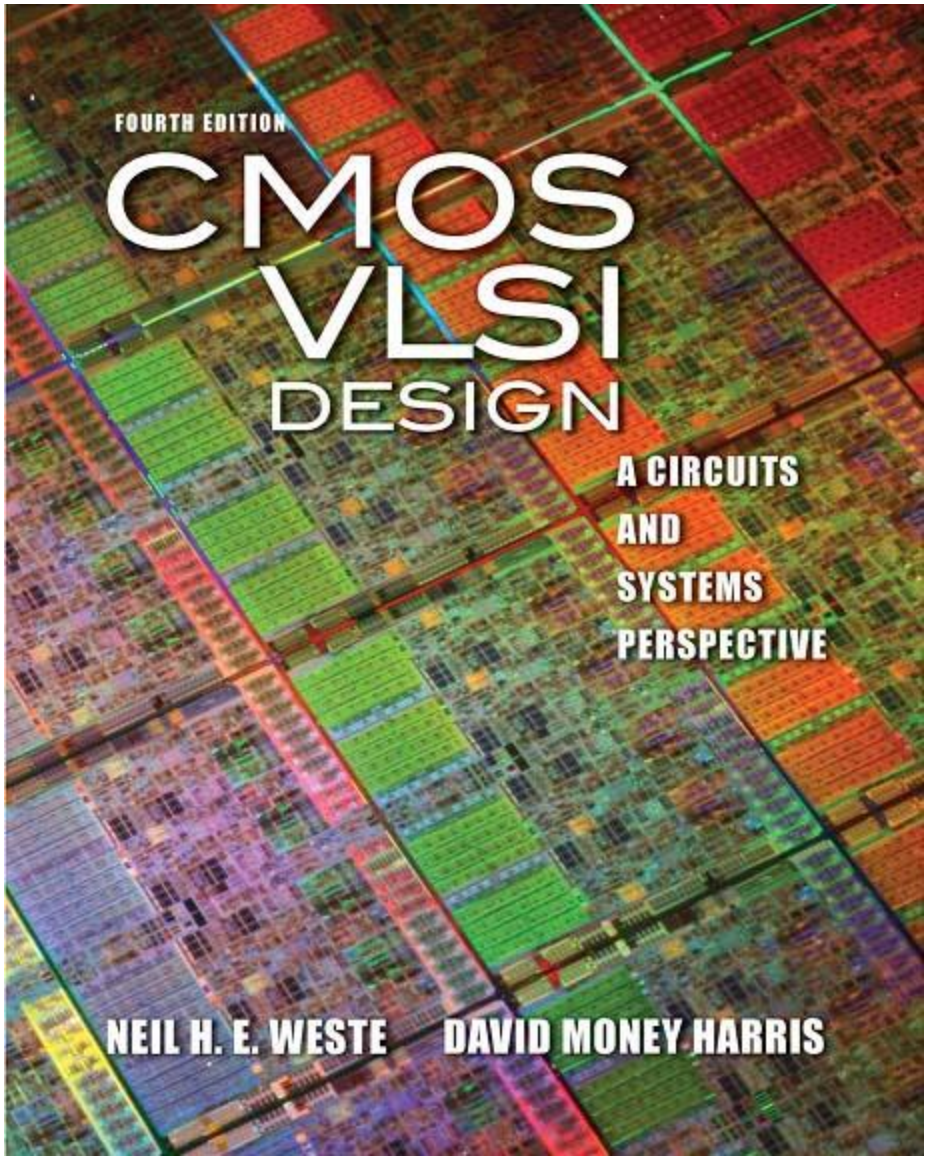


Lecture 3: CMOS Transistor Theory

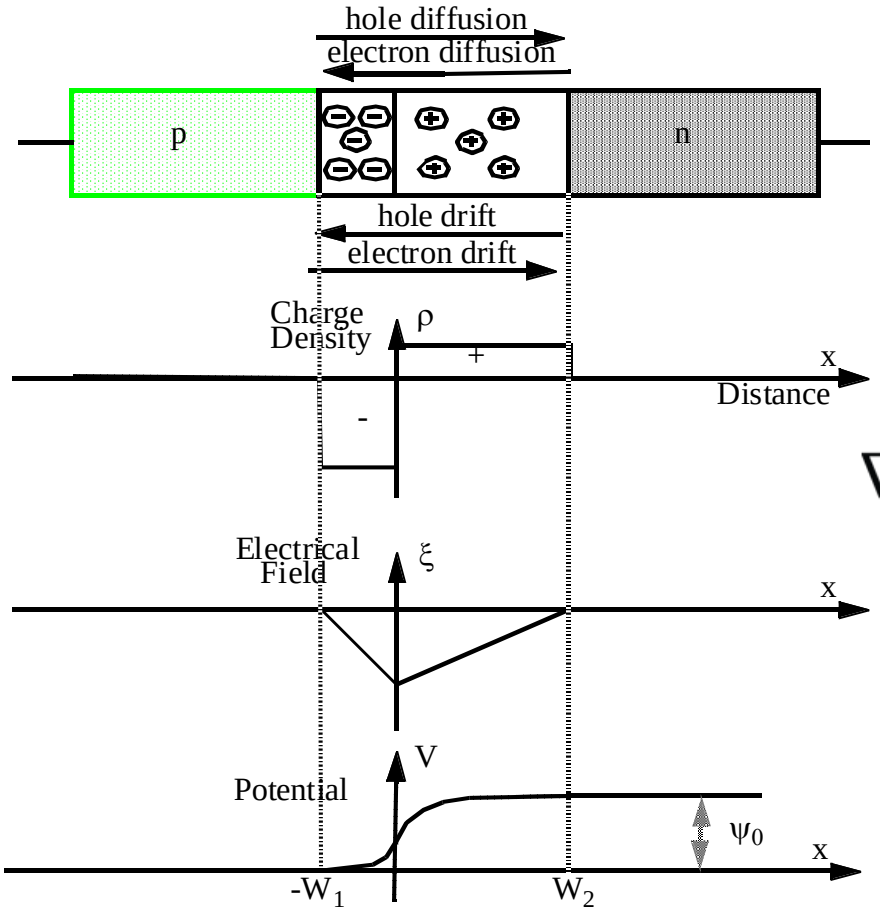
Outline

- Introduction
- MOS Capacitor
- nMOS I-V Characteristics
- pMOS I-V Characteristics
- Gate and Diffusion Capacitance

Diodes

- Diodes do not appear in CMOS digital design as separate devices.
- However, they are present as junctions and parasitic elements in all devices.
- We will use a simple 1D analysis.
- We will not concern ourselves too much with the DC behavior too much.

Depletion Region



(a) Current flow.

(b) Charge density.

$$\nabla \cdot E = \frac{\rho}{\varepsilon_0}$$

E = electric field

ρ = charge density

ε_0 = permittivity

(c) Electric field.

$$\cdot E = -\nabla \varphi$$

(d) Electrostatic potential.

DC Characteristics

$$\phi_0 = \phi_T \ln\left(\frac{N_A N_D}{n_i^2}\right)$$

$$\phi_T = \frac{kT}{q} = 25.8mV \qquad k = 1.38 \times 10^{-23} \text{J/K}$$

For example, a Si p-n junction, $N_A = 10^{18} cm^{-3}, N_D = 10^{15} cm^{-3}$

$$\phi_0 = 0.0258 \ln\left(\frac{10^{18} \times 10^{15}}{(9.65 \times 10^9)^2} V\right) = 0.774V$$

Majority Carriers: $n \approx N_D$

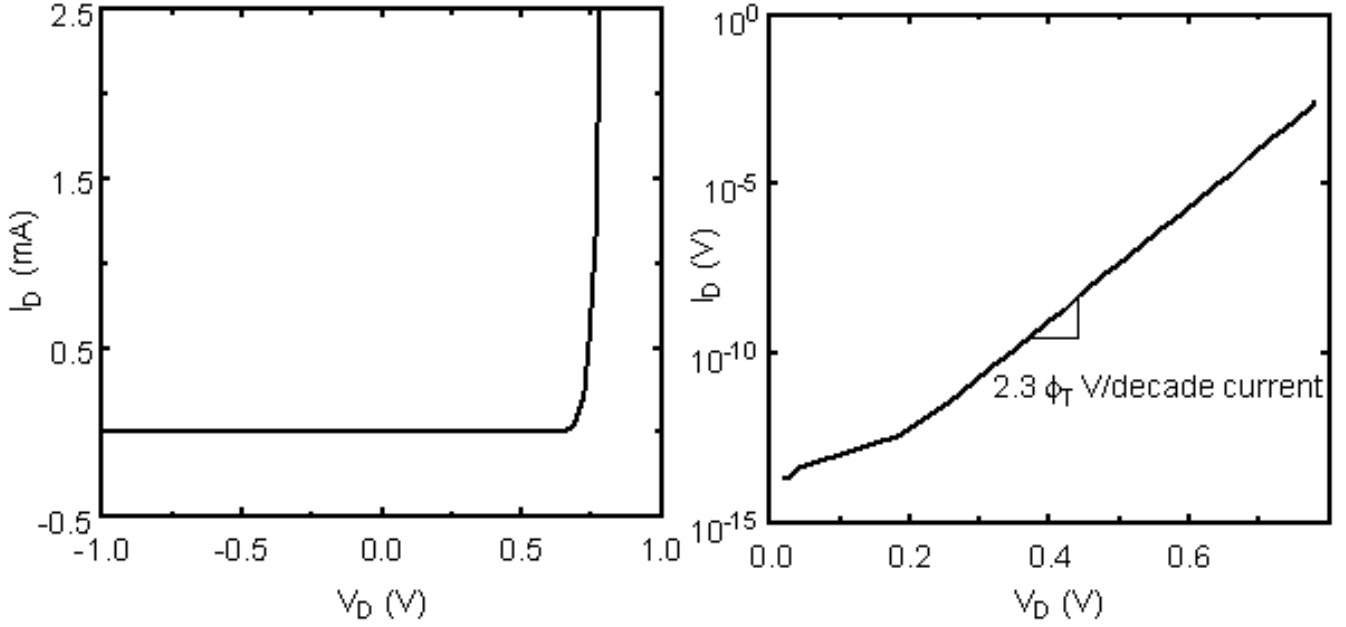
Minority Carriers: $p \approx \frac{n_i^2}{N_D}$.

Similarly, for a density of N_A ($\gg n_i$) acceptor atoms per cubic centimeter:

Majority Carriers: $p \approx N_A$

Minority Carriers: $n \approx \frac{n_i^2}{N_A}$.

Diode Current

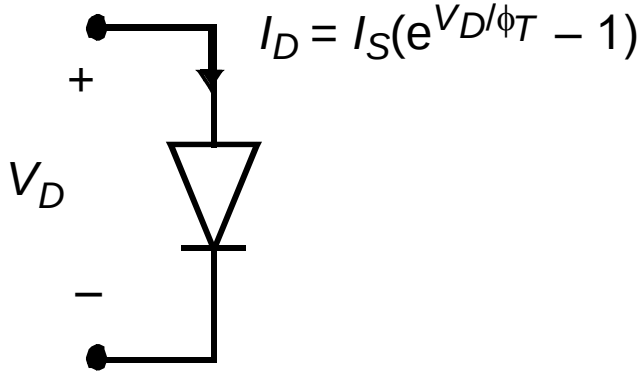


(a) On a linear scale.

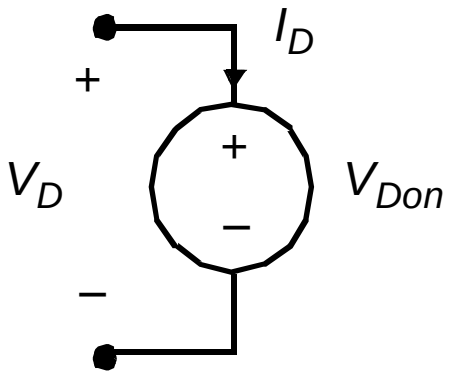
(b) On a logarithmic scale (forward bias).

$$I_D = I_S\left(e^{V_D/\phi_T} - 1\right)$$

Models for Manual Analysis



(a) Ideal diode model



(b) First-order diode model

Junction Capacitance

- in a more general form as

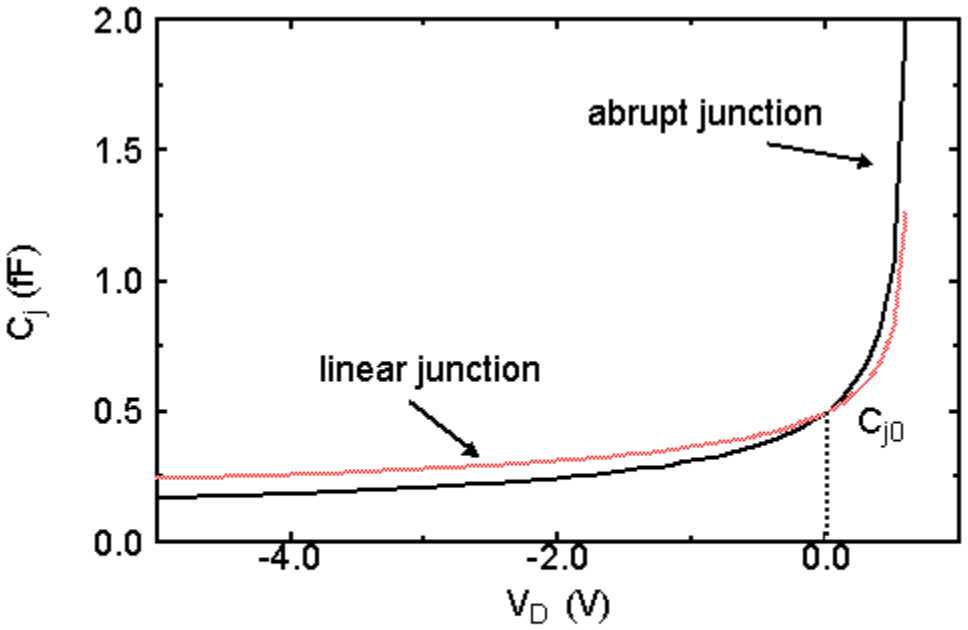
$$C_j(V) = A\sqrt{\frac{\varepsilon_{Si} q}{2}\left(\frac{N_A N_D}{N_A + N_D}\right)\left(\frac{1}{\sqrt{\phi_0 - V}}\right)}$$

- m is the gradient coefficient and is 0.5 for abrupt junctions and 1/3 for linearly graded junction profiles
- The value of the junction capacitance ultimately depends on the external bias voltage applied across the pn-junction.

$$C_{j0} = \sqrt{\frac{\varepsilon_{Si} q}{2}\left(\frac{N_A N_D}{N_A + N_D}\right)\left(\frac{1}{\phi_0}\right)}$$

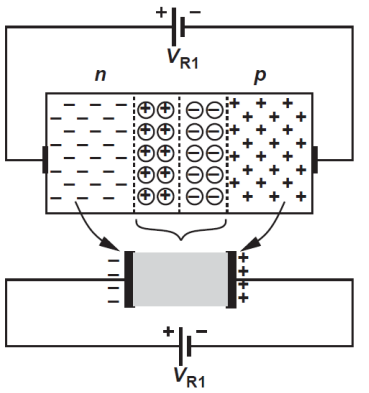
$$C_j(V) = \frac{AC_{j0}}{\left(1 - \frac{V}{\phi_0}\right)^m}$$

Junction Capacitance

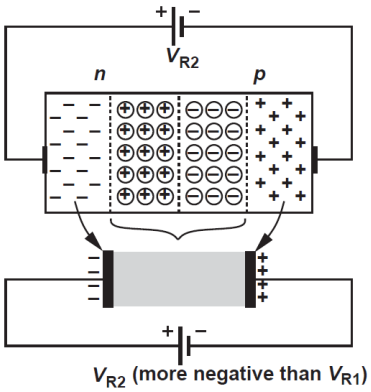


$$C_j = \frac{C_{j0}}{(1 - V_D/\phi_0)^m}$$

m = 0.5: abrupt junction
m = 0.33: linear junction



(a)



V_{R2} (more negative than V_{R1})

(b)

Junction Capacitance

- *m* is known as the *grading coefficient*.
- Keep in mind that C_j is a small signal parameter. For large signal switching, an equivalent capacitance has to be calculated as

$$C_{eq} = \frac{\Delta Q_j}{\Delta V_D} = \frac{Q(V_{high}) - Q(V_{low})}{V_{high} - V_{low}} = K_{eq} C_{j0}$$

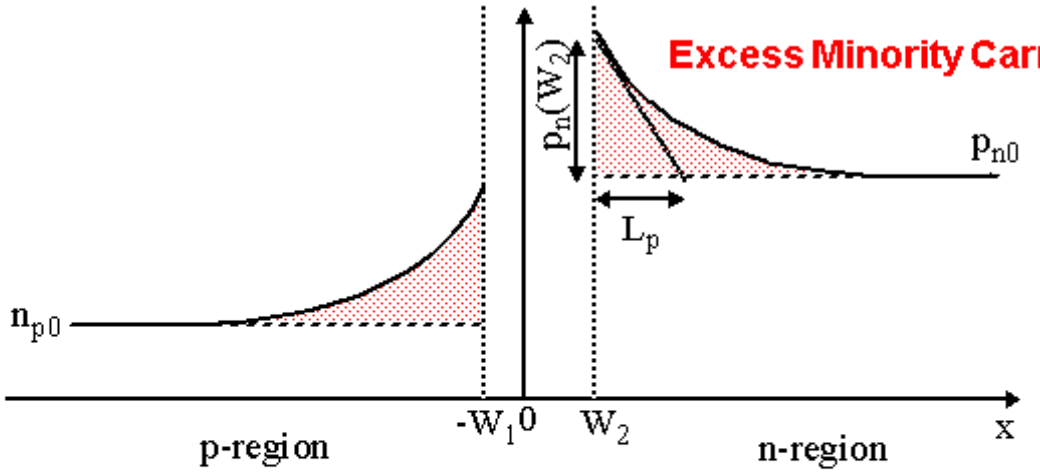
- C_{eq} has been defined such that the same amount of charge is transferred as the nonlinear model

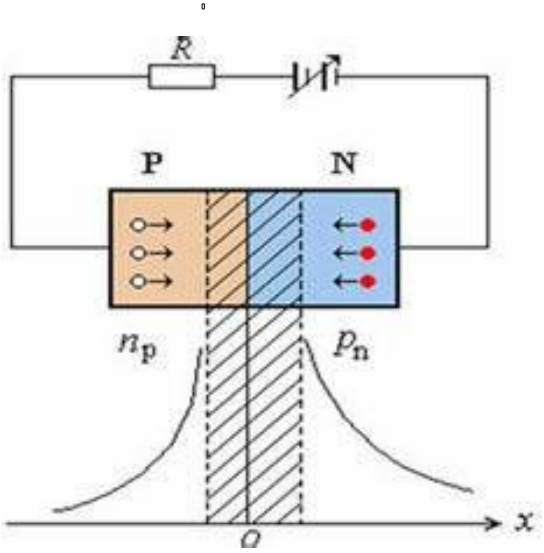
$$K_{eq} = \frac{-\phi_0^m}{(V_{high} - V_{low})(1-m)} \left[(\phi_0 - V_{high})^{(1-m)} - (\phi_0 - V_{low})^{(1-m)} \right]$$

Junction Capacitance

- As a numerical example, a diode is switched between *0* and *-2.5 V*. The diode has $C_{j0} = 2 \times 10^{-3}$ F/m^2, $A_D = 0.5\ (\mu m)^2$, $\Phi_0 = 0.64\ V$, $m = 0.5$.
- $K_{eq} = 0.622$, $C_{eq} = 1.24\ fF/(\mu m)^2$.

Diffusion Capacitance





$$C_d = \frac{dQ_D}{dV_D} = \tau_T \frac{dI_D}{dV_D} \approx \frac{\tau_T I_D}{\phi_T}$$

Diffusion Capacitance

- Effective in forward bias

$$Q_p = qA_D \int_{W_2}^{W_n} \left(p_n(x) - p_{n0}\right) dx$$

$$= qA_D \frac{\left(W_n - W_2\right) p_{n0} \left(e^{V_D / \phi_T} - 1\right)}{2}$$

$$\approx \frac{W_n^2}{2D_p} I_{Dp} = \tau_{Tp} I_{Dp}$$

$$I_D = \frac{Q_p}{\tau_{Tp}} + \frac{Q_n}{\tau_{Tn}} \equiv \frac{Q_D}{\tau_T}$$

Diffusion Capacitance

- From this lifetime analysis of excess charge,

$$C_d = \frac{dQ_D}{dV_D} = \tau_T \frac{dI_D}{dV_D} \approx \frac{\tau_T I_D}{\phi_T}$$

- Note that C_d is also a small signal capacitance

$$C_{eq} = \frac{\Delta Q_D}{\Delta V_D} = \frac{\tau_T \left(I_D\left(V_{high}\right) - I_D\left(V_{low}\right)\right)}{V_{high} - V_{low}}$$

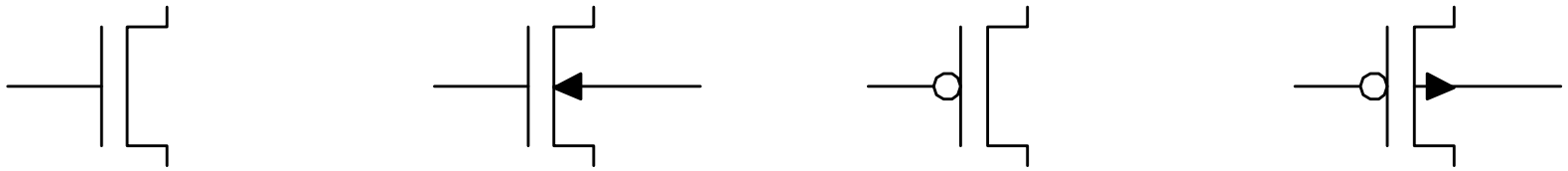
$$= \frac{C_d\left(high\right) - C_d\left(low\right)}{V_{high} - V_{low}} \phi_T$$

Other Diode Parameters

- Secondary Effects
  - Resistivity of regions outside junction
  - Breakdown voltage
  - Temperature dependence
    - Φ_T has a linear dependence
    - I_S doubles every *8˚C*
    - Overall, current doubles every *12˚C*.

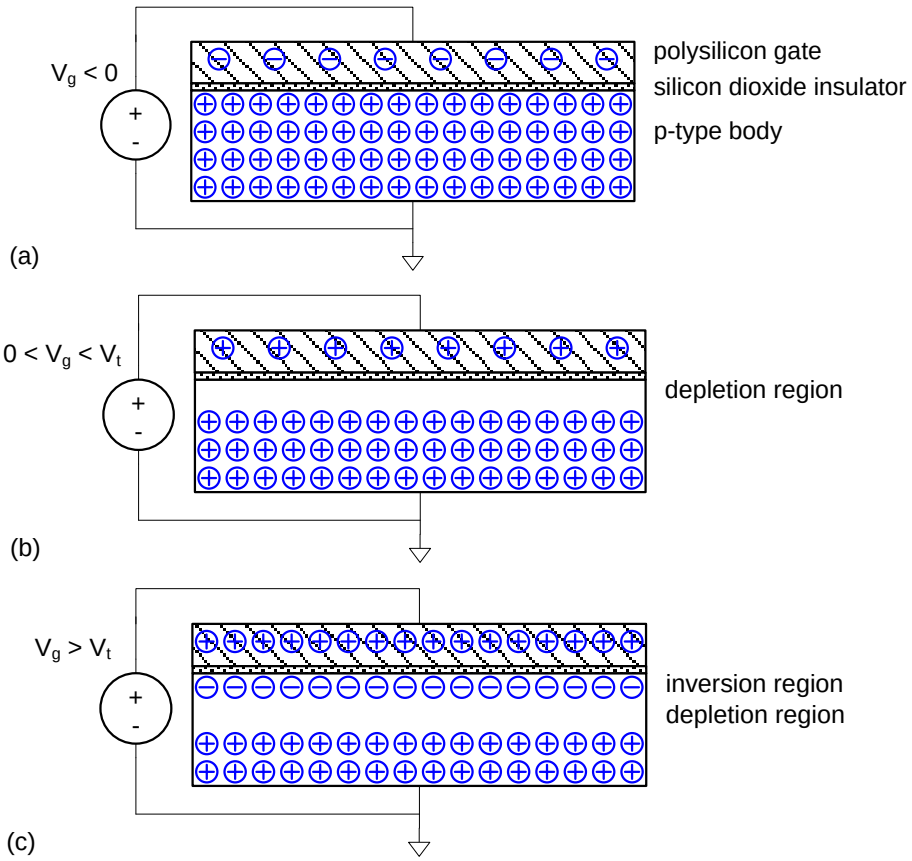
Introduction

- So far, we have treated transistors as ideal switches
- An ON transistor passes a finite amount of current
  - Depends on terminal voltages
  - Derive current-voltage (I-V) relationships
- Transistor gate, source, drain all have capacitance
  - $I = C (\Delta V/\Delta t) \rightarrow \Delta t = (C/I)\, \Delta V$
  - Capacitance and current determine speed



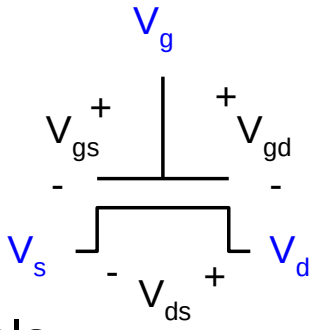
MOS Capacitor

- Gate and body form MOS capacitor
- Operating modes
  - Accumulation
  - Depletion
  - Inversion



Terminal Voltages

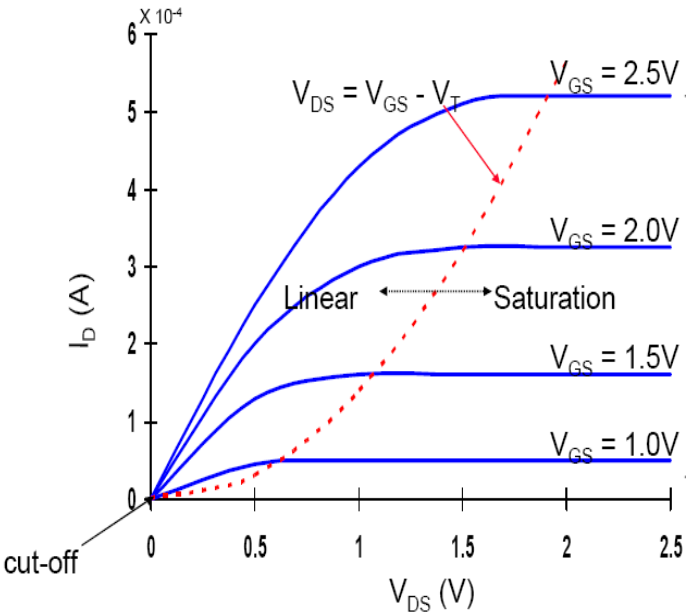
- Mode of operation depends on V_g, V_d, V_s
  - $V_{gs} = V_g - V_s$
  - $V_{gd} = V_g - V_d$
  - $V_{ds} = V_d - V_s = V_{gs} - V_{gd}$

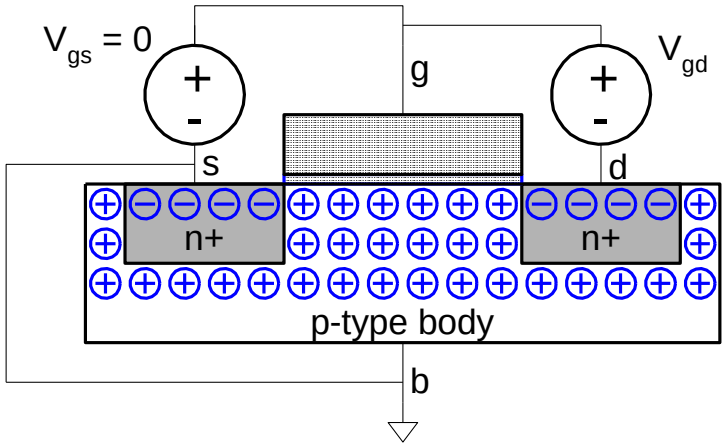


- Source and drain are symmetric diffusion terminals
  - By convention, source is terminal at lower voltage
  - Hence $V_{ds} \geq 0$
- nMOS body is grounded. First assume source voltage is 0 too.
- Three regions of operation
  - *Cutoff*
  - *Linear (Resistive)*
  - *Saturation (Active)*

nMOS Cutoff

- No channel
- $I_{ds} \approx 0$

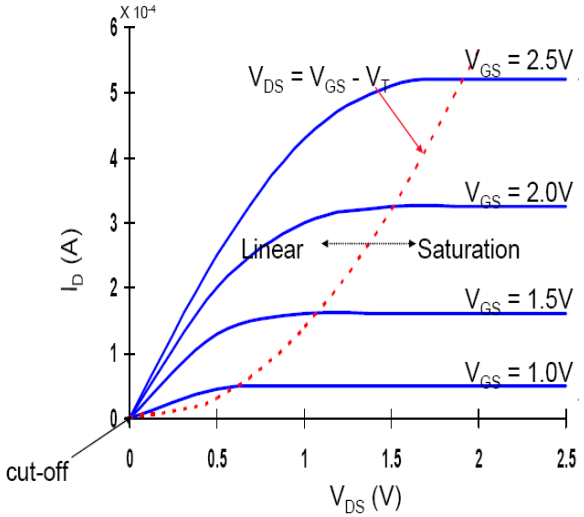


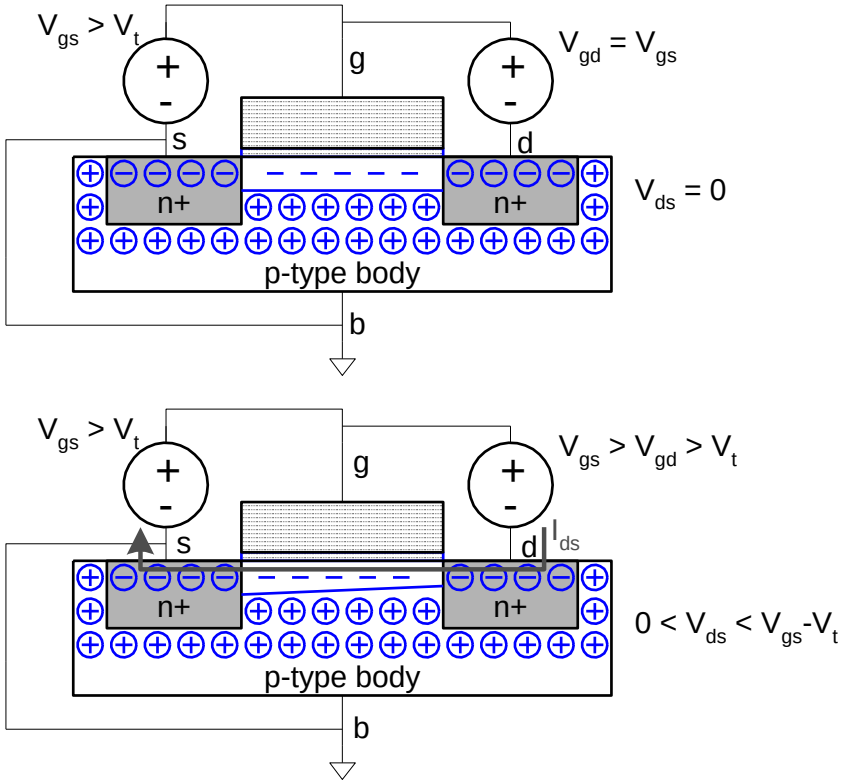


nMOS Linear

- Channel forms
- Current flows from d to s
  - e^- from s to d
- I_{ds} increases with V_{ds}
- Similar to linear resistor

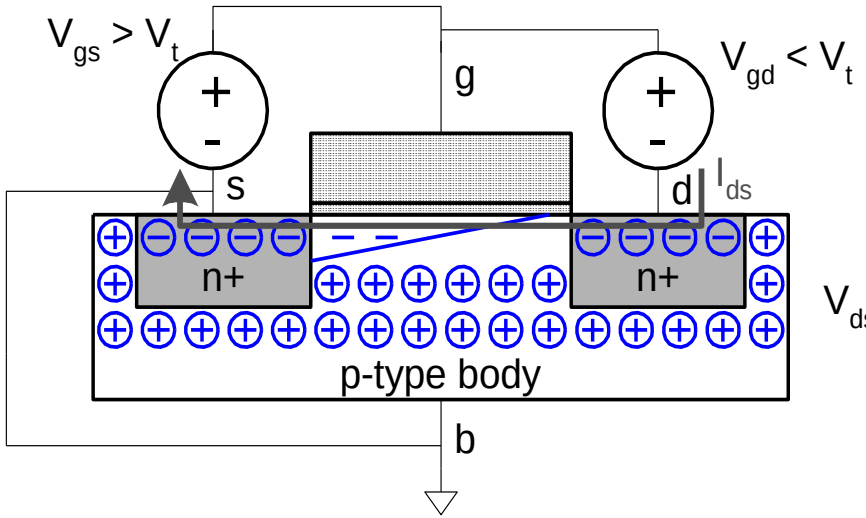
$$V_{ds} = V_d - V_s = V_{gs} - V_{gd}$$



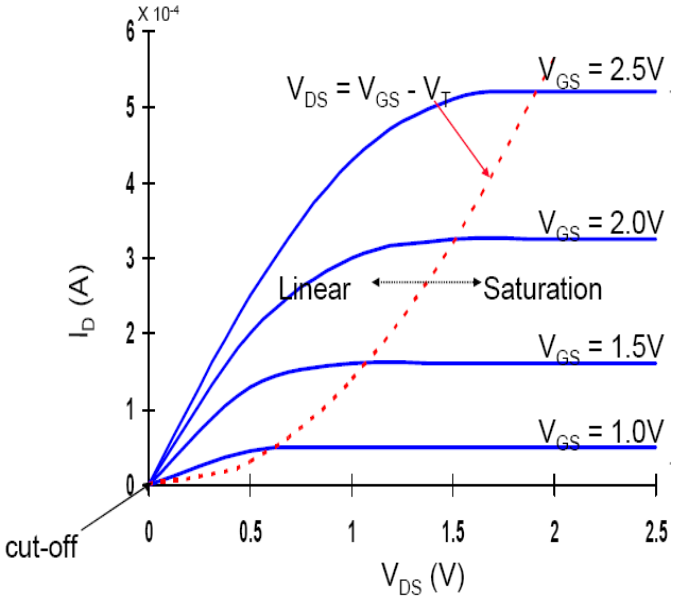


nMOS Saturation

- Channel pinches off
- I_{ds} independent of V_{ds}
- We say current *saturates*
- Similar to current source



$V_{ds} > V_{gs} - V_t$



$$V_{ds} = V_d - V_s = V_{gs} - V_{gd}$$

Mosfet basic operation

I-V Characteristics

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

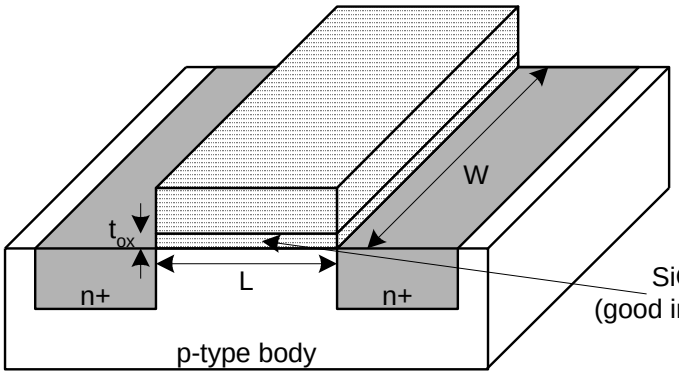
Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversions
  - Gate – oxide – channel
- $Q_{channel}$ =
-
-

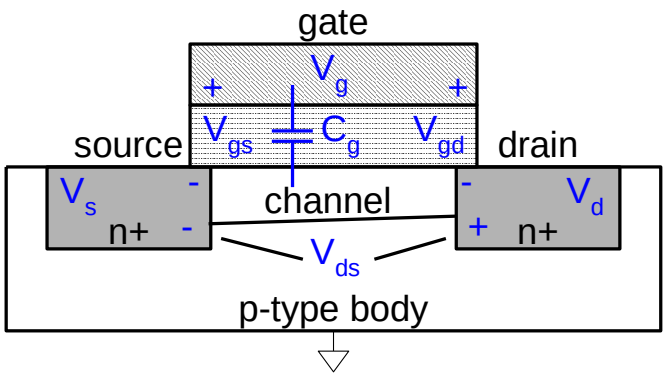
$$V_c = (V_s + V_d)/2 = V_s + V_{ds}/2$$

$$V_g - V_c = V_{gs} - V_{ds}/2$$

:



SiO_2 gate oxide
(good insulator, ε_{ox} = 3.9)



Carrier velocity

- Charge is carried by e-
- Electrons are propelled by the lateral electric field between source and drain
  - E =
- Carrier velocity *v* proportional to lateral E-field
  - *v* =
- Time for carrier to cross channel:
  - *t* =

nMOS Linear I-V

- Now we know
  - How much charge $Q_{channel}$ is in the channel
  - How much time *t* each carrier takes to cross

$$I_{ds} =$$

$$=$$

$$=$$

nMOS Saturation I-V

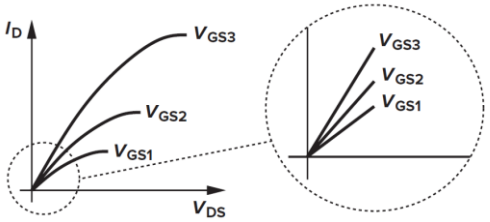
- If $V_{gd} < V_t$, channel pinches off near drain
  - When $V_{ds} > V_{dsat}$ =
- Now drain voltage no longer increases current

$$I_{ds} =$$

$$=$$

nMOS I-V Summary

- *Shockley* 1st order transistor models



Linear operation in deep triode region.

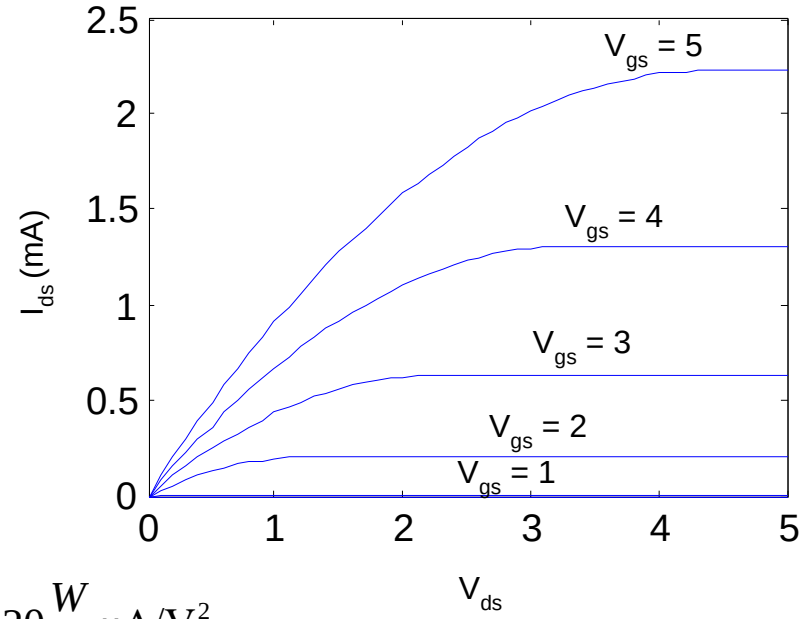
$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t \quad \text{cutoff} \\ \beta\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds} & V_{ds} < V_{dsat} \quad \text{linear} \\ \frac{\beta}{2}\left(V_{gs} - V_t\right)^2 & V_{ds} > V_{dsat} \quad \text{saturation} \end{cases}$$

$$V_{DS} \ll 2(V_{GS} - V_{TH}) \qquad I_D \approx \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH})V_{DS} \qquad R_{on} = \frac{1}{\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH})}$$

overdrive voltage

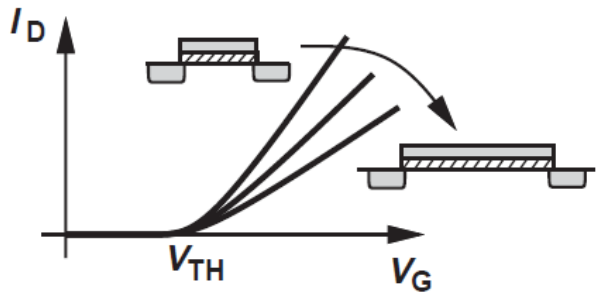
Example

- Your book will be using a 0.6 μm process
  - From AMI Semiconductor
  - t_{ox} = 100 Å
  - μ = 350 $cm^2/V*s$
  - V_t = 0.7 V
- Plot I_{ds} vs. V_{ds}
  - V_{gs} = 0, 1, 2, 3, 4, 5
  - Use W/L = 4/2 λ

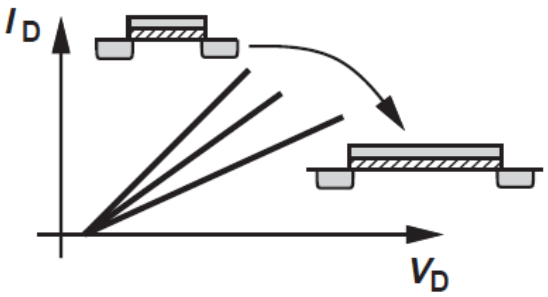


$$\beta = \mu C_{ox}\frac{W}{L} = (350)\left(\frac{3.9\times 8.85\cdot 10^{-14}}{100\cdot 10^{-8}}\right)\left(\frac{W}{L}\right) = 120\frac{W}{L}\,\mu A/V^2$$

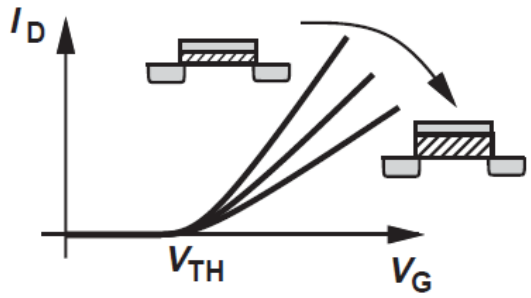
I_D-V_G and I_D-V_D characteristics



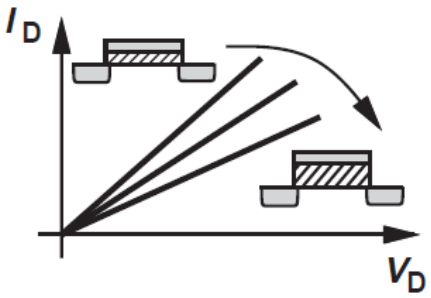
(a)



(b)



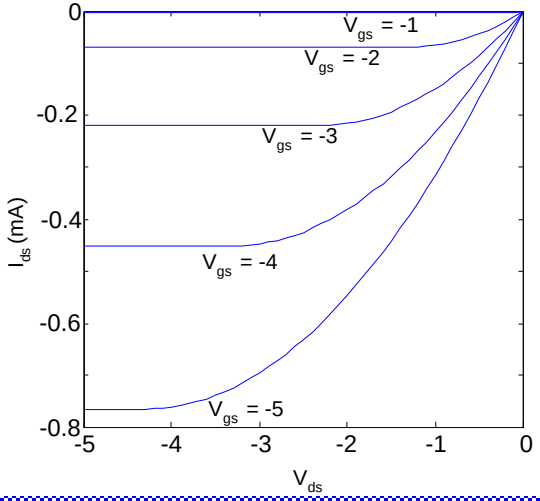
(c)



(d)

pMOS I-V

- All dopings and voltages are inverted for pMOS
  - Source is the more positive terminal
- Mobility μ_p is determined by holes
  - Typically 2-3x lower than that of electrons μ_n
  - 120 $cm^2/V \cdot s$ in AMI 0.6 μm process
- Thus pMOS must be wider to provide same current
  - In this class, assume $\mu_n / \mu_p = 2$



Capacitance

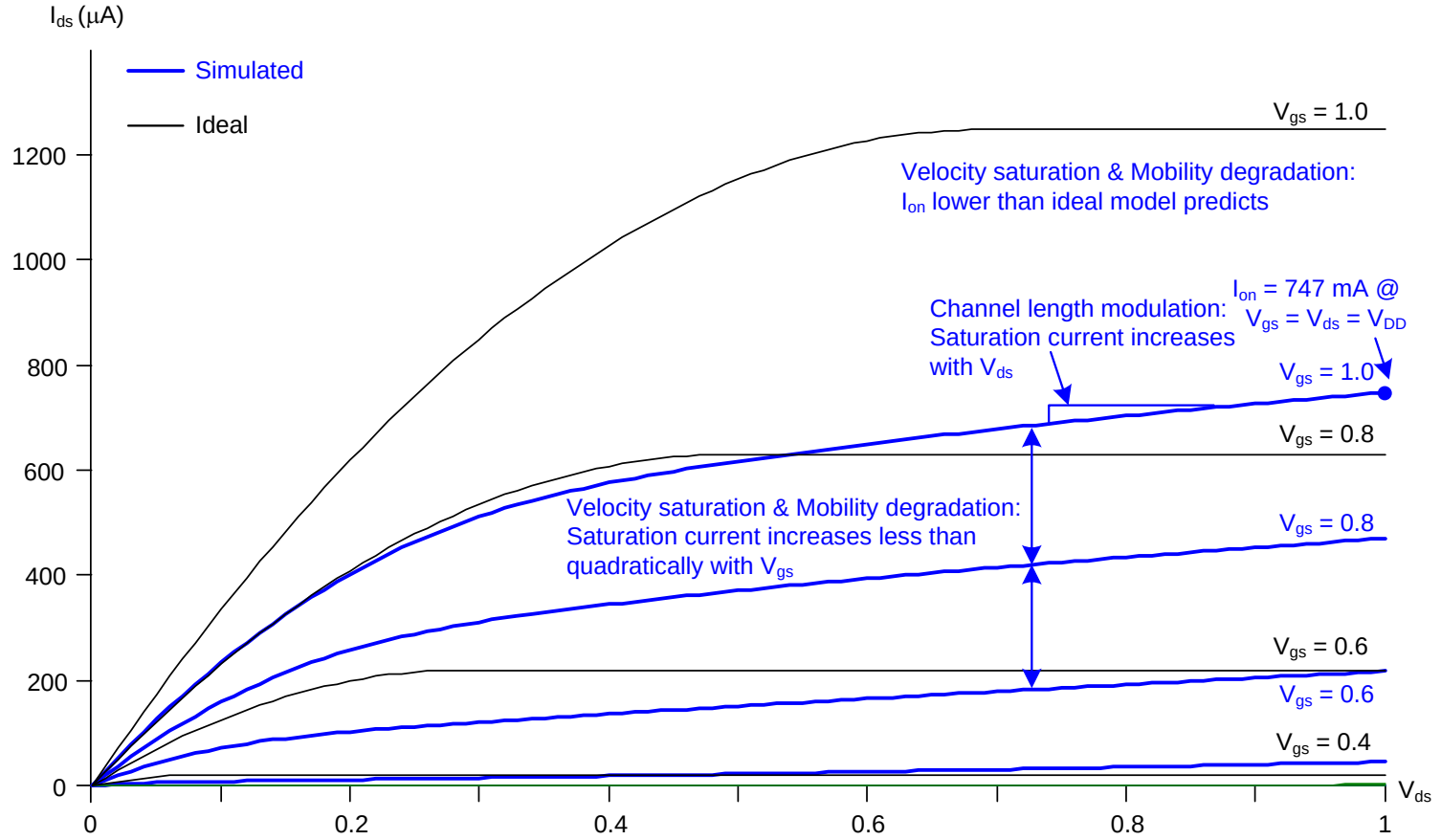
- Any two conductors separated by an insulator have capacitance
- Gate to channel capacitor is very important
  - Creates channel charge necessary for operation
- Source and drain have capacitance to body
  - Across reverse-biased diodes
  - Called diffusion capacitance because it is associated with source/drain diffusion

More Corrections

- Mobility is reduced with increasing gate voltage.
  - We will study this effect in detail later.
- Current conduction occurs below the threshold voltage.
  - We will study this effect later.
- Channel length modulation has to be corrected.
- Threshold voltage depends on W and L.
- Parasitic resistances in the source and drain
- Latchup
- Speed limit of carriers

Ideal vs. Simulated nMOS I-V Plot

- 65 nm IBM process, V_{DD} = 1.0 V



A Unified Model for Manual Analysis

The simplest model in SPICE (Level 1 or default model) uses the above equations.

Parameter	SPICE Parameter	Units	Typical Values
$\mu_n C_{ox}$	KP	A/V^2	200μ
V_{T0}	VTO	V	0.5 – 1.0
λ	LAMBDA	V^{-1}	0.05 – 0.005

$$I_D = 0 \text{ for } V_{GT} \le 0$$

$$I_D = k' \frac{W}{L}\left(V_{GT}V_{min} - \frac{V_{min}^2}{2}\right)(1 + \lambda V_{DS}) \text{ for } V_{GT} \ge 0$$

with $V_{min} = \min(V_{GT}, V_{DS}, V_{DSAT})$,

$V_{GT} = V_{GS} - V_T$,

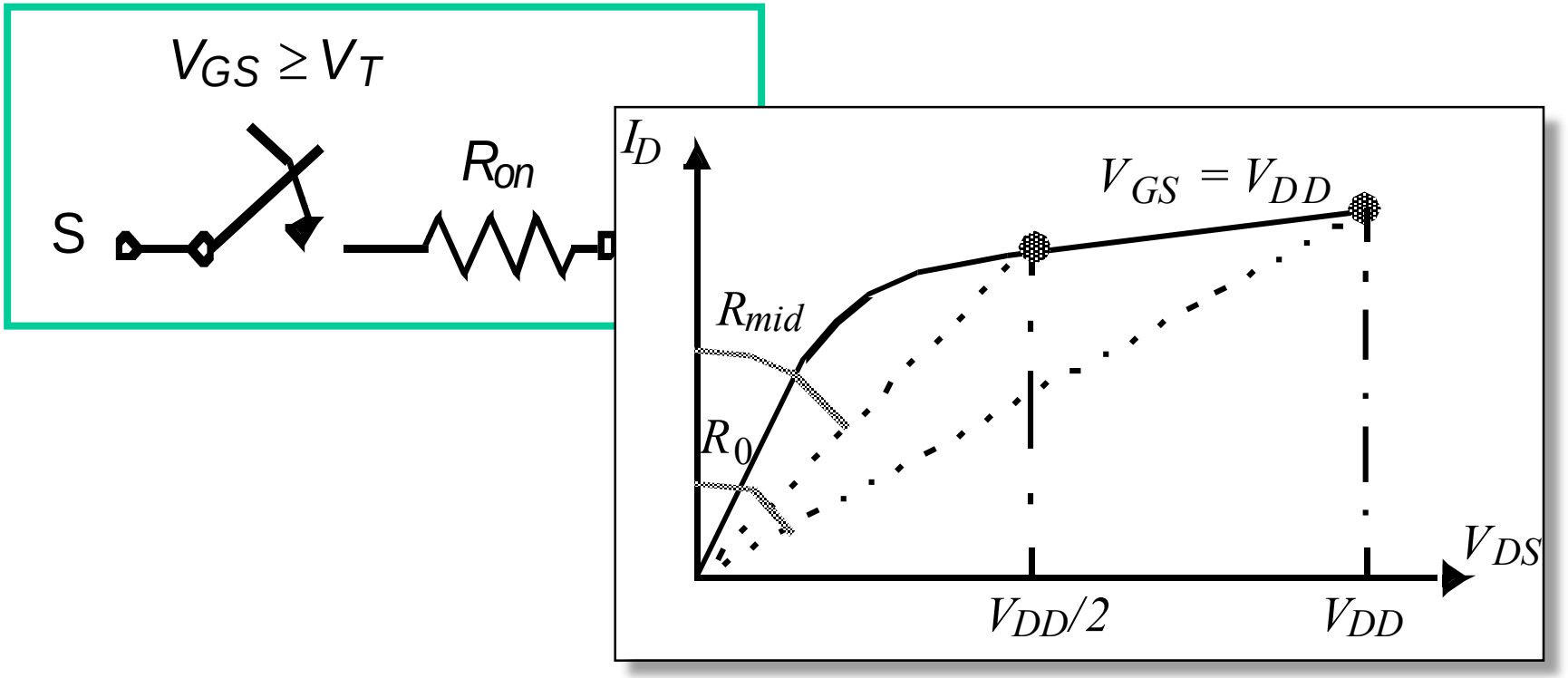
and $V_T = V_{T0} + \gamma(\sqrt{|-2\phi_F + V_{SB}|} - \sqrt{|-2\phi_F|})$

Transistor Model for Manual Analysis

Table 3.2 Parameters for manual model of generic 0.25 μm CMOS process (minimum length device).

	V_{T0} (V)	γ ($V^{0.5}$)	V_{DSAT} (V)	k' (A/V^2)	λ (V^{-1})
NMOS	0.43	0.4	0.63	115×10^{-6}	0.06
PMOS	−0.4	-0.4	-1	-30×10^{-6}	-0.1

The Transistor as a Switch



$$R_{eq} = \frac{1}{2}\left(\frac{V_{DD}}{I_{DSAT}(1+\lambda V_{DD})} + \frac{V_{DD}/2}{I_{DSAT}(1+\lambda V_{DD}/2)}\right) \approx \frac{3}{4}\frac{V_{DD}}{I_{DSAT}}\left(1-\frac{5}{6}\lambda V_{DD}\right)$$

Drain-Source Resistance

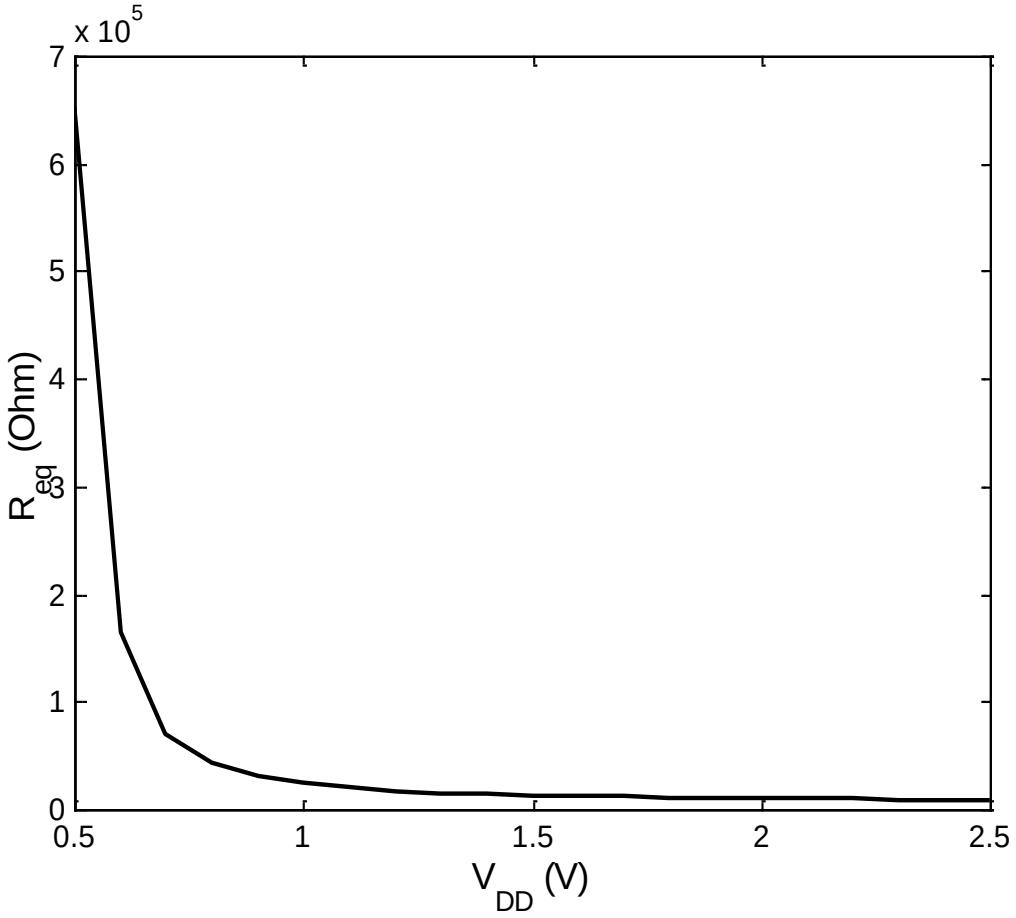
- Applying the general formula for a transistor switching from V_{DD} to $V_{DD}/2$,

$$R_{eq} \approx \frac{3}{4}\frac{V_{DD}}{I_{D,sat}}\left(1-\frac{7}{9}\lambda V_{DD}\right)$$

- Alternatively, using the endpoints and averaging,

$$R_{eq} \approx \frac{3}{4}\frac{V_{DD}}{I_{D,sat}}\left(1-\frac{5}{6}\lambda V_{DD}\right)$$

Drain-Source Resistance



Drain-Source Resistance

Table 3.3 Equivalent resistance R_{eq} (W/L= 1) of NMOS and PMOS transistors in 0.25 μm CMOS process (with $L = L_{min}$). For larger devices, divide R_{eq} by W/L.

V_{DD} (V)	1	1.5	2	2.5
NMOS (kΩ)	35	19	15	13
PMOS (kΩ)	115	55	38	31

$$R_{on} = \frac{1}{\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH})}$$

Drain-Source Resistance

- Large signal drain-source resistance is a nonlinear quantity varying across operating regions.
- One can define an equivalent resistance

$$R_{eq} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \frac{V_{ds}(t)}{I_d(t)} dt$$

- For a weakly nonlinear function,

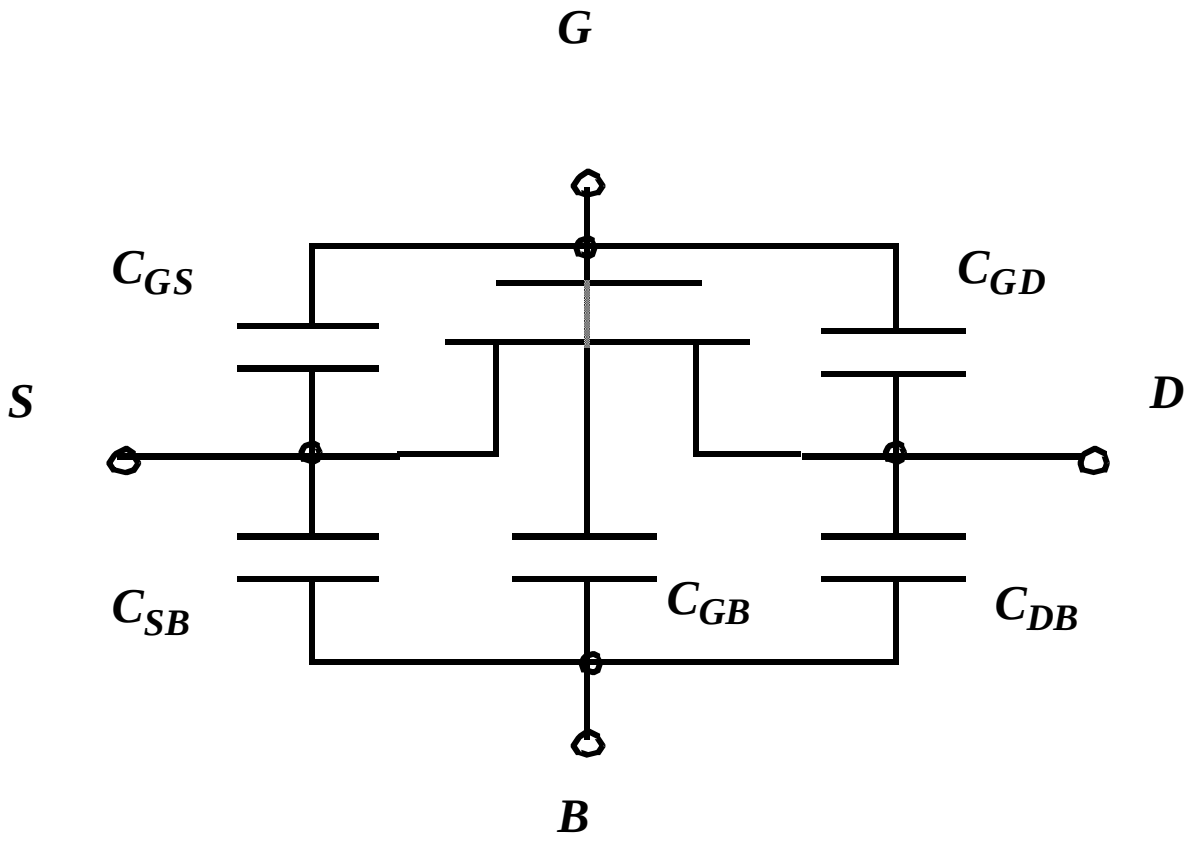
$$R_{eq} = \frac{1}{2} \left[\frac{V_{ds}(t_1)}{I_d(t_1)} + \frac{V_{ds}(t_2)}{I_d(t_2)} \right]$$

Drain-Source Resistance

- Note the following
  - R is inversely proportional to W/L
  - For $V_{DD} >> V_T + V_{D,sat}/2$, R is independent of V_{DD}.
  - When V_{DD} is close to V_T, resistance increases.

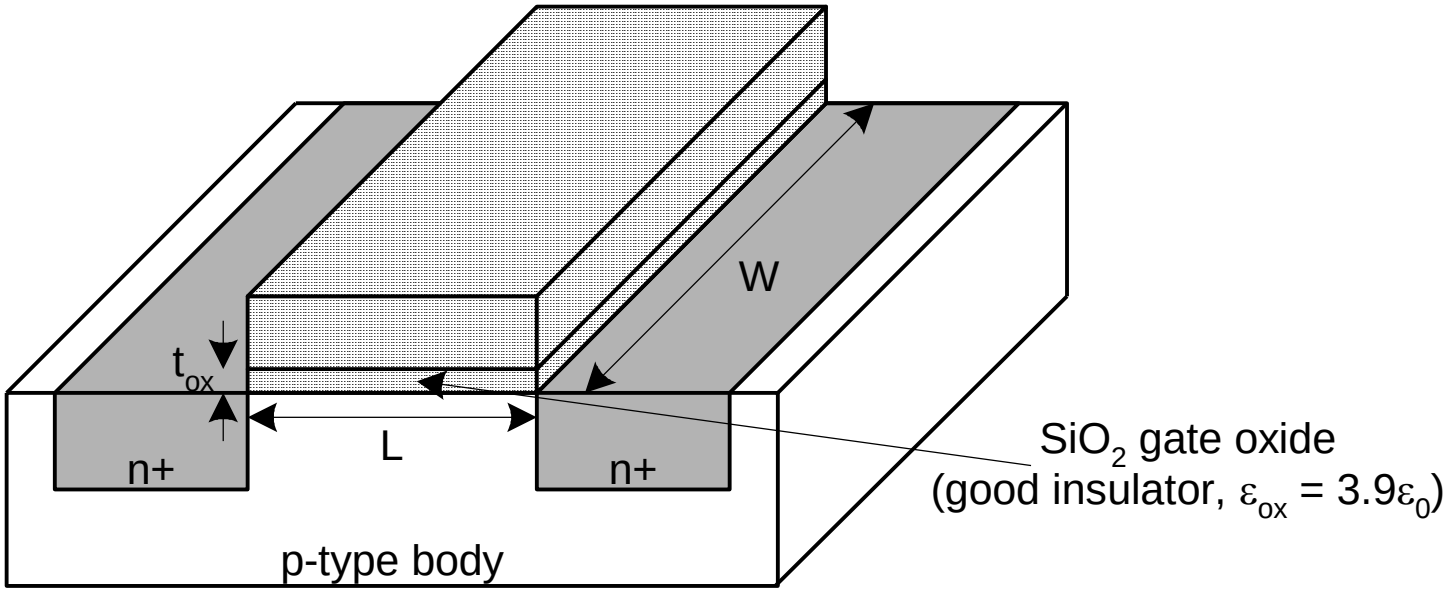
$$R_{on} = \frac{1}{\mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})}$$

MOS Capacitances



Gate Capacitance

- Approximate channel as connected to source
- $C_{gs} = \varepsilon_{ox} WL/t_{ox} = C_{ox} WL = C_{permicron} W$
- $C_{permicron}$ is typically about 2 fF/μm

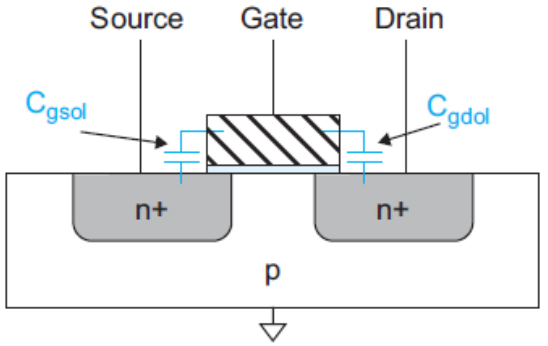


Gate Capacitance

Operation Region	C_{gb}	C_{gs}	C_{gd}
Cut-off	$C_{ox}WL_{eff}$	C_{ov}	C_{ov}
Resistive	0	$C_{ox}WL_{eff}/2 + C_{ov}$	$C_{ox}WL_{eff}/2 + C_{ov}$
Active	0	$(2/3)\ C_{ox}WL_{eff} + C_{ov}$	C_{ov}

$$C_g = C_{gs} + C_{gd} + C_{gb} \approx C_0 + 2C_{gol}W$$

Parameter	Cutoff	Linear	Saturation
C_{gb}	$\leq C_0$	0	0
C_{gs}	0	$C_0/2$	$2/3\ C_0$
C_{gd}	0	$C_0/2$	0
$C_g = C_{gs} + C_{gd} + C_{gb}$	C_0	C_0	$2/3\ C_0$

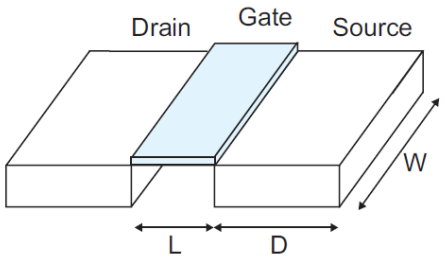


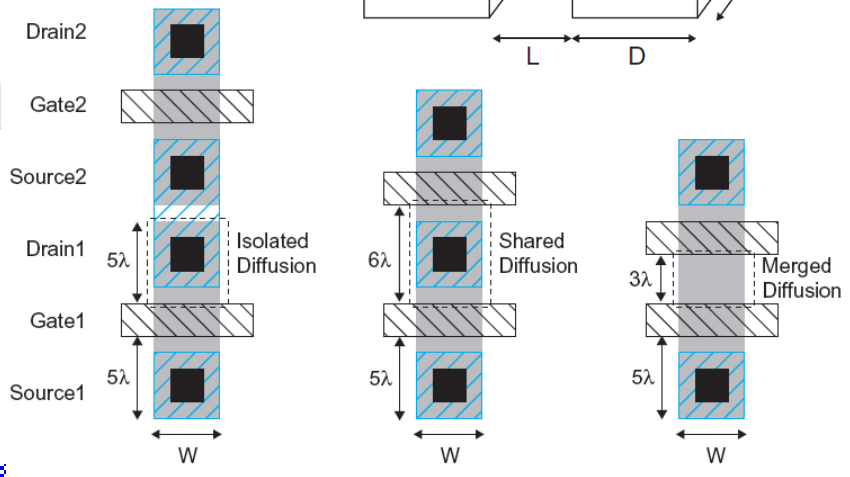
$$C_{gsol(\text{overlap})} = C_{gsol}W$$

$$C_{gdol(\text{overlap})} = C_{gdol}W$$

Diffusion Capacitance

- C_{sb}, C_{db}
- Undesirable, called *parasitic* capacitance
- Capacitance depends on area and perimeter
  - Use small diffusion nodes
  - Comparable to C_g for contacted diff
  - ½ C_g for uncontacted
  - Varies with process





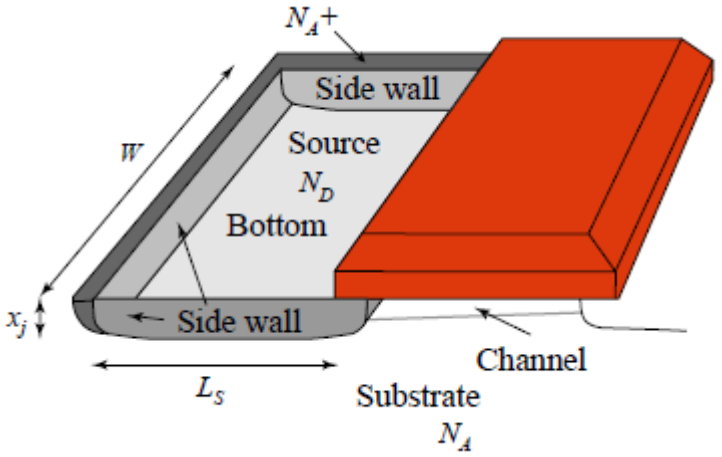
$$C_{sb} = AS \times C_{jbs} + PS \times C_{jbssw}$$

$$C_{jbs} = C_J \left(1 + \frac{V_{sb}}{\psi_0}\right)^{-M_J}$$

$$\psi_0 = v_T \ln \frac{N_A N_D}{n_i^2}$$

$$C_{jbssw} = C_{JSW} \left(1 + \frac{V_{sb}}{\psi_{SW}}\right)^{-M_{JSW}}$$

$$C_{jbsswg} = C_{JSW} \left(1 + \frac{V_{sb}}{\psi_{SWG}}\right)^{-M_{JSWG}}$$



☐ Bottom

–Area cap

–$C_{bottom} = C_j \cdot L_S \cdot W$

☐ Sidewalls

– Perimeter cap

–$C_{sw} = C_{jsw} \cdot (2LS+W)$

☐ Gate edge

–$C_{ge} = C_{jswg} \cdot W$

– Usually automatically included in the SPICE model

$$C_{diff} = C_{bottom} + C_{sw} = C_j \times AREA + C_{jsw} \times PERIMETER$$
$$= C_j L_S W + C_{jsw}(2L_S + W) + C_{jswg} \cdot W$$

Capacitance Model Summary

- **Gate-Channel Capacitance**
  - $C_{gc} \approx 0$ (|VGS| < |VT|)(off)
  - $C_{gc} = C_{gd}$ Cox·W·Leff (Linear)
    - 50% g to s, 50% g to d
  - C_{gc} = (2/3)·Cox·W·Leff (Saturation)
    - 100% g to s
- **Gate Overlap Capacitance**
  - $C_{gsol(\text{overlap})} = C_{gsol} W$
    $C_{gdol(\text{overlap})} = C_{gdol} W$ (Always)
- **Junction/Diffusion Capacitance**
  - C_{diff} = Cj·L_S·W + Cjsw·(2LS + W) + C_{jgswg}W(Always)

Capacitances in 0.25 μm CMOS Process

	C_{ox} (fF/μm^2)	C_O (fF/μm)	C_j (fF/μm^2)	m_j	ϕ_b (V)	C_{jsw} (fF/μm)	m_{jsw}	ϕ_{bsw} (V)
NMOS	6	0.31	2	0.5	0.9	0.28	0.44	0.9
PMOS	6	0.27	1.9	0.48	0.9	0.22	0.32	0.9