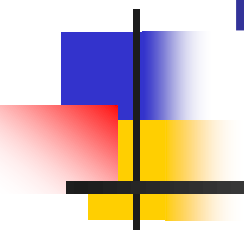


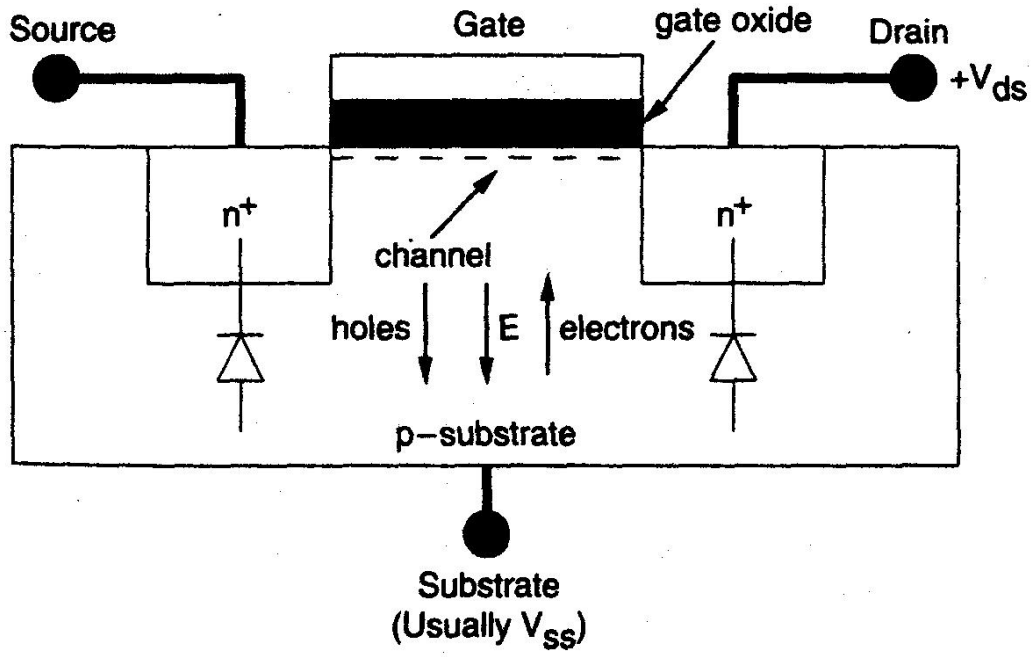
Chapter 2
Devices and Layouts

闕志達

台灣大學電機系

NMOS Enhancement Transistor

- Physical structure of an NMOS transistor.
- Gate, Source, Drain, Substrate, Channel.
- https://www.youtube.com/watch?v=tz62t-q_KEc



MOS Junction Operating Modes

- Different Vgs -- accumulation, depletion, inversion (weak and strong)

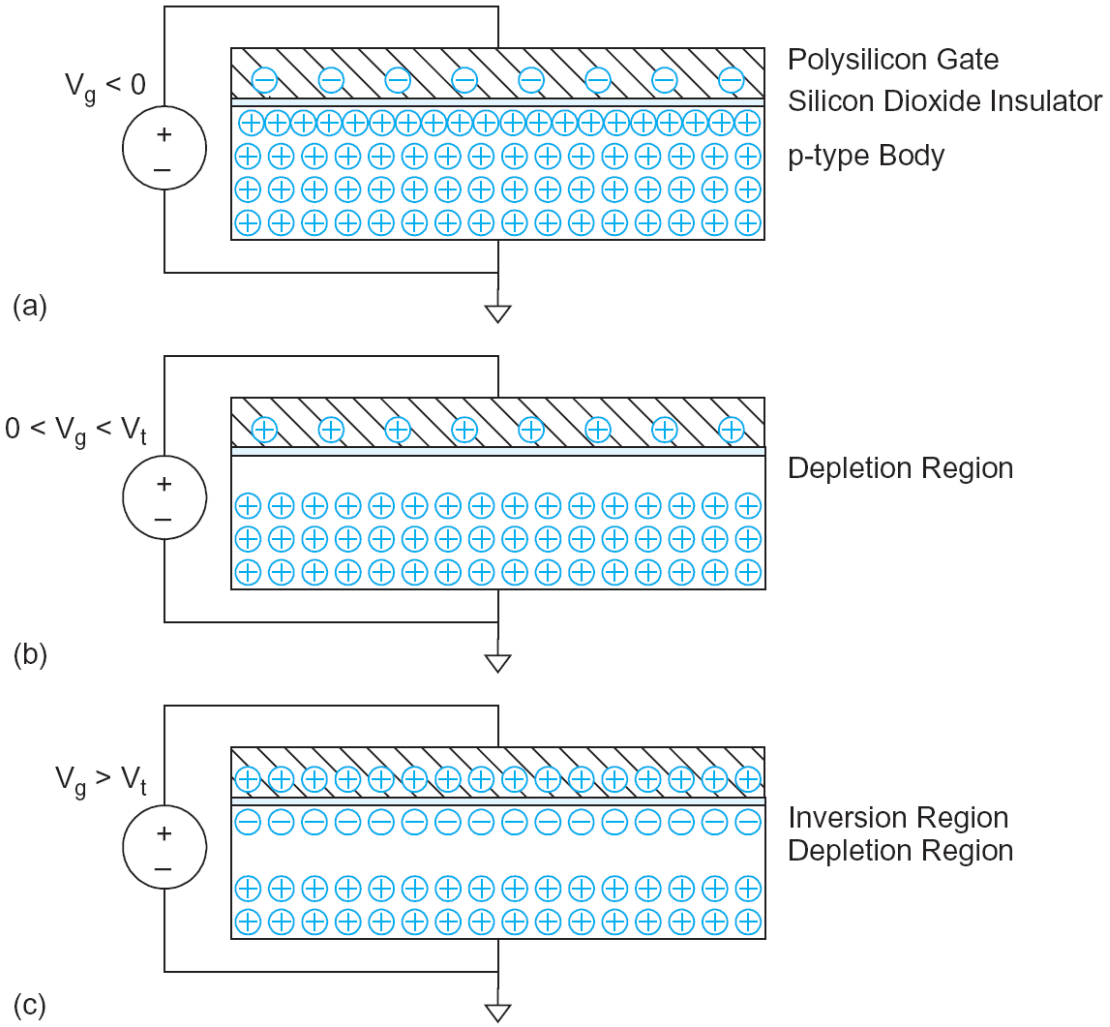
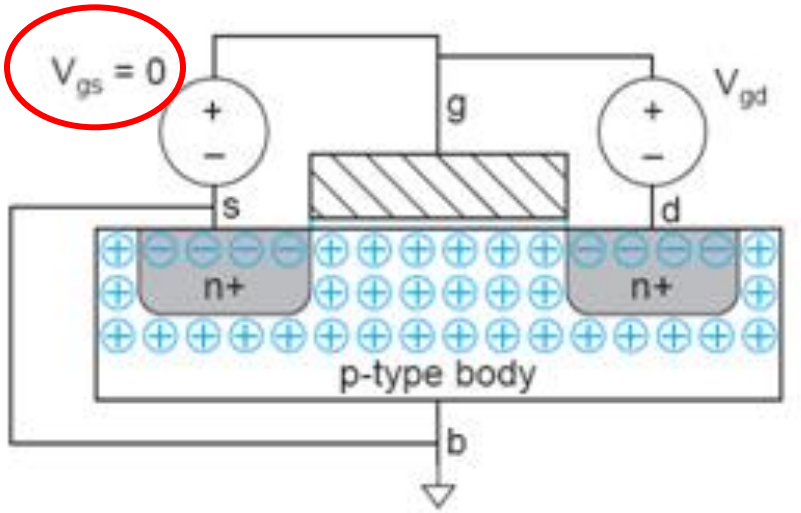


FIGURE 2.2 MOS structure demonstrating (a) accumulation, (b) depletion, and (c) inversion

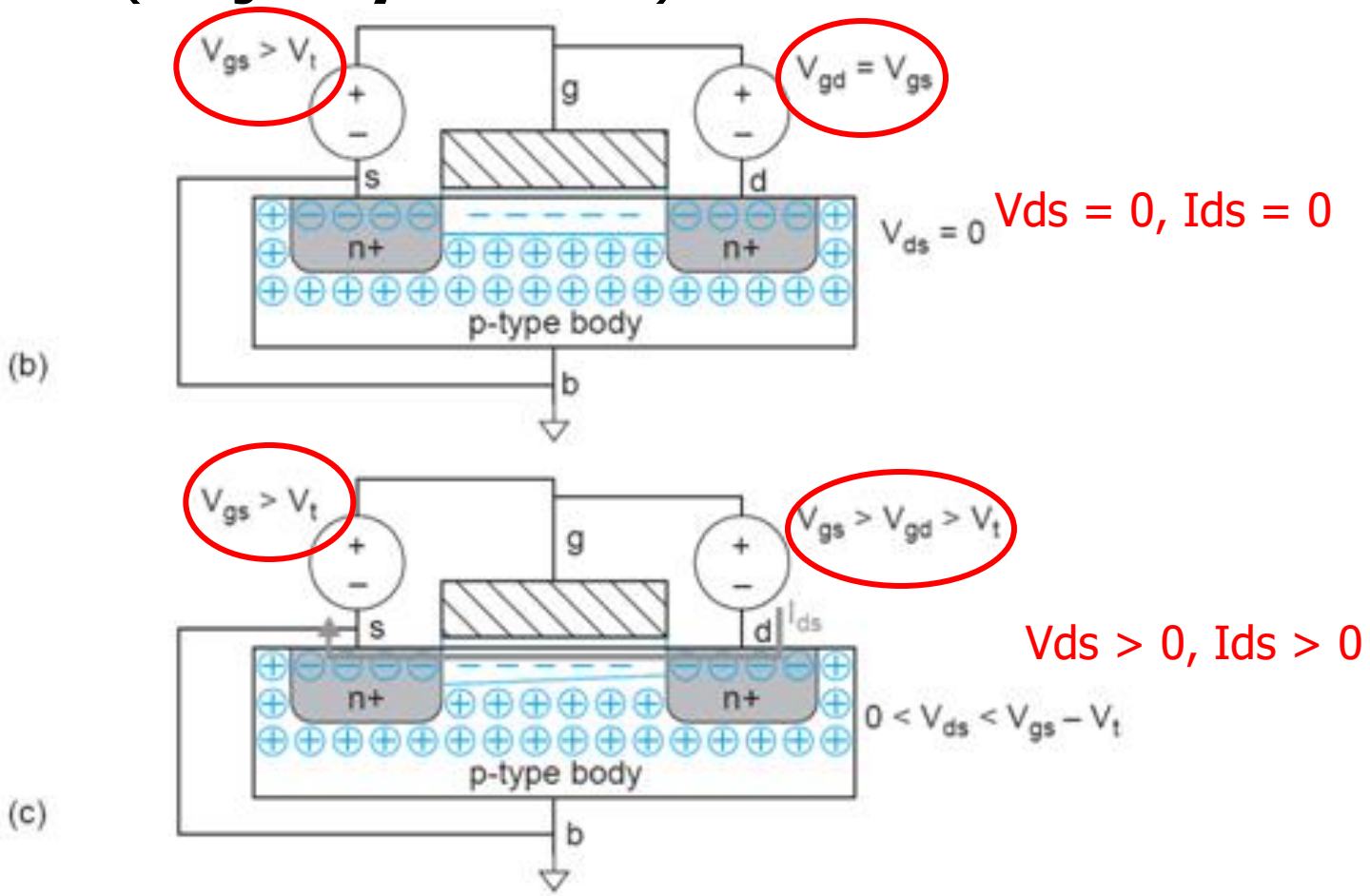
MOS Operating Regions

- Different Vgs and Vds -- Cutoff, linear, saturation
- Cutoff region: No channel, Ids = 0



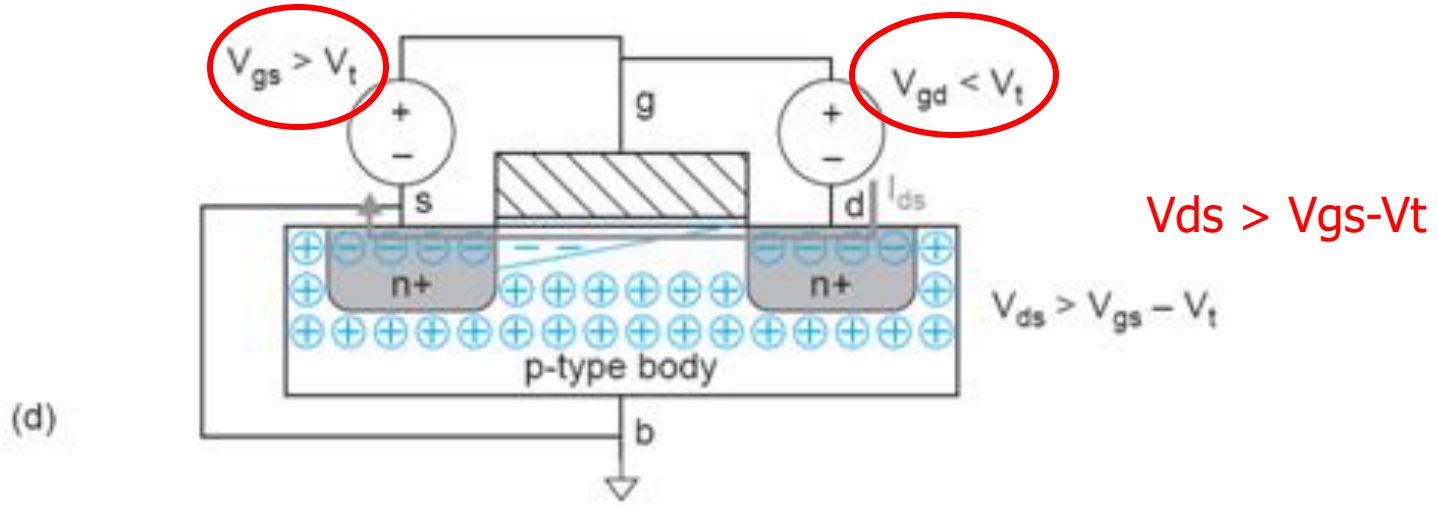
MOS Operating Regions

- Linear region: channel forms, current flow by drift of electrons (**majority carriers**).



MOS Operating Regions

- Saturation region: channel pinches off; current independent of Vds.



PMOS Transistor

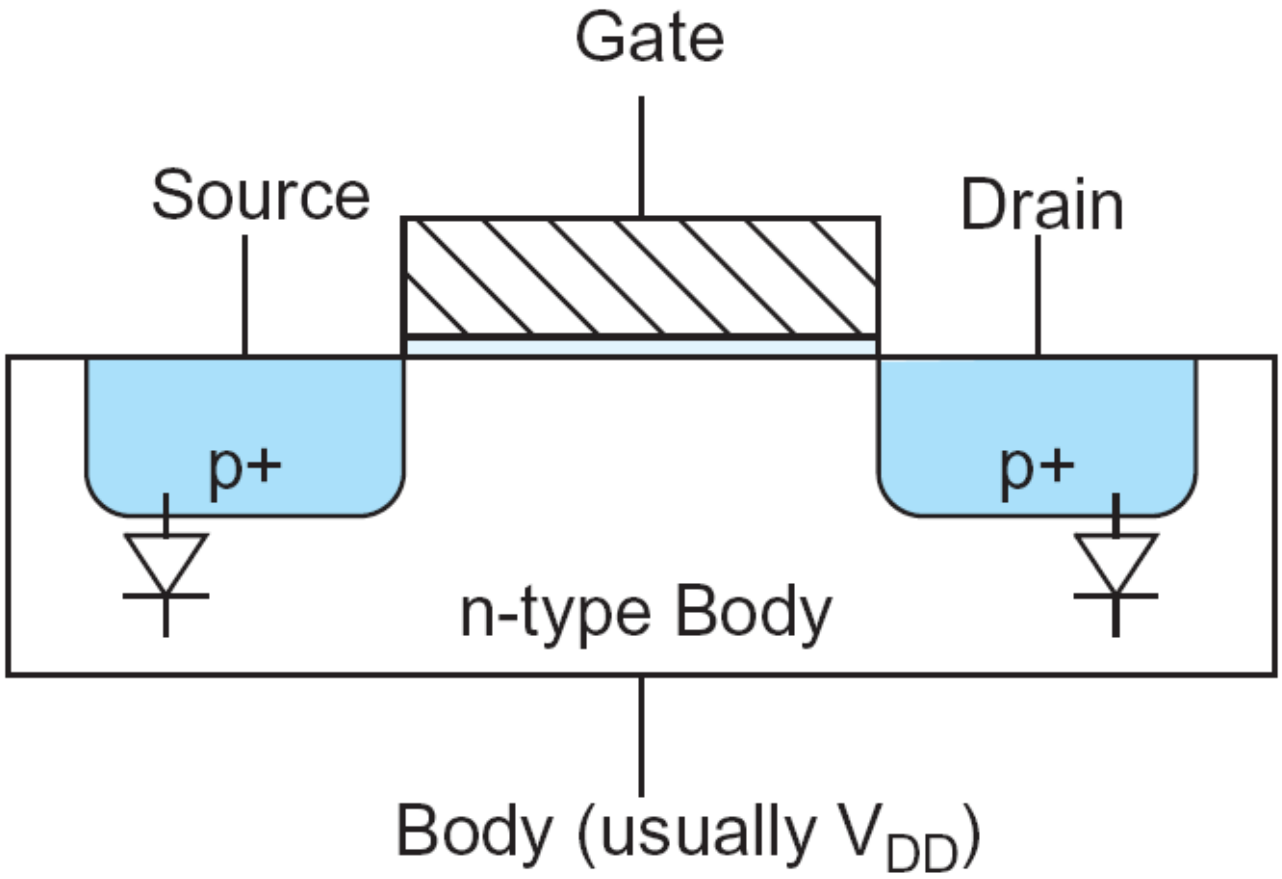
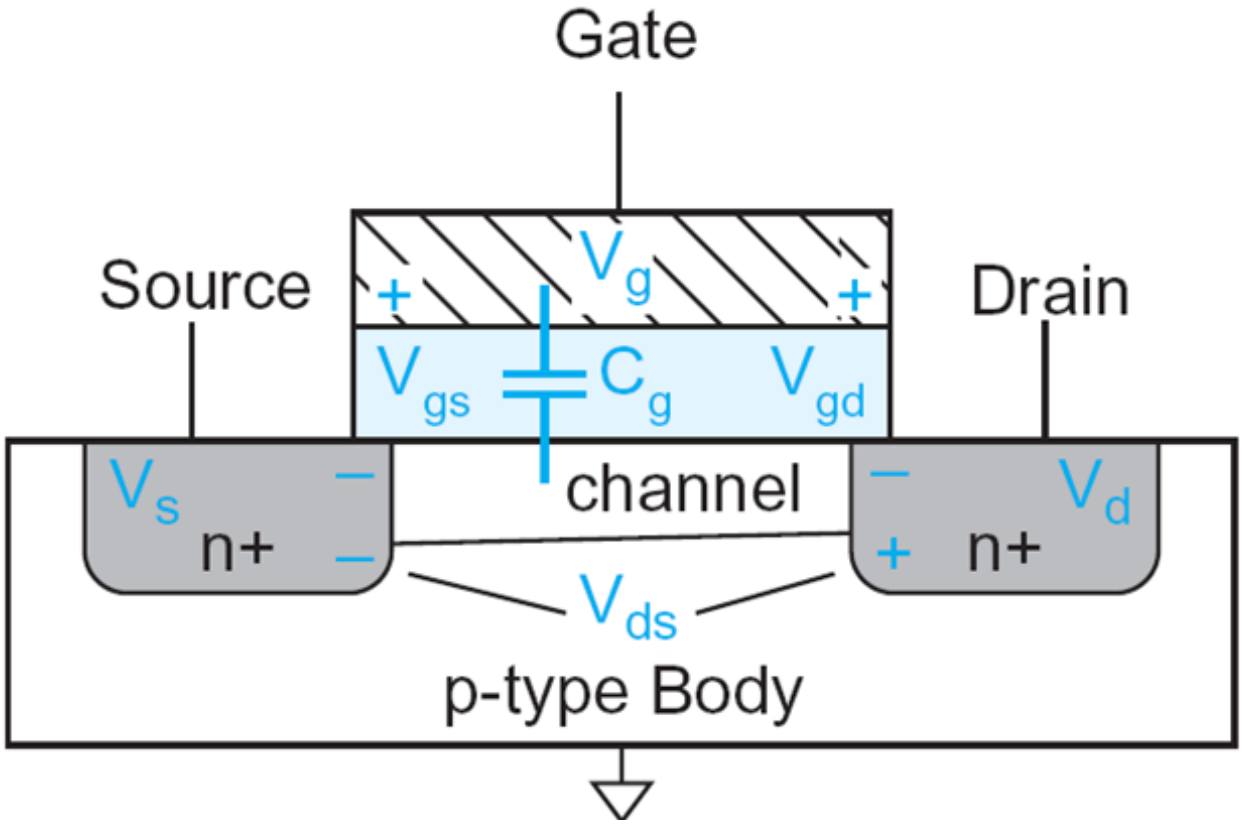


FIGURE 2.4 pMOS transistor

I-V Characteristics

- In linear region, MOS transistor has channel inversion and majority carriers conduct current in the form of drifting by electric field.
- Current is then proportional to
  - the amount of charge in the channel, and
  - velocity of carrier movement
- Speed of carriers depends on
  - electric field, and
  - effective mass

Channel Voltage



Average gate to channel potential:

$$V_{gc} = (V_{gs} + V_{gd})/2 = V_{gs} - V_{ds}/2$$

FIGURE 2.5 Average gate to channel voltage

Channel Charge

- Parallel-plate capacitor
- $Q = C\,V$
- $C = C_g = \varepsilon_{ox} WL/t_{ox} = C_{ox} WL \qquad C_{ox} = \varepsilon_{ox}/t_{ox}$
- $V = V_{gc} - V_t = V_{gs} - V_{ds}/2 - V_t$

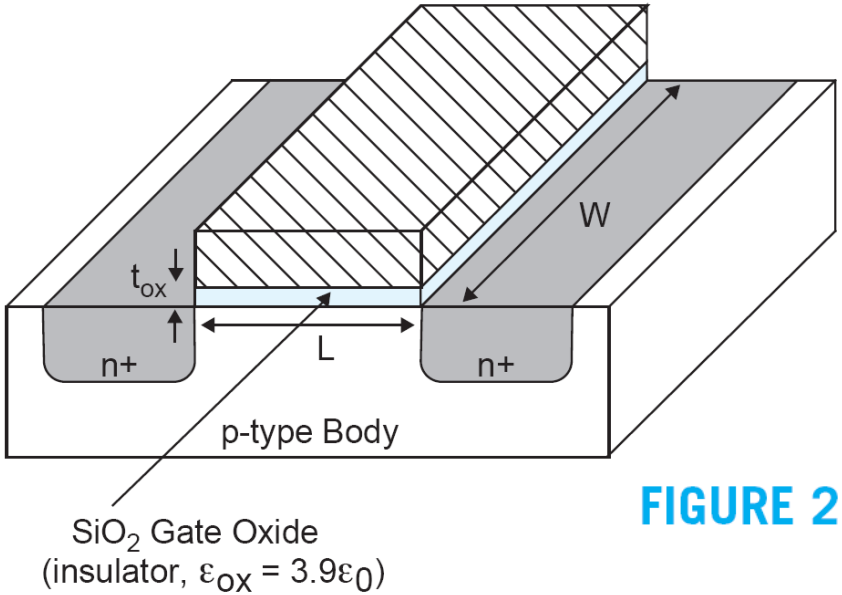


FIGURE 2.6 Transistor dimensions

Carrier Velocity

- Average velocity is related to electric filed as

$$v = \mu E,$$

μ is the mobility of the carrier.

- Electric filed is given by

$$E = V_{ds}/L$$

- Time for channel crossing

$$t = L/v$$

MOS Current in Linear Region

- $I_{ds} = Q/t$

$$= C_{ox}WL(V_{gs}\text{-}V_t\text{-}V_{ds}/2)/(L/v)$$

$$= C_{ox}WL(V_{gs}\text{-}V_t\text{-}V_{ds}/2)v/(L)$$

$$= \mu C_{ox}WL(V_{gs}\text{-}V_t\text{-}V_{ds}/2)V_{ds}/(L^2)$$

$$= \mu C_{ox}(W/L)(V_{gs}\text{-}V_t\text{-}V_{ds}/2)V_{ds}$$

$$= \beta(V_{gs}\text{-}V_t\text{-}V_{ds}/2)V_{ds} \qquad \beta = \mu C_{ox}(W/L)$$

MOS current in saturation

- When the drain voltage gets higher and $V_{gd} < V_t$, the channel pinches off.
- When $V_{ds} > V_{ds(sat)} = V_{gs} - V_t$, the current does not increase as V_{ds} increases.

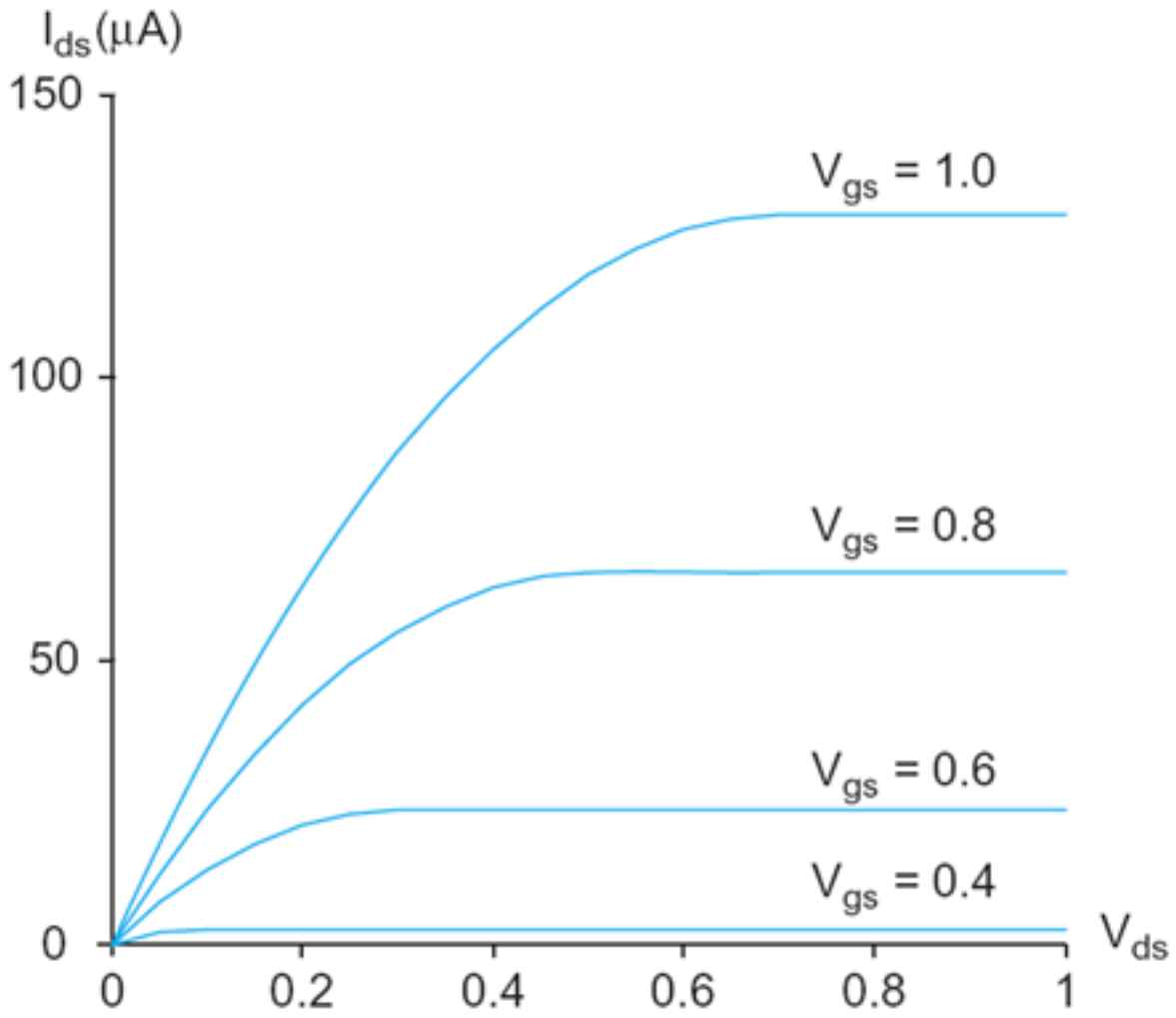
$$I_{ds} = \beta (V_{gs}\text{-}V_t\text{-}V_{ds(sat)}/2)V_{ds(sat)}$$

$$= (\beta/2)(V_{gs}\text{-}V_t)^2$$

NMOS I-V Summary

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

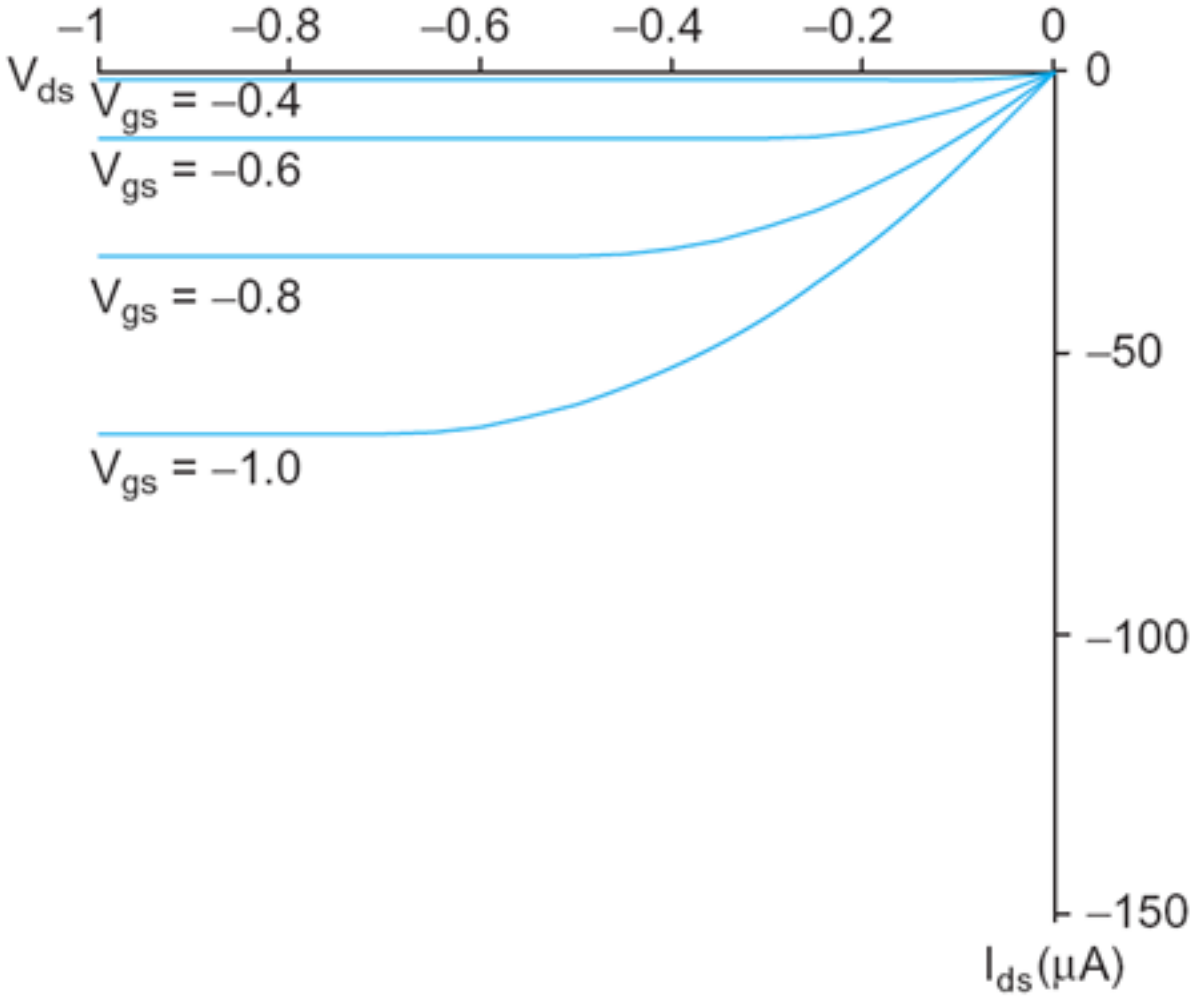
NMOS I-V Curves



PMOS Transistor

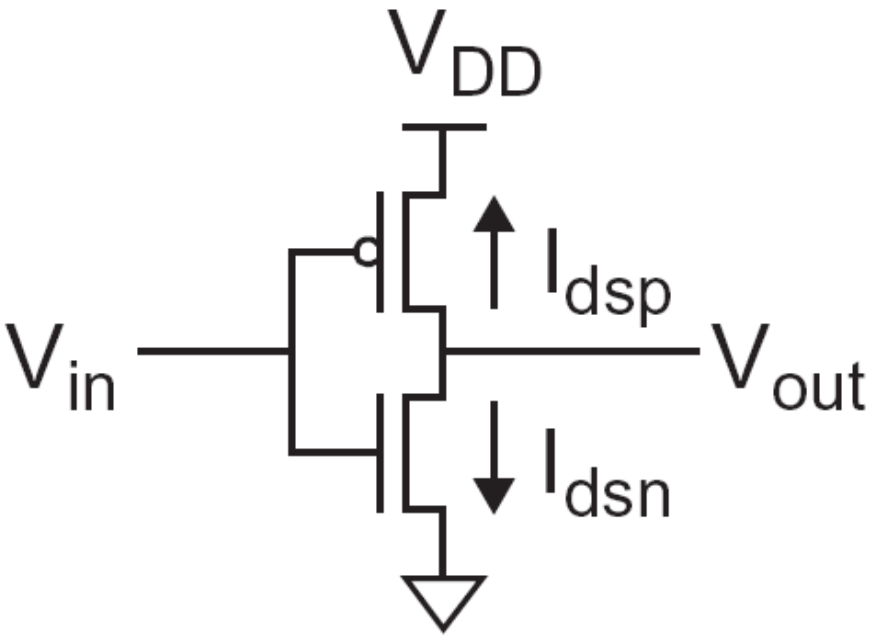
- All doping and voltages are inverted
- Mobility is determined by holes, which is about 2-3 times lower than electrons.
- 2X-3X larger (W/L) needed for PMOS to achieve same level of current.

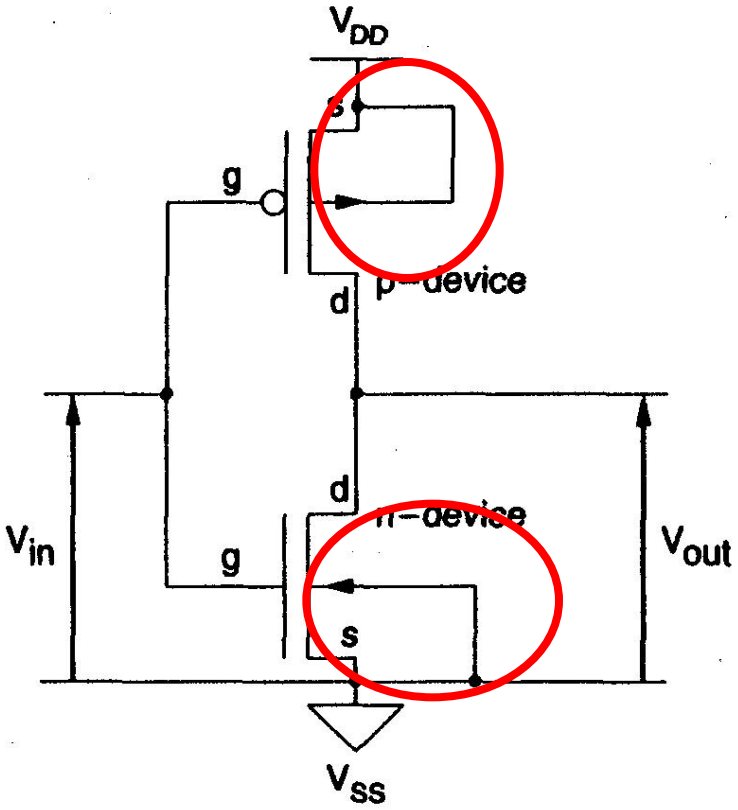
PMOS I-V Curves



CMOS Inverter

- One NMOS and one PMOS in series, both gates being controlled by the input.

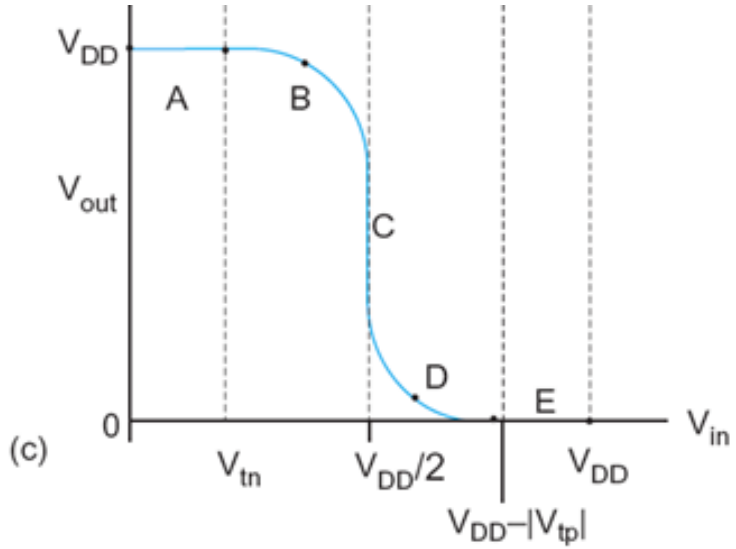




Five Regions

Table 2.3 Summary of CMOS inverter operation

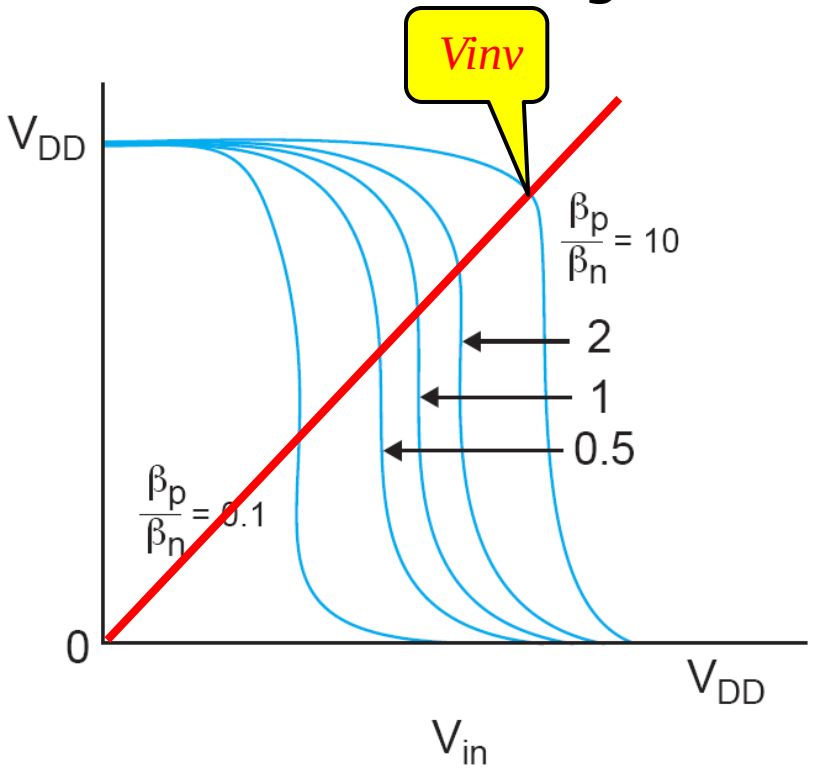
Region	Condition	p-device	n-device	Output
A	$0 \le V_{in} < V_{tn}$	linear	cutoff	$V_{out} = V_{DD}$
B	$V_{tn} \le V_{in} < V_{DD}/2$	linear	saturated	$V_{out} > V_{DD}/2$
C	$V_{in} = V_{DD}/2$	saturated	saturated	V_{out} drops sharply
D	$V_{DD}/2 < V_{in} \le V_{DD} - \lvert V_{tp}\rvert$	saturated	linear	$V_{out} < V_{DD}/2$
E	$V_{in} > V_{DD} - \lvert V_{tp}\rvert$	cutoff	linear	$V_{out} = 0$



NMOS/PMOS Size ratio

- The gate threshold voltage, V_{inv}, where $V_{in} = V_{out}$ is controlled by the β_n/β_p ratio. For a β_n/β_p ratio of one, the inverter has the same current sourcing and sinking capability, i.e., the NMOS transistor is as strong as the PMOS transistor.

$$\beta_n = \frac{\mu_n \varepsilon}{t_{ox}}\left(\frac{W_n}{L_n}\right) \qquad \beta_p = \frac{\mu_p \varepsilon}{t_{ox}}\left(\frac{W_p}{L_p}\right)$$

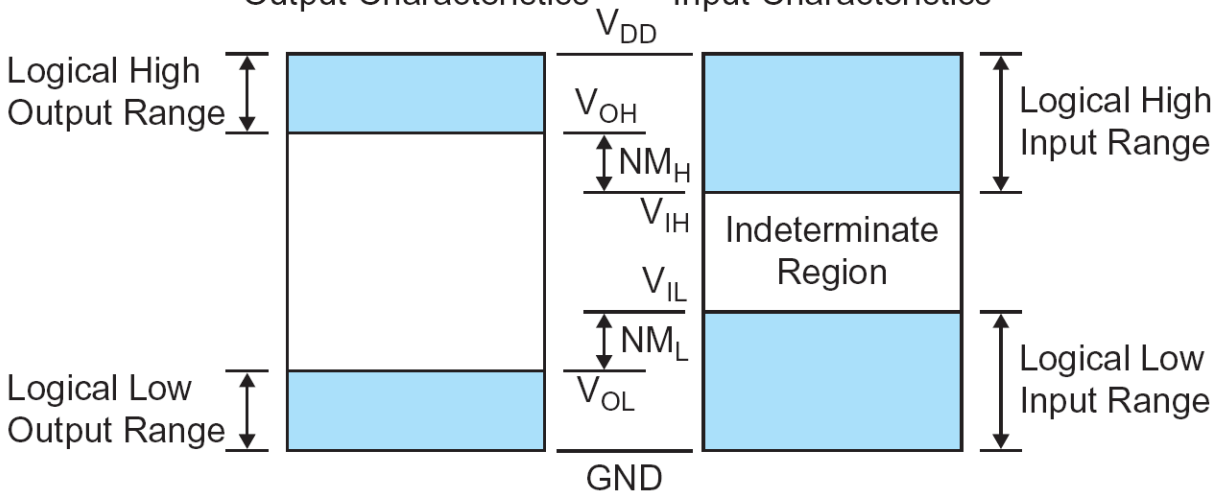


Noise Margin

- Noise margin is the margin of noise that can be inflicted on a gate output before it wrongly switches the next-stage circuit.

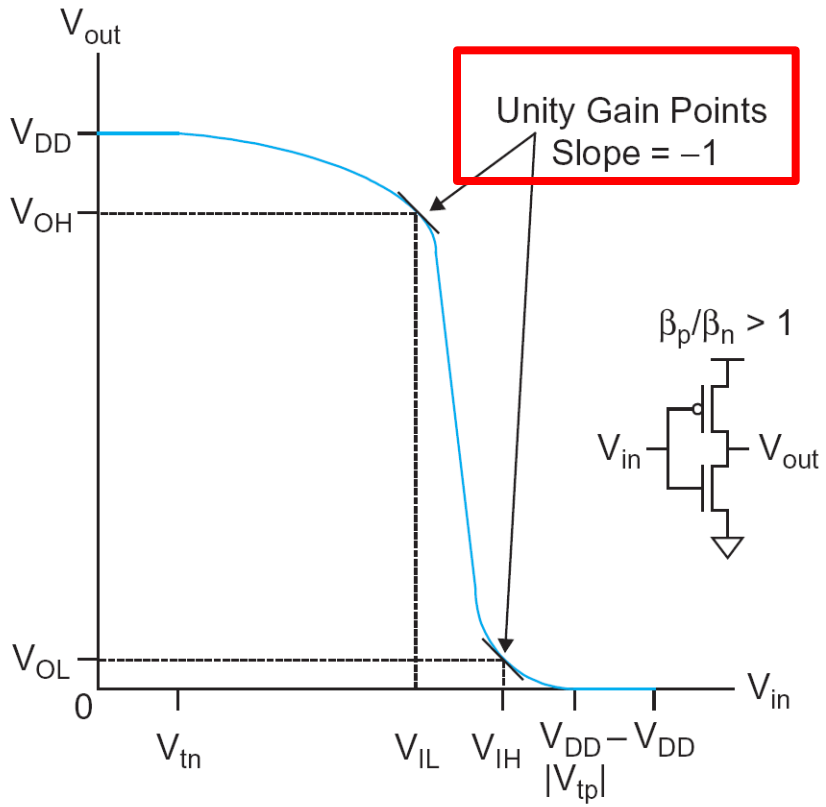
$$NM_L = V_{IL} - V_{OL} \qquad NM_H = V_{OH} - V_{IH}$$

V_{IH} = minimum HIGH input voltage
V_{IL} = maximum LOW input voltage
V_{OH} = minimum HIGH output voltage
V_{OL} = maximum LOW output voltage.



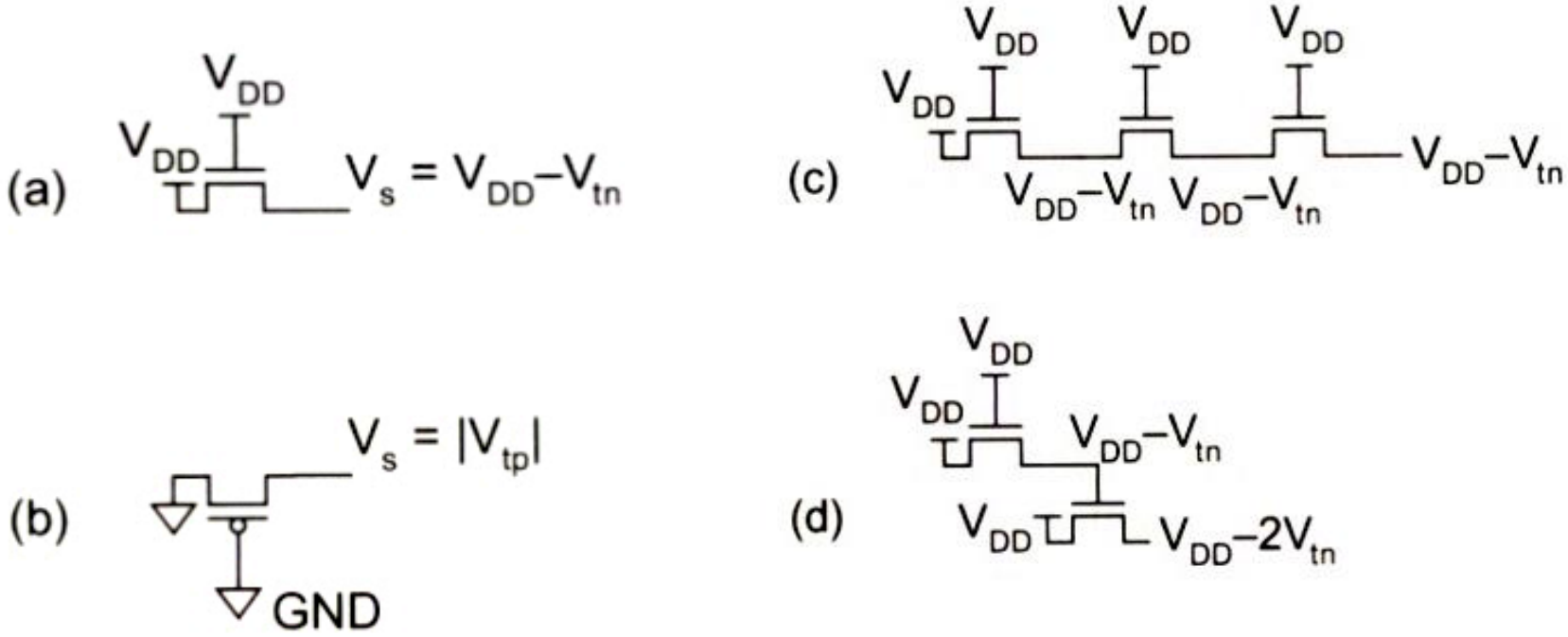
Noise Margin (cont'd)

- The determination of V_{IH}, V_{IL}, V_{OH}, V_{OL} is achieved by locating the two unity gain points (slope = -1) in the transfer characteristic of the inverter.



Pass Transistor

- Threshold drop



CMOS Layout Video

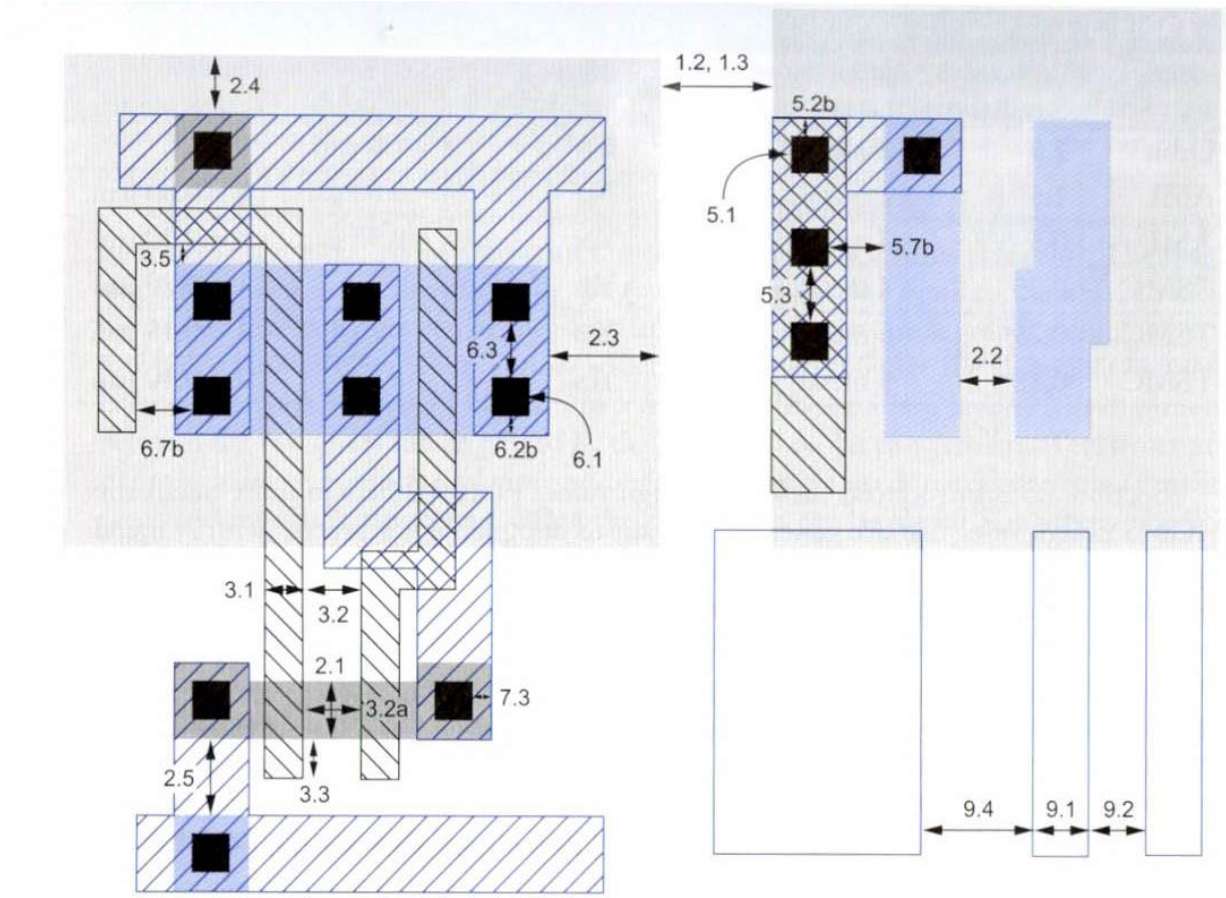
- https://www.youtube.com/watch?v=S6-3fMIdoa8

Design Rules

- Rules that specify some limitation on the precision of IC fabrication.
- Compromise between performance and yield.
- Represent a tolerance that ensures very high probability of correct fabrication.
- A layout violating some rules can still work, and vice versa.
- λ rules and μ rules. λ rules are more conservative.

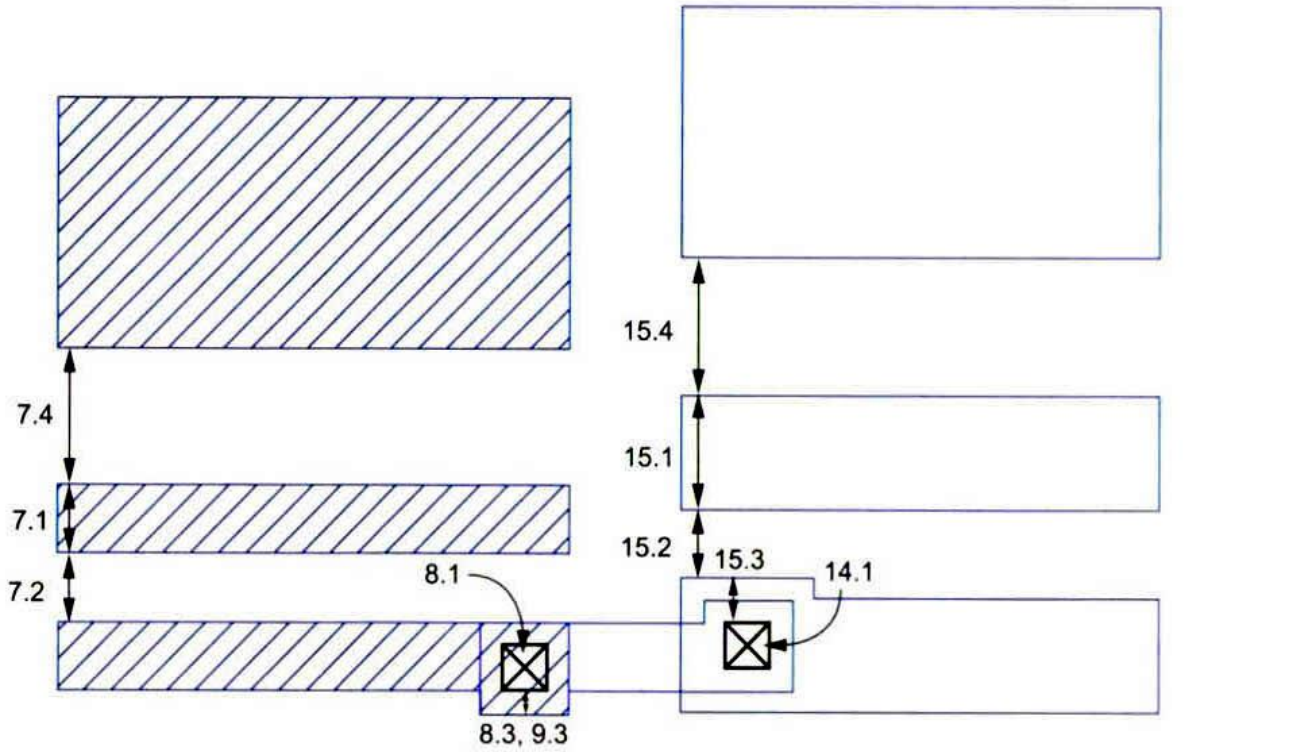
Layout Design Rules

- Design rules are rules on the width, spacing, extension, overlap of the same layer or between different layers.



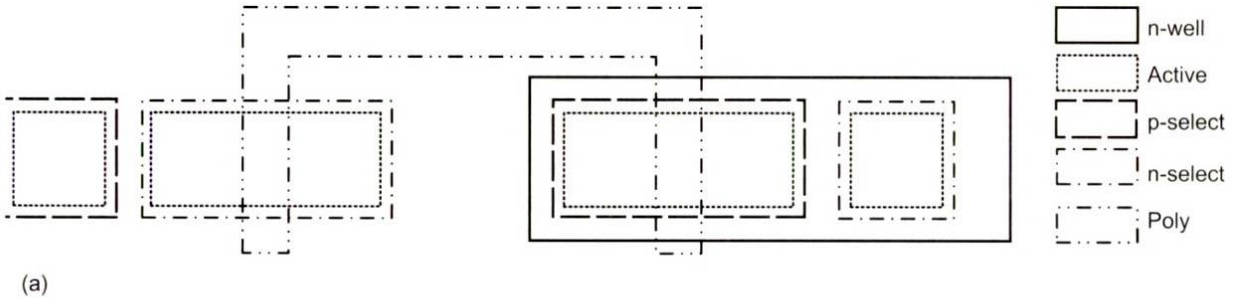
More Design Rules

- Rules for Metal 1 and Metal 3.

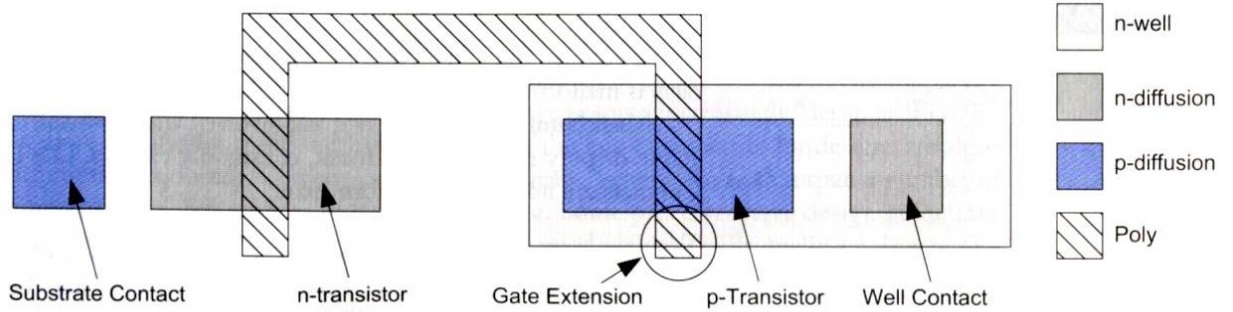


Design Rule Background

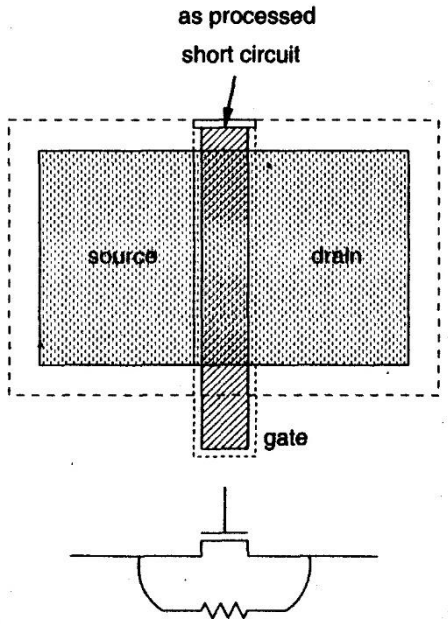
- Well rules -- Outside clearance is due to lateral diffusion. Inside clearance is to prevent bird's beak produced by field oxide encroachment.
- Transistor rules -- to prevent short-circuited MOS



(a)

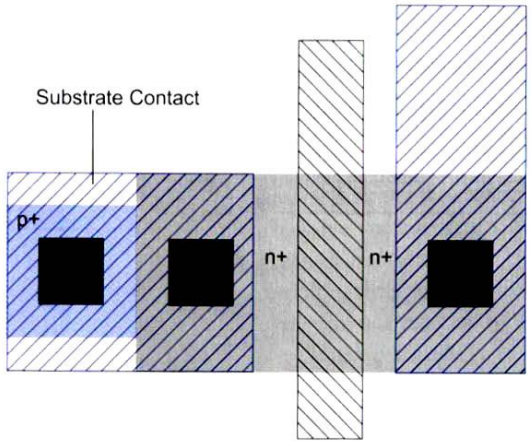


(b)

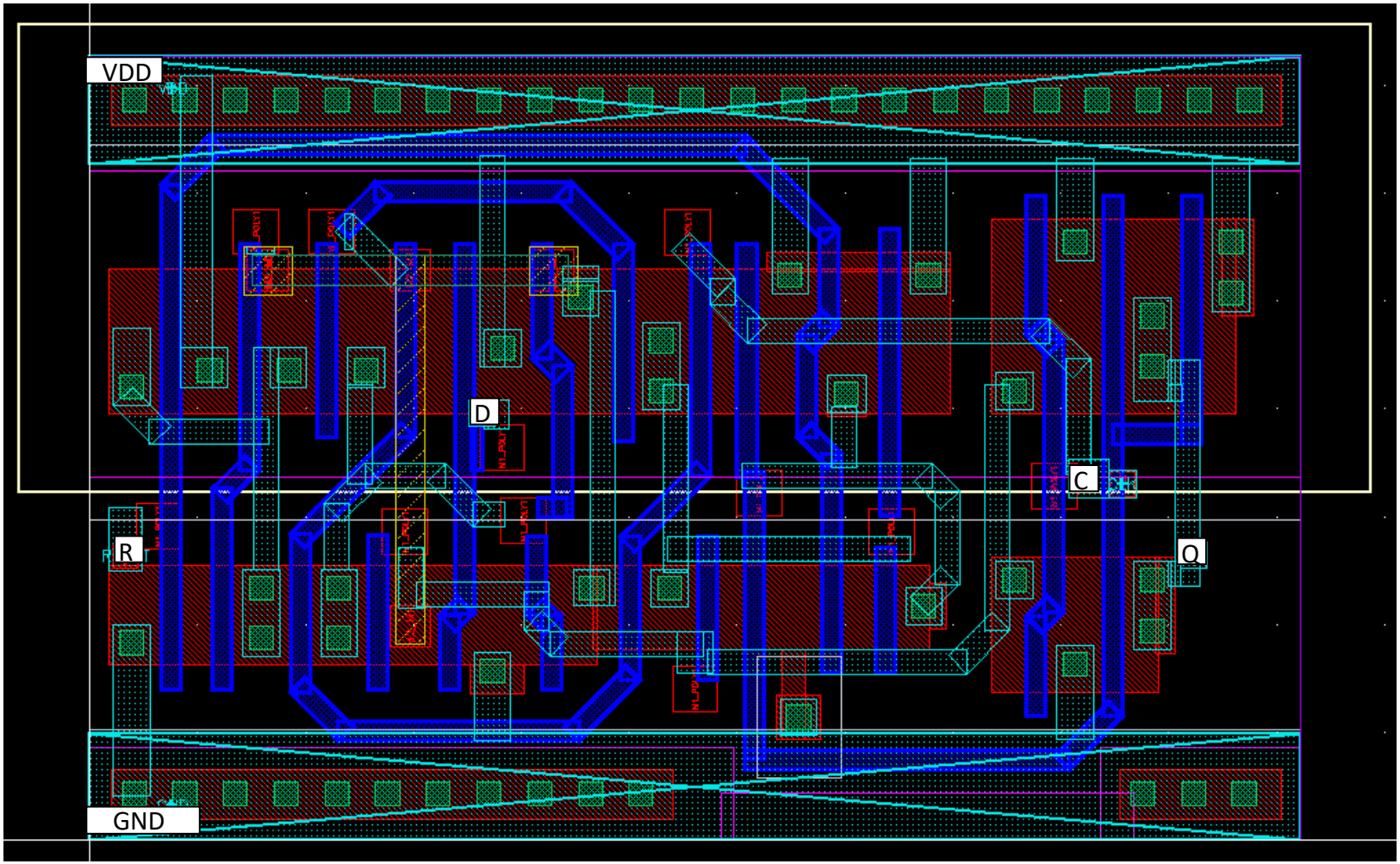


Contact/Metal/Via Rules

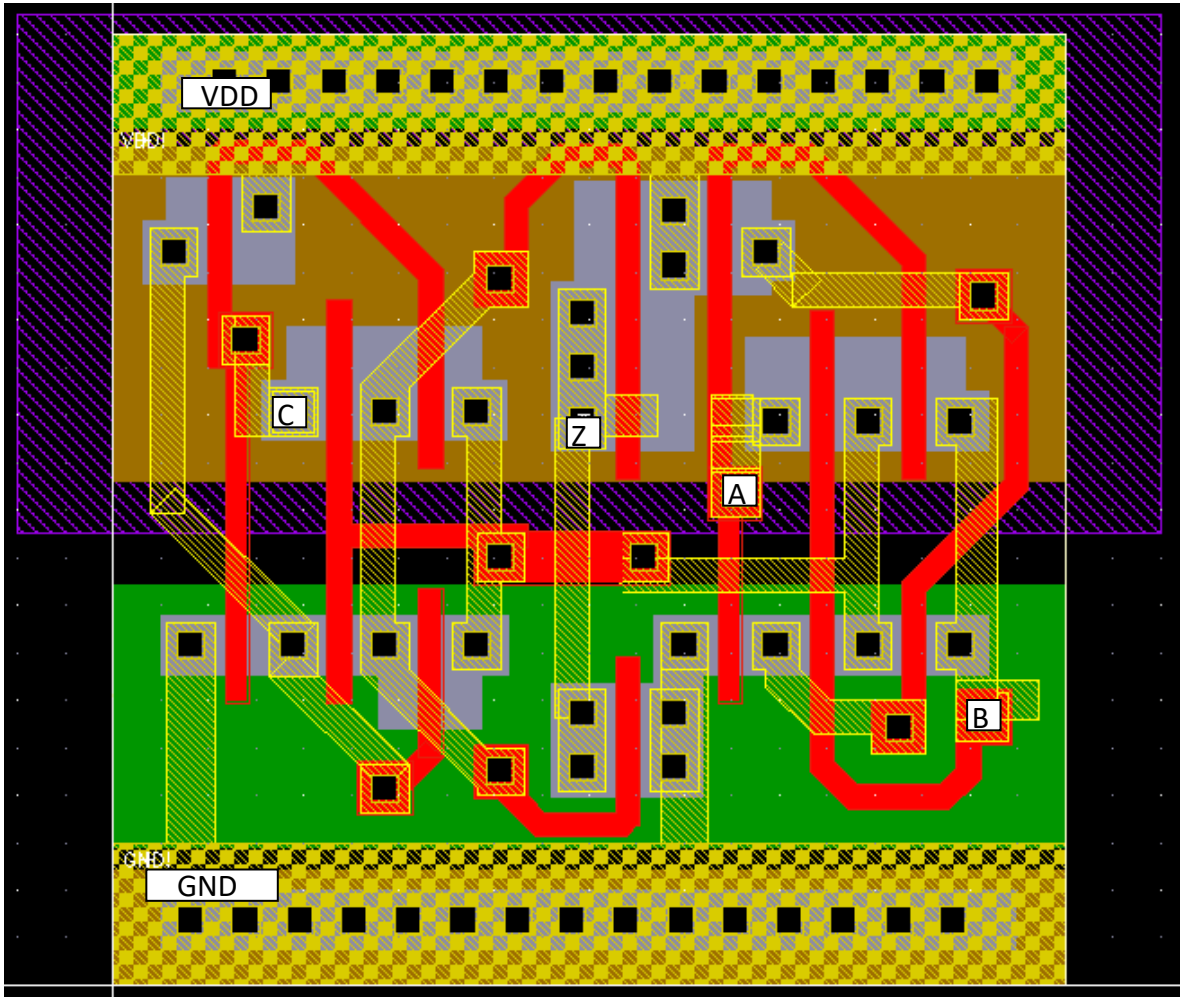
- Contacts from metal to diffusion, poly, well, substrate. Uniform sized squares. As many as possible.
- Metal spacing increases as metal width and thickness.
- Stacked via (between metals) can save area, but no straddling boundary of lower-level contact edges.



D-FF Layout

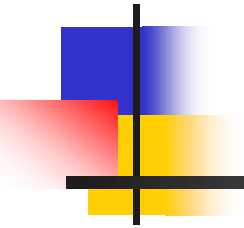


Full Adder Layout (Partial)



A micron rule set

Table 3.3	Micron design rules for 65nm process		
Layer	**Rule**	**Description**	**65 nm rule (μm)**
Well	1.1	Width	0.5
	1.2	Spacing to well at different potential	0.7
	1.3	Spacing to well at same potential	0.7
Active (diffusion)	2.1	Width	0.10
	2.2	Spacing to active	0.12
	2.3	Source/drain surround by well	0.15
	2.4	Substrate/well contact surround by well	0.15
	2.5	Spacing to active of opposite type	0.25
Poly	3.1	Width	0.065
	3.2	Spacing to poly over field oxide	0.10
	3.2a	Spacing to poly over active	0.10
	3.3	Gate extension beyond active	0.10
	3.4	Active extension beyond poly	0.10
	3.5	Spacing of poly to active	0.07
Select	4.1	Spacing from substrate/well contact to gate	0.15
	4.2	Overlap of active	0.12
	4.3	Overlap of substrate/well contact	0.12
	4.4	Spacing to select	0.20
Contact (to poly or active)	5.1, 6.1	Width (exact)	0.08
	5.2b, 6.2b	Overlap by poly or active	0.01
	5.3, 6.3	Spacing to contact	0.10
	5.4	Spacing to gate	0.07
Metal1	7.1	Width	0.09
	7.2	Spacing to well metal1	0.09
	7.3, 8.3	Overlap of contact or via	0.01
	7.4	Spacing to metal for lines wider than 0.5 μm	0.30
Via1–Via5	8.1, 14.1, …	Width (exact)	0.10
	8.2, 14.2, …	Spacing to via on same layer	0.10



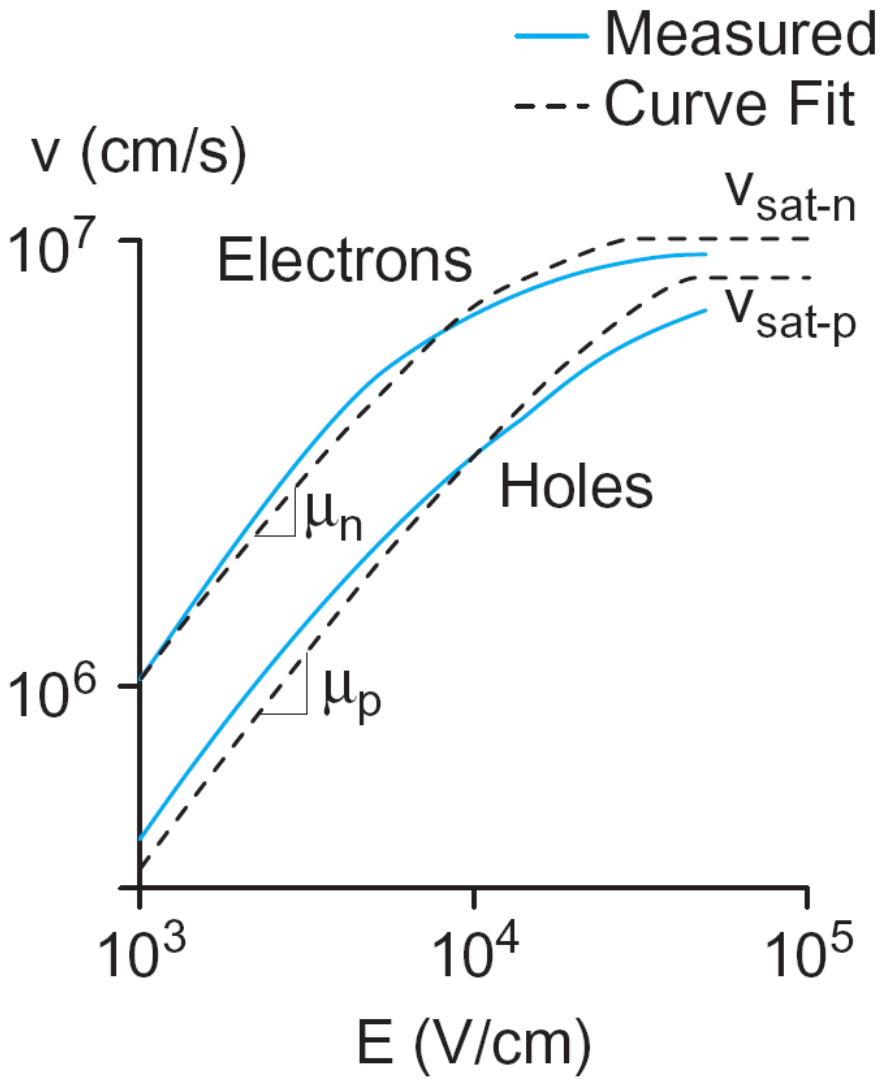
補充教材

Non-ideal I-V Effects

- Velocity Saturation/Mobility Degradation
- Channel Length Modulation
- Body Effect, other V_t Effects
- Subthreshold Conduction (Channel Leakage)
- Tunneling (Gate Leakage)
- Junction Leakage
- Temperature Dependence

Velocity Saturation

- The mobility, μ, is the ease with which carriers drift in the substrate will cease to scale linearly with electric field (E) as it becomes large.



Mobility Degradation

- Mobility for electron is greater than that for holes, hence NMOS transistors usually carry larger currents than PMOS ones.
- Mobility decreases as the temperature or the doping concentration increases.
- Mobility also decreases as the vertical electric field (V_{gs}-V_t) increases.

Channel Length Modulation

- Higher V_{ds} shortens the length of the channel due to depletion region around the drain.

$$L_{\text{eff}} = L - L_d$$

- The drain current is affected

$$I_{ds} = \frac{\beta}{2}(V_{gs} - V_t)^2 \left(1 + \frac{V_{ds}}{V_A}\right)$$

where V_A is called the Early voltage.

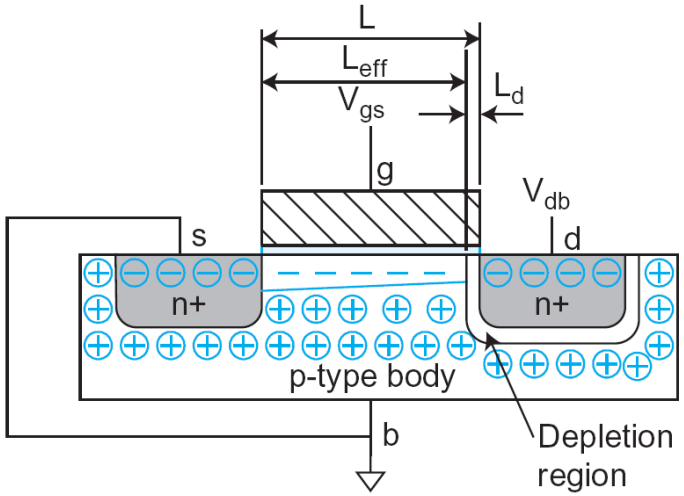


FIGURE 2.18 Depletion region shortens effective channel length

Actual I-V Characteristics

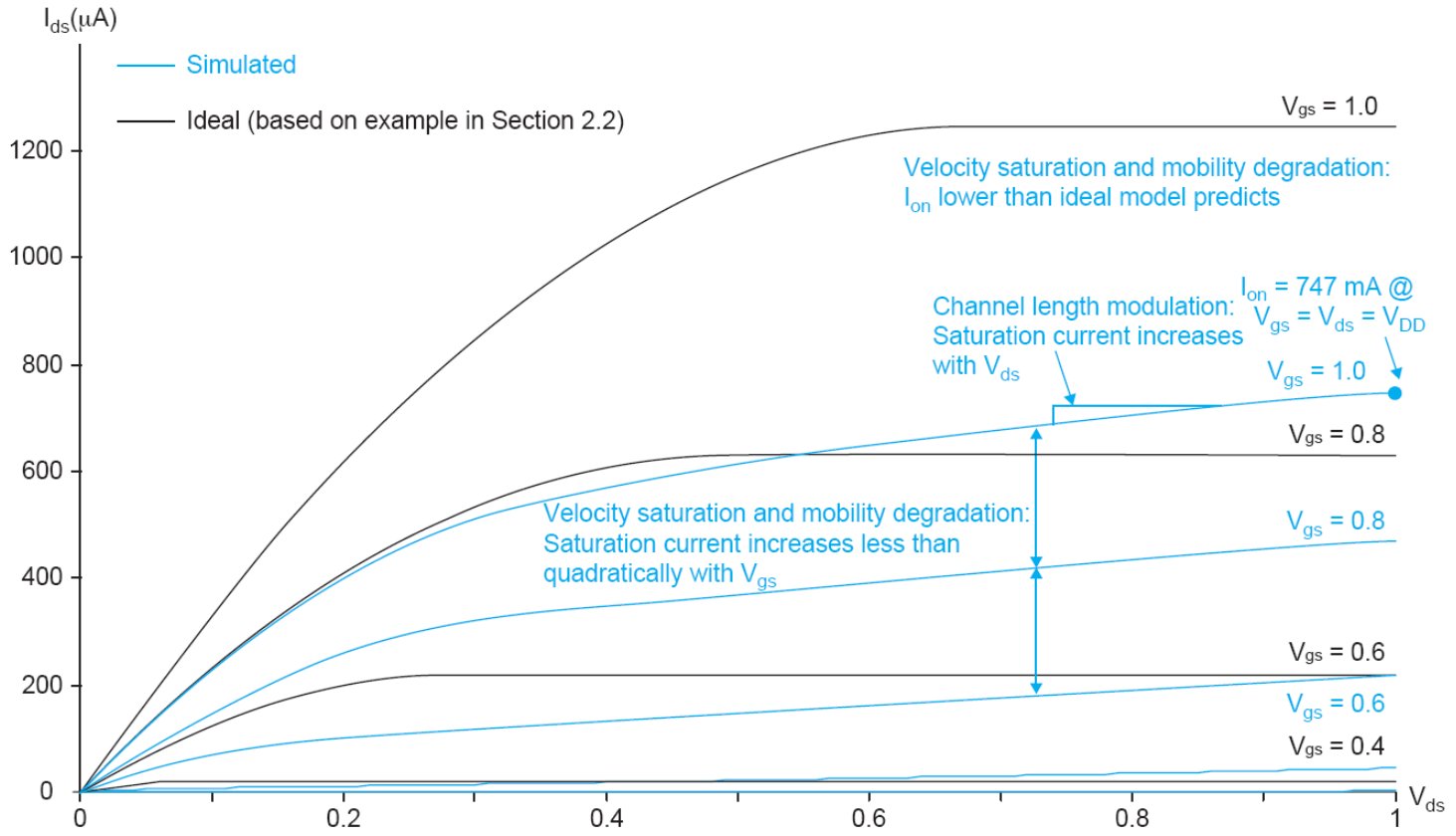


FIGURE 2.14 Simulated and ideal I-V characteristics

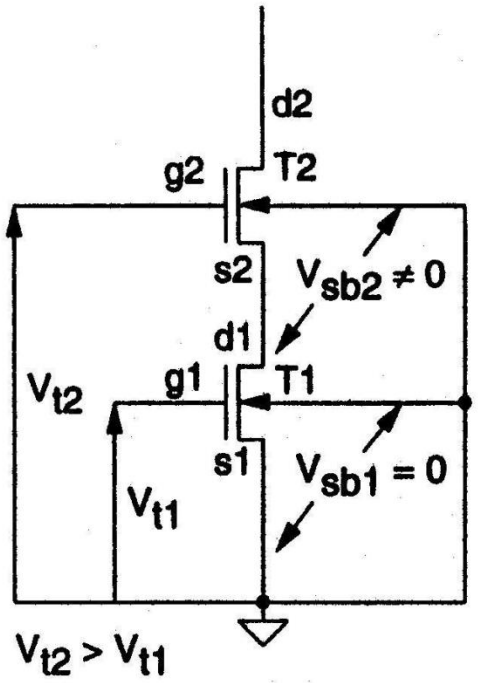
Body Effect

- Body is the fourth terminal of an MOS transistor
- An effect that influences the threshold voltage is the body effect, which is caused by nonzero source to bulk (substrate) voltage (V_{sb})

$$V_t = V_{t0} + \gamma\left(\sqrt{\phi_s + V_{sb}} - \sqrt{\phi_s}\right)$$

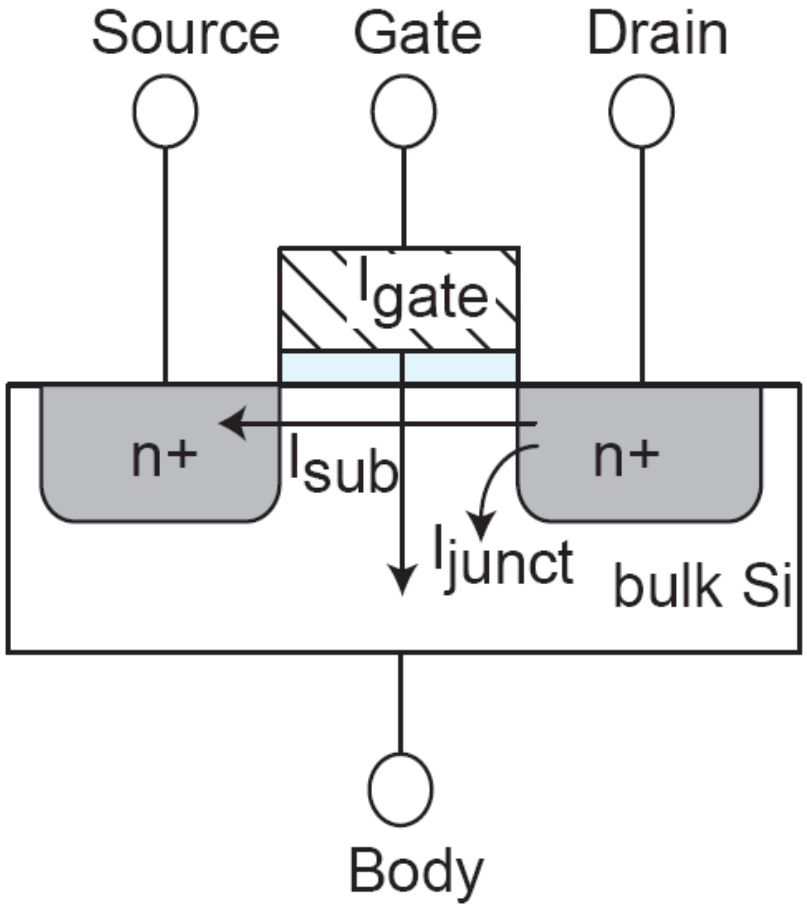
where γ is the body effect coefficient.

- For example, nonzero V_{sb2} makes $V_{t2} > V_{t1}$.



Leakage

- Three different paths



Subthreshold Conduction

- MOS transistors in the cutoff region actually conduct current except that the current is very small (on the order of 10^{-6} to 10^{-11}A) . So this region should be named subthreshold region ($V_{gs} < V_t$), and in this region the drain current depends exponentially on V_{gs}

$$I_{ds} = I_{ds0} e^{\frac{V_{gs} - V_t}{n v_T}} \left(1 - e^{\frac{-V_{ds}}{v_T}} \right)$$

- v_T is the thermal voltage and is 26mV at room temperature. V_t can be modified by body effect and DIBL effect.

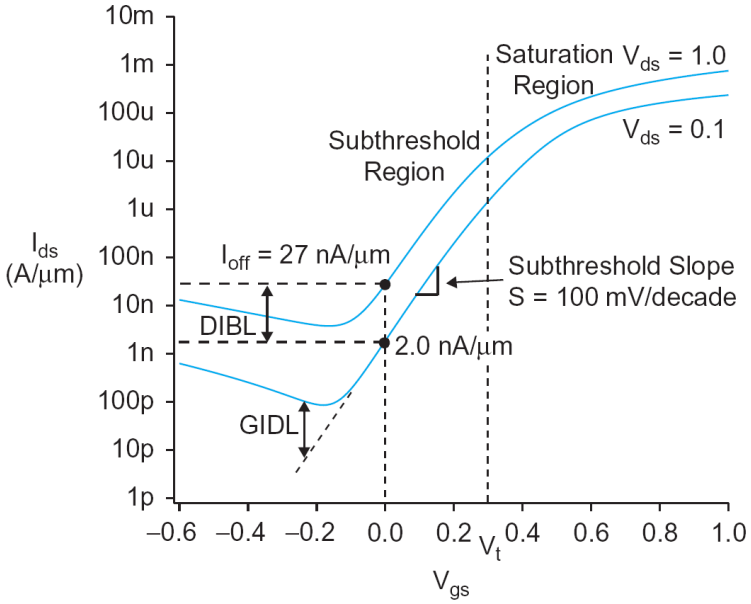


FIGURE 2.20 I-V characteristics of a 65 nm nMOS transistor at 70 °C on a log scale

Channel Leakage

- This current is also considered leakage current.
- As V_t becomes smaller in advanced technology subthrehsold conduction becomes a major problem in power consumption since it seems that all open switches in a digital IC are leaking and there are billions of such switches.
- Some technology offers option to increase V_t of selected transistors to reduce such leakage using another mask.

Tunneling

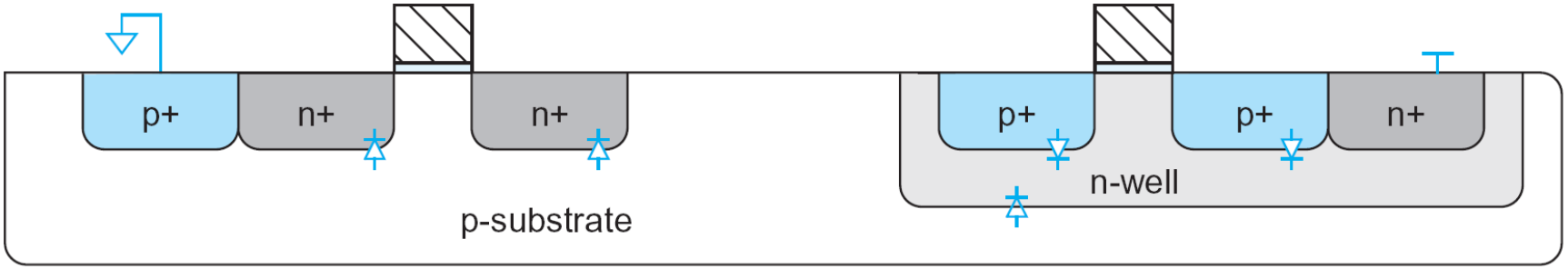
- When the gate oxide is very thin, quantum mechanics predicts that current can flow across the gate oxide by electron tunneling, causing gate leakage. Direct tunneling (low V) and FN tunneling (high V)
- For gate oxide thinner than 15-20 angstrom, tunneling current may become comparable to subthreshold leakage.
- To achieve high C_{ox} without using very thin insulator, new material with higher dielectric constant (k), e.g. Si_3N_4, has been used.
- Electrically programmable circuits, e.g., EEPROM and flash memory employ this tunneling to transport charges to and from otherwise isolated regions.

Junction Leakage

- MOS transistors are protected by depletion region formed through reversed biased diodes with leakage current.

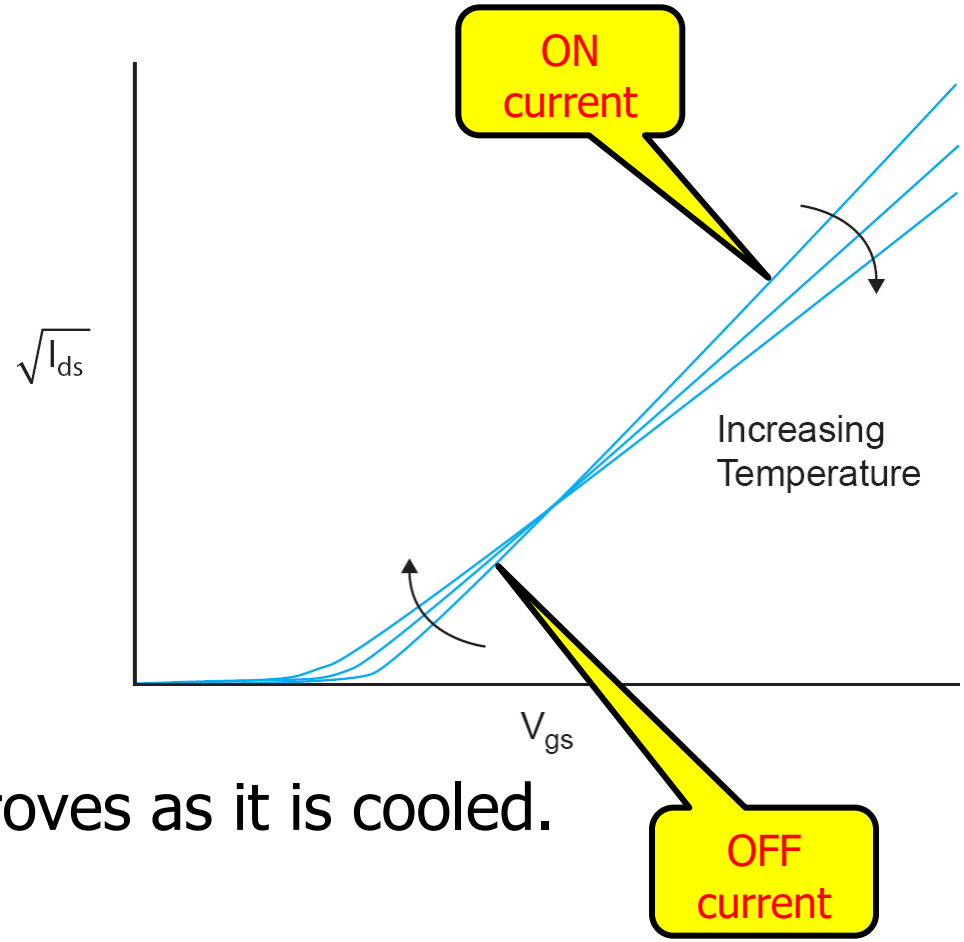
$$I_D = I_S\left(e^{\frac{V_D}{v_T}} - 1\right)$$

- This leakage has been limiting factor for node voltage storage and is now overwhelmed by subthreshold conduction due to lower threshold voltage.



Temperature Dependence

- Many MOS characteristics are influenced by temperature.
- As T ↑
  - Mobility ↓
  - Threshold voltage ↓
  - Junction leakage ↑
  - Subthreshold leakage ↑
  - OFF current ↑
  - ON current ↓



- Circuit performance improves as it is cooled.

Enhanced Features in New Process

- Multiple thresholds
  - Low V_t transistors on critical circuits; high V_t transistors elsewhere.
- Thicker oxide – for I/O transistors; 3.3V for I/O, 1.8V for core.
- High-k gate dielectric – HfO_2 (k=20), ZrO_2 (k=23) Si_3N_4 (k=6.5-7.5)
  - increase the ε (k) of the gate insulator to get more current
- FinFET (3D transistor)
- Fully Depleted Silicon on Insulator (FD-SOI)
  - https://www.youtube.com/watch?v=uvV7jcpQ7UY