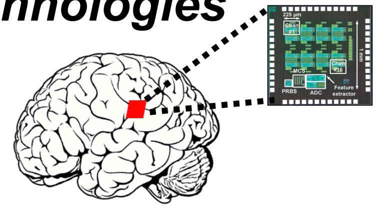
Integrated Neurotechnologies Laboratory

EPFL

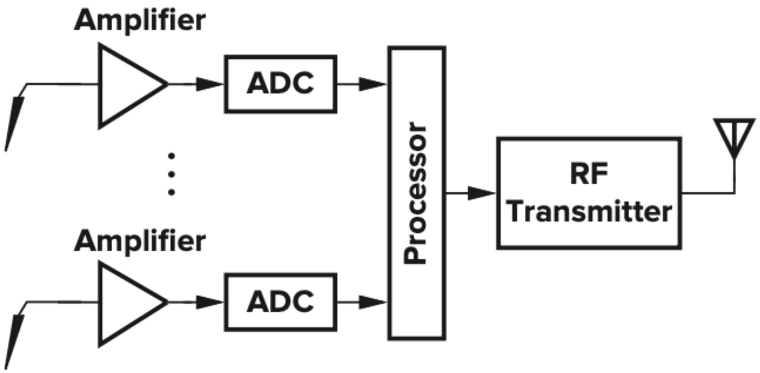


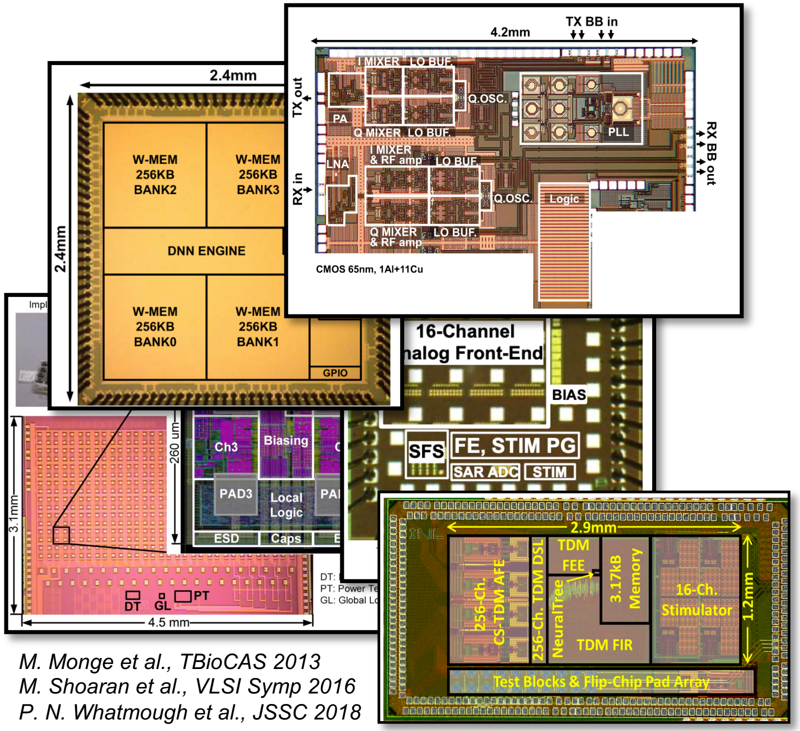
Analog IC design (EE-320), Lecture 2

Prof. Mahsa Shoaran

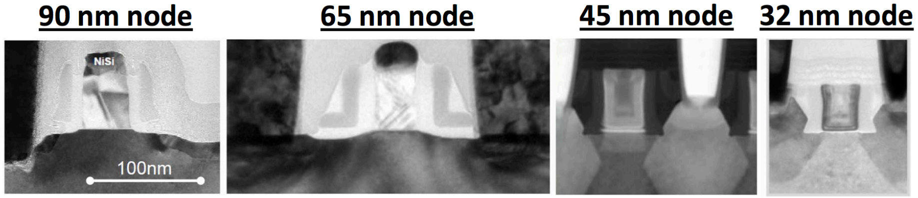
Institute of Electrical and Micro Engineering, School of Engineering, EPFL

Review: CMOS Analog Design



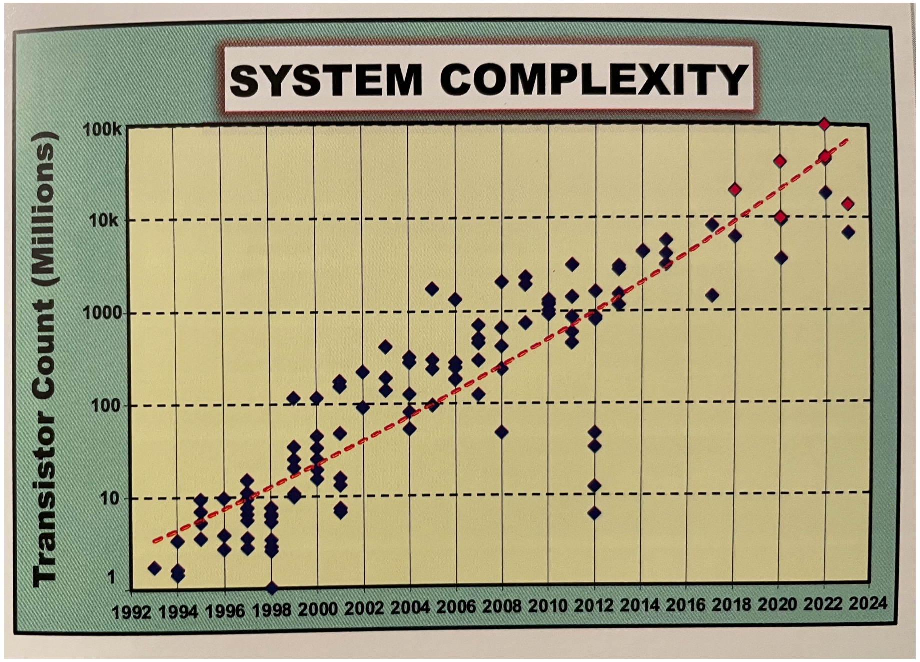


M. Monge et al., TBioCAS 2013
M. Shoaran et al., VLSI Symp 2016
P. N. Whatmough et al., JSSC 2018
R. Wu et al., JSSC 2017
U. Shin et al., JSSC 2022

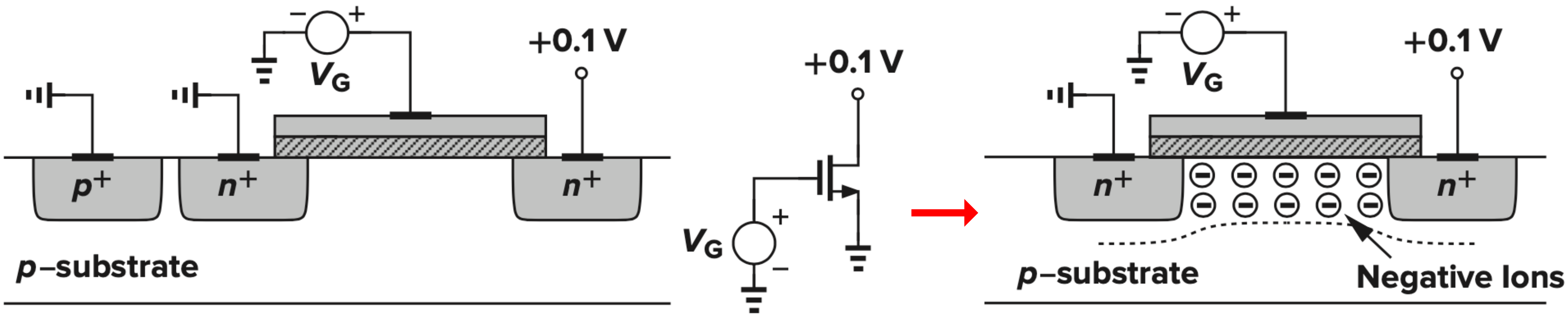


courtesy V. Moroz, IEDM

Microprocessor Transistor Count, Moore's Law



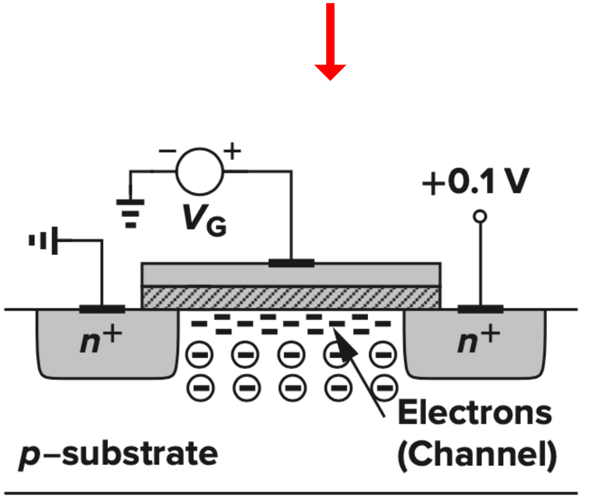
Review: MOS threshold voltage



$$\boxed{V_{TH} = \Phi_{MS} + 2\Phi_F + \frac{Q_{dep}}{C_{ox}}}$$

$$\Phi_F = (kT/q)\ln(N_{sub}/n_i)$$

$$Q_{dep} = \sqrt{4q\epsilon_{si}|\Phi_F|N_{sub}}$$



Q_{dep} : charge in the depletion region

N_{sub}: *doping density of the substrate*

C_{ox} : the gate-oxide capacitance per unit area

n_i : density of electrons in undoped silicon

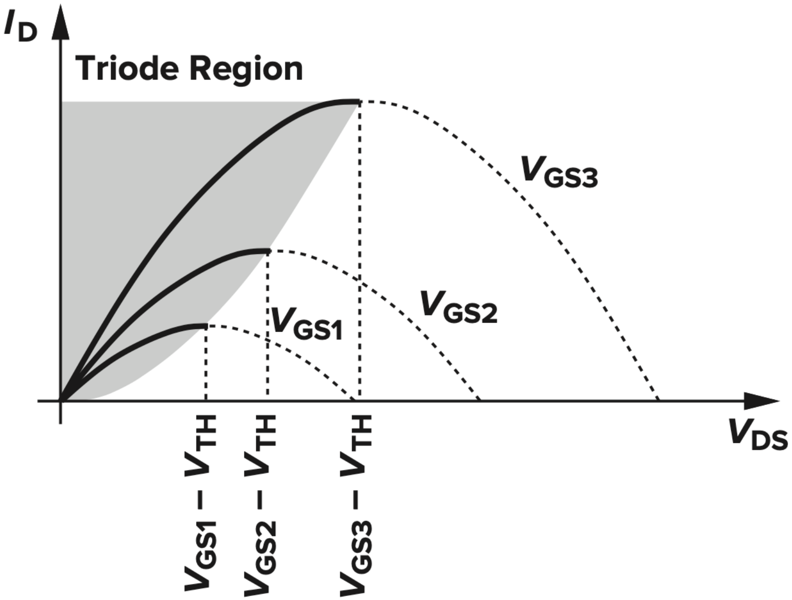
ε_{si} : the dielectric constant of silicon

Φ_{MS} : the difference between the work functions of the poly (gate) and substrate

Review: I_D vs. V_{DS} in the triode region

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

- ***The overdrive voltage:*** $V_{GS} - V_{TH}$
- ***The aspect ratio:*** *W/L*



$$\boxed{V_{DS} \leq V_{GS} - V_{TH}}$$

$\downarrow$ or $V_{GD} > V_{TH}$

- "**triode**" region (or **linear** region)

Review: I_D vs. V_{DS} in the triode region

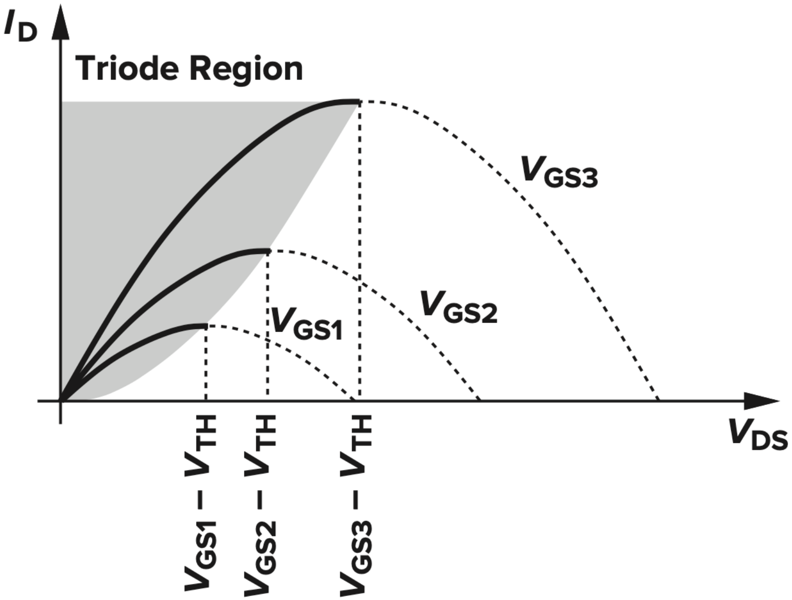
$$\boxed{I_D = \mu_n C_{ox} \frac{W}{L}\left[(V_{GS} - V_{TH})V_{DS} - \frac{1}{2}V_{DS}^2\right]}$$

- ***The overdrive voltage:*** $V_{GS} - V_{TH}$
- ***The aspect ratio:*** *W/L*

$$\boxed{V_{DS} \leq V_{GS} - V_{TH}}$$

↓ or $V_{GD} > V_{TH}$

- "**triode**" region (or **linear** region)

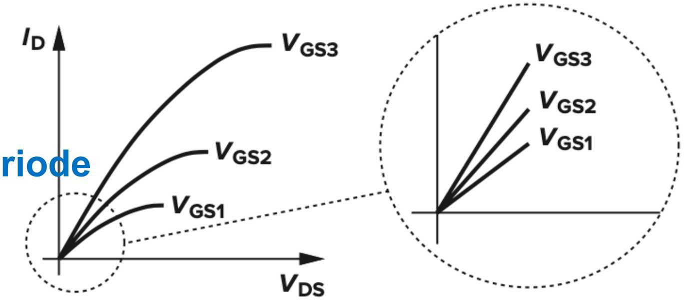


$$V_{DS} \ll 2(V_{GS} - V_{TH})$$

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

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

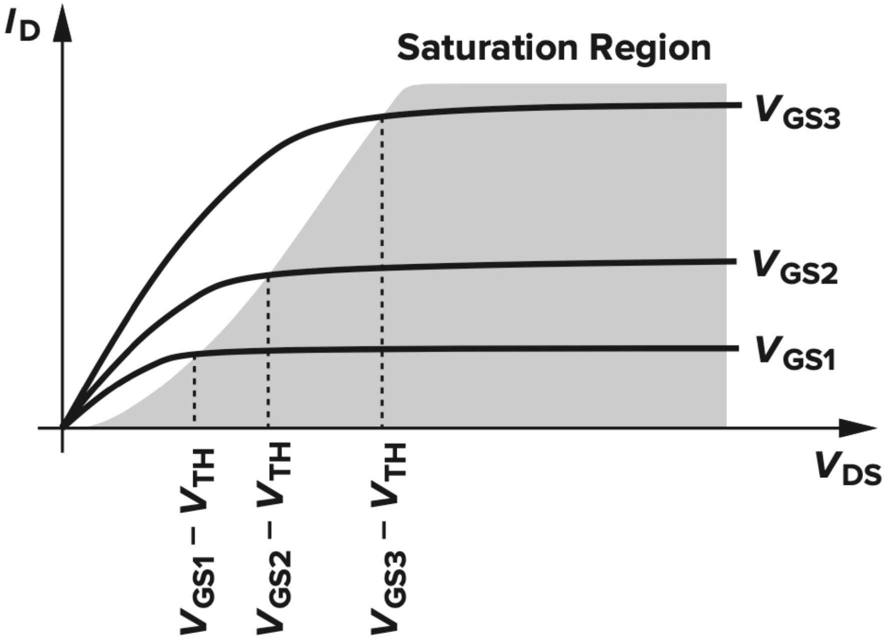
Deep triode region



MOS in Saturation

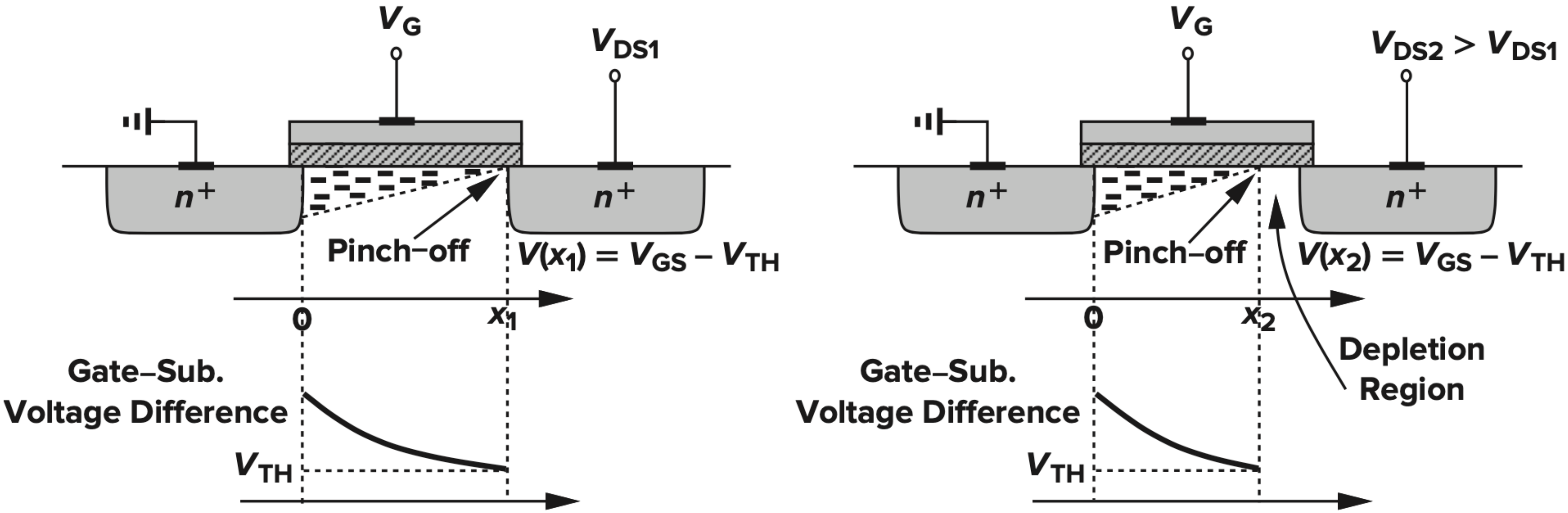
$$\boxed{V_{DS} > V_{GS} - V_{TH}}$$

- I_D becomes relatively **constant**, and we say that the device operates in the "**saturation**" region.



MOS in Saturation: Pinch-off behavior

$$\boxed{V_{DS} > V_{GS} - V_{TH}}$$

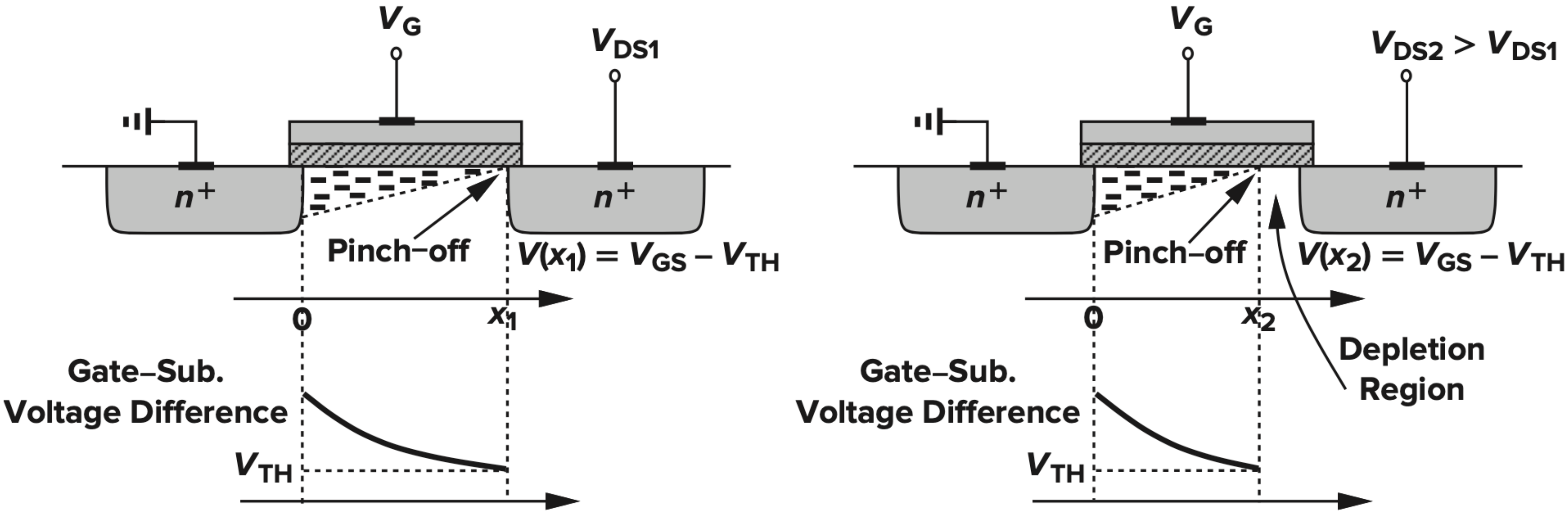


- Recall that the local density of the inversion-layer charge is proportional to:

$$V_{GS} - V(x) - V_{TH}$$

MOS in Saturation: Pinch-off behavior

$$\boxed{V_{DS} > V_{GS} - V_{TH}}$$



- Recall that the local density of the inversion-layer charge is proportional to:

$$V_{GS} - V(x) - V_{TH}$$

- if V(x) → $V_{GS} - V_{TH}$, then $Q_d(x)$ drops to zero. So, if V_{DS} is slightly **greater** than $V_{GS} - V_{TH}$, the inversion layer stops at x ≤ L, the channel is "**pinched off**"

MOS in Saturation: I/V Characteristics

- As the electrons approach the pinch-off point (where $Q_d \rightarrow 0$), their velocity rises tremendously ($v = I/Q_d$)

- Upon passing the pinch-off point, the electrons simply shoot through the depletion region and arrive at the drain terminal

- Q_d the density of *mobile* charge: the integral from $x = 0$ to $x = L'$, where L' is the point at which Q_d drops to zero (e.g., x_2), and from $V(x) = 0$ to $V(x) = V_{GS} - V_{TH}$

Recall: derivation of I/V Characteristics

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

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

MOS in Saturation: I/V Characteristics

- As the electrons approach the pinch-off point (where $Q_d \rightarrow 0$), their velocity rises tremendously ($v = I/Q_d$)
- Upon passing the pinch-off point, the electrons simply shoot through the depletion region and arrive at the drain terminal
- Q_d the density of *mobile* charge: the integral from $x = 0$ to $x = L'$, where L' is the point at which Q_d drops to zero (e.g., x_2), and from $V(x) = 0$ to $V(x) = V_{GS} - V_{TH}$

*A "**square-law**" behavior:*

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

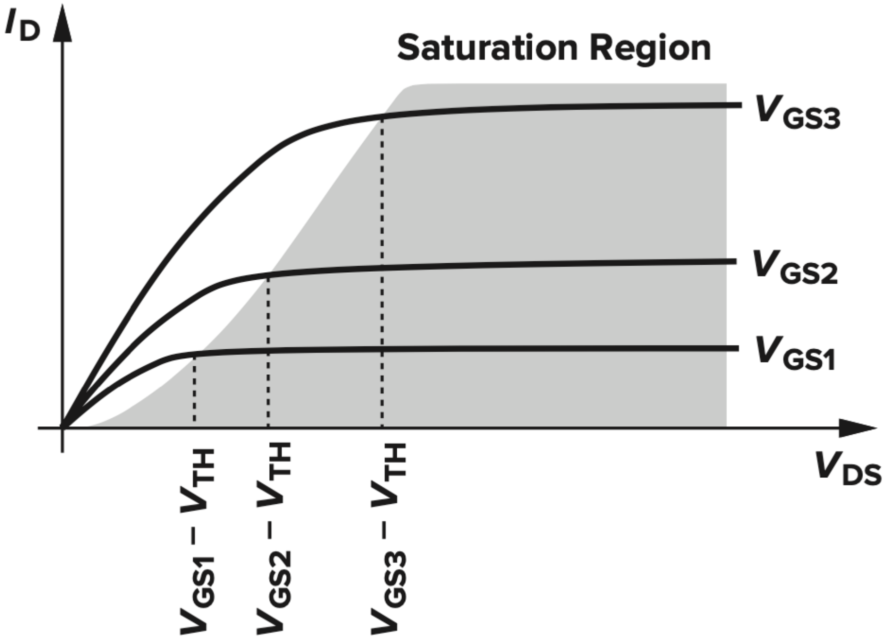
$$\boxed{V_{DS} > V_{GS} - V_{TH}}$$

or $V_{GD} < V_{TH}$

MOS in Saturation

$$\boxed{V_{DS} > V_{GS} - V_{TH}}$$

- I_D becomes relatively **constant**, and we say that the device operates in the "**saturation**" region.



MOS in Saturation: I/V Characteristics

- As the electrons approach the pinch-off point (where $Q_d \rightarrow 0$), their velocity rises tremendously ($v = I/Q_d$)

- Upon passing the pinch-off point, the electrons simply shoot through the depletion region and arrive at the drain terminal

- Q_d the density of *mobile* charge: the integral from $x = 0$ to $x = L'$, where L' is the point at which Q_d drops to zero (e.g., x_2), and from $V(x) = 0$ to $V(x) = V_{GS} - V_{TH}$

*A "**square-law**" behavior:*

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

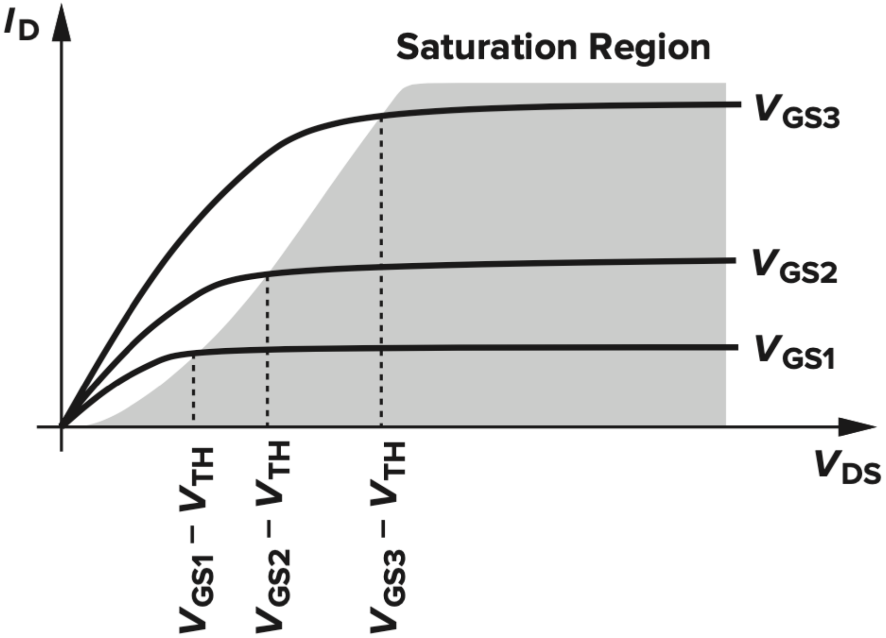
$$\boxed{V_{DS} > V_{GS} - V_{TH}}$$

or $V_{GD} < V_{TH}$

$$V_{GS} = \sqrt{\frac{2I_D}{\mu_n C_{ox}\frac{W}{L'}}} + V_{TH}$$

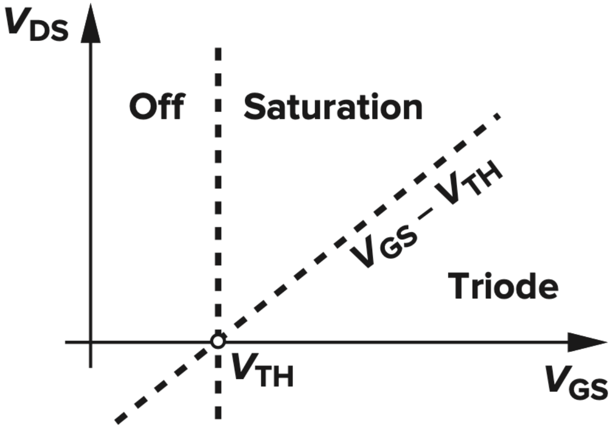
MOS in Saturation

$$\boxed{V_{DS} > V_{GS} - V_{TH}}$$



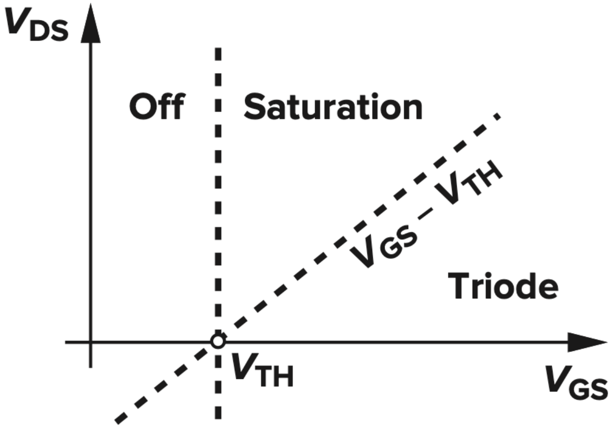
MOS in Saturation

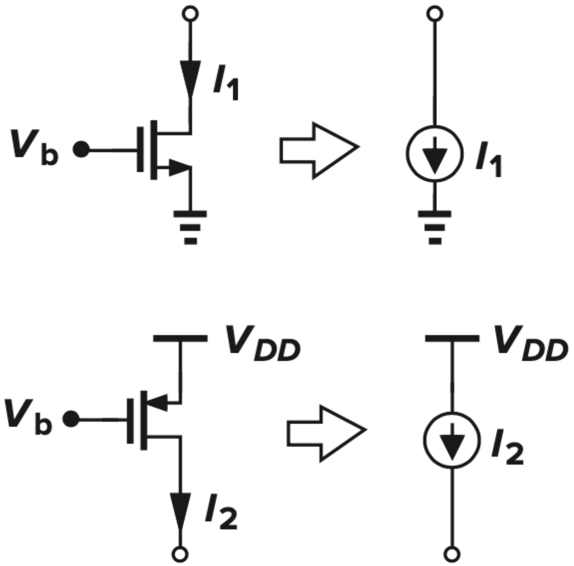
- The drain-source voltage must be **equal to** or **greater** than **overdrive** voltage
  - $\mathbf{V_{D,sat} = V_{GS} - V_{TH}}$, where $\mathbf{V_{D,sat}}$ denotes the **minimum** V_{DS} necessary for operation in **saturation**
  - the **larger** the $\mathbf{V_{D,sat}}$, the **less headroom** is available



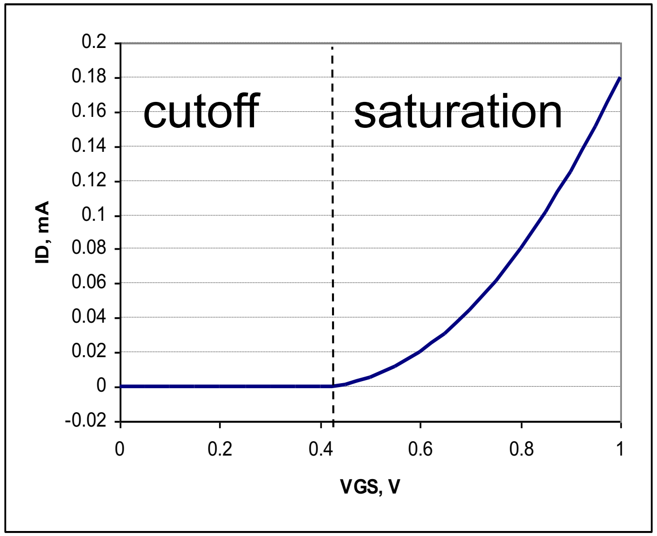
MOS in Saturation

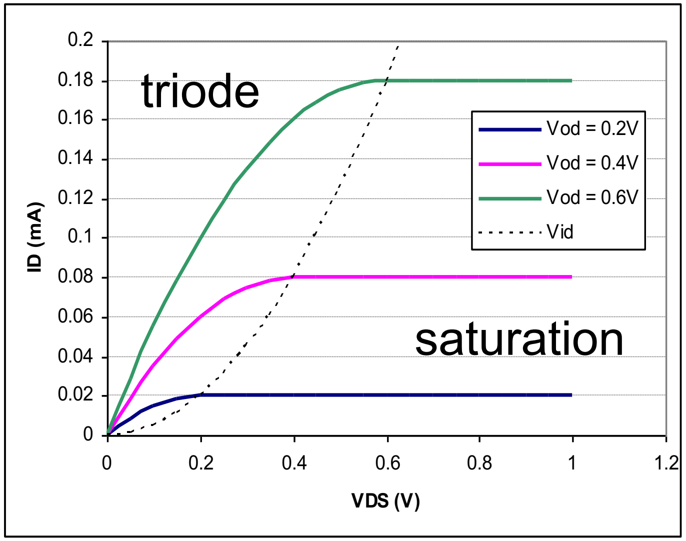
- The drain-source voltage must be **equal to** or **greater** than **overdrive** voltage
  - **$V_{D,sat}$ = V_{GS} − V_{TH}**, where **$V_{D,sat}$** denotes the **minimum** V_{DS} necessary for operation in **saturation**
  - the **larger** the **$V_{D,sat}$,** the **less headroom** is available





MOS I/V in different regions





I_D for PMOS transistor, regimes of operation

- We assume that I_D flows from the drain to the source, whereas holes flow in the reverse direction: a negative sign needed

- Note that V_{GS}, V_{DS}, V_{TH}, and $V_{GS} - V_{TH}$ are **negative** for a PMOS that is **turned on**
- Mobility of holes ~1/2 of electrons → **PMOS: lower "current driving" capability**

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

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

I_D for PMOS transistor, regimes of operation

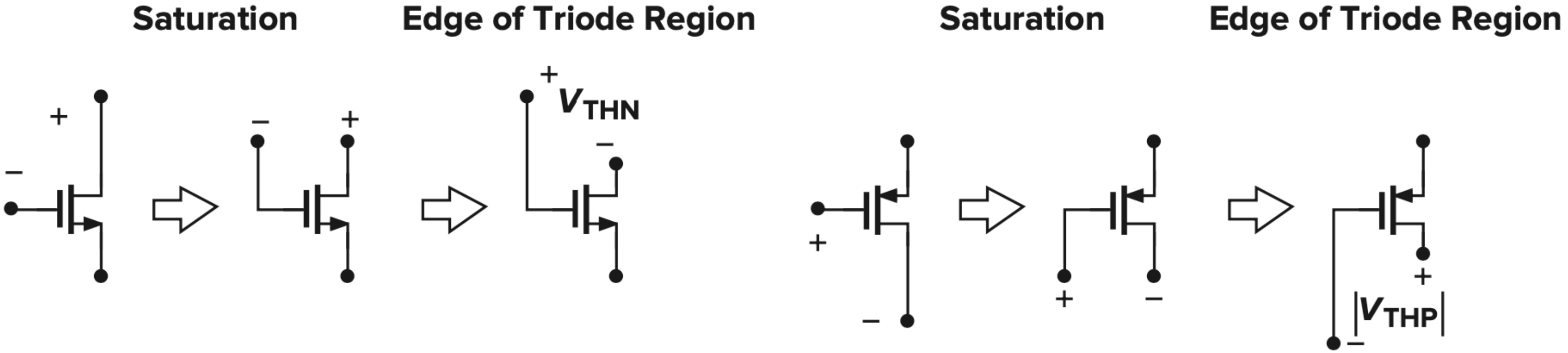
- We assume that I_D flows from the drain to the source, whereas holes flow in the reverse direction: a negative sign needed

- Note that V_{GS}, V_{DS}, V_{TH}, and $V_{GS} - V_{TH}$ are **negative** for a PMOS that is **turned on**
- Mobility of holes ~1/2 of electrons → **PMOS: lower "current driving" capability**

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

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

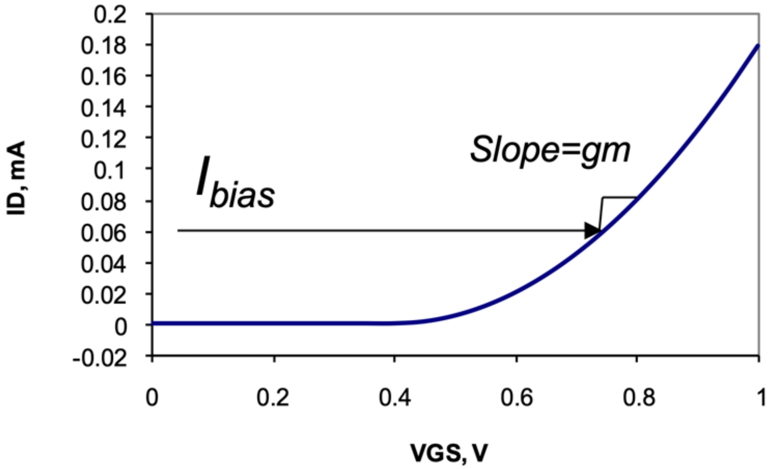
PMOS:
saturation: $V_{GD} > V_{TH}$
triode: $V_{GD} < V_{TH}$
V_{TH} is negative



MOS transconductance (g_m) in saturation

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{VDS\text{ const.}} = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH})$$

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



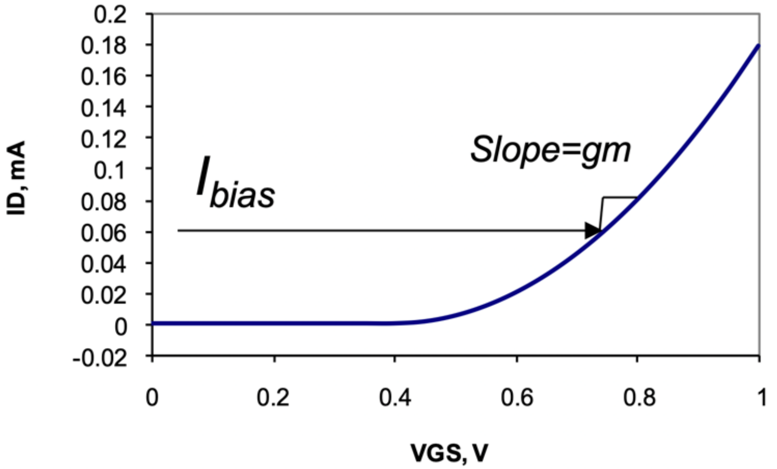
MOS transconductance (g_m) in saturation

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{VDS \text{ const.}} = \mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH})$$

or

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

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



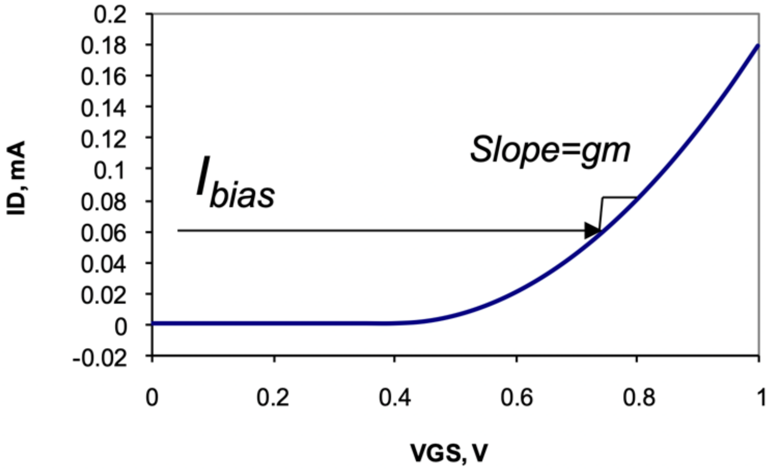
MOS transconductance (g_m) in saturation

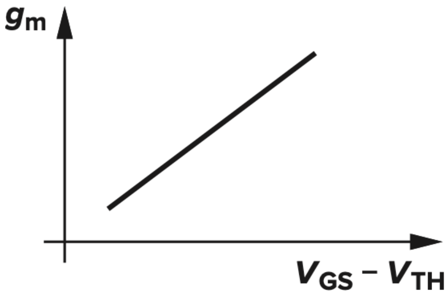
$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{VDS\ \text{const.}} = \mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{TH})$$

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

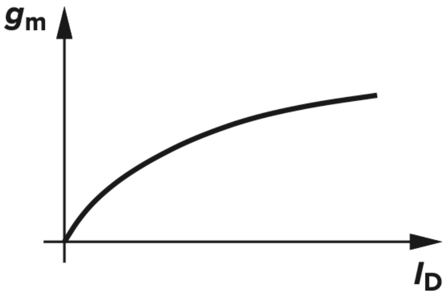
or

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

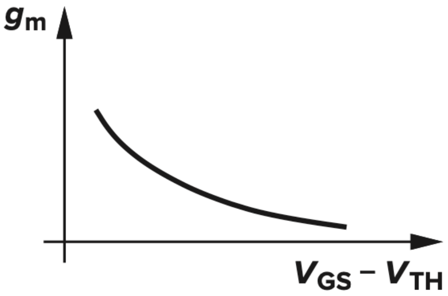




W/L Constant



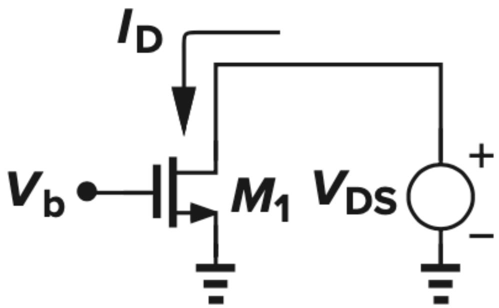
W/L Constant



I_D Constant

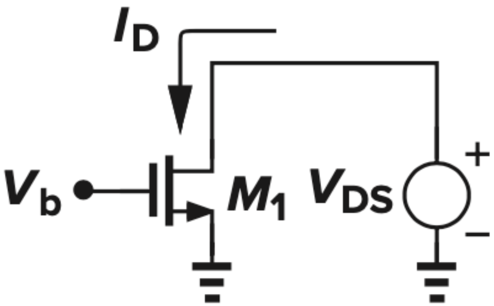
Example: MOS transconductance (g_m) in triode?

- Plot g_m as a function of V_{DS}?



Example: MOS transconductance (g_m) in triode?

- Plot g_m as a function of V_{DS}?

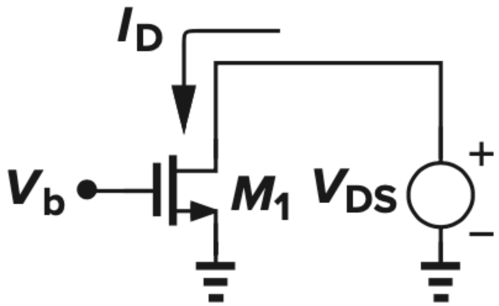


$$g_m = \frac{\partial}{\partial V_{GS}} \left\{ \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \left[2(V_{GS} - V_{TH}) V_{DS} - V_{DS}^2 \right] \right\}$$

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

Example: MOS transconductance (g_m) in triode?

- Plot g_m as a function of V_{DS}?



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

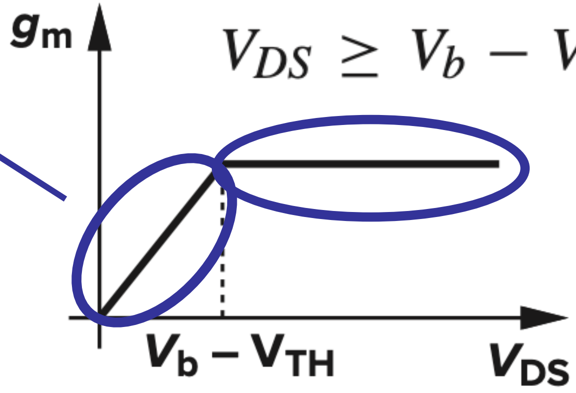
$$= \mu_n C_{ox} \frac{W}{L} V_{DS}$$

triode

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

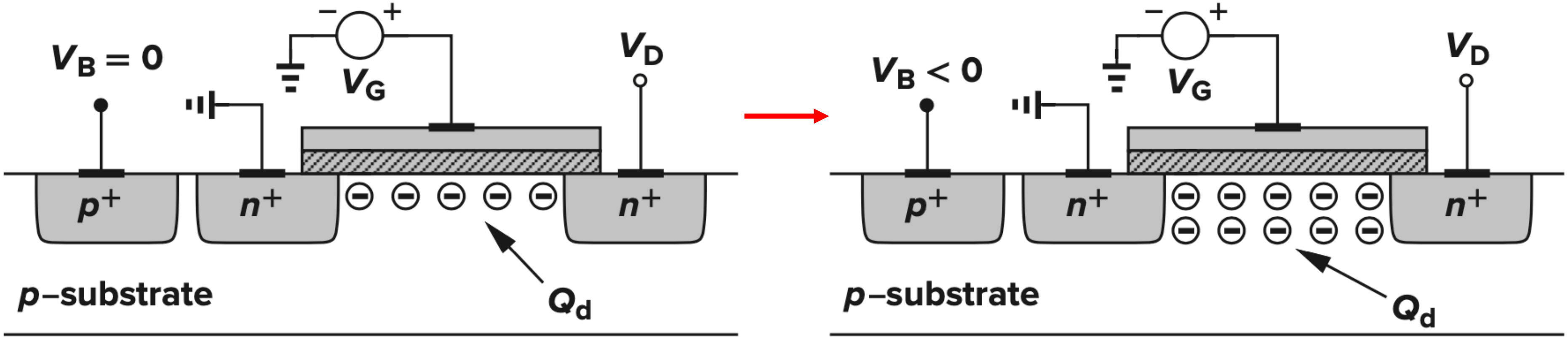
saturation

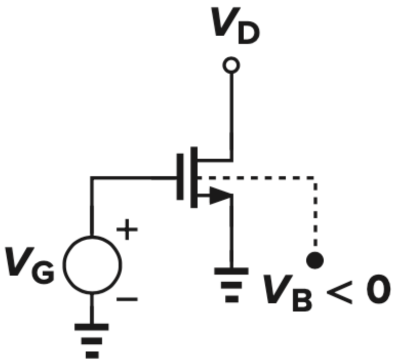
$$V_{DS} \geq V_b - V_{TH}$$



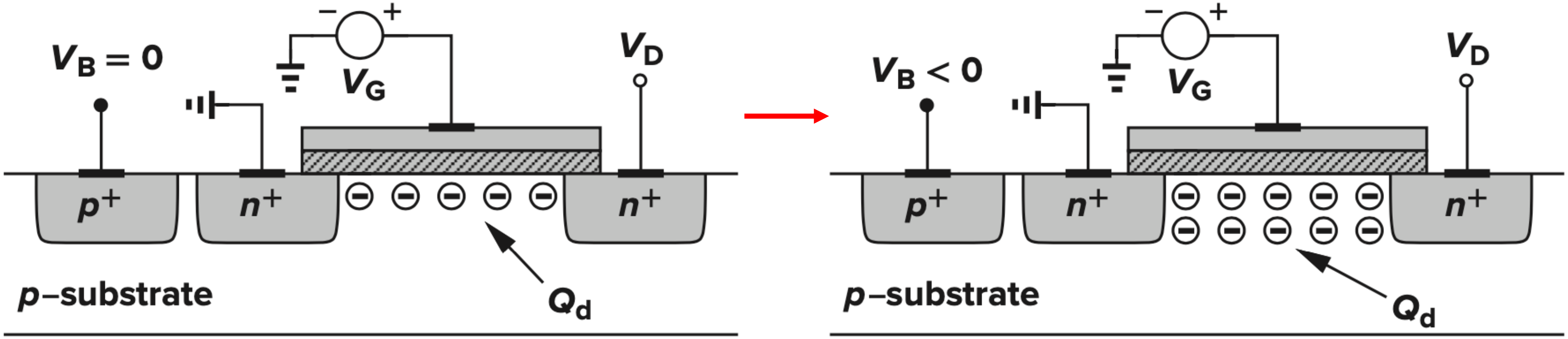
- **g_m drops in the triode region**
- so we usually employ MOS in **saturation** for **amplifying** the signal

Second-order effects: 1) Body Effect

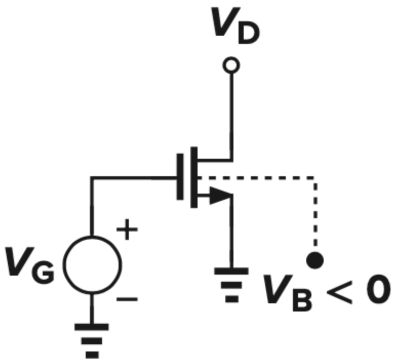




Second-order effects: 1) Body Effect



- As V_B becomes more negative, threshold voltage increases:



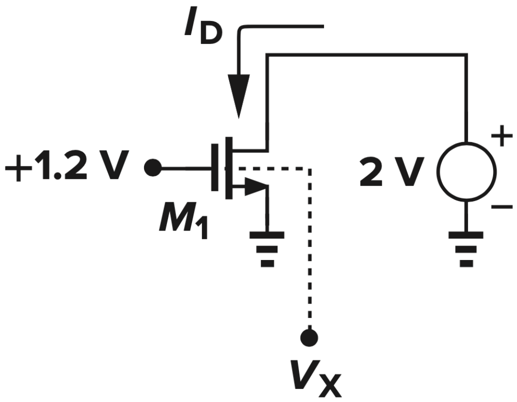
$$\boxed{V_{TH} = V_{TH0} + \gamma\left(\sqrt{2\Phi_F + V_{SB}} - \sqrt{|2\Phi_F|}\right)}$$

$$\gamma = \sqrt{2q\epsilon_{si}N_{sub}}/C_{ox}$$

V_{SB} is the source-bulk potential difference

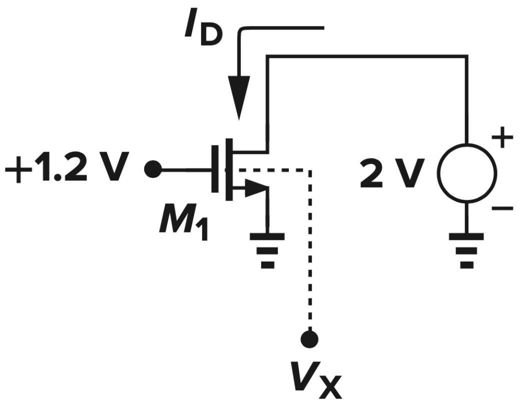
Example: Body Effect

- Plot the drain current if V_X varies from $-\infty$ to 0. Assume $V_{TH0} = 0.3\text{V}$, $\gamma = 0.4\text{V}^{1/2}$, and $2\Phi_F = 0.7\text{V}$.



Example: Body Effect

- Plot the drain current if V_X varies from $-\infty$ to 0. Assume $V_{TH0} = 0.3\text{V}$, $\gamma = 0.4\text{V}^{1/2}$, and $2\Phi_F = 0.7\text{V}$.

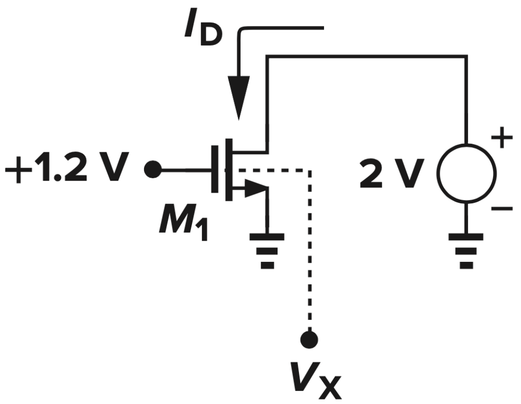


$$\boxed{V_{TH} = V_{TH0} + \gamma\left(\sqrt{2\Phi_F + V_{SB}} - \sqrt{|2\Phi_F|}\right)}$$

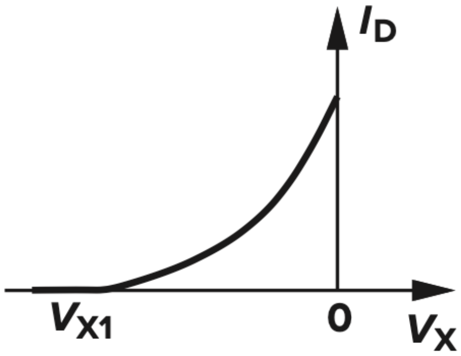
$$1.2\ \text{V} = 0.3 + 0.4\left(\sqrt{0.7 - V_{X1}} - \sqrt{0.7}\right) \longrightarrow V_{X1} = -8.83\ \text{V}$$

Example: Body Effect

- Plot the drain current if V_X varies from $-\infty$ to 0. Assume $V_{TH0} = 0.3\text{V}$, $\gamma = 0.4\text{V}^{1/2}$, and $2\Phi_F = 0.7\text{V}$.



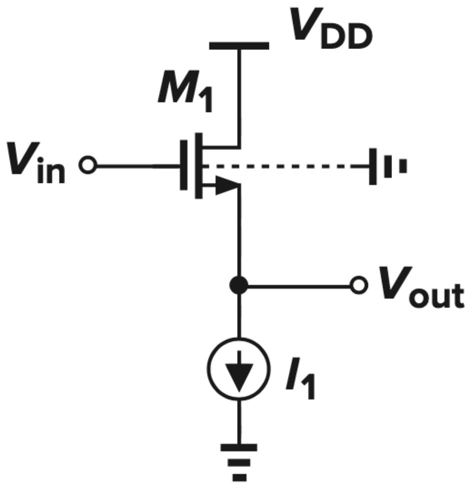
$$1.2\ \text{V} = 0.3 + 0.4\left(\sqrt{0.7 - V_{X1}} - \sqrt{0.7}\right)$$

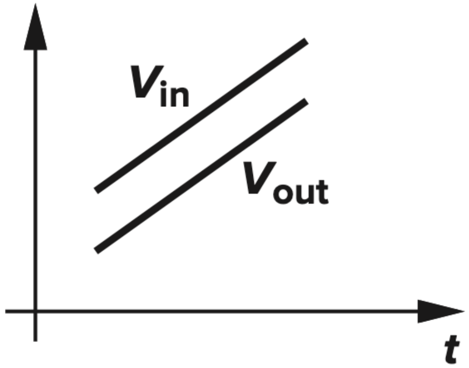


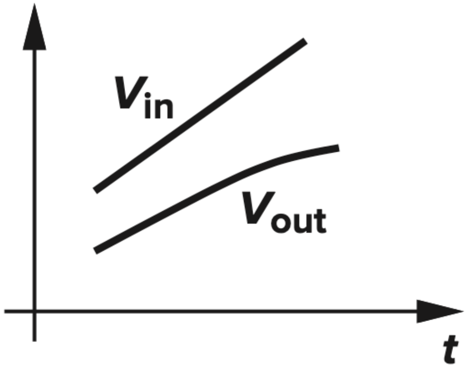
for $V_{X1} < V_X < 0$

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left[V_{GS} - V_{TH0} - \gamma\left(\sqrt{2\Phi_F - V_X} - \sqrt{2\Phi_F}\right)\right]^2$$

Body Effect







$$I_1 = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{in} - V_{out} - V_{TH})^2$$

$$\boxed{V_{TH} = V_{TH0} + \gamma\left(\sqrt{2\Phi_F + V_{SB}} - \sqrt{|2\Phi_F|}\right)}$$

- if V_{SB} becomes negative, V_{TH} decreases

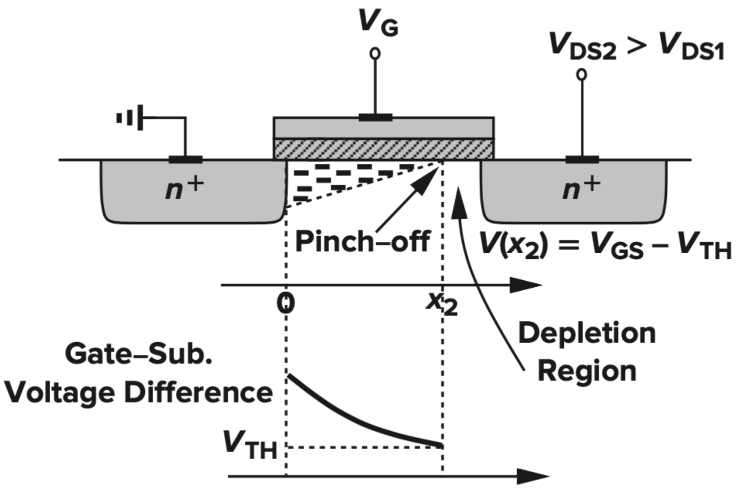
Second-order effects: 2) Channel-Length Modulation

- The actual length of the channel gradually decreases as the potential difference between the gate and the drain decreases >> **L′ is a function of V_{DS}**

$$L' = L - \Delta L$$

$$1/L' \approx (1 + \Delta L/L)/L$$

$$\boxed{\Delta L/L = \lambda V_{DS}}$$



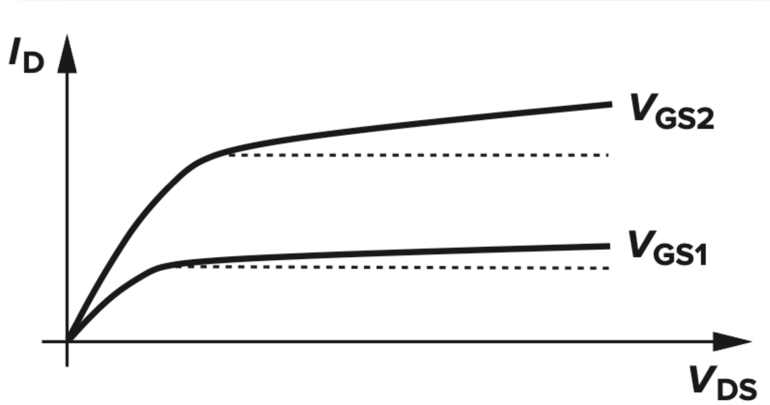
Second-order effects: 2) Channel-Length Modulation

- The actual length of the channel gradually decreases as the potential difference between the gate and the drain decreases >> **L′ is a function of V_{DS}**

$$L' = L - \Delta L$$

$$\boxed{\Delta L / L = \lambda V_{DS}}$$

$$1/L' \approx (1 + \Delta L/L)/L$$



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

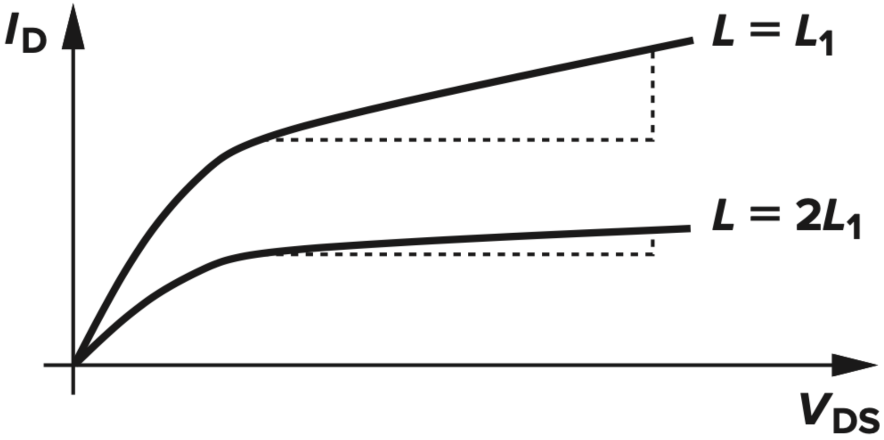
$$\downarrow$$

$$\boxed{I_D \approx \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})^2 \underbrace{(1 + \lambda V_{DS})}_{}}$$

$$\boxed{\begin{aligned} g_m &= \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})(1 + \lambda V_{DS}) \\ &= \sqrt{2\mu_n C_{ox}(W/L) I_D (1 + \lambda V_{DS})} \end{aligned}}$$

Channel-Length Modulation

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



Second-order effects: 3) Subthreshold Conduction

- In reality, for $V_{GS} \approx V_{TH}$, a “weak” inversion layer exists and some current flows from D to S. Even for $V_{GS} < V_{TH}$, I_D exhibits an exponential dependence on V_{GS}

Second-order effects: 3) Subthreshold Conduction

- In reality, for $V_{GS} \approx V_{TH}$, a “weak” inversion layer exists and some current flows from D to S. Even for $V_{GS} < V_{TH}$, I_D exhibits an exponential dependence on V_{GS}

- This “subthreshold conduction” can be formulated for roughly $V_{DS} > 100$ mV as:

$$I_D = I_0 \exp \frac{V_{GS}}{\xi V_T}$$

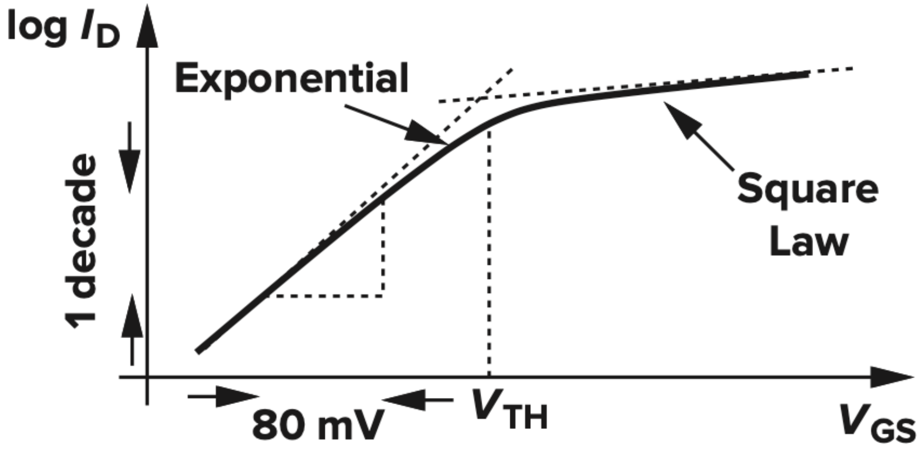
where I_0 is proportional to W/L, $\xi > 1$ is a nonideality factor, and $V_T = kT/q$

- The device operates in “**weak inversion**” (for $\mathbf{V_{GS}>V_{TH}}$ in **strong inversion**)

- Except for ξ, **similar to the exponential $\mathbf{I_C/V_{BE}}$ of a BJT**

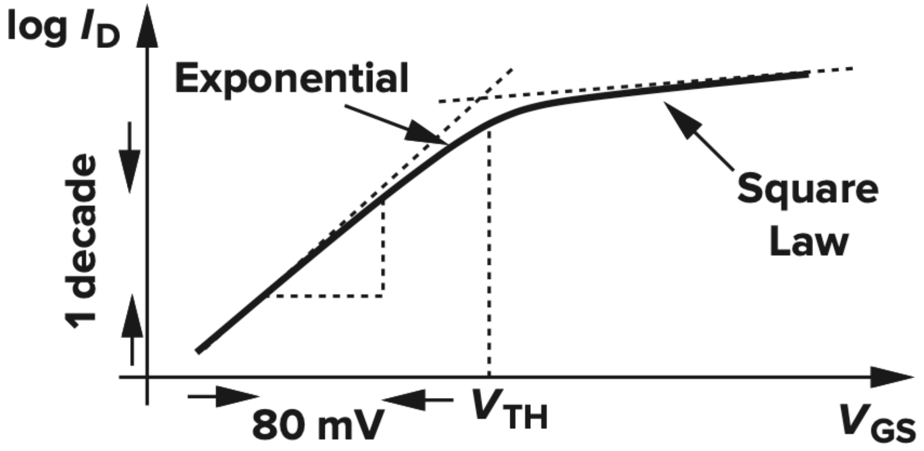
Subthreshold Conduction

- Observation: V_{GS} must decrease by **~80 mV** for I_D to decrease by **one decade**



Subthreshold Conduction

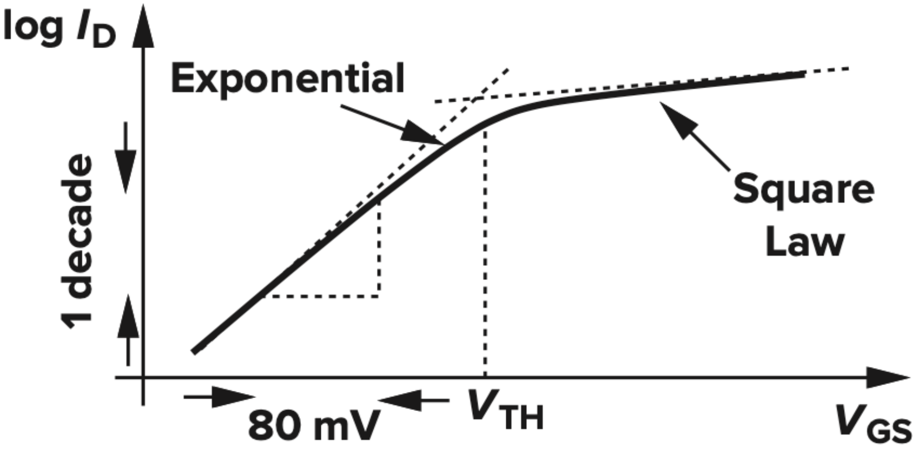
- Observation: V_{GS} must decrease by **~80 mV** for I_D to decrease by **one decade**



$$I_D = I_0 \exp \frac{V_{GS}}{\xi V_T} \longrightarrow g_m = I_D/(\xi V_T)$$

Subthreshold Conduction

- Observation: V_{GS} must decrease by **~80 mV** for I_D to decrease by **one decade**



$$\frac{I_D}{\xi V_T} = \frac{2I_D}{(V_{GS} - V_{TH})_1}$$

transition point?

$$(V_{GS} - V_{TH})_1 = 2\xi V_T$$

$$I_D = I_0 \exp \frac{V_{GS}}{\xi V_T} \longrightarrow g_m = I_D/(\xi V_T)$$

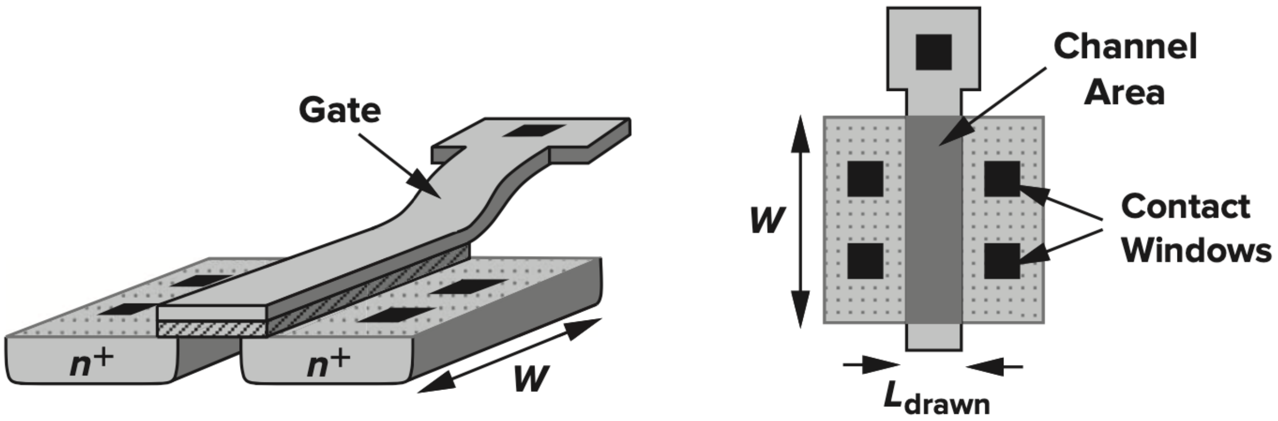
- Subthreshold operation only with a **large** device **width** or **low drain current**: the **speed** of subthreshold circuits severely limited

MOS Device Layout

- The **gate** polysilicon and the **source** and **drain** terminals must be tied to **metal** (aluminum) wires as **interconnects** with low resistance and capacitance
- One or more “contact windows” opened in each region, filled with metal, and connected to the upper metal wires
- To minimize the capacitance of S and D, the area of each junction minimized

MOS Device Layout

- The **gate** polysilicon and the **source** and **drain** terminals must be tied to **metal** (aluminum) wires as **interconnects** with low resistance and capacitance
- One or more “contact windows” opened in each region, filled with metal, and connected to the upper metal wires
- To minimize the capacitance of S and D, the area of each junction minimized



MOS Device Layout

- The layout determined by both the electrical properties of the device and the "**design rules**" imposed by the **technology**
- ***W*/*L*** is chosen to set the transconductance or other circuit parameters while the **minimum *L*** is dictated by the process
- In addition to the gate, the source and drain areas must be defined properly