel plate capacitor while operating in inversion
  - Gate is “plate” on top
  - Oxide in middle is dielectric
  - Channel is other “plate”

Q = ?

Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel

Q_{channel} = CV

What’s C?
Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel
- \( Q_{\text{channel}} = CV \)
- \( C = C_g = \frac{\varepsilon_{\text{ox}}WL}{t_{\text{ox}}} = C_{\text{ox}}WL \)
- \( V = V_{\text{gc}} - V_t \)
- \( C_{\text{ox}} = \frac{\varepsilon_{\text{ox}}}{t_{\text{ox}}} \)
What is $V_{gc}$? (The Simple Version)

- On the left of channel, its $V_{gs}$
- On the right of channel, $V_{gs} = V_{ds} + V_{gd}$ or $V_{gd} = V_{gs} - V_{ds}$
- Let's compute the average: $(V_{gs} + (V_{gs} - V_{ds})) / 2 = V_{gs} - V_{ds}/2$

Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel
- $Q_{channel} = CV$
- $C = C_g = \varepsilon_{ox}WL/t_{ox} = C_{ox}WL$
- $V = V_{gc} - V_t = (V_{gs} - V_{ds}/2) - V_t$
- $Q = C_{ox}WL * ((V_{gs} - V_{ds}/2) - V_t)$
- $C_{ox} = \varepsilon_{ox} / t_{ox}$
Now Let's Compute Carrier Velocity

- Charge is carried by e-
- Electrons move along source-drain E-field lines toward + side
- Carrier velocity $v$ proportional to lateral E-field between source and drain
- $v = \mu E$, $\mu$ called mobility

Si Electron and Hole Mobility

Electron Mobility 2-3X Hole Mobility

Significant Temperature Sensitivity (mobility drops)
Carrier velocity

- Charge is carried by e-
- Carrier velocity $v$ proportional to lateral E-field between source and drain
  - $v = \mu E$
  - $E = \frac{V_{ds}}{L}$
- Time for carrier to cross channel: $t = \frac{L}{v}$
**Carrier velocity**

- Charge is carried by e-.
- Carrier velocity \( v \) proportional to lateral E-field between source and drain.
- \( v = \mu E \)
- \( E = V_{ds}/L \)
- Time for carrier to cross channel:
  - \( t = L / v \)

\[ t_{ox} \]

\[ L \]

**Linear Region: Putting It All Together**

- \( I_{ds} = Q_{channel} / t \)
  - \( Q_{channel} = CV \)
  - \( C = \varepsilon_{ox} WL/t_{ox} \)
  - \( V = V_{gs} - V_{t} - V_{ds}/2 \)
  - \( t = L / v \)
  - \( v = \mu E \)
  - \( E = V_{ds}/L \)
- Thus: \( I_{ds} = (\varepsilon_{ox}WL/t_{ox})(V_{gs} - V_{t} - V_{ds}/2)/(L/(\mu V_{ds}/L)) \)
- Or: \( I_{ds} = (\varepsilon_{ox}^{*}\mu t_{ox}^{*}(W/L))(V_{gs} - V_{t} - V_{ds}/2) V_{ds} \)
  - \( \beta V_{GT} \)

- Or: \( I_{ds} = \beta(V_{GT} - V_{ds}/2) V_{ds} \)

Valid for:
- \( V_{gs} > V_{t} \)
- But \( V_{ds} \) relatively small

Where:
- \( \beta = C_{ox}^{*} \mu^{*}(W/L) \)
- \( V_{GT} = V_{gs} - V_{t} \)
**Saturation Region**

- **V_{dsat}: Drain Saturation Voltage**
  - V_{ds} value where channel no longer inverted in vicinity of drain
  - Saturation typically when V_{ds} = V_{GT} = V_{gs} - V_{t}

- Thus, V_{dsat} = V_{GT} = V_{gs} - V_{t}

- Substituting in prior equation: I_{dsat} = \beta (V_{GT} - V_{dsat}/2)\times V_{dsat}

- Or: I_{dsat} = \beta V_{GT}^2 / 2

*Where:*
- C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}
- \beta = (\varepsilon_{ox} \times \mu / t_{ox}) \times (W/L) = C_{ox} \times \mu \times W / L
- V_{GT} = V_{gs} - V_{t}
- V_{dsat} = V_{GT}

---

**Summary: Long Channel Model**

*William Shockley 1st order transistor models*
*1952 A Unipolar Field Effect Transistor*

\[ I_{ds} = \begin{cases} 
0, & V_{gs} < V_{Cutoff} \\
(\beta(V_{GT} - V_{ds}/2)\times V_{ds}, & V_{gs} > V_{i} \text{ and } V_{ds} < V_{dsat} \text{ Linear}
\\
\beta V_{GT}^2 / 2, & V_{ds} > V_{dsat} \text{ Saturation} 
\end{cases} \]

*Where:*
- C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}
- \beta = (\varepsilon_{ox} \times \mu / t_{ox}) \times (W/L) = C_{ox} \times \mu \times W / L
- V_{GT} = V_{gs} - V_{t}
- V_{dsat} = V_{GT}
Knobs for Designers to Change

- Typically logic designer has no control over
  - $\varepsilon_{ox}$: thin ox permittivity
  - $t_{ox}$: thickness of oxide
  - $\mu$: mobility
  - $V_t$: threshold voltage
- Only knobs left to change
  - Overall $V_{dd}$: but usually selected at system level
  - $L$: length – but there is a minimum ($2\lambda$)
- $W$: width

FIGURE 2.16 Comparison of $\alpha$-power law model with simulated transistor behavior
Example: Doubling W Approximately Doubles Current

![Graph showing the relationship between Vds (Volts), W = Length, and Vds (Volts), W = 2^x Length.]

Supplementary Material
More Careful Accounting for Channel Voltage Variation

\[ V_{GS} > V_T \quad V_{DS} \]

\[ dQ = \varepsilon W \frac{dx}{t_{ox}} \left( V_{GS} - V_T - V(x) \right) \]

\[ \tau = \frac{dx}{\mu E} = \frac{dx}{\mu \frac{dV}{dx}} = \frac{dx^2}{\mu dV} \]

\[ I = \frac{dQ}{\tau} = \frac{\mu \varepsilon W}{t_{ox}} \left( V_{GS} - V_T - V(x) \right) \frac{dV}{dx} \]

Accounting for Channel Voltage Variation

\[ V_{GS} > V_T \quad V_{DS} \]

\[ I = \int_0^t \mu \varepsilon W \frac{V_{GS} - V_T}{t_{ox}} \left( \frac{V_{DS}}{2} \right) dV \]

\[ I_{IL} = \frac{\mu \varepsilon W}{t_{ox}} \left( \frac{(V_{GS} - V_T) V_{DS}^2}{2} \right) \]

\[ I = \frac{\mu \varepsilon W}{t_{ox} L} \left( V_{GS} - V_T \right) \frac{V_{DS}^2}{2} \]

The Shockley Equation

MOSFETs-A  CMOS VLSI Design  Slide 53

MOSFETs-A  CMOS VLSI Design  Slide 54
**Re-writing the Schockley Equation**

\[ V_{GS} > V_T \]

\[ I = \frac{\mu e W}{L} \left( (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right) \]

\[ I = k' \frac{W}{L} \left( (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right) \]

\[ I = \beta \left( (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right) \]

---

**nMOS I-V Summary**

*(The Same Equations as Before)*

- **William Shockley** 1st order transistor models
  - 1952 *A Unipolar Field Effect Transistor*

\[ I_d = \begin{cases} 
0 & V_{gs} < V_t \quad \text{cutoff} \\
\beta \left( V_{gs} - V_t - \frac{V_{ds}}{2} \right) V_{ds} & V_{ds} < V_{dsat} \quad \text{linear} \\
\frac{\beta}{2} \left( V_{gs} - V_t \right)^2 & V_{ds} > V_{dsat} \quad \text{saturation}
\end{cases} \]