ELEC 401: Analog CMOS Integrated Circuit Design

Background

Shahriar Mirabbasi

Department of Electrical and Computer Engineering
University of British Columbia
shahriar@ece.ubc.ca

Technical contributions of Pedram Lajevardi in revising the slides is greatly acknowledged.

Background

1. Suggested Reading
2. Structure of MOS Transistors
3. Threshold Voltage
4. Long-Channel Current Equations
5. Regions of Operation
6. Transconductance
7. Second-Order Effects
8. Short-Channel Effects
9. MOS Layout
10. Device Capacitances
11. Small-signal Models
12. Circuit Impedance
13. Equivalent Transconductance

Suggested Reading

- Most of the material in this set are based on

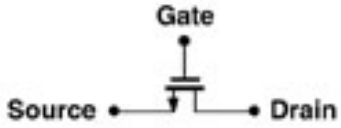
  Chapters 2, 16, and 17 of the Razavi's book: Design of Analog CMOS Integrated Circuits

Transistor

- Transistor stands for …
- Transistor are semiconductor devices that can be classified as
  - Bipolar Junction Transistors (BJTs)
  - Field Effect Transistors (FETs)
    - Depletion-Mode FETs or (e.g., JFETs)
    - Enhancement-Mode FETs (e.g., MOSFETs)

Simplistic Model

- MOS transistors have three terminals: Gate, Source, and Drain

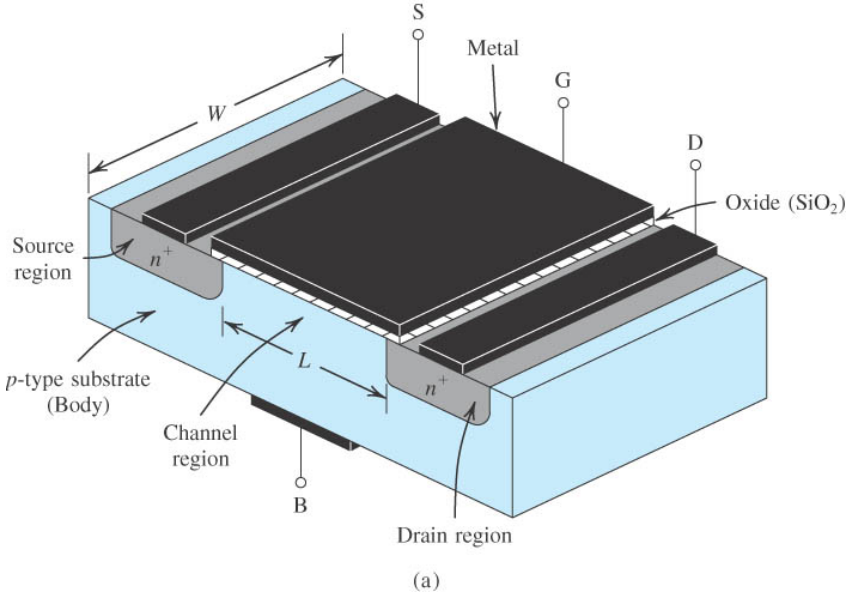


- The voltage of the Gate terminal determines the type of connection between Source and Drain (Short or Open).
- Thus, MOS devices behave like a switch

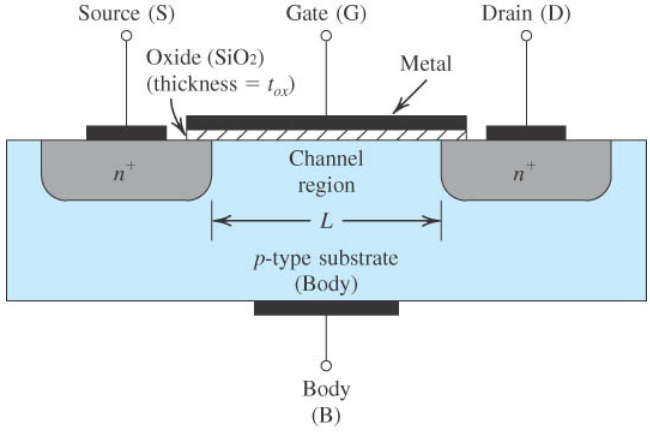
	NMOS	PMOS
V_G high	Device is ON D is shorted to S	Device is OFF D & S are disconnected
V_G low	Device is OFF D & S are disconnected	Device is ON D is shorted to S

Physical Structure - 1

- Source and Drain terminals are identical except that Source provides charge carriers, and Drain receives them.
- MOS devices have in fact 4 terminals:
  - Source, Drain, Gate, Substrate (bulk)



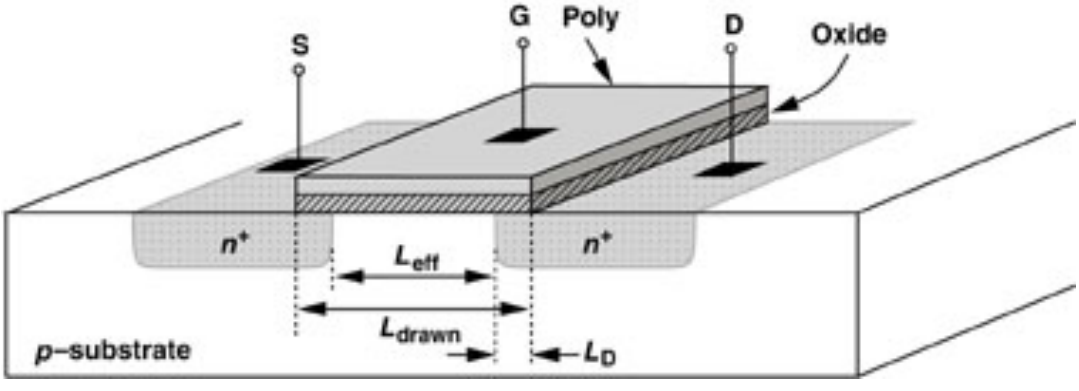
(a)



(b)

Physical Structure - 2

- Charge Carriers are electrons in NMOS devices, and holes in PMOS devices.
- Electrons have a higher mobility than holes
- So, NMOS devices are faster than PMOS devices
- We rather to have a p-type substrate?!



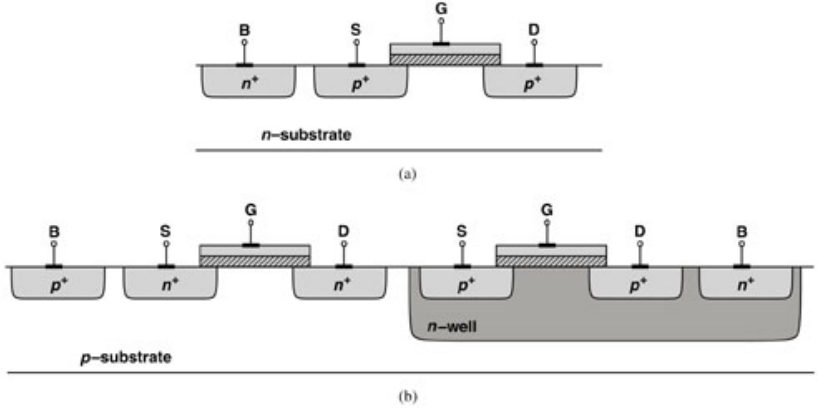
L_D: Due to Side Diffusion

Poly-silicon used instead of Metal for fabrication reasons

- Actual length of the channel (L_{eff}) is less than the length of gate

Physical Structure - 3

- N-wells allow both NMOS and PMOS devices to reside on the same piece of die.



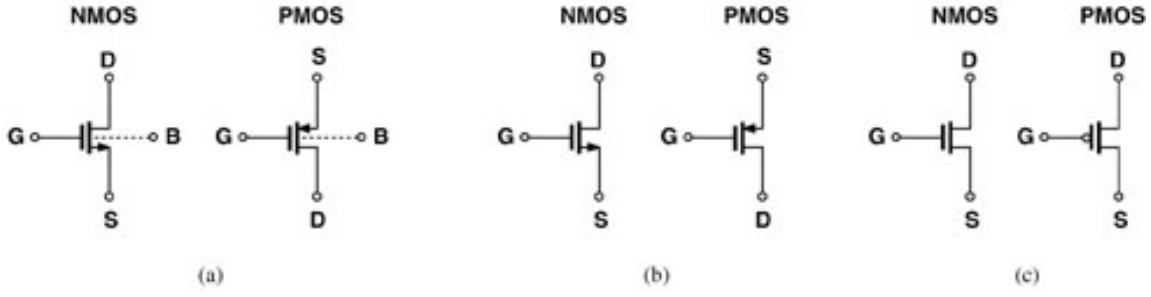
- As mentioned, NMOS and PMOS devices have 4 terminals:

  Source, Drain, Gate, Substrate (bulk)

- In order to have all PN junctions reverse-biased, substrate of NMOS is connected to the most negative voltage, and substrate of PMOS is connected to the most positive voltage.

Physical Structure - 4

- MOS transistor Symbols:



- In NMOS Devices: $Source \xrightarrow{electron} Drain$

  Current flows from Drain to Source

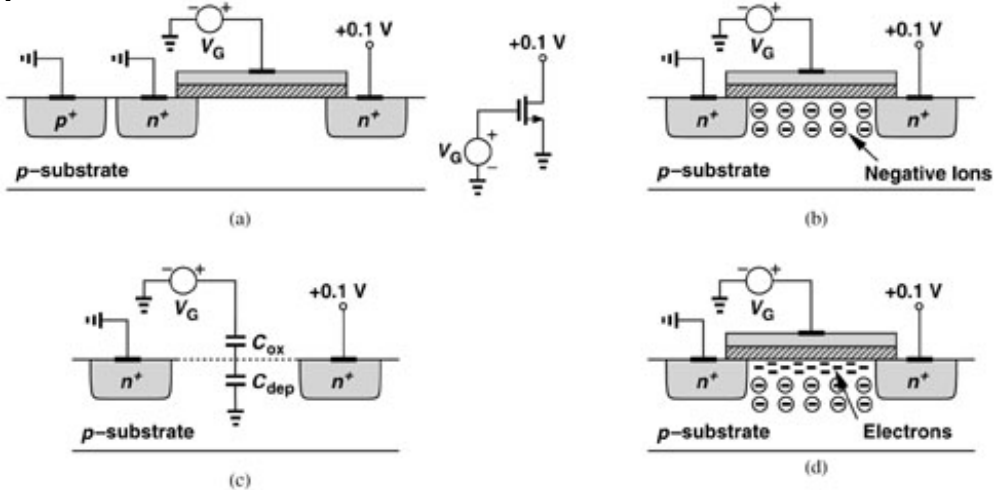
- In PMOS Devices: $Source \xrightarrow{hole} Drain$

  Current flows from Source to Drain

- Current flow determines which terminal is Source and which one is Drain. Equivalently, source and drain can be determined based on their relative voltages.

Threshold Voltage - 1

- Consider an NMOS: as the gate voltage is increased, the surface under the gate is depleted. If the gate voltage increases more, free electrons appear under the gate and a conductive channel is formed.



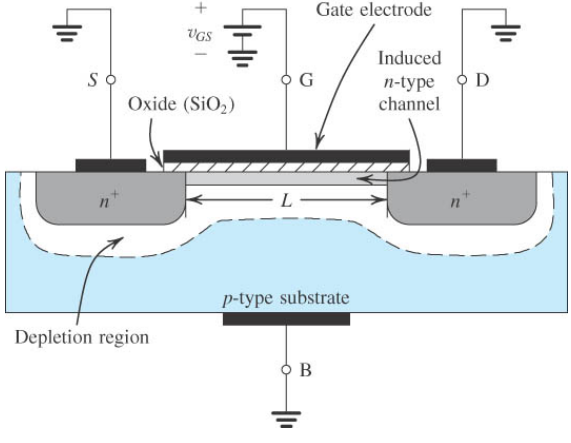
(a) An NMOS driven by a gate voltage, (b) formation of depletion region, (c) onset of inversion, and (d) channel formation

- As mentioned before, in NMOS devices charge carriers in the channel under the gate are electrons.

Threshold Voltage - 2

- Intuitively, the threshold voltage is the gate voltage that forces the interface (surface under the gate) to be completely depleted of charge (in NMOS the interface is as much n-type as the substrate is p-type)

- Increasing gate voltage above this threshold (denoted by V_{TH} or V_t) induces an inversion layer (conductive channel) under the gate.



Threshold Voltage - 3

Analytically:

$$V_{TH} = \Phi_{MS} + 2 \cdot \left|\Phi_F\right| + \frac{\left|Q_{dep}\right|}{C_{ox}}$$

Where:

Φ_{MS} = Built-in Potential = $\Phi_{gate} - \Phi_{Silicon}$

= the difference between the work functions of the polysilicon gate and the silicon substrate

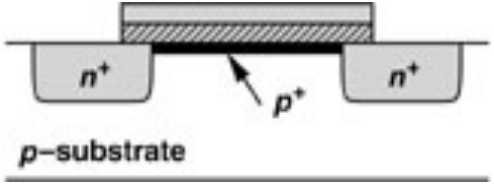
Φ_F = Work Function (electrostatic potential) = $\frac{K \cdot T}{q} \cdot \ln\left(\frac{N_{sub}}{n_i}\right)$

Q_{dep} = Charge in the depletion region = $\sqrt{4 \cdot q \cdot \varepsilon_{si} \cdot \left|\Phi_F\right| \cdot N_{sub}}$

Threshold Voltage - 4

- In practice, the “native” threshold value may not be suited for circuit design, e.g., V_{TH} may be zero and the device may be on for any positive gate voltage.

- Typically threshold voltage is adjusted by ion implantation into the channel surface (doping P-type material will increase V_{TH} of NMOS devices).

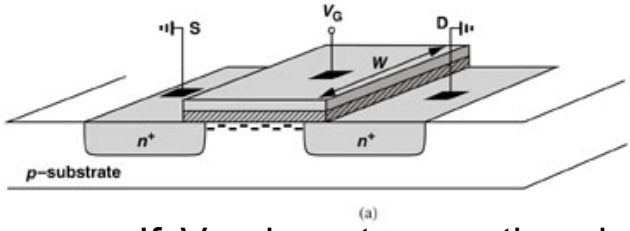


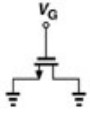
- When V_{DS} is zero, there is no horizontal electric field present in the channel, and therefore no current between the source to the drain.

- When V_{DS} is more than zero, there is some horizontal electric field which causes a flow of electrons from source to drain.

Long Channel Current Equations - 1

- The voltage of the surface under the gate, V(x), depends on the voltages of Source and Drain.
- If V_{DS} is zero, $V_D = V_S = V(x)$. The charge density Q_d (unit C/m) is uniform.

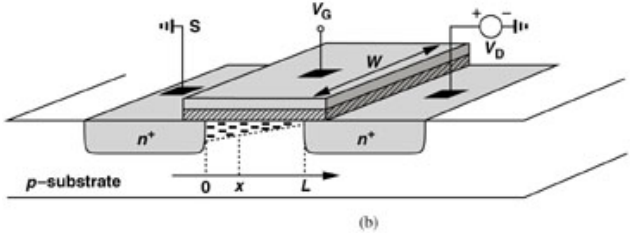


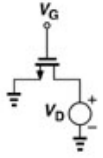


$$Q_d = \frac{-Q}{L} = \frac{-C \cdot V}{L} = \frac{-(C_{ox} WL) \cdot (V_{GS} - V_{TH})}{L}$$

$$Q_d = -WC_{ox}(V_{GS} - V_{TH})$$

- If V_{DS} is not zero, the channel is tapered, and V(x) is not constant. The charge density depends on x.





$$Q_d(x) = -WC_{ox}(V_{GS} - V(x) - V_{TH})$$

Long Channel Current Equations - 3

- Current :

$$I = \frac{dQ}{dt} = \frac{dQ}{dx} \times \frac{dx}{dt} = Q_d \cdot velocity$$

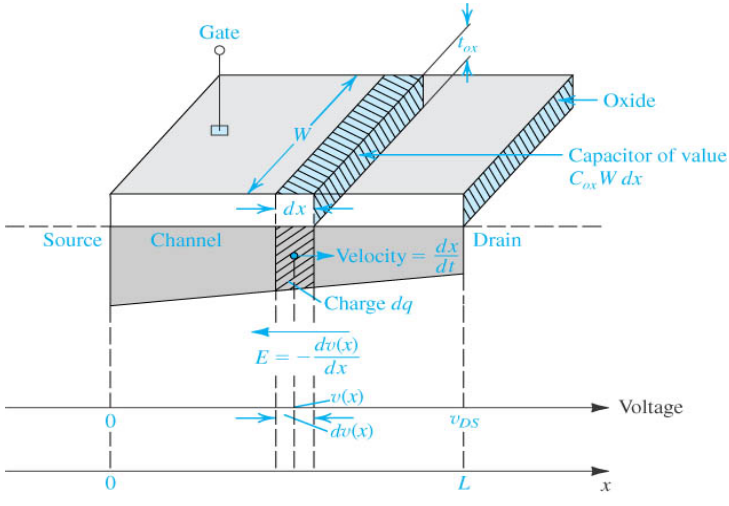
  - Velocity in terms of V(x):

$$velocity = \mu \cdot E \quad , \quad E = -\frac{dV}{dx}$$

$$\rightarrow velocity = (\mu \cdot \frac{-dV(x)}{dx})$$

  - Q_d in terms of V(x):

$$Q_d(x) = -WC_{ox}(V_{GS} - V(x) - V_{TH})$$



- Current in terms of V(x):

$$I_D = WC_{ox}[V_{GS} - V(x) - V_{TH}]\mu_n \frac{dV(x)}{dx}$$

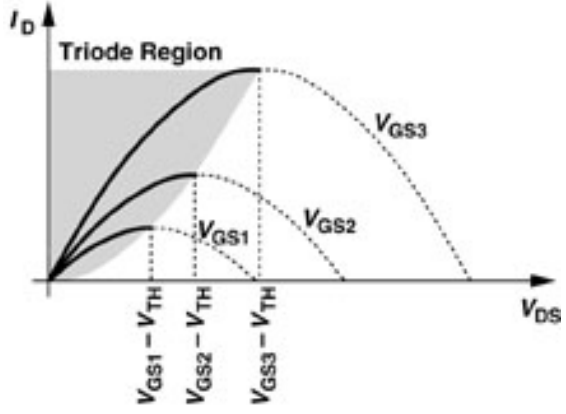
$$\int_{x=0}^{L} I_D dx = \int_{V=0}^{V_{DS}} WC_{ox}\mu_n[V_{GS} - V(x) - V_{TH}]dV$$

- Long-channel current equation:

$$I_D = \mu_n C_{ox} \frac{W}{L}[(V_{GS} - V_{TH})V_{DS} - \frac{1}{2}V_{DS}^2]$$

Long Channel Current Equations - 4

- If $V_{DS} \leq V_{GS}-V_{TH}$ we say the device is operating in triode (or linear) region.
- Current in Triode Region: $$I_D = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left[(V_{GS} - V_{TH}) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right]$$



- Terminology:

$$Aspect\ Ratio = \frac{W}{L}$$

$$Overdrive\quad Voltage = Effective\quad Voltage = V_{GS} - V_{TH} = V_{eff}$$

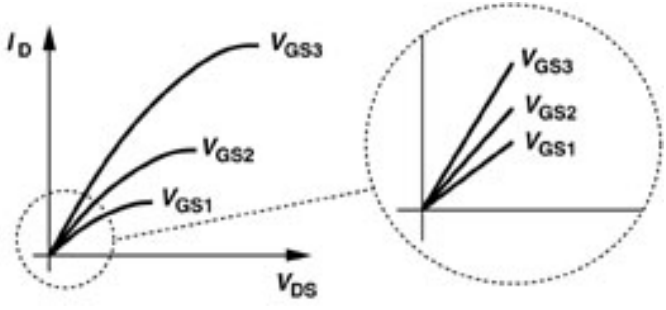
Long Channel Current Equations - 5

- For very small V_{DS} (deep Triode Region):
  - I_D can be approximated to be a linear function of V_{DS}.
  - The device resistance will be independent of V_{DS} and will only depend on V_{eff}.
  - The device will behave like a variable resistor

$$If\ V_{DS} << 2(V_{GS} - V_{TH}):$$

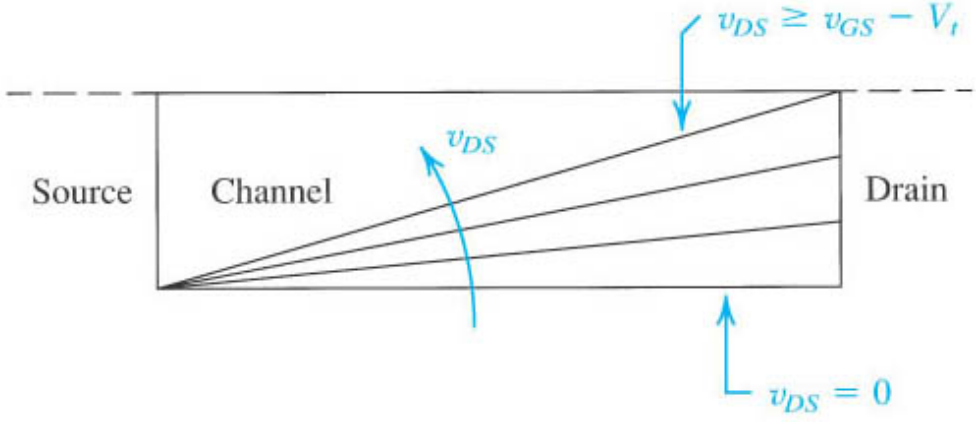
$$I_D = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) \cdot V_{DS}$$

$$R_{ON} = \frac{V_{DS}}{I_D} = \frac{1}{\mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})}$$



Long Channel Current Equations - 6

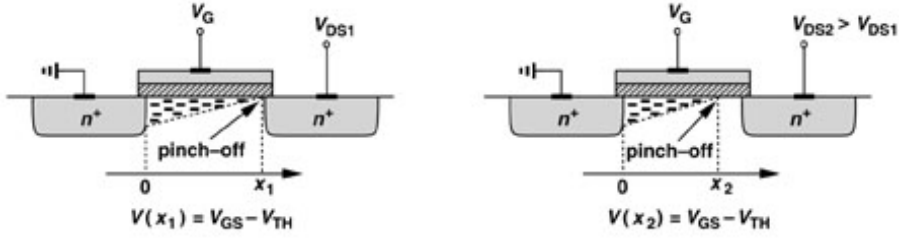
- Increasing V_{DS} causes the channel to acquire a tapered shape. Eventually, as V_{DS} reaches $V_{GS} - V_{TH}$ the channel is pinched off at the drain. Increasing V_{DS} above $V_{GS} - V_{TH}$ has little effect (ideally, no effect) on the channel's shape.



- When V_{DS} is more than $V_{GS} - V_{TH}$ the channel is pinched off, and the horizontal electric field produces a current.

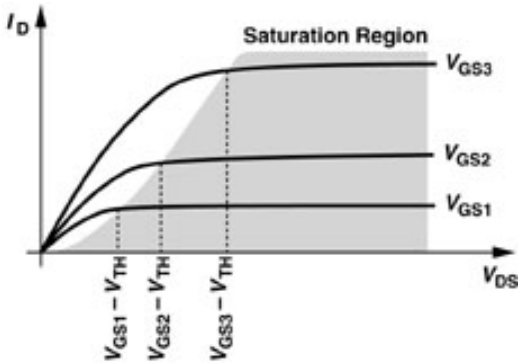
Long Channel Current Equations - 7

- If $V_{DS} > V_{GS} - V_{TH}$, the transistor is in saturation (active) region, and the channel is pinched off.



$$\int_{x=0}^{L'} I_D dx = \int_{V=0}^{V_{GS}-V_{TH}} WC_{ox}\mu_n [V_{GS} - V(x) - V_{TH}]dV$$

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})^2$$



- Let's, for now, assume that L'=L. The fact that L' is not equal to L is a second-order effect known as channel-length modulation.
- Since I_D only depends on V_{GS}, MOS transistors in saturation can be used as current sources.

Long Channel Current Equations - 8

- Current Equation for NMOS:

$$I_D = I_{DS} = \begin{cases} 0 \quad ; if\ V_{GS} < V_{TH}\ (Cut-off) \\ \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left(V_{GS} - V_{TH}\right) \cdot V_{DS} \quad ; if\ V_{GS} > V_{TH}, V_{DS} << 2(V_{GS} - V_{TH})\ (Deep\ Triode) \\ \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left[\left(V_{GS} - V_{TH}\right) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2\right] \quad ; if\ V_{GS} > V_{TH}, V_{DS} < V_{GS} - V_{TH}\ (Triode) \\ \frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})^2 \quad ; if\ V_{GS} > V_{TH}, V_{DS} > V_{GS} - V_{TH}\ (Saturation) \end{cases}$$

Long Channel Current Equations - 9

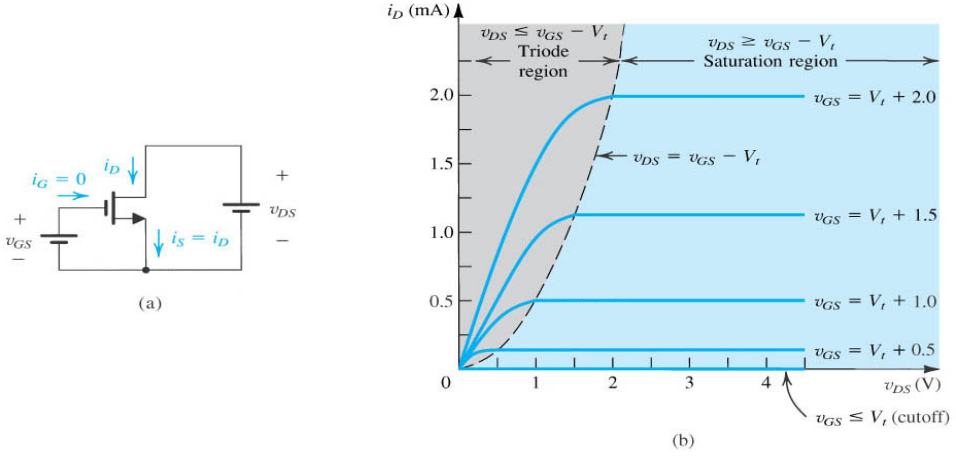
- Current Equation for PMOS:

$$I_D = I_{SD} = \begin{cases} 0 \quad ; if\ V_{SG} < |V_{TH}|\ (Cut-off) \\ \mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot \left(V_{SG} - |V_{TH}|\right) \cdot V_{SD} \quad ; if\ V_{SG} > |V_{TH}|, V_{SD} << 2(V_{SG} - |V_{TH}|)\ (Deep\ Triode) \\ \mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot \left[\left(V_{SG} - |V_{TH}|\right) \cdot V_{SD} - \frac{1}{2} \cdot V_{SD}^2\right] \quad ; if\ V_{SG} > |V_{TH}|, V_{SD} < V_{SG} - |V_{TH}|\ (Triode) \\ \frac{1}{2} \cdot \mu_p \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{SG} - |V_{TH}|)^2 \quad ; if\ V_{SG} > |V_{TH}|, V_{SD} > V_{SG} - |V_{TH}|\ (Saturation) \end{cases}$$

Regions of Operation - 1

- Regions of Operation:

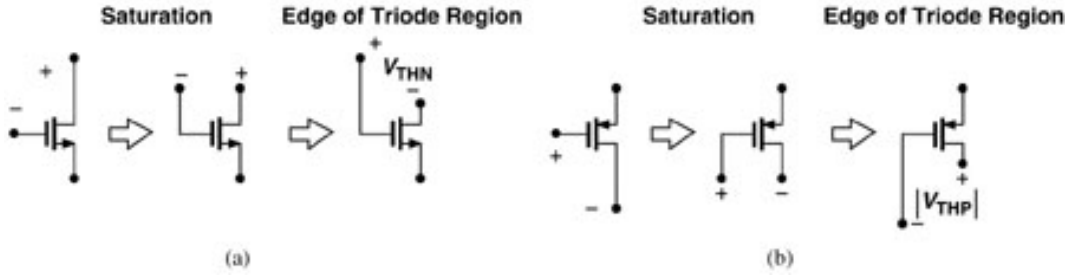
  Cut-off, triode (linear), and saturation (active or pinch-off)



- Once the channel is pinched off, the current through the channel is almost constant. As a result, the I-V curves have a very small slope in the pinch-off (saturation) region, indicating the large channel resistance.

Regions of Operation - 2

- The following illustrates the transition from pinch-off to triode region for NMOS and PMOS devices.



- For NMOS devices:

  If V_D increases (V_G Const.), the device will go from Triode to Pinch-off.

  If V_G increases (V_D Const.), the device will go from Pinch-off to Triode.

** In NMOS, as V_{DG} increases the device will go from Triode to Pinch-off.

- For PMOS devices:

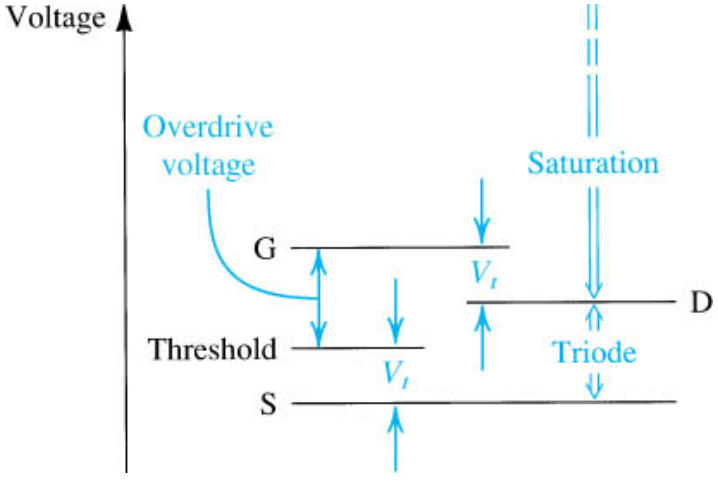
  If V_D decreases (V_G Const.), the device will go from Triode to Pinch-off.

  If V_G decreases (V_D Const.), the device will go from Pinch-off to Triode.

** In PMOS, as V_{GD} increases the device will go from Pinch-off to Triode.

Regions of Operation - 3

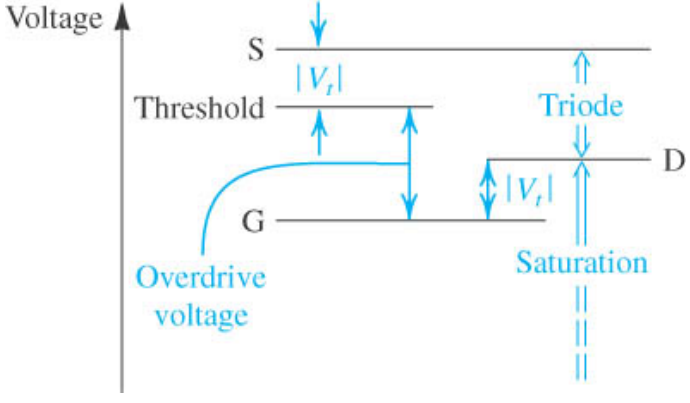
- NMOS Regions of Operation:



- Relative levels of the terminal voltages of the enhancement-type NMOS transistor for different regions of operation.

Regions of Operation - 4

- PMOS Regions of Operation:

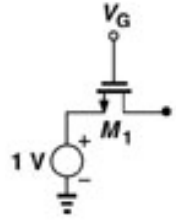


- The relative levels of the terminal voltages of the enhancement-type PMOS transistor for different regions of operation.

Regions of Operation - 5

Example:

For the following circuit assume that V_{TH}=0.7V.



(a)

- When is the device on?

- What is the region of operation if the device is on?

- Sketch the on-resistance of transistor M_1 as a function of V_G.

Transconductance - 1

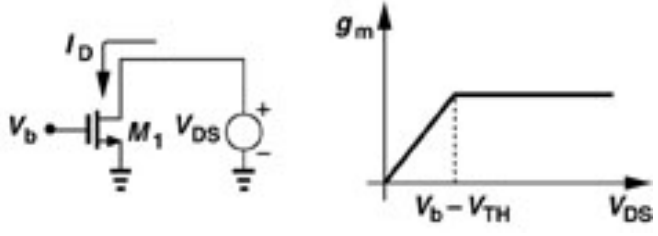
- The drain current of the MOSFET in saturation region is ideally a function of gate-overdrive voltage (effective voltage). In reality, it is also a function of V_{DS}.
- It makes sense to define a figure of merit that indicates how well the device converts the voltage to current.
- Which current are we talking about?
- What voltage is in the designer's control?
- What is this figure of merit?

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS}=Const.}$$

Transconductance - 2

Example:

Plot the transconductance of the following circuit as a function of V_{DS} (assume V_b is a constant voltage).



- Transconductance in triode:

$$g_m = \frac{\partial}{\partial V_{GS}}\left(\mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left[\left(V_{GS} - V_{TH}\right) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2\right]\right)\Bigg|_{V_{DS} = Const.}$$

$$= \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot V_{DS}$$

- Transconductance in saturation:

$$g_m = \frac{\partial}{\partial V_{GS}}\left(\frac{1}{2} \cdot \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH})^2\right)\Bigg|_{V_{DS} = Const.}$$

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

- Moral: Transconductance drops if the device enters the triode region.

Transconductance - 3

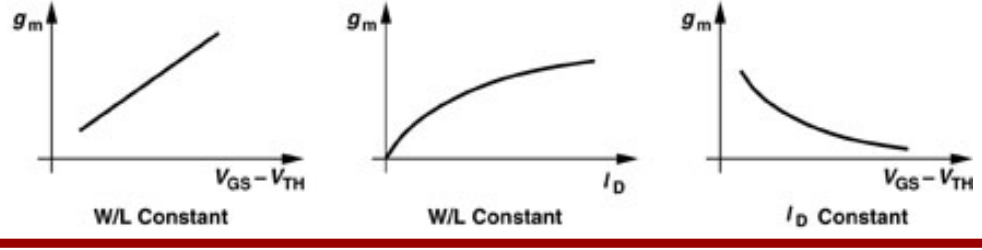
- Transconductance, g_m, in saturation:

$$g_m = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) = \sqrt{2\mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot I_D} = \frac{2 \cdot I_D}{V_{GS} - V_{TH}}$$

- If the aspect ratio is constant: g_m depends linearly on (V_{GS} - V_{TH}).
  Or g_m depends on square root of I_D.

- If I_D is constant: g_m is inversely proportional to (V_{GS} - V_{TH}).
  Or g_m depends on square root of the aspect ratio.

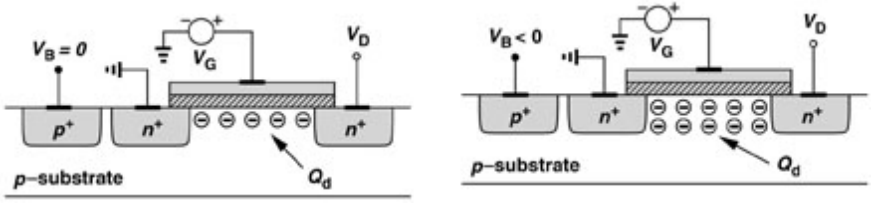
- If the overdrive voltage is constant: g_m depends linearly on I_D.
  Or g_m depends linearly on the aspect ratio.



Second-Order Effects (Body Effect)

Substrate Voltage:

- So far, we assumed that the bulk and source of the transistor are at the same voltage (V_B=V_S).
- If $V_B > V_s$, then the bulk-source PN junction will be forward biased, and the device will not operate properly.
- If $V_B < V_s$,
  - the bulk-source PN junction will be reverse biased.
  - the depletion region widens, and Q_{dep} increases.
  - V_{TH} will be increased (Body effect or Backgate effect).



- It can be shown that (what is the unit for γ ?):

$$V_{TH} = V_{TH0} + \gamma \cdot \left(\sqrt{\left|2 \cdot \Phi_F + V_{SB}\right|} - \sqrt{\left|2 \cdot \Phi_F\right|} \right) \quad \text{where } \gamma = \frac{\sqrt{2 \cdot q \cdot \varepsilon_{si} \cdot N_{sub}}}{C_{ox}}$$

Body Effect - 2

Example:

Consider the circuit below (assume the transistor is in the active region):

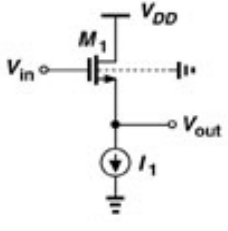
- If body-effect is ignored, V_{TH} will be constant, and I_1 will only depend on $V_{GS1}=V_{in}-V_{out}$. Since I_1 is constant, $V_{in}-V_{out}$ remains constant.

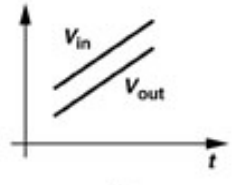
$$V_{in} - V_{out} - V_{TH} = C = Const. \rightarrow V_{in} - V_{out} = V_{TH} + C = D = Conts.$$

- In general, I_1 depends on $V_{GS1}- V_{TH} =V_{in}-V_{out}-V_{TH}$ (and with body effect V_{TH} is not constant). Since I_1 is constant, $V_{in}-V_{out}-V_{TH}$ remains constant:

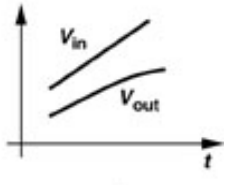
$$V_{in} - V_{out} - V_{TH} = C = Const. \rightarrow V_{in} - V_{out} = V_{TH} + C$$

- As V_{out} increases, V_{SB1} increases, and as a result V_{TH} increases. Therefore, $V_{in}-V_{out}$ Increases.





No Body Effect

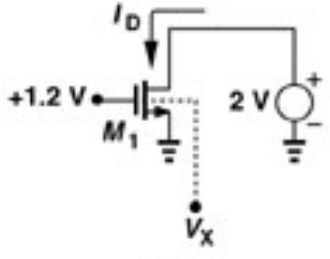


With Body Effect

Body Effect - 3

Example:

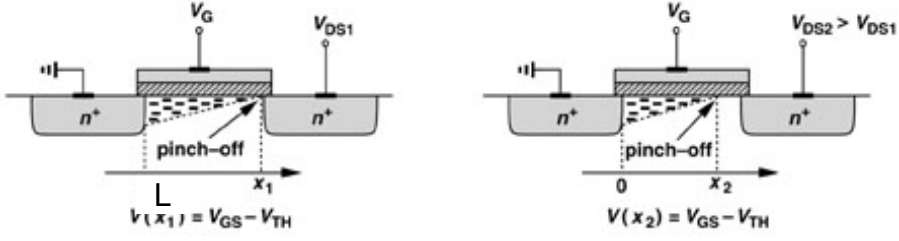
For the following Circuit sketch the drain current of transistor M_1 when V_X varies from $-\infty$ to 0. Assume V_{TH0}=0.6V, γ=0.4V$^{1/2}$, and $2\Phi_F$=0.7V.



(a)

Channel Length Modulation - 1

- When a transistor is in the saturation region ($V_{DS} > V_{GS} - V_{TH}$), the channel is pinched off.



- The drain current is $I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})^2$ where $L' = L - \Delta L$

$$\frac{1}{L'} = \frac{1}{L - \Delta L} = \frac{1}{L}\cdot\frac{1}{1 - \Delta L/L} \approx \frac{1}{L}\cdot\left(1 + \Delta L/L\right)$$

$L' = L\left(1 - \frac{\Delta L}{L}\right)$

- Assuming $\Delta L/L = \lambda \cdot V_{DS}$ we get: $\frac{1}{L'} \approx \frac{1}{L}\cdot\left(1 + \Delta L/L\right) = \frac{1}{L}\cdot\left(1 + \lambda \cdot V_{DS}\right)$

- The drain current is $I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})^2 \approx \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})^2 \cdot (1 + \lambda \cdot V_{DS})$

Want to make small

- As I_D actually depends on both V_{GS} and V_{DS}, MOS transistors are not ideal current sources (why?).

if trying to create a current source

Channel Length Modulation - 2

- λ represents the relative variation in effective length of the channel for a given increment in V_{DS}.
- For longer channels λ is smaller, i.e., $\lambda \propto 1/L$
- Transconductance: $$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS}=Const.}$$

  In Triode:
  $$g_m = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot V_{DS}$$

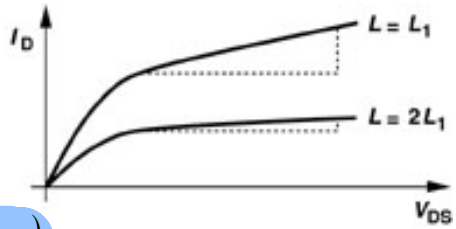
  In Saturation (ignoring channel length modulation):
  $$g_m = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) = \sqrt{2\mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot I_D} = \frac{2 \cdot I_D}{V_{GS} - V_{TH}}$$

  In saturation with channel length modulation:
  $$g_m = \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot (V_{GS} - V_{TH}) \cdot (1 + \lambda \cdot V_{DS}) = \sqrt{2\mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot I_D \cdot (1 + \lambda \cdot V_{DS})} = \frac{2 \cdot I_D}{V_{GS} - V_{TH}}$$
- The dependence of I_D on V_{DS} is much weaker than its dependence on V_{GS}.

Channel Length Modulation - 3

Example:

Given all other parameters constant, plot I_D-V_{DS} characteristic of an NMOS for $L=L_1$ and $L=2L_1$



- In Triode Region: $I_D \approx \mu_n \cdot C_{ox} \cdot \frac{W}{L} \cdot \left[\left(V_{GS} - V_{TH} \right) \cdot V_{DS} - \frac{1}{2} \cdot V_{DS}^2 \right]$

  Therefore : $\frac{\partial I_D}{\partial V_{DS}} \propto \frac{W}{L}$

- In Saturation Region: $I_D \approx \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 \cdot (1 + \lambda \cdot V_{DS})$

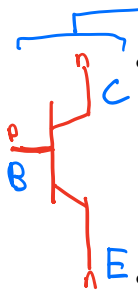
  So we get : $\frac{\partial I_D}{\partial V_{DS}} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 \cdot \lambda$

  Therefore : $\frac{\partial I_D}{\partial V_{DS}} \propto \frac{W \cdot \lambda}{L} \propto \frac{W}{L^2}$

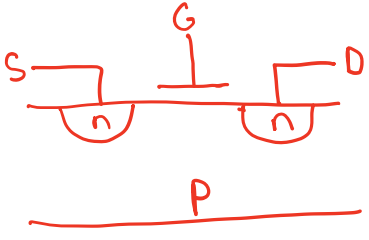
- Changing the length of the device from L_1 to $2L_1$ will flatten the I_D-V_{DS} curves (slope will be divided by two in triode and by four in saturation).
- Increasing L will make a transistor a better current source, while degrading its current capability.
- Increasing W will improve the current capability.

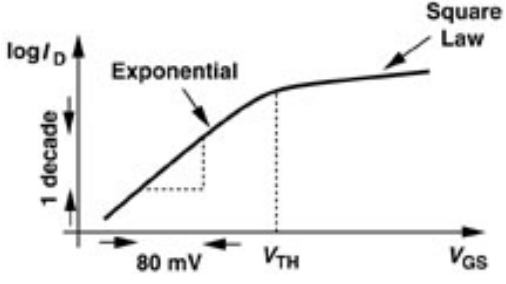
Sub-threshold Conduction

- If $V_{GS} < V_{TH}$, the drain current is not zero.
- The MOS transistors behave similar to BJTs.



- In BJT: $I_C = I_S \cdot e^{\frac{V_{BE}}{V_T}}$





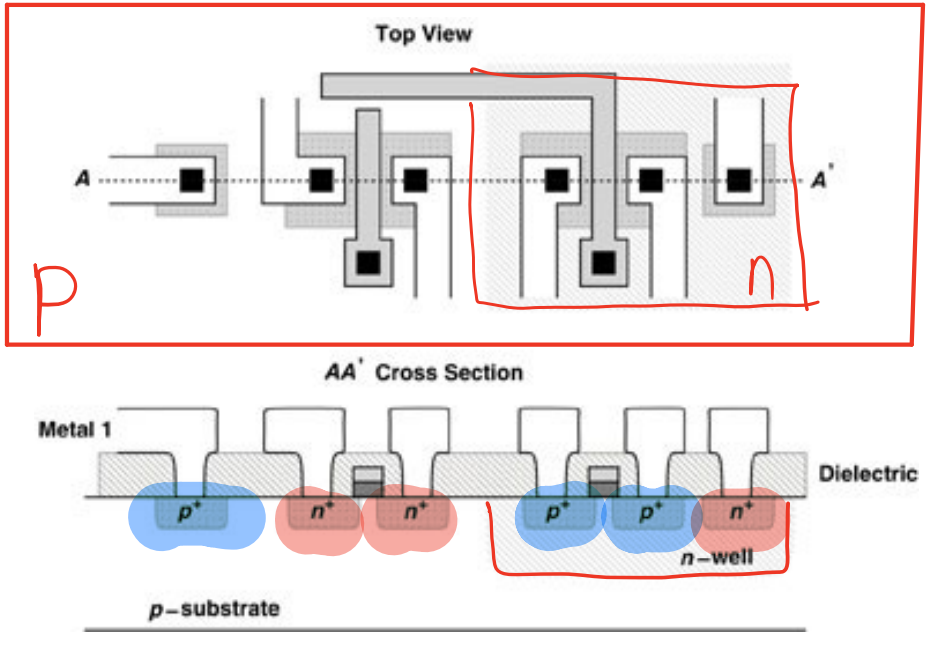
- In MOS: $I_D = I_0 \cdot e^{\frac{V_{GS}}{\zeta \cdot V_T}}$

↳ Not a physical connection like BJT. Capacitor connected

- As shown in the figure, in MOS transistors, the drain current drops by one decade for approximately each 80mV of drop in V_{GS}.
- In BJT devices the current drops faster (one decade for approximately each 60mv of drop in V_{GS}).
- This current is known as sub-threshold or weak-inversion conduction.

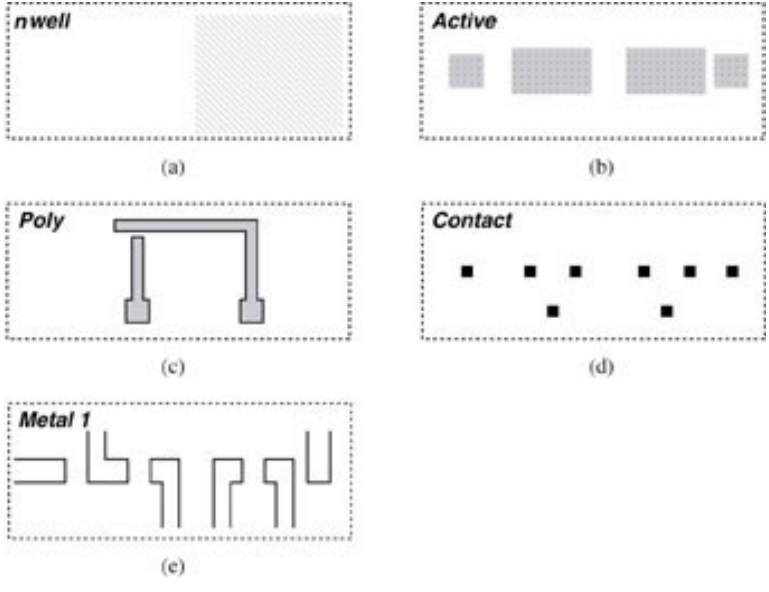
CMOS Processing Technology

- Top and side views of a typical CMOS process



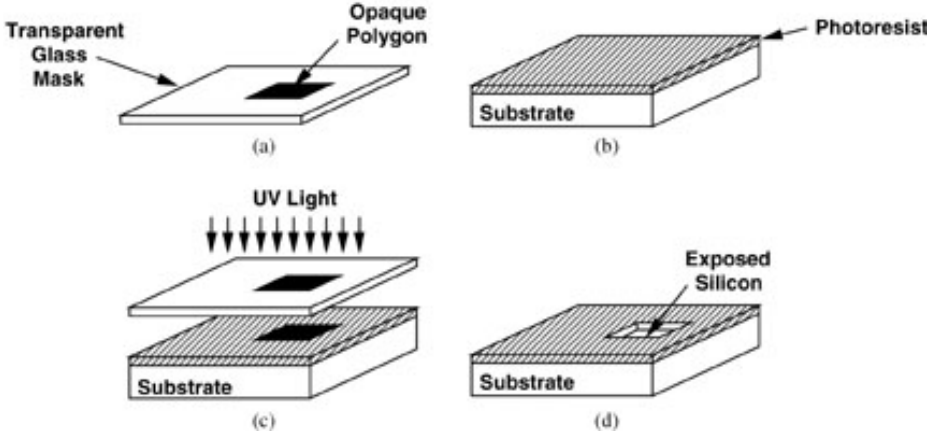
CMOS Processing Technology

- Different layers comprising CMOS transistors

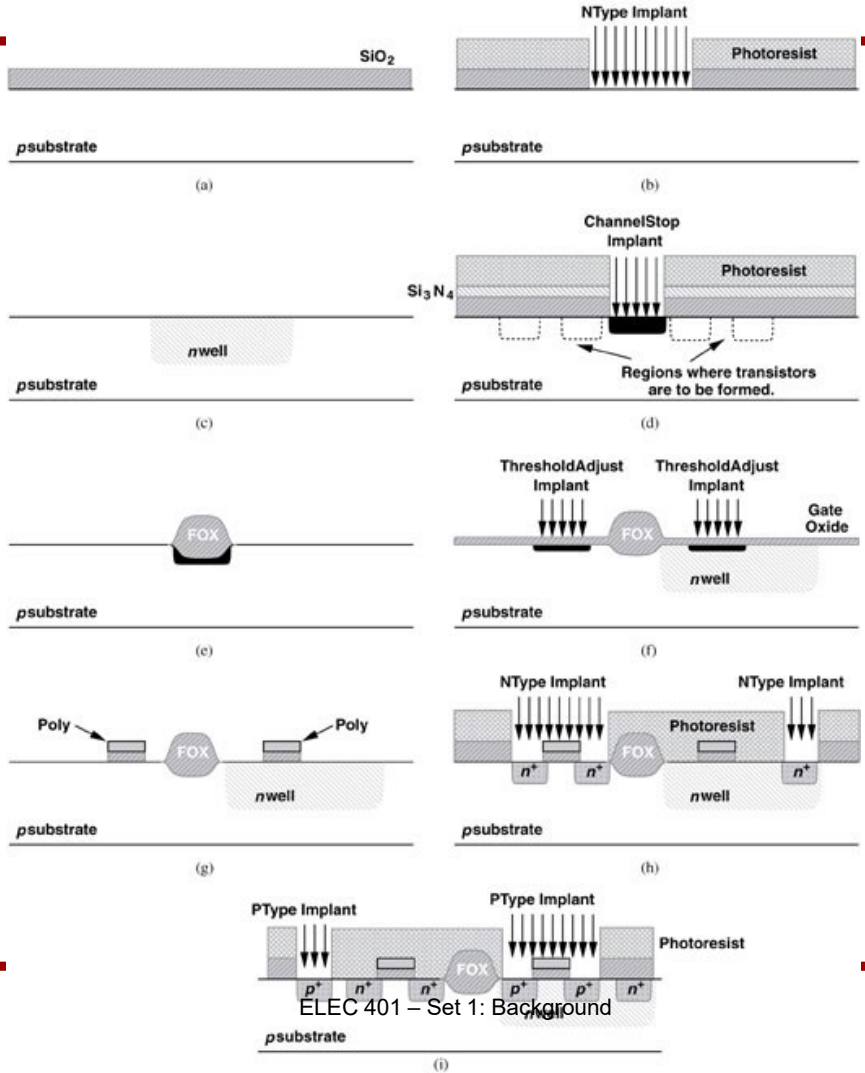


Photolithography (Lithography)

- Used to transfer circuit layout information to the wafer

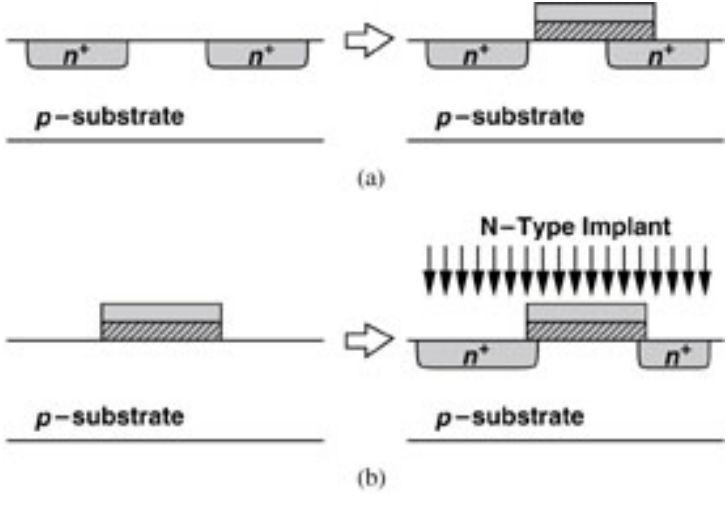


Typical Fabrication Sequence



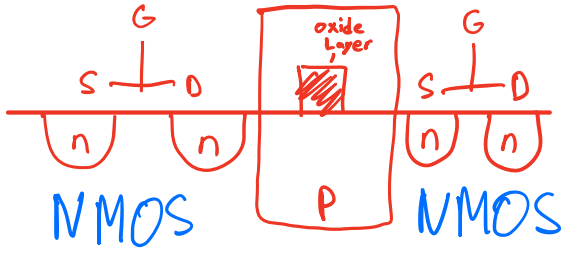
Self-Aligned Process

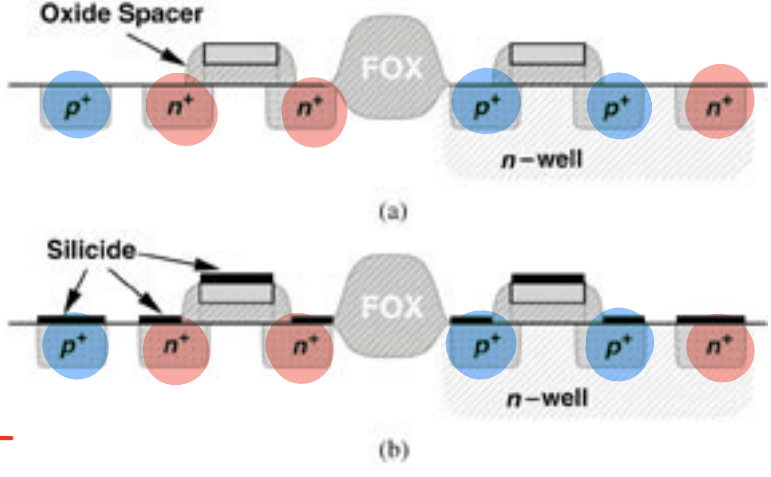
- Why source and drain junctions are formed after the gate oxide and polysilicon layers are deposited?
  - Why is Gate not metal?



Back-End Processing

- Oxide spacers and silicide

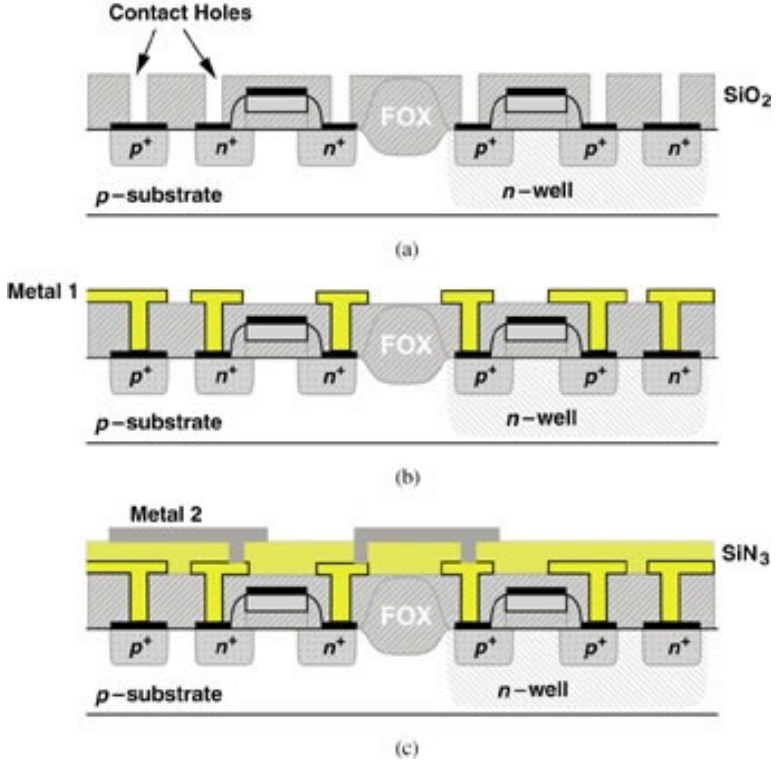




Shallow Trench Isolation

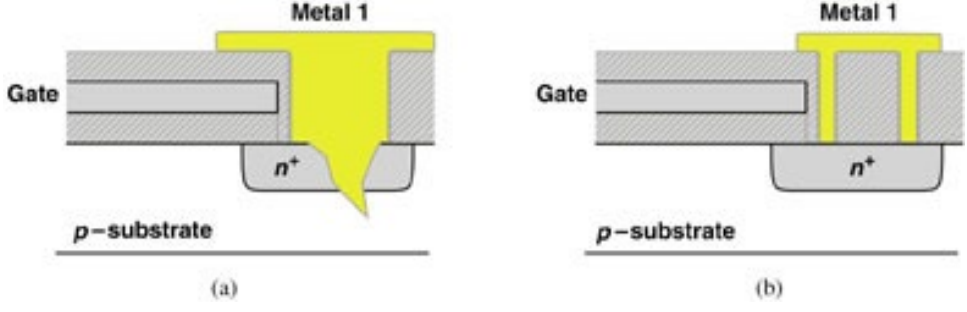
Back-End Processing

- Contact and metal layers fabrication



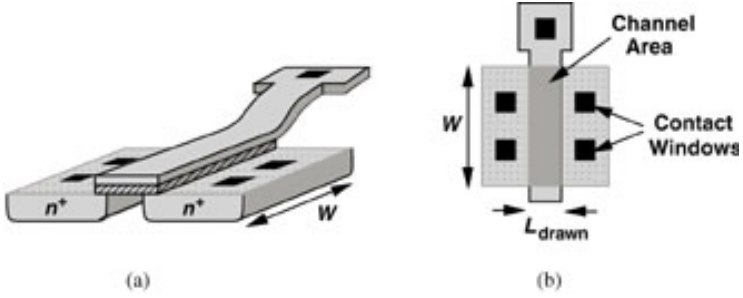
Back-End Processing

- Large contact areas should be avoided to minimize the possibility of spiking



MOS Layout - 1

- It is beneficial to have some insight into the layout of the MOS devices.

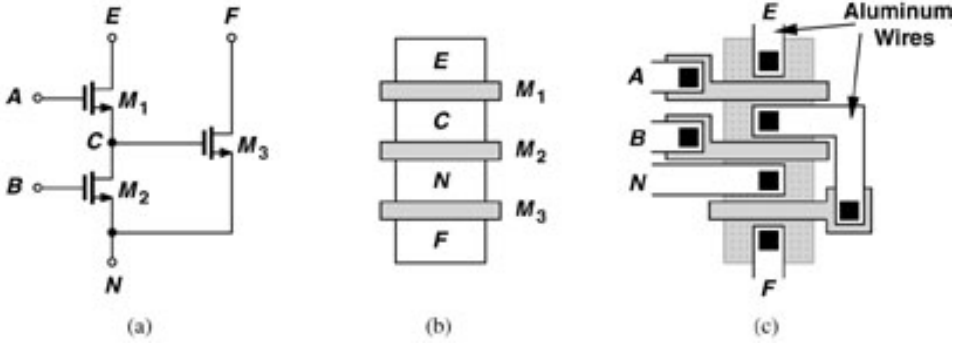


- When laying out a design, there are many important parameters we need to pay attention to such as: drain and source areas, interconnects, and their connections to the silicon through contact windows.
- Design rules determine the criteria that a circuit layout must meet for a given technology. Things like, minimum length of transistors, minimum area of contact windows, …

MOS Layout - 2

Example:

Figures below show a circuit with a suggested layout.

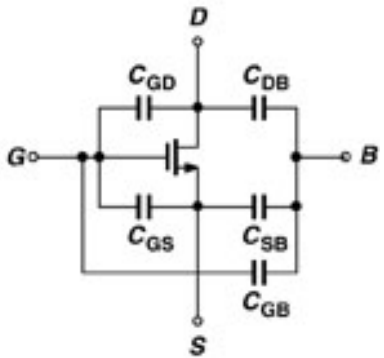


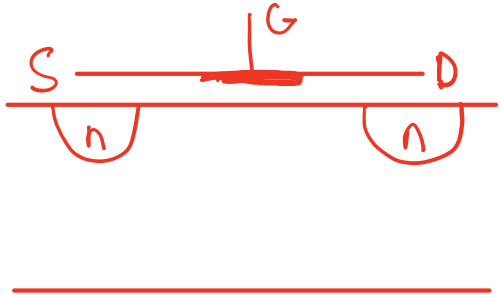
- The same circuit can be laid out in different ways, producing different electrical parameters (such as different terminal capacitances).

Device Capacitances - 1

- The quadratic model determines the DC behavior of a MOS transistor.
- The capacitances associated with the device are important when studying the AC behavior of that device.
- There is a capacitance between any two terminals of a MOS transistor. So there are 6 capacitances in total.
- The Capacitance between Drain and Source is usually negligible (C_{DS}=0).

$$\binom{4}{2} = \frac{4!}{2!(4-2)!} = 6$$





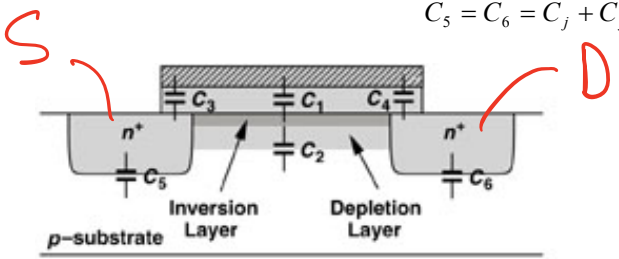
- These capacitances will depend on the region of operation (Bias values).

Device Capacitances - 2

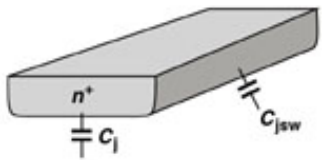
- The following will be used to calculate the capacitances between terminals:
  1. Oxide Capacitance: $C_1 = W \cdot L \cdot C_{ox}$, $C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}}$
  2. Depletion Capacitance: $C_2 = C_{dep} = W \cdot L \cdot \sqrt{\frac{q \cdot \varepsilon_{si} \cdot N_{sub}}{4 \cdot \Phi_F}}$
  3. Overlap Capacitance: $C_3 = C_4 = C_{ov} = W \cdot L_D \cdot C_{ox} + C_{fringe}$
  4. Junction Capacitance:
     - Sidewall Capacitance: C_{jsw}
     - Bottom-plate Capacitance: C_j

$$C_{jun} = \frac{C_{j0}}{\left[1 + \frac{V_R}{\Phi_B}\right]^m}$$

$$C_5 = C_6 = C_j + C_{jsw}$$



(a)



(b)

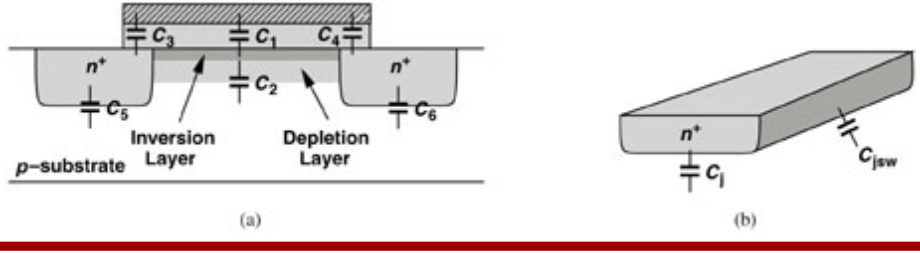
Device Capacitances - 3

In Cut-off:

1. C_{GS}: is equal to the overlap capacitance. $C_{GS} = C_{ov} = C_3$
2. C_{GD}: is equal to the overlap capacitance. $C_{GD} = C_{ov} = C_4$
3. C_{GB}: is equal to $C_{gate\text{-}channel} = C_1$ in series with $C_{channel\text{-}bulk} = C_2$.
4. C_{SB}: is equal to the junction capacitance between source and bulk.
5. C_{DB}: is equal to the junction capacitance between source and bulk.

$$C_{SB} = C_5$$

$$C_{DB} = C_6$$

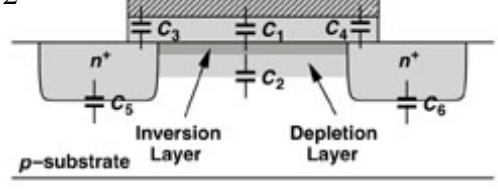


Device Capacitances - 4

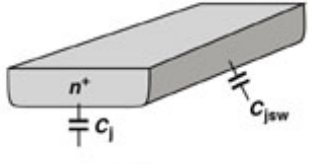
In Triode:

- The channel isolates the gate from the substrate. This means that if V_G changes, the charge of the inversion layer are supplied by the drain and source as long as V_{DS} is close to zero. So, C_1 is divided between gate and drain terminals, and gate and source terminals, and C_2 is divided between bulk and drain terminals, and bulk and source terminals.

1. C_{GS}: $C_{GS} = C_{ov} + \frac{C_1}{2}$
2. C_{GD}: $C_{GD} = C_{ov} + \frac{C_1}{2}$
3. C_{GB}: the channel isolates the gate from the substrate. $C_{GB} = 0$
4. C_{SB}: $C_{SB} = C_5 + \frac{C_2}{2}$
5. C_{DB}: $C_{DB} = C_6 + \frac{C_2}{2}$



(a)



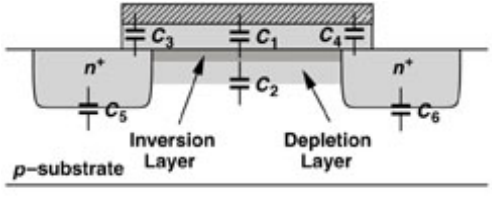
(b)

Device Capacitances - 5

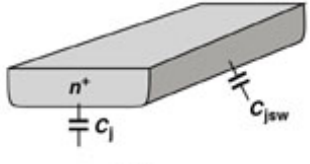
In Saturation:

- The channel isolates the gate from the substrate. The voltage across the channel varies which can be accounted for by adding two equivalent capacitances to the source. One is between source and gate, and is equal to two thirds of C_1. The other is between source and bulk, and is equal to two thirds of C_2.

1. C_{GS}: $C_{GS} = C_{ov} + \frac{2}{3}C_1$
2. C_{GD}: $C_{GD} = C_{ov}$
3. C_{GB}: the channel isolates the gate from the substrate. $C_{GB} = 0$
4. C_{SB}: $C_{SB} = C_5 + \frac{2}{3}C_2$
5. C_{DB}: $C_{DB} = C_6$



(a)

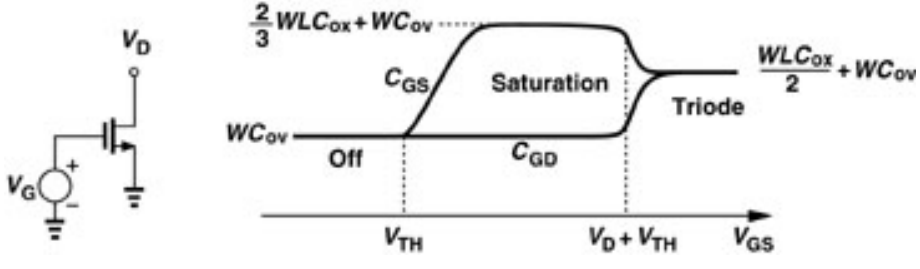


(b)

Device Capacitances - 6

- In summary:

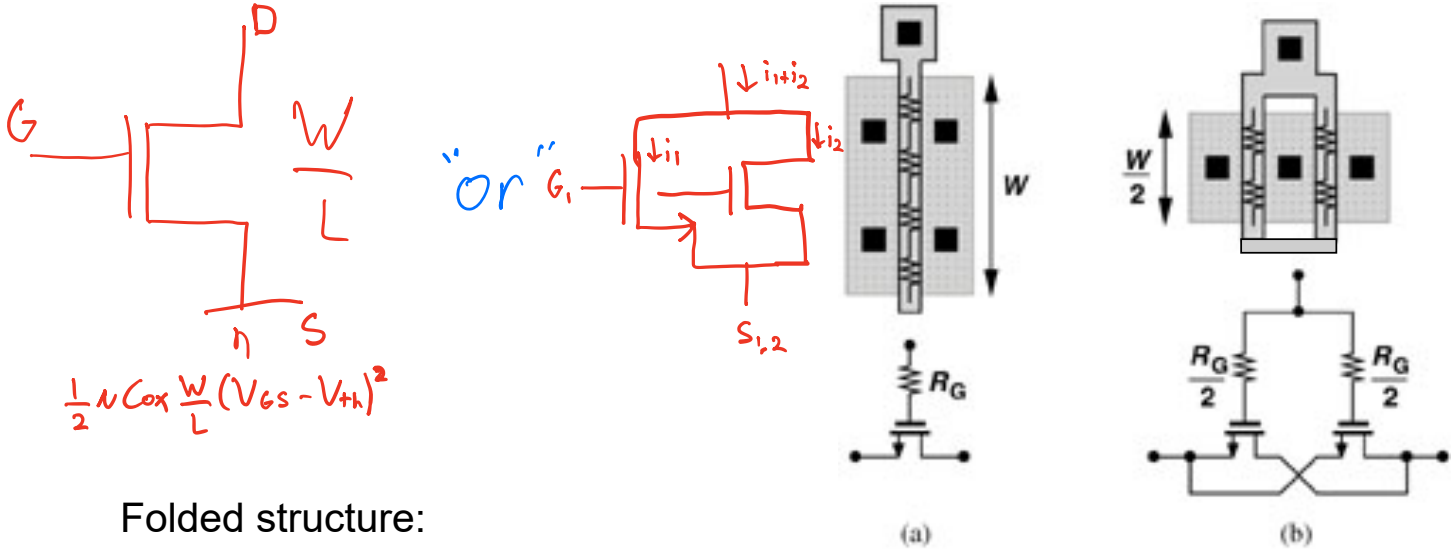
	Cut-off	Triode	Saturation
C_{GS}	C_{ov}	$C_{ov}+\frac{C_1}{2}$	$C_{ov}+\frac{2}{3}C_1$
C_{GD}	C_{ov}	$C_{ov}+\frac{C_1}{2}$	C_{ov}
C_{GB}	$\frac{C_1 \cdot C_2}{C_1+C_2} \langle C_{GB} \langle C_1$	0	0
C_{SB}	C_5	$C_5+\frac{C_2}{2}$	$C_5+\frac{2}{3}C_2$
C_{DB}	C_6	$C_6+\frac{C_2}{2}$	C_6



Importance of Layout

Example (Folded Structure):

Calculate the gate resistance of the circuits shown below.

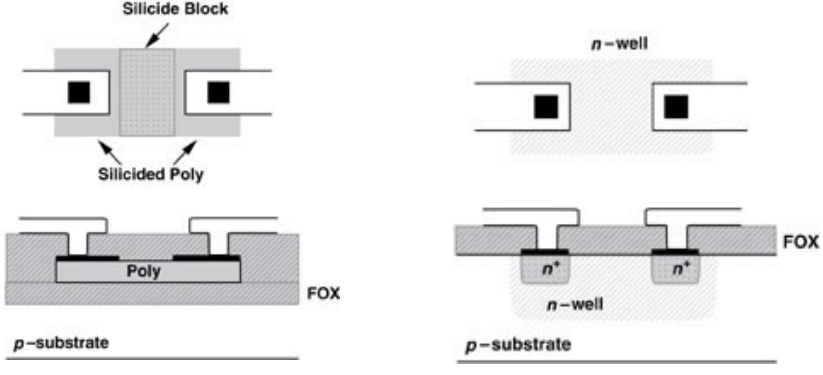


Folded structure:

- Decreases the drain capacitance
- Decreases the gate resistance
- Keeps the aspect ratio the same

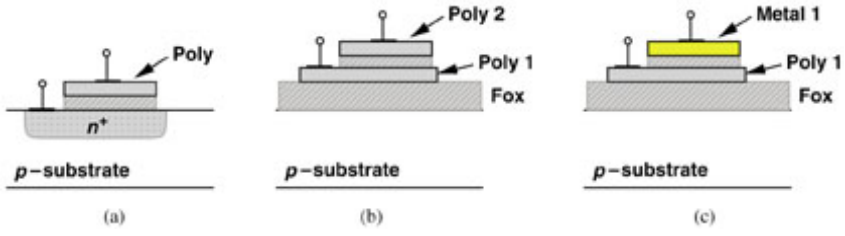
Passive Devices

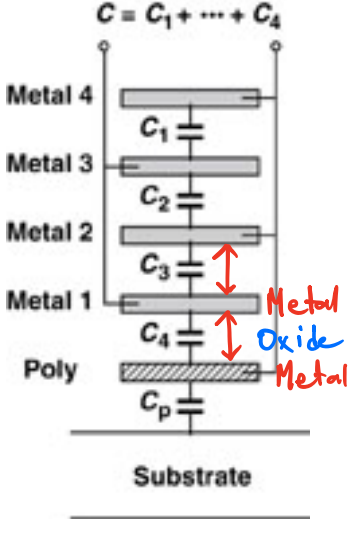
- Resistors



Passive Devices

- Capacitors:





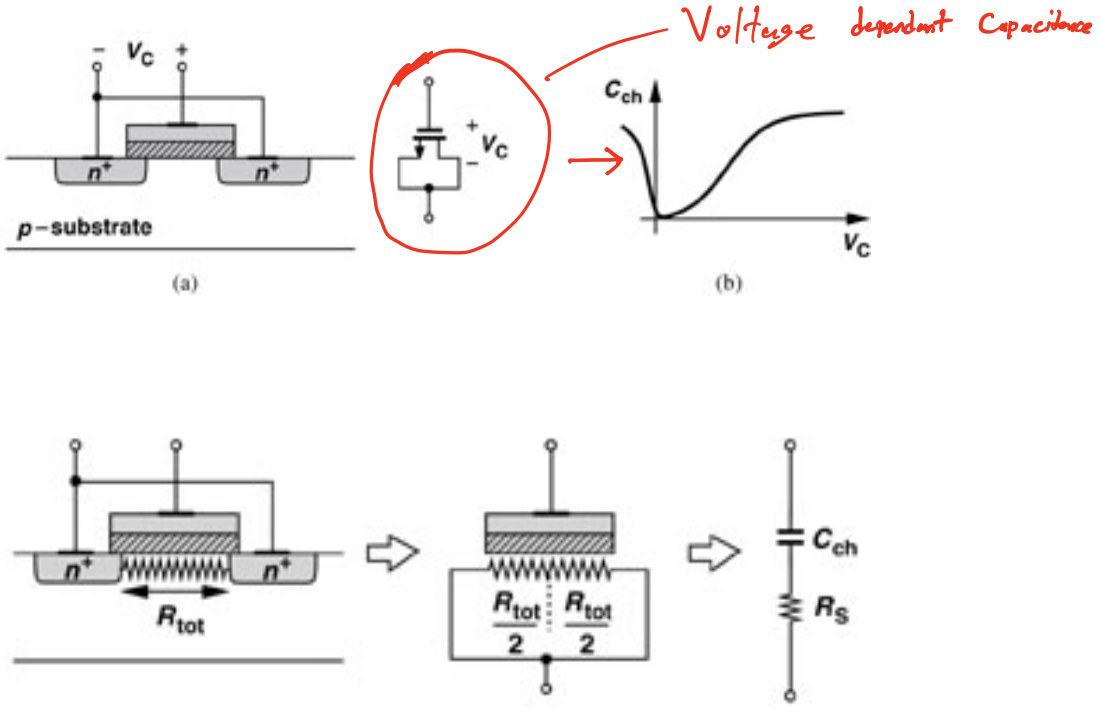
$\varepsilon \cdot \frac{A}{D}$

$\varepsilon_r \varepsilon_o$

MOM Capacitor

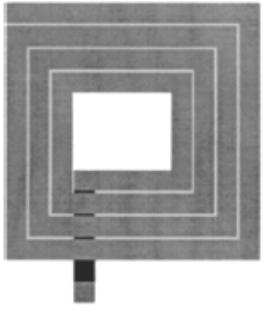
Passive Devices

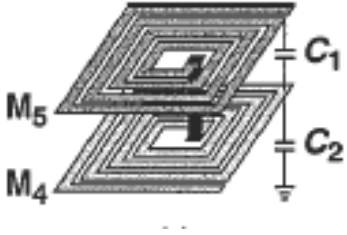
- Capacitors

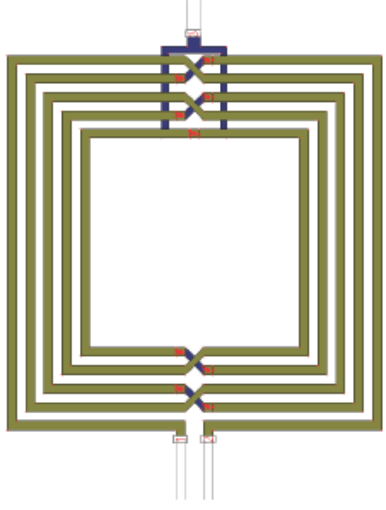


Passive Devices

- Inductors

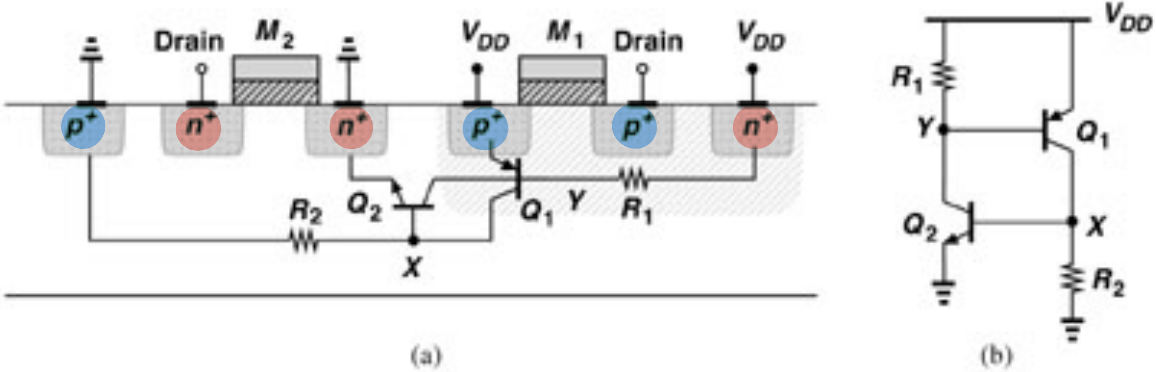






Latch-Up

- Due to parasitic bipolar transistors in a CMOS process



Small Signal Models - 1

- Small signal model is an approximation of the large-signal model around the operation point.
- In analog circuits most MOS transistors are biased in saturation region.
- In general, I_D is a function of V_{GS}, V_{DS}, and V_{BS}. We can use this Taylor series approximation:

Ignore if Δx small

$$\text{Taylor Expansion}: I_D = I_{D0} + \frac{\partial I_D}{\partial V_{GS}} \cdot \Delta V_{GS} + \frac{\partial I_D}{\partial V_{DS}} \cdot \Delta V_{DS} + \frac{\partial I_D}{\partial V_{BS}} \cdot \Delta V_{BS} + \text{second order terms}$$

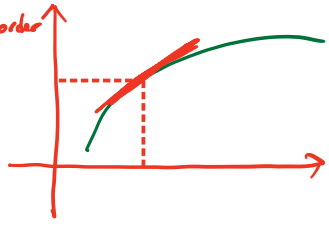
$$\Delta I_D \approx \frac{\partial I_D}{\partial V_{GS}} \cdot \Delta V_{GS} + \frac{\partial I_D}{\partial V_{DS}} \cdot \Delta V_{DS} + \frac{\partial I_D}{\partial V_{BS}} \cdot \Delta V_{BS} = g_m \cdot \Delta V_{GS} + \frac{\Delta V_{DS}}{r_o} + g_{mb} \cdot \Delta V_{BS}$$

$y = f(x)$

$$y = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 \cdots$$

(Δx)

If small, we can ignore higher order terms in TSE



Linearize non-linear fcu. by approximating with a tangent

Small Signal Models - 2

- Current in Saturation: $I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS}-V_{TH})^2 \approx \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{TH})^2\cdot(1+\lambda\cdot V_{DS})$
- Taylor approximation: $\Delta I_D \approx \frac{\partial I_D}{\partial V_{GS}}\cdot\Delta V_{GS} + \frac{\partial I_D}{\partial V_{DS}}\cdot\Delta V_{DS} + \frac{\partial I_D}{\partial V_{BS}}\Delta V_{BS}$

  g_m $\quad 1/r_o$ $\quad g_{mb}$
- Partial Derivatives:

S ← D

$\frac{\Delta I_D}{\Delta V_{DS}}$

$$\frac{\partial I_D}{\partial V_{GS}} = \mu_n\cdot C_{ox}\cdot\frac{W}{L}\cdot(V_{GS}-V_{TH})\cdot(1+\lambda\cdot V_{DS}) = g_m$$

$$\frac{\partial I_D}{\partial V_{DS}} = \frac{1}{2}\cdot\mu_n\cdot C_{ox}\cdot\frac{W}{L}\cdot(V_{GS}-V_{TH})^2\cdot\lambda \approx I_D\cdot\lambda = \frac{1}{r_o}$$

$$\frac{\partial I_D}{\partial V_{BS}} = \frac{\partial I_D}{\partial V_{TH}}\cdot\frac{\partial V_{TH}}{\partial V_{BS}} = \left[-\mu_n\cdot C_{ox}\cdot\frac{W}{L}\cdot(V_{GS}-V_{TH})\cdot(1+\lambda\cdot V_{DS})\right]\cdot\left[-\frac{\gamma}{2\sqrt{|2\cdot\Phi_F+V_{SB}|}}\right]$$

$$= -g_m\cdot\left[-\frac{\gamma}{2\sqrt{|2\cdot\Phi_F+V_{SB}|}}\right] = g_m\cdot\eta = g_{mb}$$

Bulk Transconductance

Small Signal Models - 3

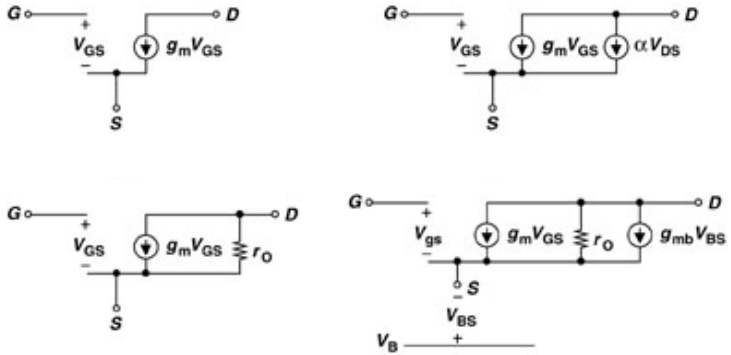
- Small-Signal Model:

$$i_D = g_m \cdot v_{GS} + \frac{v_{DS}}{r_o} + g_{mb} \cdot v_{BS}$$

Can Cancel

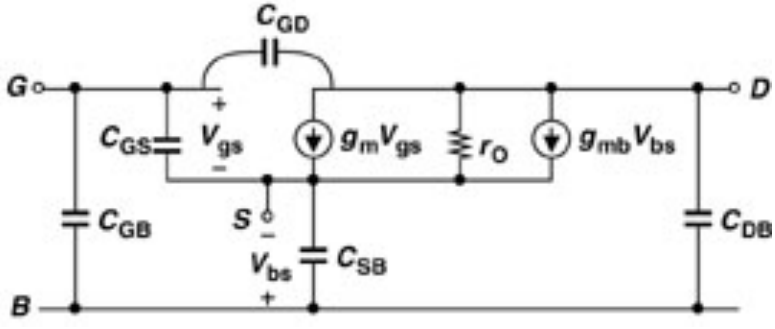
$$r_o = \frac{1}{\lambda I_D}$$

- Terms, $g_m v_{GS}$ and $g_{mb} v_{BS}$, can be modeled by dependent sources. These terms have the same polarity: increasing v_G, has the same effect as increasing v_B.
- The term, v_{DS}/r_o can be modeled using a resistor as shown below.



Small Signal Models - 4

- Complete Small-Signal Model with Capacitances:



- Small signal model including all the capacitance makes the intuitive (qualitative) analysis of even a few-transistor circuit difficult!

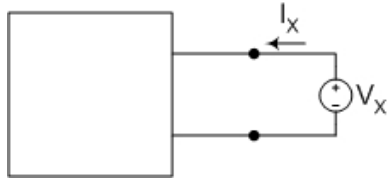
- Typically, CAD tools are used for accurate circuit analysis

- For intuitive analysis we try to find a simplest model that can represent the role of each transistor with reasonable accuracy.

Circuit Impedance - 1

- It is often useful to determine the impedance of a circuit seen from a specific pair of terminals.

- The following is the recipe to do so:
  1. Connect a voltage source, V_X, to the port.
  2. Suppress all independent sources.
  3. Measure or calculate I_X.

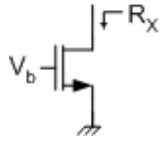


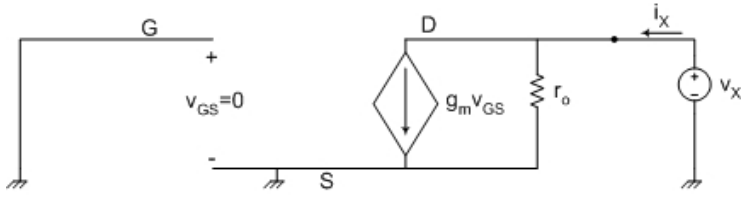
$$R_X = \frac{V_X}{I_X}$$

Circuit Impedance - 2

Example:

- Find the small-signal impedance of the following current sources.
- We draw the small-signal model, which is the same for both circuits, and connect a voltage source as shown below:





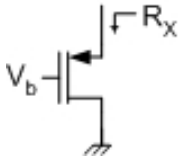
$$i_X = \frac{v_X}{r_o} + g_m \cdot v_{GS} = \frac{v_X}{r_o}$$

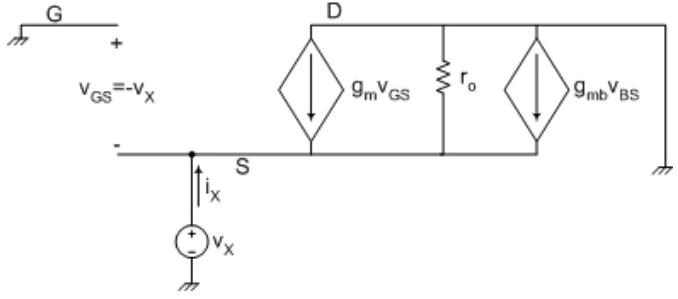
$$R_X = \frac{v_X}{i_X} = r_o$$

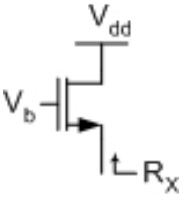
Circuit Impedance - 3

Example:

- Find the small-signal impedance of the following circuits.
- We draw the small-signal model, which is the same for both circuits, and connect a voltage source as shown below:







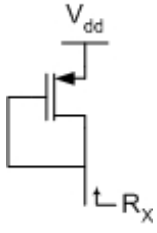
$$i_X = \frac{v_X}{r_o} - g_m \cdot v_{GS} - g_{mb} \cdot v_{BS} = \frac{v_X}{r_o} + g_m \cdot v_X + g_{mb} \cdot v_X$$

$$R_X = \frac{v_X}{i_X} = \frac{1}{\frac{1}{r_o} + g_m + g_{mb}} = r_o \left\| \frac{1}{g_m} \right\| \frac{1}{g_{mb}}$$

Circuit Impedance - 4

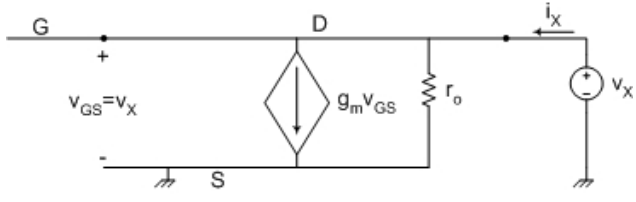
Example:

- Find the small-signal impedance of the following circuit. This circuit is known as the diode-connected load, and is used frequently in analog circuits.



- We draw the small-signal model and connect the voltage source as shown below:

$$i_X = \frac{v_X}{r_o} + g_m \cdot v_{GS} = \frac{v_X}{r_o} + g_m \cdot v_X = v_X \cdot \left(\frac{1}{r_o} + g_m \right)$$



$$R_X = \frac{v_X}{i_X} = \frac{1}{\frac{1}{r_o} + g_m} = r_o \left\| \frac{1}{g_m} \right.$$

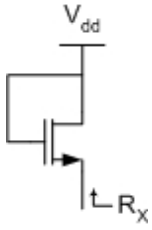
- If channel length modulation is ignored ($r_o = \infty$) we get:

$$R_X = r_o \left\| \frac{1}{g_m} \right. = \infty \left\| \frac{1}{g_m} \right. = \frac{1}{g_m}$$

Circuit Impedance - 5

Example:

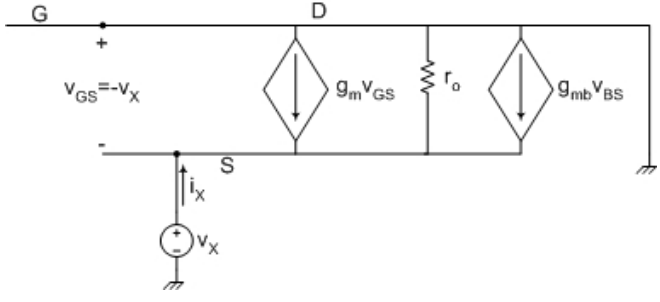
- Find the small-signal impedance of the following circuit. This circuit is a diode-connected load with body effect.



$$i_X = \frac{v_X}{r_o} - g_m \cdot v_{GS} - g_{mb} \cdot v_{BS} = \frac{v_X}{r_o} + g_m \cdot v_X + g_{mb} \cdot v_X$$

$$= v_X \cdot \left(\frac{1}{r_o} + g_m + g_{mb} \right)$$

$$R_X = \frac{v_X}{i_X} = \frac{1}{\frac{1}{r_o} + g_m + g_{mb}} = r_o \left\| \frac{1}{g_m + g_{mb}} = r_o \right\| \frac{1}{g_m} \left\| \frac{1}{g_{mb}} \right.$$



- If channel length modulation is ignored ($r_o = \infty$) we get:

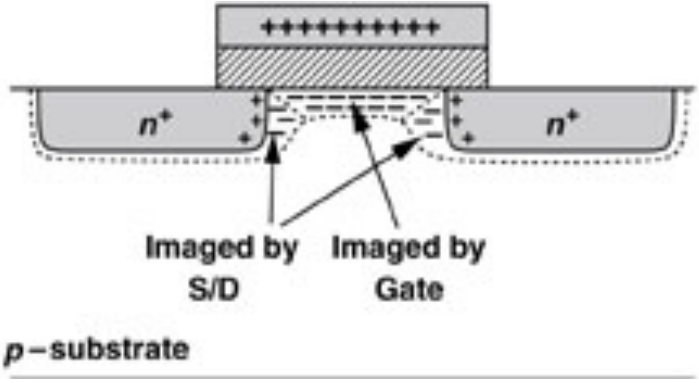
$$R_X = r_o \left\| \frac{1}{g_m + g_{mb}} = \infty \right\| \frac{1}{g_m + g_{mb}} = \frac{1}{g_m + g_{mb}} = \frac{1}{g_m} \left\| \frac{1}{g_{mb}} \right.$$

Short-Channel Effects

- Threshold Reduction
  - Drain-induced barrier lowering (DIBL)
- Mobility degradation
- Velocity saturation
- Hot carrier effects
  - Substrate current
  - Gate current
- Output impedance variation

Threshold Voltage Variation in Short Channel Devices

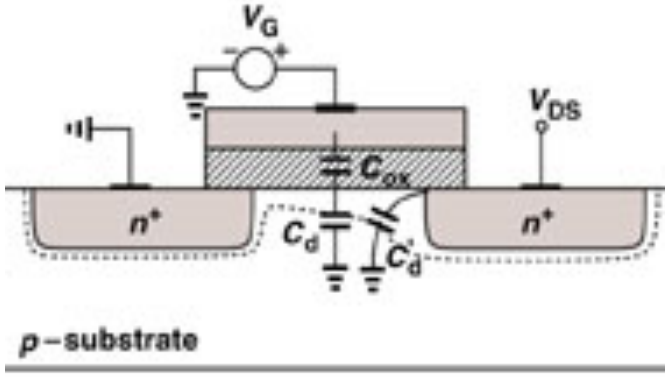
- The Threshold of transistors fabricated on the same chip decreases as the channel length decreases.



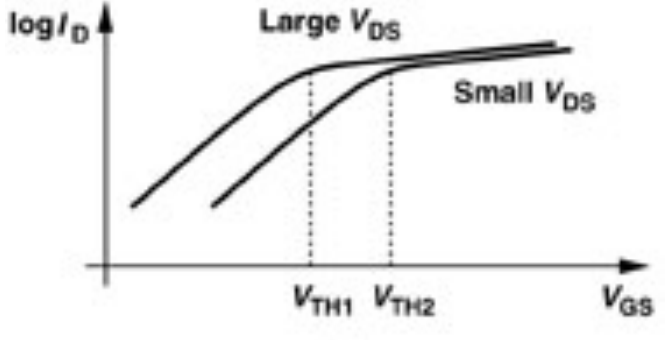
- Intuitively, the extent of depletion regions associated with drain and source in the channel area, reduces the immobile charge that must be imaged by the charge on the gate.

Drain-Induced Barrier Lowering (DIBL)

When the channel is short, the drain voltage increases the channel surface potential, lowering the barrier to flow charge from source (think of increased electric field) and therefore, decreasing the threshold.



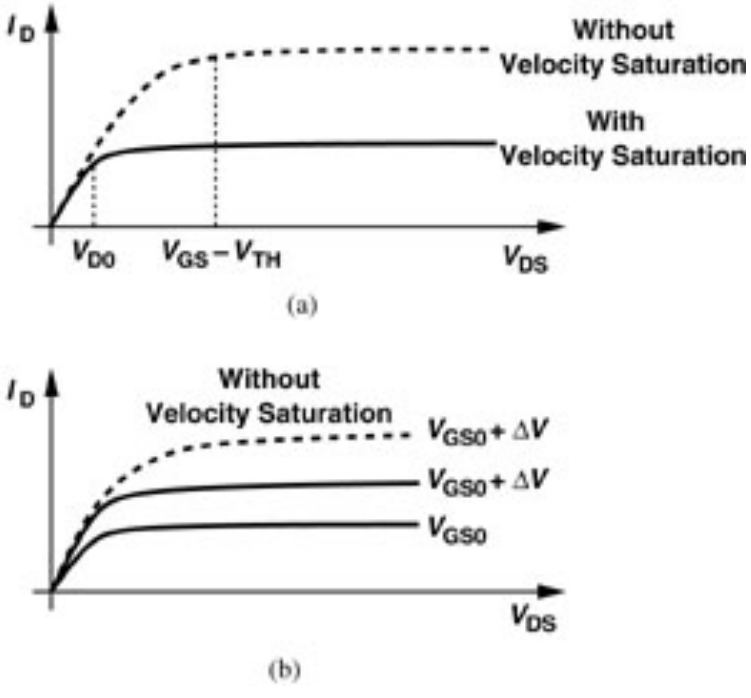
(a)



(b)

Effects of Velocity Saturation

- Due to drop in mobility at high electric fields



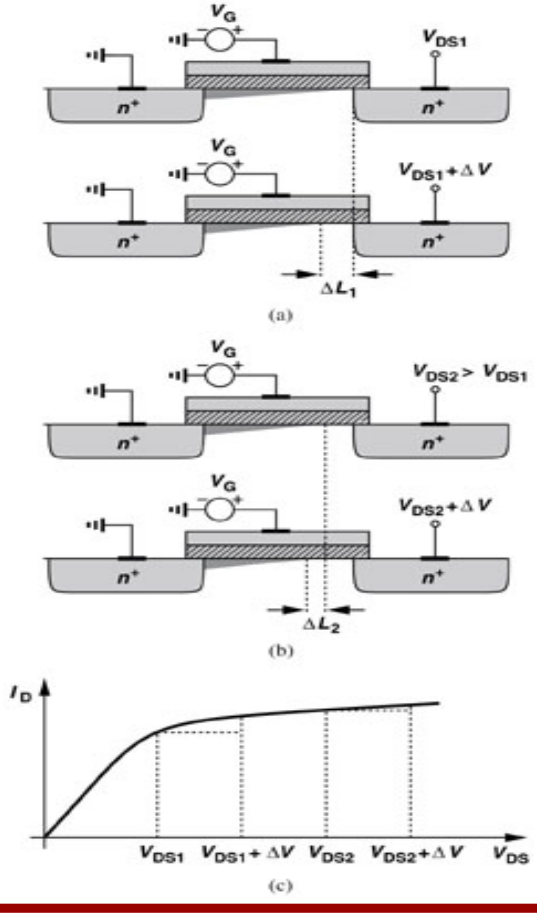
- (a) Premature drain current saturation and (b) reduction in g_m

Hot Carrier Effects

- Short channel devices may experience high lateral drain-source electric field
- Some carriers that make it to drain have high velocity (called “hot” carriers)
- “Hot” carriers may “hit” silicon atoms at high speed and cause impact ionization
- The resulting electron and holes are absorbed by the drain and substrate causing extra drain-substrate current
- Really “hot” carriers may be injected into gate oxide and flow out of gate causing gate current!

Output Impedance Variation

Recall the definition of λ.



Output Impedance Variation in Short-Channel Devices