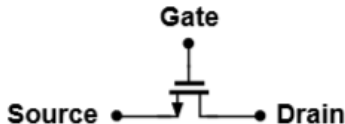


Chapter 2: MOS Device Physics

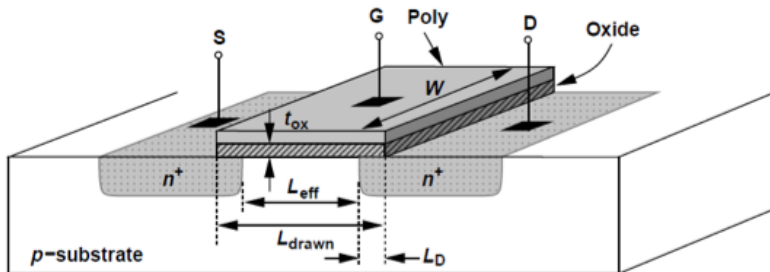
- 2.1 General Considerations**
- 2.2 MOS I/V Characteristics**
- 2.3 Second-Order Effects**
- 2.4 MOS Device Models**
- 2.5 Appendix A: FinFETs**
- 2.6 Appendix B: Behavior of MOS Device as a Capacitor**

MOSFET as a Switch



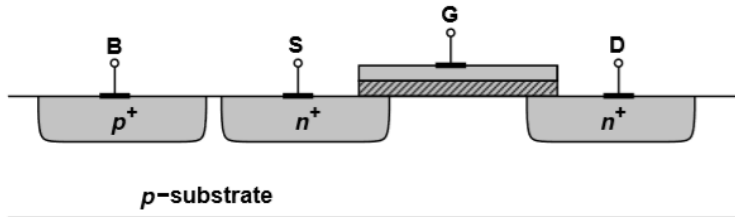
- **When gate voltage is high, device is on.**
- **Source and drain are interchangeable.**
- **But,**
 - **At what gate voltage does the device turn on?**
 - **How much is the resistance between S and D?**
 - **What limits the speed of the device?**

MOSFET Structure



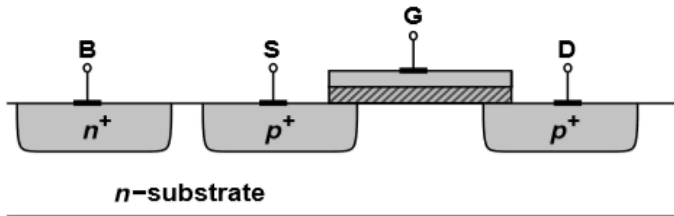
- n -type MOS (NMOS) has n -doped source (S) and drain (D) on p -type substrate (“bulk” or “body”).
- S/D junctions “side-diffuse” during fabrication so that effective length $L_{eff} = L_{drawn} - 2L_D$.
- Typical values are $L_{eff} \approx 10$ nm and $t_{ox} \approx 15$ Å.
- The S terminal provides charge carriers and the D terminal collects them.
- As voltages at the three terminals changes, the source and drain may exchange roles.

MOSFET Structure



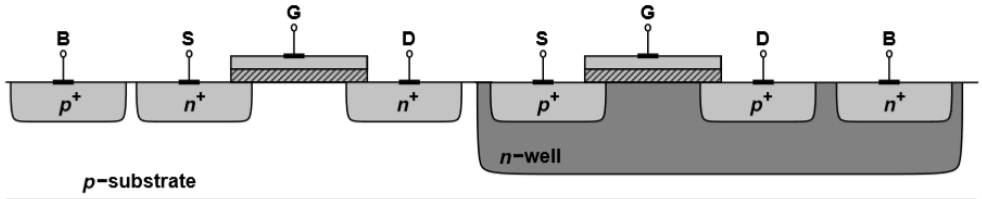
- MOSFETs actually have *four* terminals.
- Substrate potential greatly influences device characteristics.
- Typically S/D junction diodes are reversed-biased and the NMOS substrate is connected to the most negative supply in the system.

MOSFET Structure



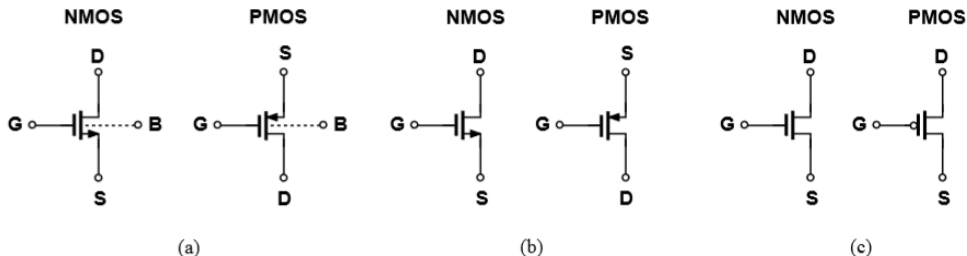
- **PMOS** is obtained by inverting all of the doping types (including the substrate).

MOSFET Structure



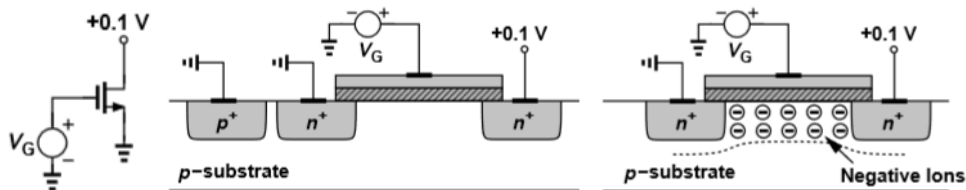
- In complementary MOS (CMOS) technologies both NMOS (NFET) and PMOS (PFET) are needed and fabricated on the same wafer.
- In today's CMOS, the PMOS is fabricated in an n -well, where the n -well is tied to the most positive supply voltage.

MOS Symbols



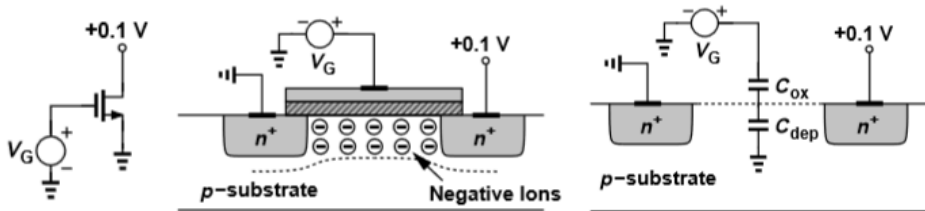
- Substrate is denoted by “B” (bulk).
- PMOS source is positioned on top since it has a higher potential than the gate.
- Most circuits have NMOS and PMOS bulk tied to ground and VDD , respectively, so we tend to omit the connections (b,c).
- Digital circuits tend to incorporate “switch” symbols (c).

Threshold Voltage



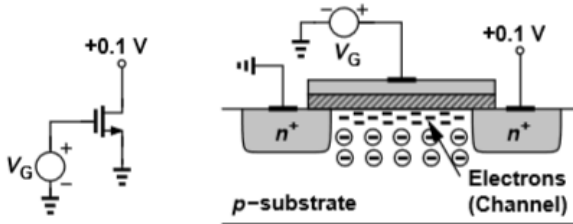
- As V_G increases from zero, holes in p -substrate are repelled leaving negative ions behind to form a depletion region.
- There are no charge carriers, so no current flow.

Threshold Voltage



- Increasing V_G further increases the width of the depletion region and the potential at the oxide-silicon interface.
- Structure resembles voltage divider consisting of gate-oxide capacitor and depletion region capacitor in series.

Threshold Voltage



- When interface potential reaches sufficiently positive value, electrons flow from the source to the interface and eventually to the drain.
- This creates a channel of charge carriers (inversion layer) beneath the gate oxide.
- The value of V_G at which the inversion layer occurs is the threshold voltage (V_{TH}).

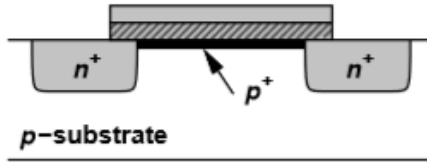
Threshold Voltage

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

$$\Phi_F = (kT/q) \ln(N_{sub}/n_i), \quad Q_{dep} = \sqrt{4q\epsilon_{si}|\Phi_F|N_{sub}}.$$

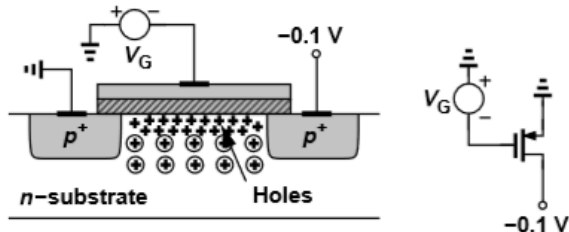
- **Where**
 - **Φ_{MS}** is the difference between the work functions of the polysilicon gate and the silicon substrate.
 - **k** is Boltzmann's constant.
 - **q** is the electron charge.
 - **N_{sub}** is the doping density of the substrate.
 - **n_i** is the density of electrons in undoped silicon.
 - **Q_{dep}** is the charge in the depletion region.
 - **C_{ox}** is the gate oxide capacitance per unit area.
 - **ϵ_{si}** is the dielectric constant of silicon.

Threshold Voltage



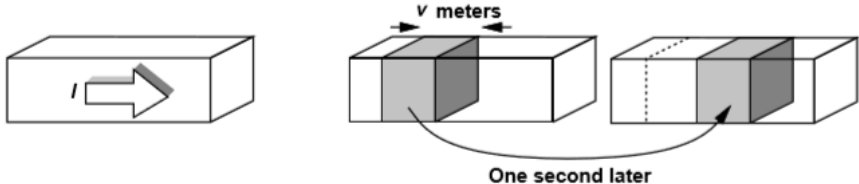
- In practice, threshold voltage is adjusted by implanting dopants into the channel area during device fabrication.
- For NMOS, adding a thin sheet of p^+ increases the gate voltage necessary to deplete the region.

Threshold Voltage



- Turn-on phenomena in PMOS is similar to that of NMOS but with all polarities reversed.
- If the gate-source voltage becomes sufficiently *negative*, an inversion layer consisting of holes is formed at the oxide-silicon interface, providing a conduction path between source and drain.
- PMOS threshold voltage is negative.

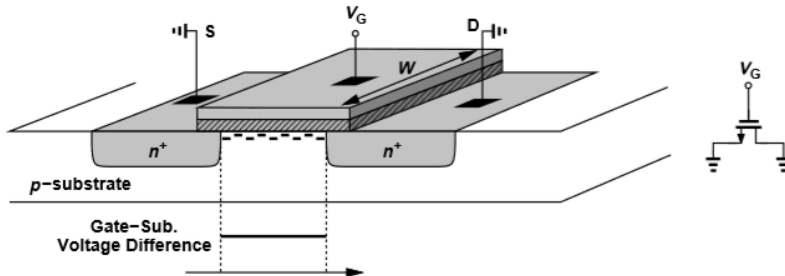
Derivation of I/V Characteristics



$$I = Q_d \cdot v$$

- **Where**
 - Q_d is the mobile charge density along the direction of current I .
 - v is the charge velocity.

Derivation of I/V Characteristics

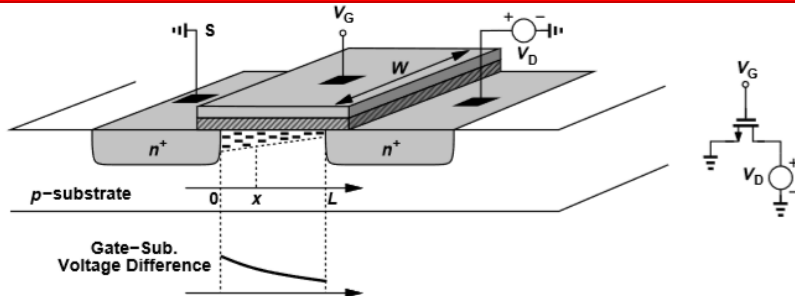


- Onset of inversion occurs at $V_{GS} = V_{TH}$.
- Inversion charge density produced by gate oxide capacitance is proportional to $V_{GS} - V_{TH}$ since for $V_{GS} \geq V_{TH}$, charge placed on the gate must be mirrored by charge in the channel, yielding a uniform channel charge density:

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

- Where WC_{ox} is the total capacitance per unit length.

Derivation of I/V Characteristics



- Channel potential varies from zero at the source to V_D at the drain.
- Local voltage *difference* between the gate and the channel varies from V_G to V_G - V_D.
- Charge density now varies with respect to x :

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

where $V(x)$ is the channel potential at x.

Derivation of I/V Characteristics

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

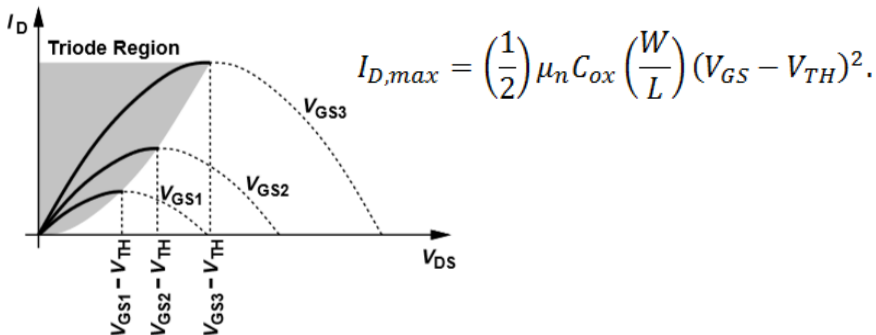
- **Since**

- $I = Q_d \cdot v.$
- $v = \mu E.$
- $E(x) = -dV/dx.$
- $Q_d(x) = W C_{ox} [V_{GS} - V(x) - V_{TH}].$
- $I_D = W C_{ox} [V_{GS} - V(x) - V_{TH}] \mu_n (dV(x)/dx).$
- $$\int_{x=0}^L I_D dx = \int_{V=0}^{V_{DS}} W C_{ox} \mu_n [V_{GS} - V(x) - V_{TH}] dV.$$

- **A negative sign is added because the charge carriers are negative for NMOS.**

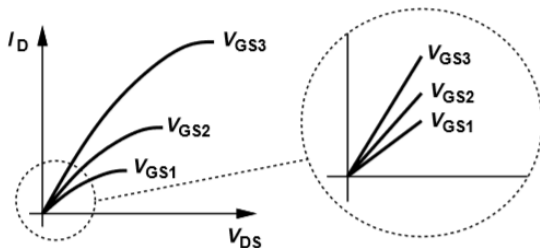
Derivation of I/V Characteristics

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



- $V_{GS} - V_{TH}$ is known as the “overdrive voltage.”
- W/L is known as the “aspect ratio.”
- If $V_{DS} \leq V_{GS} - V_{TH}$, we say the device is operating in the “triode region.”

Derivation of I/V Characteristics



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

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

- In this case, the drain current is a linear function of V_{DS} so the path from source to drain can be represented by a linear resistor:

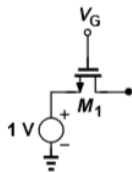
$$R_{on} = \frac{1}{\mu_n C_{ox} \left(\frac{W}{L} \right) (V_{GS} - V_{TH})}.$$

Derivation of I/V Characteristics



- If $V_{DS} \ll 2(V_{GS} - V_{TH})$, the device is operating in “deep triode region.”
- In this region, a MOSFET can operate as a resistor whose value is controlled by the overdrive voltage.
- Unlike bipolar transistors, a MOS device may be on even if it carries no current.

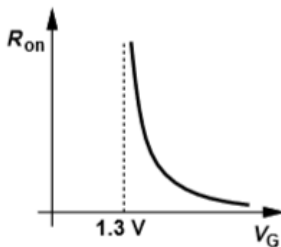
Derivation of I/V Characteristics



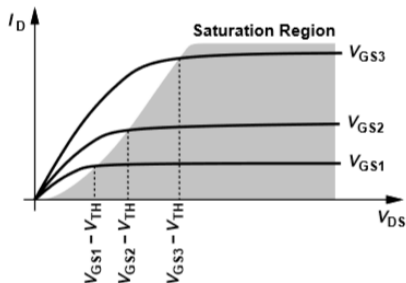
For example, given the topology on the left and that

- $\mu_n C_{ox} = 50 \mu\text{A}/\text{V}^2$
- $W/L = 10$
- $V_{TH} = 0.3 \text{ V}$

$$R_{on} = \frac{1}{50 \mu\text{A}/\text{V}^2 \times 10 (V_G - 1 \text{ V} - 0.3 \text{ V})}.$$

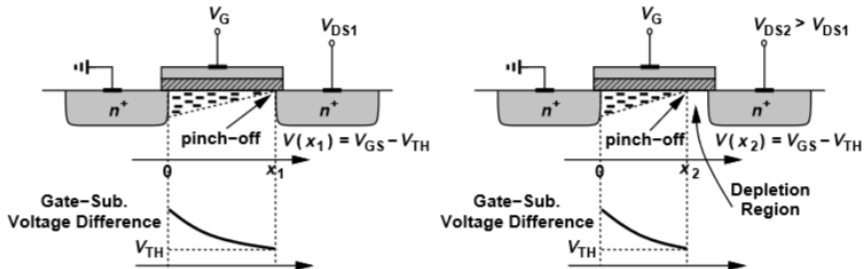


Derivation of I/V Characteristics



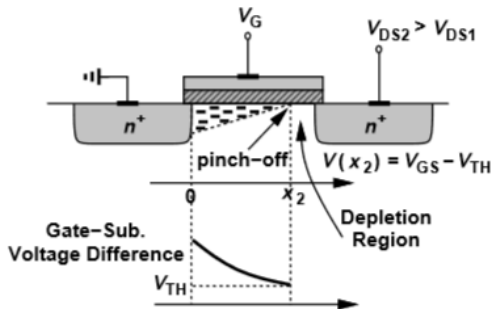
- In reality, if $V_{DS} > V_{GS} - V_{TH}$, I_D becomes relatively constant and we say that the device operates in “saturation region.”
- $V_{D,sat} = V_{GS} - V_{TH}$ denotes the minimum V_{DS} necessary for operation in saturation.

Derivation of I/V Characteristics



- If V_{DS} is slightly larger than $V_{GS} - V_{TH}$, the inversion layer stops at $x \leq L$, and the channel becomes “pinched off.”
- As V_{DS} increases, the point at which Q_D equals zero gradually moves towards the source.
- At some point along the channel, the local potential difference between the gate and the oxide-silicon interface is not sufficient to support an inversion layer.

Derivation of I/V Characteristics



- Electron velocity ($v = I/Q_d$) rises tremendously as they approach the pinch-off point (where $Q_d \rightarrow 0$) and shoot through the depletion region near the drain junction and arrive at the drain terminal.

Derivation of I/V Characteristics

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

- **Since the integral becomes**

$$\int_{x=0}^{x=x_2=L'} I_D dx = \int_{V=0}^{V=V_{GS}-V_{TH}} W C_{ox} \mu_n [V_{GS} - V(x) - V_{TH}] dV .$$

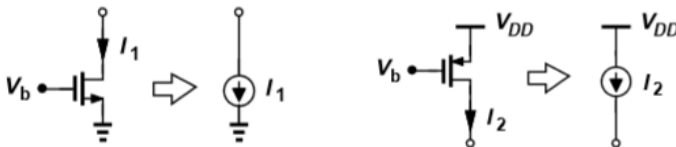
$$I_D = \frac{1}{2} \mu_n C_{ox} \left(\frac{W}{L'} \right) (V_{GS} - V_{TH})^2 .$$

- **ID is relatively independent of VDS if L' remains close to L.**
- **The device exhibits a “square-law” behavior.**

Derivation of I/V Characteristics

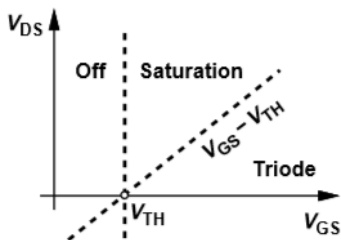
- **For PMOS devices, the equations become**
 - $I_D = -\mu_p C_{ox} \frac{W}{L} \left[(V_{GS} - V_{TH}) V_{DS} - \frac{1}{2} V_{DS}^2 \right]$.
 - $I_D = -\frac{1}{2} \mu_p C_{ox} \frac{W}{L'} (V_{GS} - V_{TH})^2$.
- **The negative sign shows up due to the assumption that drain current flows from drain to source, whereas holes in a PMOS flow in the reverse direction.**
- **V_{GS}, V_{DS}, V_{TH}, and V_{GS}-V_{TH} are negative for a PMOS transistor that is turned on.**
- **Since the mobility of holes is about 1/2 the mobility of electrons, PMOS devices suffer from lower “current drive” capability.**

Derivation of I/V Characteristics

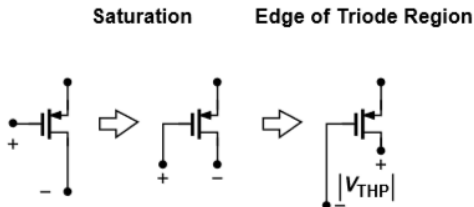
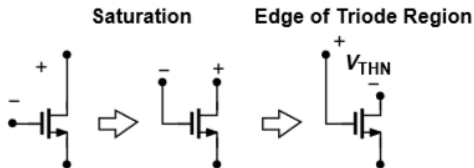


- A saturated MOSFET can be used as a current source connected between the drain and the source.
- NMOS current sources inject current into ground while PMOS current sources draws current from V_{DD} .

Derivation of I/V Characteristics



- $V_{DS} = V_{GS} - V_{TH} = V_{D,sat}$ is the line between saturation and triode region.
- For a given V_{DS} , the device eventually leaves saturation as V_{GS} increases.
- The drain is defined as the terminal with a higher (lower) voltage than the source for an NMOS (PMOS).

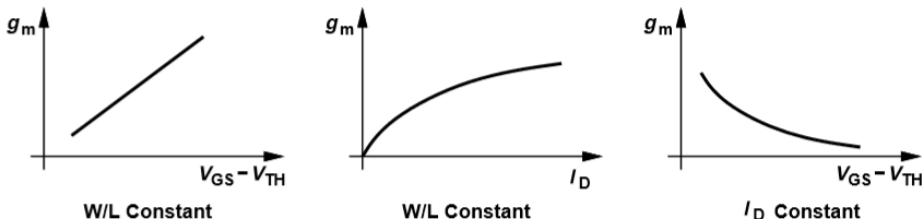


MOS Transconductance

$$\begin{aligned} g_m &= \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS} \text{ const.}} = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH}) \\ &= \sqrt{2\mu_n C_{ox} \frac{W}{L} I_D} = \frac{2I_D}{V_{GS} - V_{TH}}. \end{aligned}$$

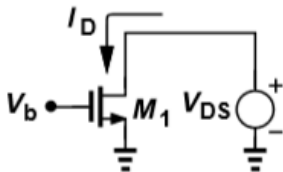
- **Transconductance (usually defined in the saturation region) is defined as the change in drain current divided by the change in the gate-source voltage.**
- **gm represents the sensitivity of the device since a high value implies a small change in VGS will result in a large change in ID.**
- **Transconductance in saturation region is equal to the inverse of Ron in the deep triode region.**

MOS Transconductance



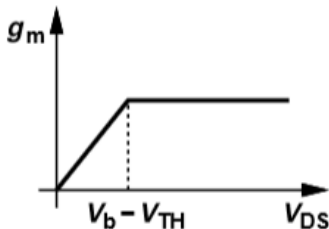
- Each expression for transconductance is useful in studying its behavior.
- Drain current and overdrive voltage are *bias* values.
- If a small signal is applied to a device with defined bias values, we assume the signal amplitude is small enough that the variation in transconductance is negligible.

MOS Transconductance



- To find the transconductance for the topology on the left with respect to V_{DS} ,
 - So long as $V_{DS} \geq V_b - V_{TH}$, M_1 is in saturation, so I_D is relatively constant, and therefore so is g_m .
 - When M_1 enters triode region ($V_{DS} < V_b - V_{TH}$),

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

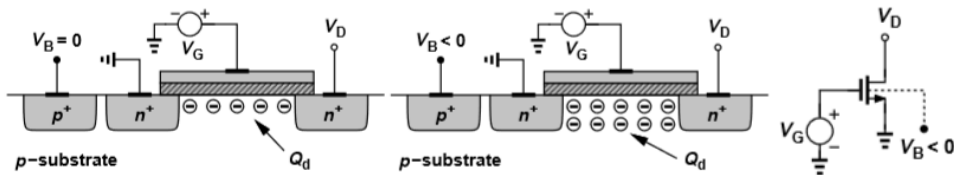


MOS Transconductance

- **For PMOS,**

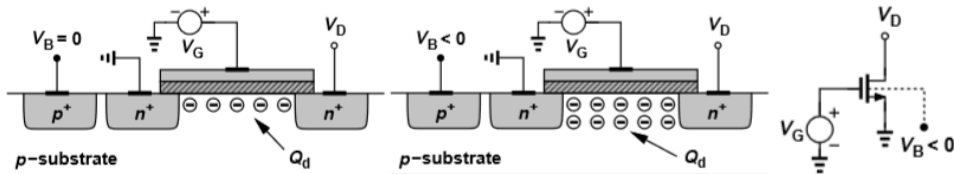
$$\begin{aligned}g_m &= -\mu_p C_{ox}(W/L)(V_{GS} - V_{TH}) \\&= -2I_D/(V_{GS} - V_{TH}) \\&= \sqrt{2\mu_p C_{ox}(W/L)I_D}.\end{aligned}$$

Second-Order Effects



- Originally, with the bulk of an NMOS tied to ground, the threshold voltage was defined as
$$V_{TH} = \Phi_{MS} + 2\Phi_F + \frac{Q_{dep}}{C_{ox}}.$$
- Decreasing the bulk voltage (V_B) increases the number of holes attracted to the substrate connection, which leaves a larger negative charge behind and makes the depletion region wider, increasing Q_d and thus increasing V_{TH} .
- This is known as the “body effect” or “back-gate effect.”

Second-Order Effects



- With body effect, the expression which characterizes the dependence of threshold voltage on the bulk voltage is

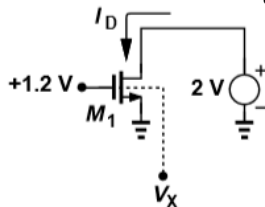
$$V_{TH} = V_{TH0} + \gamma \left(\sqrt{2\Phi_F + V_{SB}} - \sqrt{|2\Phi_F|} \right).$$

- Where,

$$V_{TH0} = \Phi_{MS} + 2\Phi_F + \frac{Q_{dep}}{C_{ox}}.$$

$$\gamma = \sqrt{2q\epsilon_{si}N_{sub}/C_{ox}} \text{ denotes the body effect coefficient.}$$

Second-Order Effects



- For example, let's find the drain current as bulk voltage varies from negative infinity to 0 given the topology on the left and that

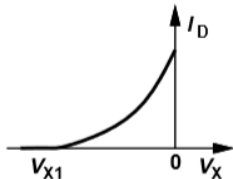
- $V_{TH0} = 0.3 \text{ V}$.
- $\gamma = 0.4 \text{ V}^{1/2}$.
- $2\Phi_F = 0.7 \text{ V}$.

- If V_X is sufficiently negative, V_{TH} of M_1 exceeds 1.2 V and the device is off, therefore , where

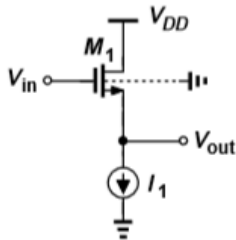
$$V_{X1} < V_X < 0$$

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

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



Second-Order Effects

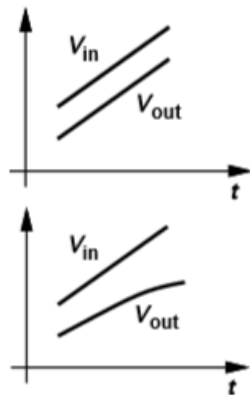


Body effect manifests itself whenever the source voltage varies with respect to the bulk potential.

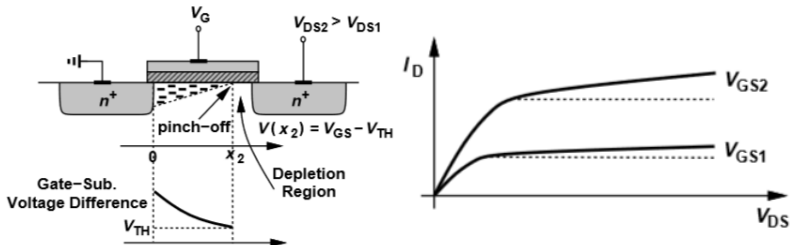
Given the topology on the left and first ignoring body effect, as V_{in} varies, V_{out} follows the input because the drain current remains equal to I_1 , where

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

- **With body effect, as $V_{in,out}$ become more positive, V_{SB} increases, which increases V_{TH} and thus $V_{in} - V_{out}$ must increase to maintain a constant I_D .**



Second-Order Effects



- Originally, when the device was in saturation region, drain current was characterized by

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

- The actual length of the channel ($L' = L - \Delta L$) is a function of V_{DS} , which is an effect called “channel length modulation.”
- $1/L' \approx (1 + \Delta L/L)/L$, and $\Delta L/L = \lambda V_{DS}$, where λ is the channel-length modulation coefficient” gives us

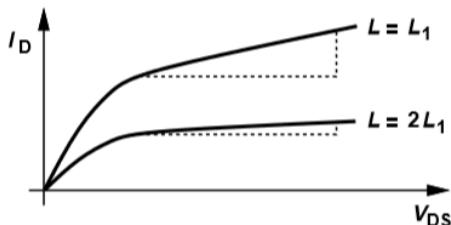
$$I_D \approx \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS}).$$

Second-Order Effects

- With the effect of channel length modulation, the expressions derived for transconductance of the device that need modification are

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

Second-Order Effects



- **Knowing that**

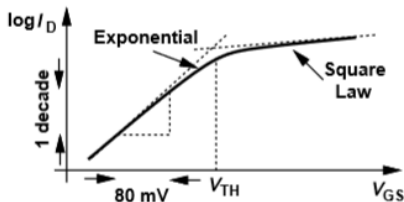
- $I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS}),$
- $\lambda \propto 1/L,$

and keeping all other parameters constant, we can see that if the length L is doubled, the slope of I_D vs. V_{DS} is divided by *four*.

- **This is due to**

$$\partial I_D / \partial V_{DS} \propto \lambda / L \propto 1/L^2$$

Second-Order Effects



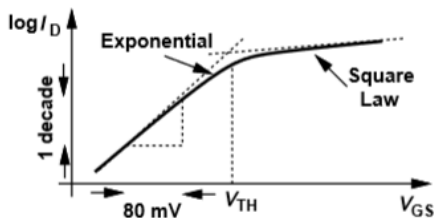
- MOSFETs do not turn off abruptly when $V_{GS} < V_{TH}$, but actually there is a “weak” inversion layer that exists and finite current flows from drain to source with an exponential dependence on V_{GS} .
- This effect is called “subthreshold conduction.”
- When $V_{GS} < V_{TH}$,

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

where $I_0 \propto W/L$, ξ is the ideality factor, and $V_T = kT/q$.

- Here the device operates in “weak inversion.”

Second-Order Effects



- To examine MOSFET behavior as the drain “current density,” I_D/W varies, we must consider the equations for both strong and weak inversion:

- $$I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 \Rightarrow V_{GS} = \sqrt{\frac{2I_D}{\mu_n C_{ox} W/L}} + V_{TH}.$$

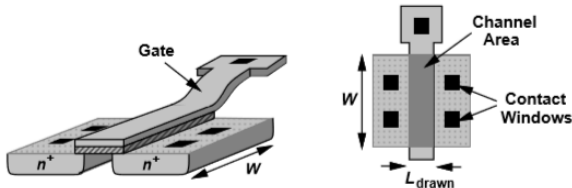
- $$I_D = \alpha \frac{W}{L} \exp \frac{V_{GS}}{\xi V_T} \Rightarrow V_{GS} = \xi V_T \ln \frac{I_D}{\alpha W/L}.$$

- For a given current and w/L , we must obtain V_{GS} from both expressions and select the lower value.
- If I_D remains constant and W increases, V_{GS} falls and the device goes from strong to weak inversion.

Second-Order Effects

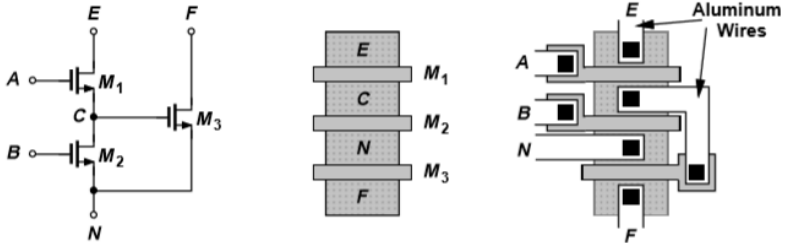
- **MOSFETs experience undesirable effects if terminal voltage differences exceed certain limits, e.g.**
 - **If V_{GS} is too high, the gate oxide breaks down irreversibly, damaging the transistor.**
 - **In short channel devices, excessively large V_{DS} widens depletion region around the drain so that it touches the depletion region around the source, creating a very large drain current (an effect called “punchthrough”).**

MOS Device Layout



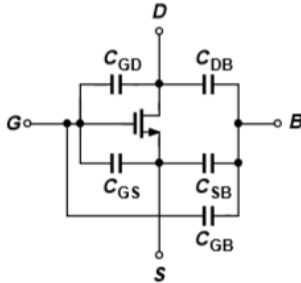
- The gate polysilicon and the source and drain terminals must be tied to metal (aluminum) wires that serve as interconnects with low resistance and capacitance.
- This is accomplished with “contact windows” which are filled with metal and connected to the upper metal wires.
- To minimize the capacitance of the source and drain, the total area of each junction must be minimized.

MOS Device Layout



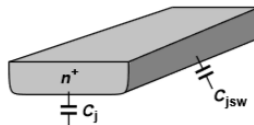
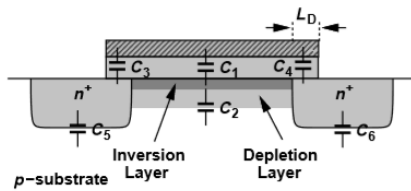
- Since M_1 and M_2 share the same S/D junctions at node C and M_2 and M_3 do the same at node N , we can lay them out as shown above.
- Since the gate polysilicon of M_3 cannot be directly tied to the source material of M_1 , a metal interconnect is necessary.

MOS Device Capacitances



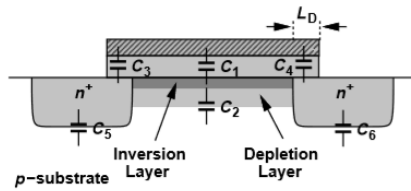
- To better predict high-frequency behavior, it is necessary to consider device capacitances.
- Capacitance exists between every two of the four terminals, and their values depend on the bias conditions of the transistor.

MOS Device Capacitances



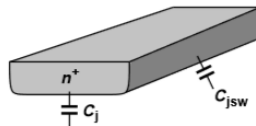
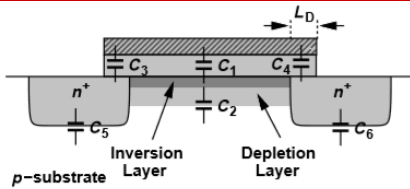
- **Capacitances include**
 - The oxide capacitance between the gate and the channel C_1 .
 - The depletion capacitance between the channel and the substrate C_2 .
 - The capacitance due to the overlap of the gate poly with the source and drain areas C_3 and C_4 .
 - The junction capacitance between the source/drain areas and the substrate C_5 and C_6 .

MOS Device Capacitances



- $C_1 = WLC_{ox}$.
- $C_2 = WL\sqrt{q\epsilon_{si}N_{sub}/(4\Phi_F)}$.
- **Due to fringing electric fields, C_3 and C_4 cannot be written as $WLD C_{ox}$; rather we must find the overlap capacitance per unit width (C_{ov}) and multiply that value with W .**

MOS Device Capacitances



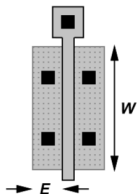
- **C5 and C6 are decomposed into two components:**
 - **The bottom-plate capacitance associated with the bottom of the junction, C_j:**

$$C_j = C_{j0} / [1 + V_R / (\Phi_B)]^m,$$

where **V_R** is the reverse voltage across the junction, **Φ_B** is the junction built-in potential, and **m** is a power typically in the range of 0.3 and 0.4.

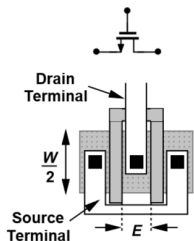
- **The sidewall capacitance due to the perimeter of the junction, C_{jsw}.**

MOS Device Capacitances



- For example calculating the source and drain junction capacitance of the topology on the left,

$$C_{DB} = C_{SB} = WEC_j + 2(W + E)C_{jsw}.$$



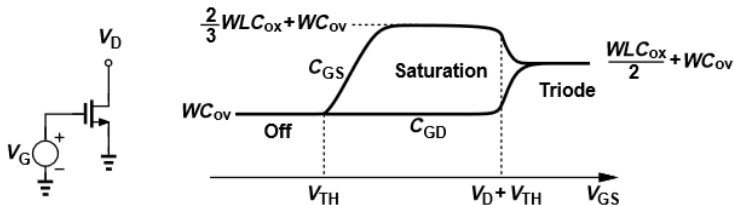
- Calculating the source and drain junction capacitance of the second topology on the left,

$$C_{DB} = \frac{W}{2}EC_j + 2\left(\frac{W}{2} + E\right)C_{jsw}$$

$$\begin{aligned} C_{SB} &= 2\left[\frac{W}{2}EC_j + 2\left(\frac{W}{2} + E\right)C_{jsw}\right] \\ &= WEC_j + 2(W + 2E)C_{jsw}. \end{aligned}$$

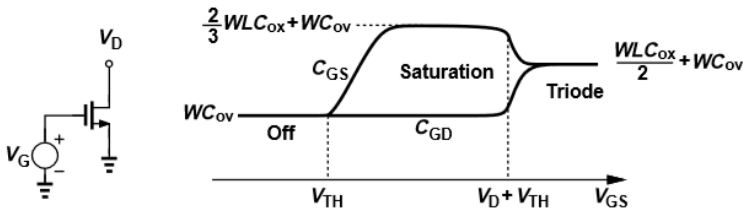
- We assumed the total source/drain perimeter $2(W+E)$ is multiplied by C_{jsw} .

MOS Device Capacitances



- Considering capacitance values in different operating regions, if the device is off,
 - $C_{GD} = C_{GS} = C_{ov}W$.
 - $C_{GB} = (WLC_{ox})C_d / (WLC_{ox} + C_d)$, where L is the effective length, $C_d = WL\sqrt{q\epsilon_{si}N_{sub}/(4\Phi_F)}$, and $\epsilon_{si} = \epsilon_{r,si} \times \epsilon_0 = 11.8 \times (8.85 \times 10^{-14}) \text{ F/cm}$.
 - The values of C_{SB} and C_{DB} are a function of the source and drain voltages with respect to the substrate.

MOS Device Capacitances

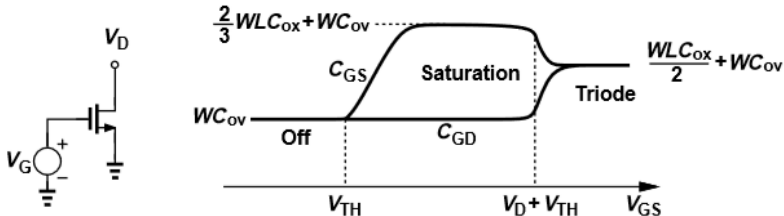


- In deep triode region, the source/drain have approximately equal voltages, so the gate-channel capacitance WLC_{ox} is divided equally between the gate-source terminals and the gate-drain terminals, which results in

$$C_{GD} = C_{GS} = WLC_{ox}/2 + WC_{ov}.$$

- CGB is usually neglected in triode and saturation regions because the inversion layer acts as a “shield” between the gate and the bulk, so if V_G varies, the charge is supplied by the source/drain rather than the bulk.

MOS Device Capacitances



- If the device is in saturation region, C_{GD} will be roughly equal to WC_{ov} .
- The varying potential difference between gate and channel cause nonuniform vertical electric field in the gate oxide while going from source to drain, which results in C_{GS} being

$$C_{GS} = 2WL_{eff}C_{ox}/3 + WC_{ov}.$$

MOS Device Capacitances



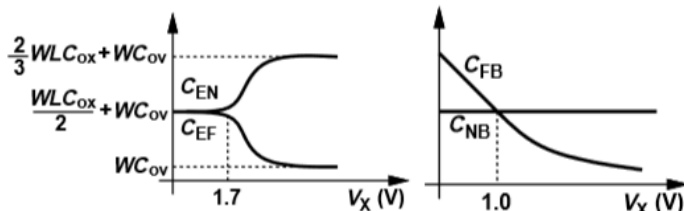
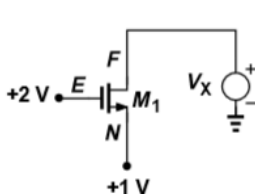
- Lets sketch the capacitance of M1 on the left as V_X varies from 0 to 3 V assuming $V_{TH} = 0.3$ V and $\lambda = \gamma = 0$.

- For $V_X \approx 0$, M1 is in triode region, so CFB is maximum, and

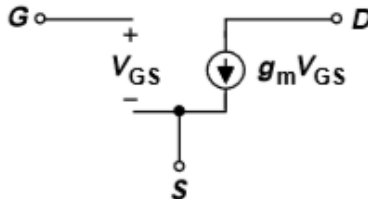
$$C_{EN} \approx C_{EF} = (1/2)WLC_{ox} + WC_{ov}.$$

- As V_X exceeds 1 V, the roles of the source/drain are exchanged, bringing M1 out of the triode region for $V_X \geq 2$ V – 0.3 V.

- C_{NB} is independent of V_X .

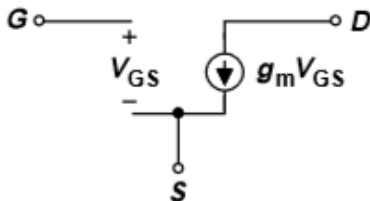


MOS Small-Signal Model



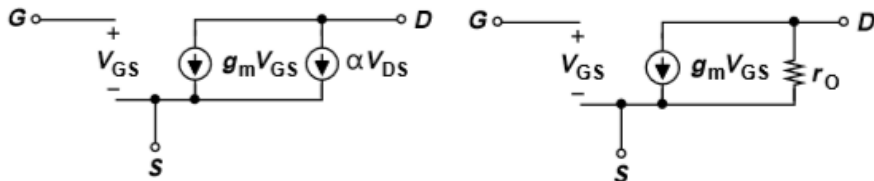
- If perturbation in bias conditions are small, a “small-signal” model can be used to simplify calculations (derived for saturation region).
- In order to derive the small-signal model, we
 - Apply certain bias voltages to the terminals of the device.
 - Increment the potential difference between two of the terminals while the other terminal voltages remain constant.
 - Measure the resulting change in all terminal currents.

MOS Small-Signal Model



- By changing the voltage between two terminals by $\Delta V = V_{GS}$ and then measuring a current change $\Delta I = g_m V_{GS}$, we can model the effect by a voltage-dependent current source.
- Above is the small-signal model of an ideal MOSFET.

MOS Small-Signal Model



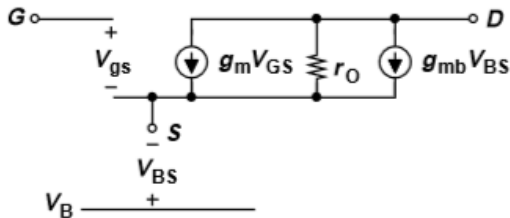
- Due to channel-length modulation, drain current also varies with V_{DS} , but a current source whose value linearly depends on the voltage across it is equivalent to a linear resistor:

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

• it is usually much larger than r_{DS}

$$\approx \frac{1 + \lambda V_{DS}}{\lambda I_D} \approx \frac{1}{\lambda I_D} \quad \text{in voltage gain or most}$$

MOS Small-Signal Model



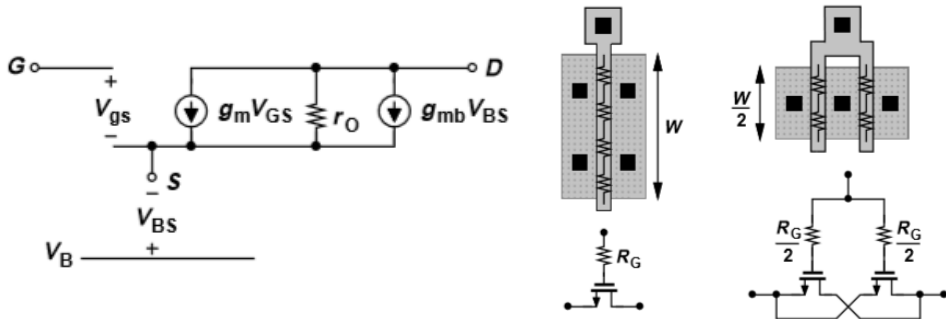
- Due to body effect, bulk potential influences V_{TH} and hence gate-source overdrive.
- With all other terminals held at a constant voltage, the bulk behaves as a second gate since the drain current is a function of the bulk voltage given by $g_{mb}V_{BS}$, where

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH}) \left(-\frac{\partial V_{TH}}{\partial V_{BS}} \right).$$

$$\frac{\partial V_{TH}}{\partial V_{BS}} = -\frac{\partial V_{TH}}{\partial V_{SB}} = -\frac{\gamma}{2} (2\Phi_F + V_{SB})^{-1/2}.$$

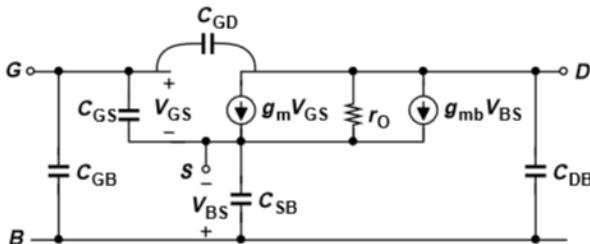
$$g_{mb} = g_m \frac{\gamma}{2\sqrt{2\Phi_F + V_{SB}}} = \eta g_m, \quad \eta = g_{mb}/g_m.$$

MOS Small-Signal Model



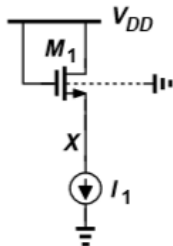
- The small-signal model above is adequate for most low-frequency small-signal analyses.
- In reality, each terminal exhibits a finite ohmic resistance due to resistivity of the material (and contacts), but proper layout can minimize these resistances.
 - Folding reduces the gate resistance by a factor of four.

MOS Small-Signal Model

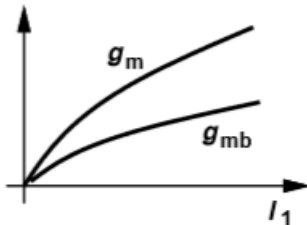


- The complete MOS small-signal model not only includes channel-length modulation and body effect, but also the capacitances between each terminal.

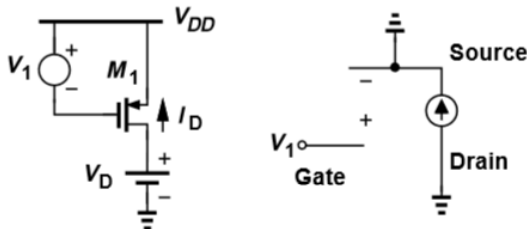
MOS Small-Signal Model



- To sketch g_m and g_{mb} of M_1 on the left as a function of bias current I_1 ,
 - $g_m = \sqrt{2\mu_n C_{ox}(W/L)I_D} \Rightarrow g_m \propto \sqrt{I_1}$.
 - g_{mb} dependence on I_1 is less straight forward, but as I_1 increases, V_X decreases and so does V_{SB} .



MOS Small-Signal Model



- The derivation of the small-signal model for PMOS yields the *exact* same model as for NMOS.
- The model shows the voltage-dependent current source pointing *upward*, giving the (wrong) impression that the direction of the current in the PMOS is opposite of that in NMOS.

MOS Spice Models

NMOS Model

LEVEL = 1	VTO = 0.7	GAMMA = 0.45	PHI = 0.9
NSUB = 9e+14	LD = 0.08e-6	UO = 350	LAMBDA = 0.1
TOX = 9e-9	PB = 0.9	CJ = 0.56e-3	CJSW = 0.35e-11
MJ = 0.45	MJSW = 0.2	CGDO = 0.4e-9	JS = 1.0e-8

PMOS Model

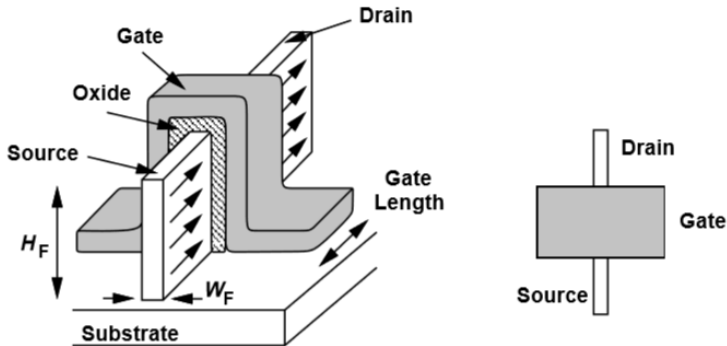
LEVEL = 1	VTO = -0.8	GAMMA = 0.4	PHI = 0.8
NSUB = 5e+14	LD = 0.09e-6	UO = 100	LAMBDA = 0.2
TOX = 9e-9	PB = 0.9	CJ = 0.94e-3	CJSW = 0.32e-11
MJ = 0.5	MJSW = 0.3	CGDO = 0.3e-9	JS = 0.5e-8

- **Simulators such as SPICE and Cadence need accurate models for each device.**
- **Above is the simplest MOS SPICE model, known as “Level 1,” and provide typical values for each parameter corresponding to 0.5-μm technology.**

NMOS versus PMOS Devices

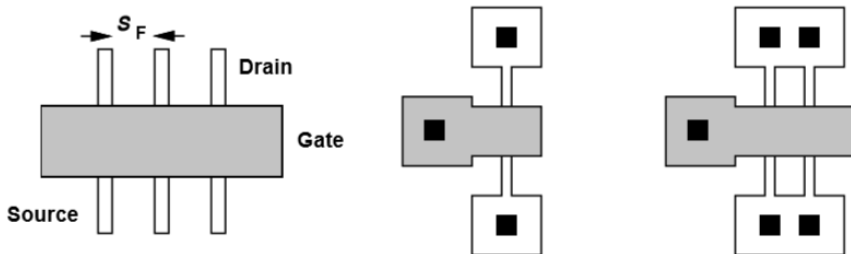
- **PMOS devices are quite inferior to NMOS in most CMOS technology.**
- **Lower mobility of holes ($\mu_p C_{ox} \approx 0.5 \mu_n C_{ox}$) yield lower current drive and conductance.**
- **NMOS exhibit higher output resistance, providing more ideal current sources and higher gain in amplifiers.**
- **As such is it preferred to incorporate NMOS rather than PMOS wherever possible.**

FinFETs



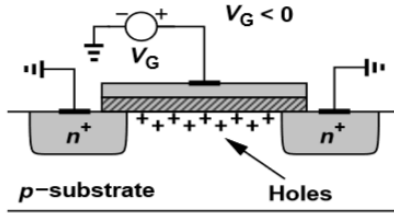
- FinFETs have three-dimensional geometry and exhibit superior performance as channel lengths fall below ~ 20 nm.
- Here, $W = W_F + 2H_F$, but since H_F is not under the circuit designer's control and W_F , impacts device imperfections, there are only discrete values for transistor width.

FinFETs



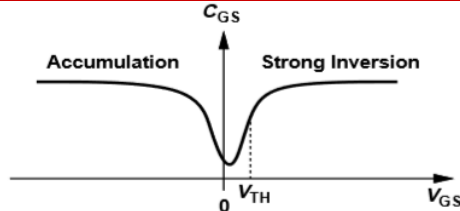
- Spacing between fins, S_F , also plays a significant role in performance and is typically fixed.
- Due to small dimensions of the intrinsic FinFET, the gate and S/D contacts must be placed away from the core of the device.

Behavior of MOS Devices as a Capacitor



- Recall that if source, drain, and bulk are grounded and the gate voltage rises, an inversion layer begins to form for $V_{GS} \approx V_{TH}$, and the device operates in the subthreshold region for $0 < V_{GS} < V_{TH}$.
- The transistor can be considered a two-terminal device and we can examine its capacitance for different gate voltages.

Behavior of MOS Devices as a Capacitor



- A very negative V_{GS} causes holes to be attracted in the substrate to the oxide interface, and say the MOSFET operates in the “accumulation region” with unit area capacitance of C_{ox} .
- As V_{GS} rises, the density of holes at the interface falls, a depletion region begins to form under the oxide, and the device enters weak inversion, causing capacitance to consist of a series combination of C_{ox} and C_{dep} .
- As V_{GS} exceeds V_{TH} , the oxide-silicon interface sustains a channel and the unit area capacitance returns to C_{ox} .