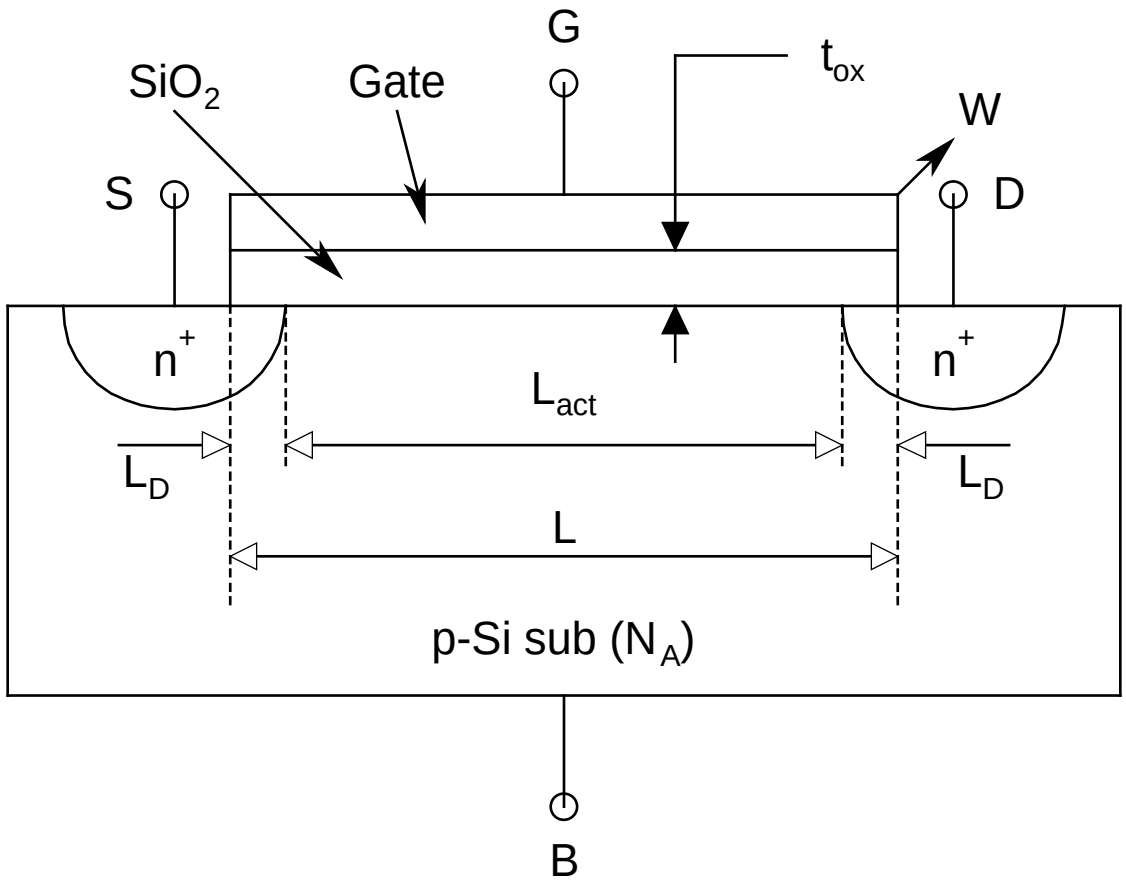
METAL-OXIDE-SEMICONDUCTOR FIELD-EFFECT TRANSISTOR (MOSFET)

- ***Extremely popular device*** - has almost pushed BJTs out of the market
- ***Three-Layer Device*** (***Metal***, ***Dielectric***, and ***Semiconductor***)
- ***Four-Terminal Device*** [***Drain*** (D), ***Source*** (S), ***Gate*** (G), and ***Body/Substrate*** (B)]
- ***Current*** through two terminals (D and S) can be ***controlled*** by the ***voltages*** applied at G and B
  - ***Voltage Controlled Device***

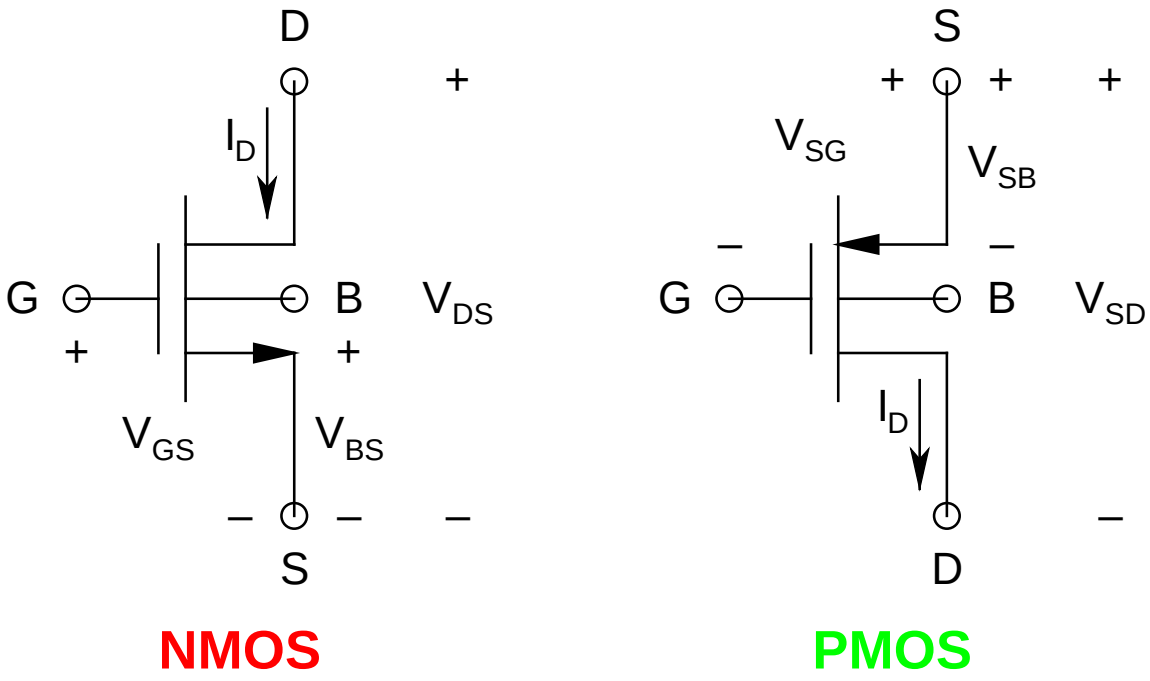
- ***Unipolar device***
  - Either ***electrons*** or ***holes*** participate in ***current conduction***
- ***Active device***
  - Capable of producing ***voltage/current/power gain***
- ***Two basic usage***:
  - **Amplification** (***Analog Circuits***)
  - **Switching** (***Digital Circuits***)
- ***Two Types*: *NMOS*** and ***PMOS***

NMOS Structure



- ***Technology Parameters***:
  - ***Channel Length*** (L)
  - ***Channel Width*** (W)
  - ***Oxide Thickness*** (t_{ox})
  - ***Substrate Doping*** (N_A)
- L_D: ***Lateral overlap*** between G and S/D
- ***Actual*** channel length: $L_{act} = L - 2L_D$
- For now, we will assume $L_D = 0$
  - $L_{act} = L$

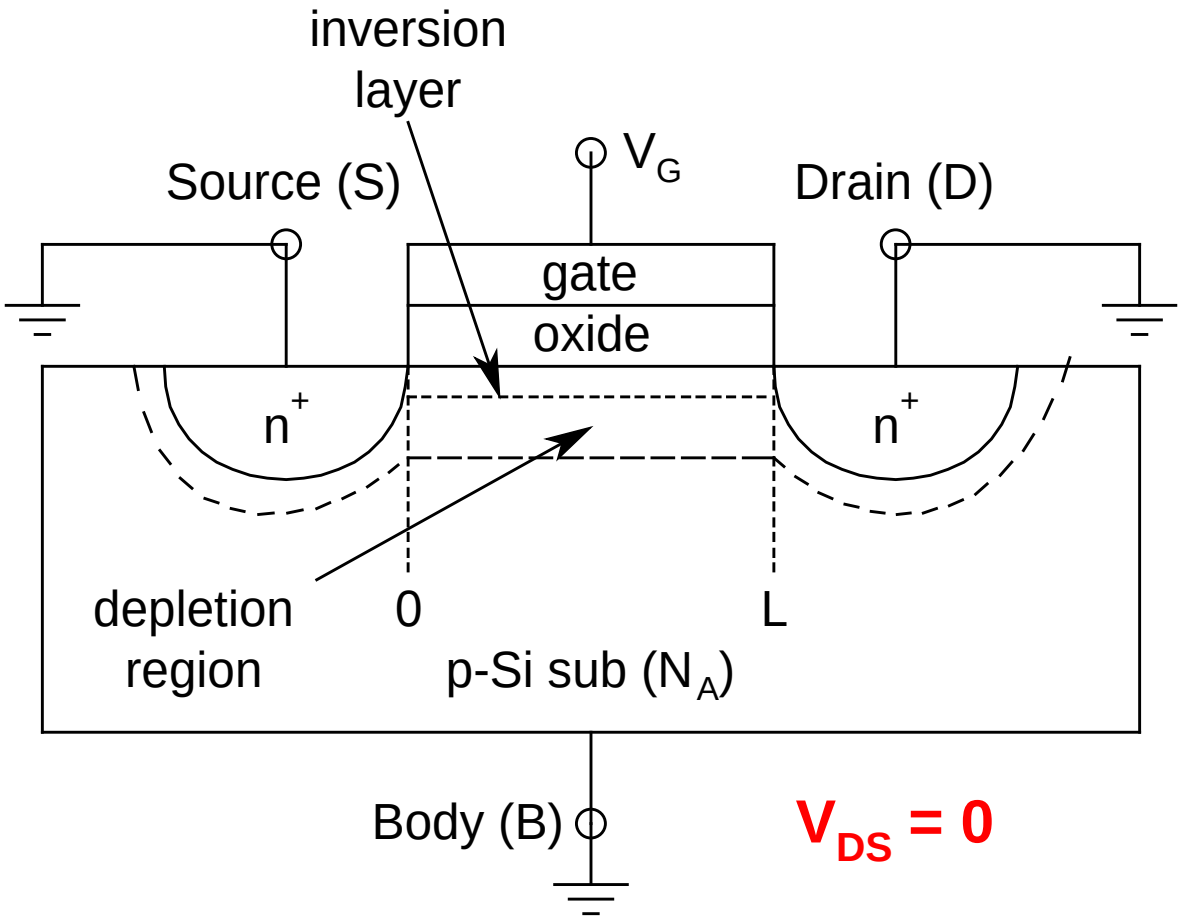
Symbols and Current-Voltage Conventions



- ***Voltage Convention***:
  - ***NMOS***: V_{GS} (***gate-source voltage***), V_{DS} (***drain-source voltage***), V_{BS} (***body-source voltage***)
  - ***PMOS***: V_{SG} (***source-gate voltage***), V_{SD} (***source-drain voltage***), V_{SB} (***source-body voltage***)
- ***Current Convention***:
  - ***NMOS***: I_D (***drain current***) flows into the ***drain terminal*** and exits from the ***source terminal***
  - ***PMOS***: I_D flows into the ***source terminal*** and exits from the ***drain terminal***

- ***Gate is DC isolated by the insulator***
  - ***Gate Current $I_G = 0$***
  - ***Tremendous advantage***!
- ***Same current I_D flows through the device***
- ***Extremely compact device***
  - ***Saves a lot of area***
- ***Reversible device***:
  - ***D and S terminals are determined by their bias states***

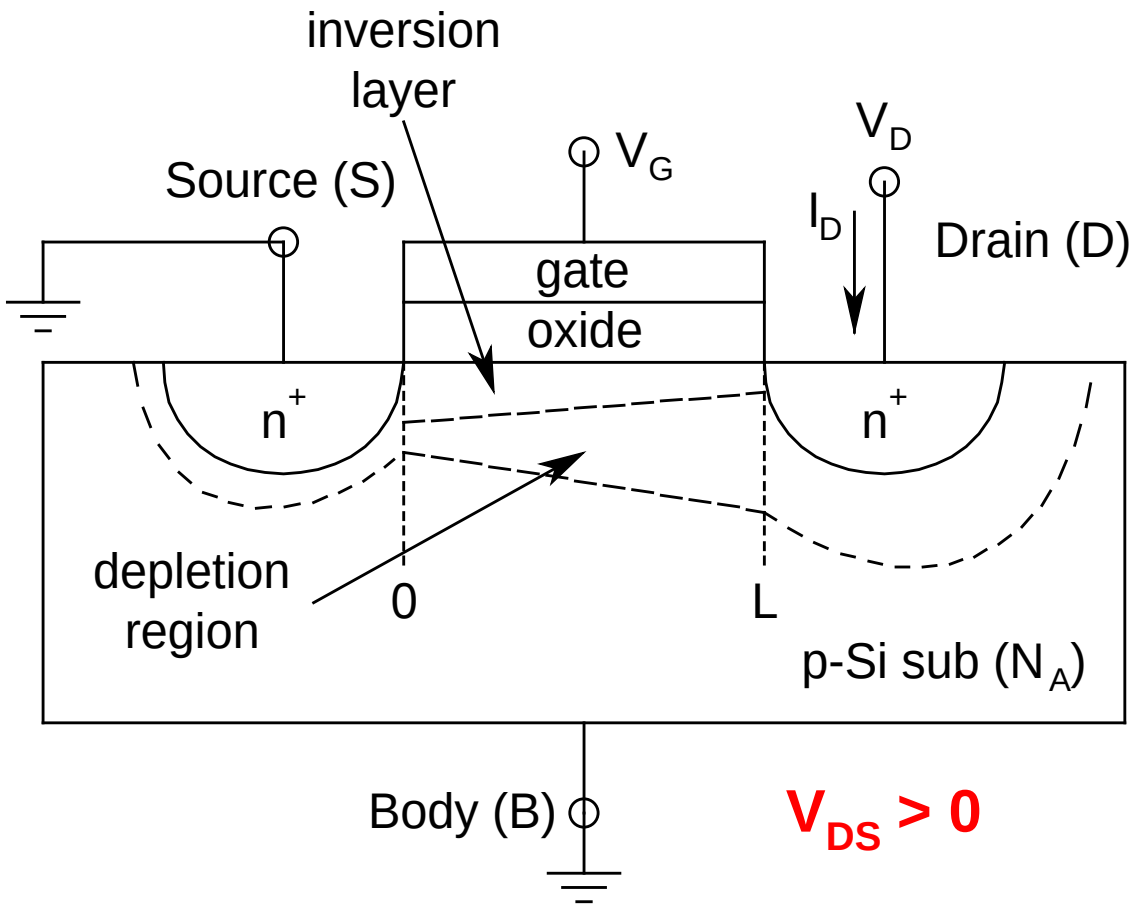
Operation


inversion
layer
Source (S)
V_G
Drain (D)
gate
oxide
n^+
n^+
depletion
region
0
L
p-Si sub (N_A)
Body (B)
$V_{DS} = 0$

- The structure is ***similar*** to an ***n^+pn^+ BJT***
- However, ***BJT action*** is ***not possible*** due to ***large channel length*** (L)
- The way to make the device ***conduct*** is to form a ***layer of electrons*** between S and D
  - Known as ***Inversion Layer***
- Then, if a ***bias*** is applied ***between S and D***, then ***inversion layer electrons*** will move ***towards the higher potential*** due to ***drift***
  - A ***current*** would result

- Consider $V_S = V_D = V_B = V_G = 0$
  - Device is ***off*** and ***no current flows***
- Note that the ***structure*** is similar to a ***capacitor***
- Now, as V_G is made ***positive***, initially it will ***repel holes from surface towards bulk***, ***uncovering ionized acceptor atoms there***
  - ***Formation of a depletion layer***
- There will be ***depletion layers*** around ***SB and DB junctions*** as well

- As V_G is kept on ***increasing***, the ***depletion charge*** will keep on ***increasing***
- At a ***certain value*** of V_G (= V_{GS}), a ***layer of electrons*** will ***appear at the surface***
- This ***particular value*** of V_{GS} is known as the ***threshold voltage*** V_{TN}
- Still ***no current*** would flow, since $V_{DS} = 0$
- ***SB and DB junctions*** must remain ***either at zero bias or reverse bias*** all the time
  - V_{SB} and $V_{DB} \geq 0$


inversion
layer
Source (S)
V_G
V_D
I_D
Drain (D)
gate
oxide
n^+
n^+
depletion
region
0
L
p-Si sub (N_A)
Body (B)
$V_{DS} > 0$

- With $V_{DS} > 0$, ***inversion layer electrons*** will ***move*** towards the ***higher potential***, i.e., D
  - The ***drain current*** I_D would ***flow from D to S***
- ***Note***:
  - The ***depletion charge*** would ***increase*** as we move ***towards the D*** (since the ***DB junction*** is ***more reverse biased***)
    - The ***inversion charge*** would ***decrease*** as we move ***towards the D***
    - For ***sufficiently high*** V_{DS}, it may ***disappear altogether***

Body Effect

- The ***threshold voltage*** V_{TN} is a ***function*** of the ***SB voltage*** V_{SB}
- As $V_{SB}\uparrow$, the ***SB junction depletion charge*** would ***increase***
  - For the ***same*** V_{GS}, ***inversion charge*** would ***decrease*** (to maintain ***charge balance***)
  - Thus, to ***restore*** the ***original level*** of ***inversion***, V_{GS} has to be ***increased***
  - Implies that V_{TN} has ***increased***

- Expressed as:

$$V_{TN} = V_{TN0} + \gamma\left(\sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F}\right)$$

$$V_{TN0} = V_{TN}\big|_{V_{SB}=0} = \textbf{\textit{Zero back-bias threshold voltage}}$$

$$\gamma = \frac{\sqrt{2q\varepsilon_s N_A}}{C'_{ox}} = \textbf{\textit{Body-effect coefficient}}$$

$$C'_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} = \textbf{\textit{Oxide capacitance per unit area}}$$

$$\phi_F = V_T \ln\left(\frac{N_A}{n_i}\right) = \textbf{\textit{Bulk potential}}\ (\sim 0.3 - 0.45\ \text{V})$$

Current-Voltage Relation

- For $V_{GS} > V_{TN}$ and ***small*** V_{DS}:

$$I_D = k_N \left(V_{GT} V_{DS} - V_{DS}^2 / 2 \right)$$

$$V_{GT} = V_{GS} - V_{TN} = \textit{Gate overdrive}$$

$$k_N = \left(W/L \right) k'_N$$

$$= \textit{Device transconductance parameter}$$

$$W/L = \textit{Aspect ratio}$$

$k'_N = \mu_n C'_{ox}$

$=$ ***Process transconductance parameter***

μ_n = ***Channel electron mobility***

- For ***small*** V_{DS}, the V_{DS}^2 term can be ***neglected***
  - I_D changes ***linearly*** with V_{DS}
    - ***Linear*** (or ***Non-Saturation***) Region
- As $V_{DS}\uparrow$, the ***restraining*** effect of V_{DS}^2 term$\uparrow$
  - ***Rate of increase*** of I_D with V_{DS} ***slows down***

- For ***inversion channel*** to exist at the ***D end***, ***V_{GD} must be > V_{TN}***
  - ***V_{DS} must be < V_{GT}***
- When ***$V_{DS} = V_{GT}$***, the ***channel*** is said to be ***pinched-off*** at the ***D end***, and ***I_D does not increase any more***
- This value of V_{DS} is known as the ***drain-to-source saturation voltage*** $V_{DS,sat}$
  - $V_{DS,sat} = V_{GT}$

- For $V_{DS} > V_{DS,sat}$, the ***mode of operation*** is known as ***saturation***
- ***Drain current in saturation***:

$$I_D = \frac{k_N}{2} V_{GT}^2$$

  - ***Obtained from the non-saturation I_D expression by substituting $V_{DS} = V_{GT}$***
  - ***Note that I_D is independent of V_{DS}***
- ***Above equations are valid for $V_{GT} > 3V_T$***

  (~ ***80 mV at room temperature***)

The Complete LEVEL 0 Model

$$I_D = k_N\left(V_{GT}V_{DS} - V_{DS}^2/2\right)$$

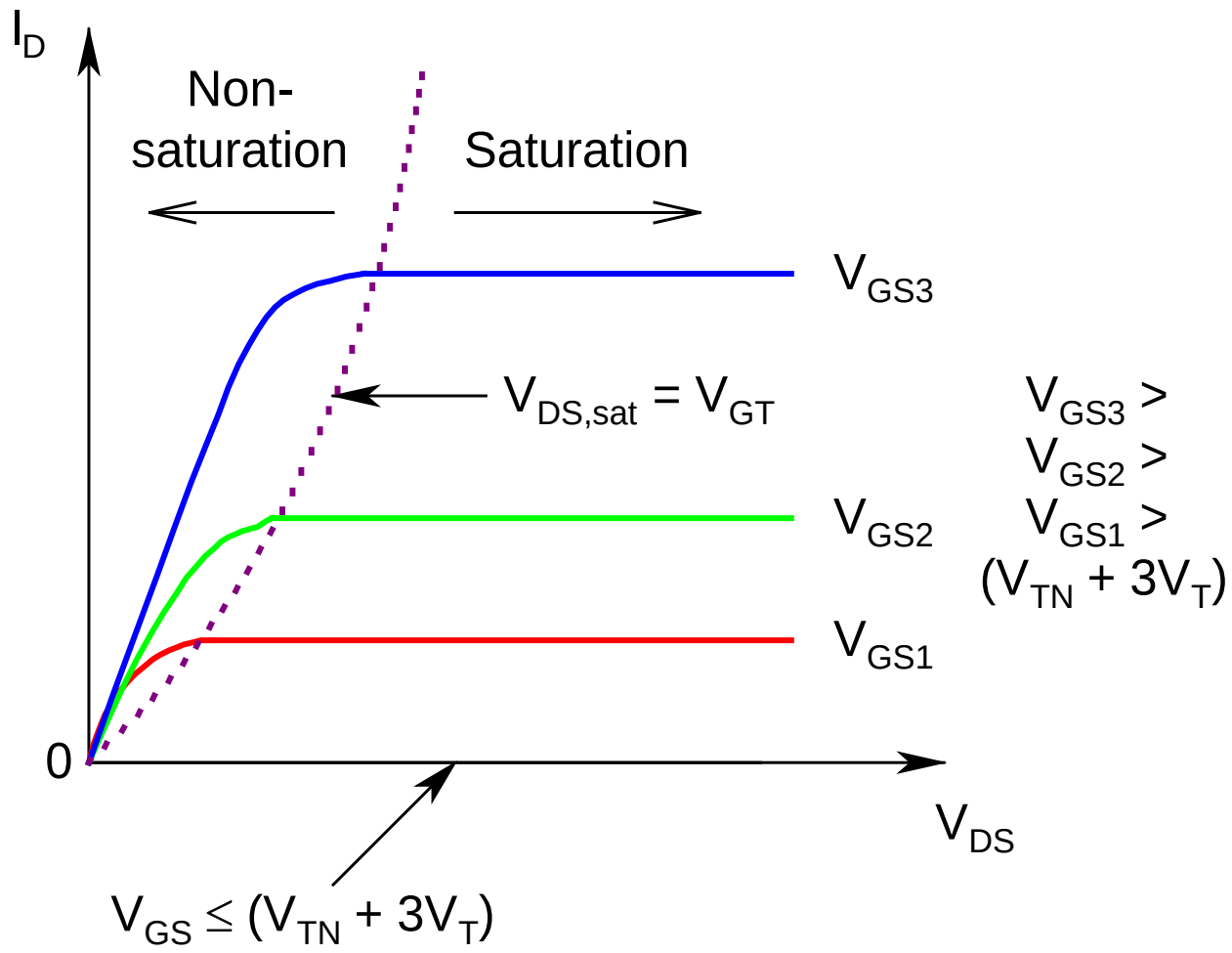
$$\left(\textbf{\textit{linear region}} \text{ - } V_{GT} > 3V_T, V_{DS} < V_{GT}\right)$$

$$= \left(k_N/2\right)V_{GT}^2$$

$$\left(\textbf{\textit{saturation region}} \text{ - } V_{GT} > 3V_T, V_{DS} \geq V_{GT}\right)$$

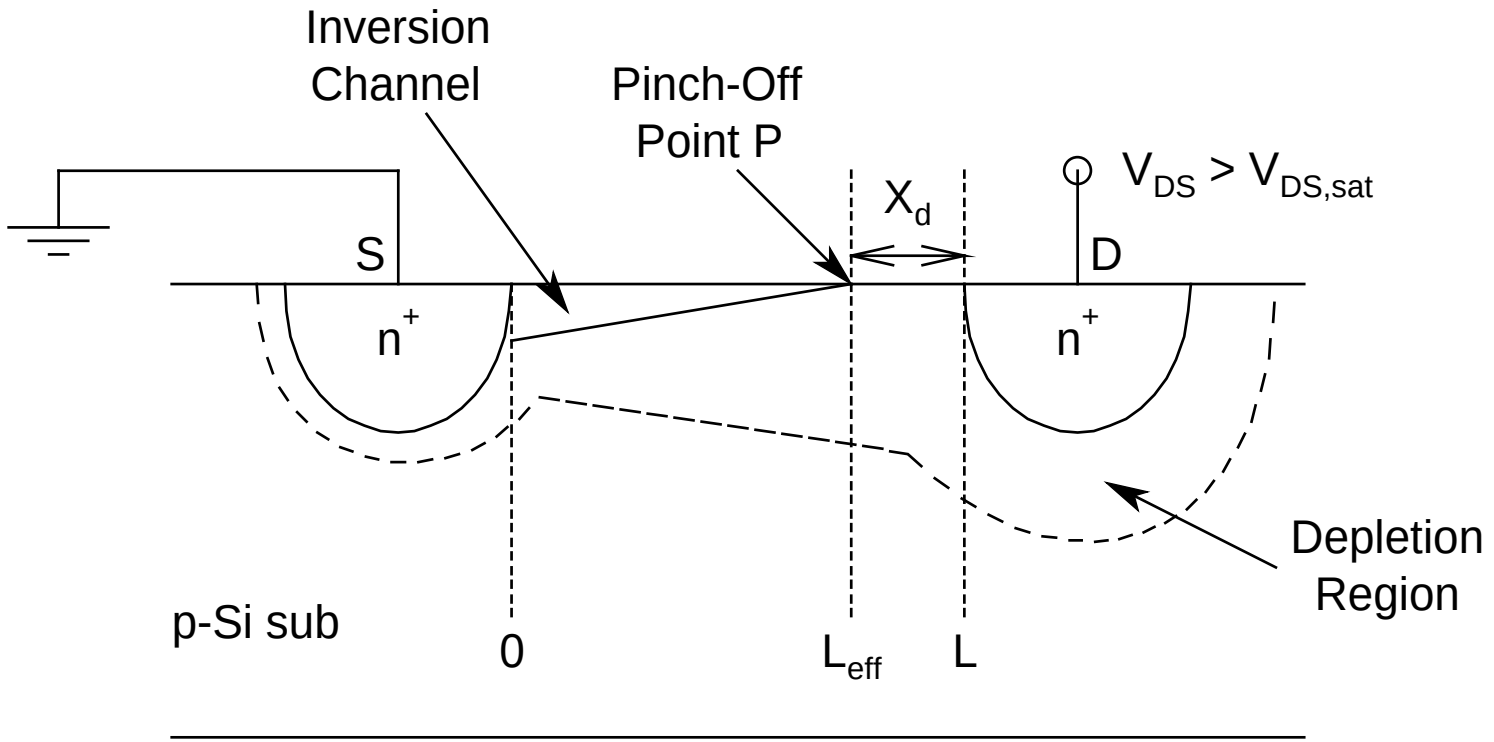
$$= 0$$

$$\left(\textbf{\textit{cutoff region}} \text{ - } V_{GT} \leq 3V_T, \textbf{\textit{any}}\ \boldsymbol{V_{DS}}\right)$$



I_D-V_{DS} Characteristics

Channel Length Modulation (CLM)



$X_d \rightarrow$ length of the pinched-off region

- For $V_{DS} = V_{DS,sat}$, ***pinch-off point P at D end***
- For $V_{DS} > V_{DS,sat}$, ***P moves towards source***
- ***Effective channel length reduces*** from L to $L_{eff} = L - X_d$
  - X_d = ***pinch-off region/drain region/saturation region length***
- ***Excess voltage*** $(V_{DS} - V_{DS,sat})$ ***drops across*** X_d

- ***Reduction of effective channel length causes an increase in current***
  - ***Channel length modulation***
- With $V_{DS}\uparrow$, $X_d\uparrow$, $L_{eff}\downarrow$, and $I_D\uparrow$
  - ***No real current saturation***
- Thus, ***saturated drain current***:

$$I_{D,sat} = \left(k'_N/2\right)\left(W/L_{eff}\right)V_{GT}^2$$
$$= \left(k_N/2\right)V_{GT}^2\left(1+\lambda V_{DS}\right)$$

- λ = ***Channel length modulation parameter***

$$= \frac{1}{L}\frac{dX_d}{dV_{DS}}$$

  - ***Function of L and N_A***
  - ***Higher L and N_A => Lower λ***
  - ***Typical values of λ*** may range from ***close to 0*** to ***as high as 0.1-0.3*** V^{-1}
  - ***Very similar to V_A for BJTs***

- This gives ***LEVEL 1 model*** (also known as ***Shichman-Hodges model***) for MOSFETs:

$$I_D = k_N\left[V_{GT}V_{DS} - V_{DS}^2/2\right](1+\lambda V_{DS})$$

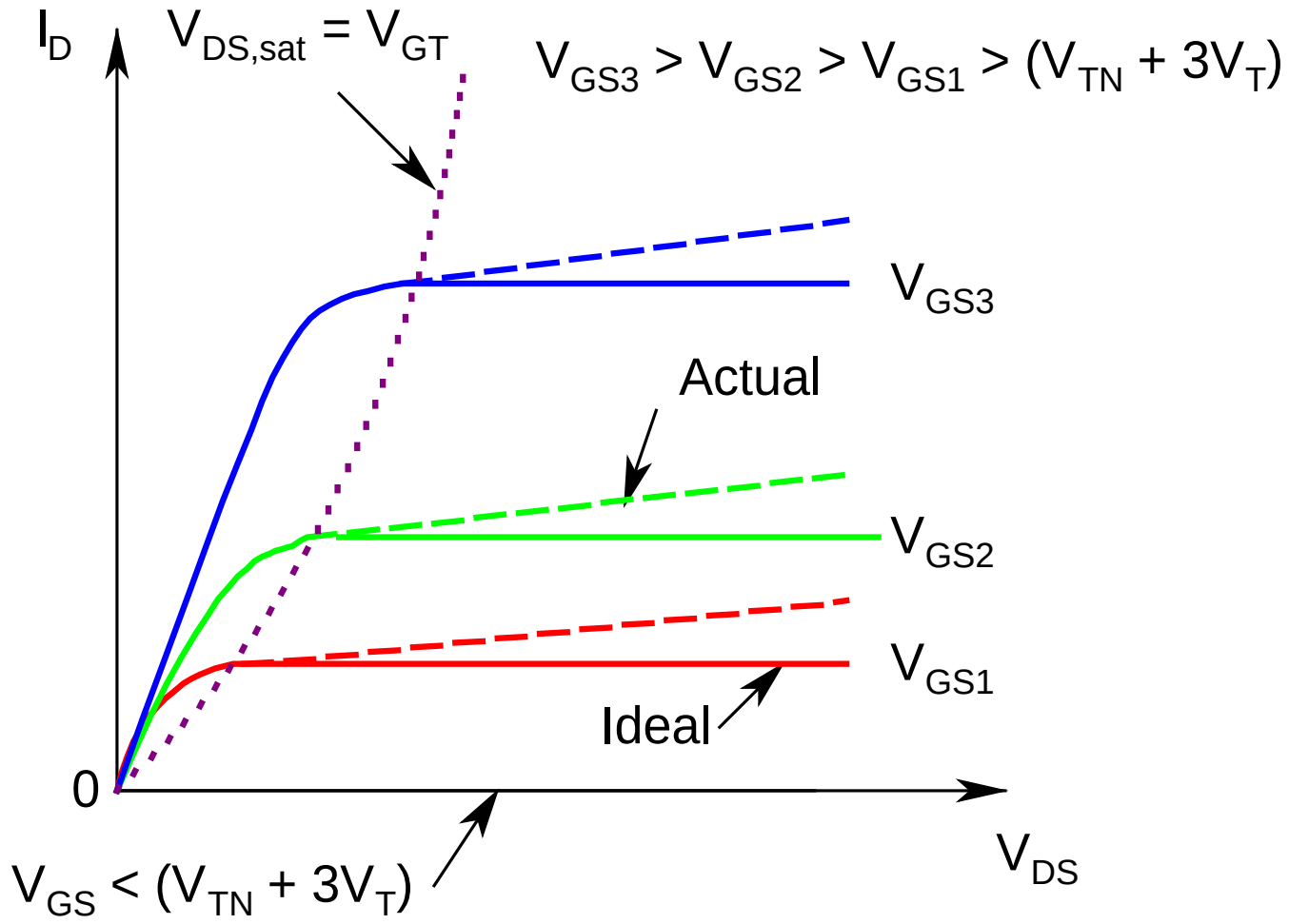
(***linear region*** - $V_{GT} > 3V_T,\ V_{DS} < V_{GT}$)

$$= (k_N/2)V_{GT}^2(1+\lambda V_{DS})$$

(***saturation region*** - $V_{GT} > 3V_T,\ V_{DS} \geq V_{GT}$)

$$= 0$$

(***cutoff region*** - $V_{GT} \leq 3V_T,$ any V_{DS})



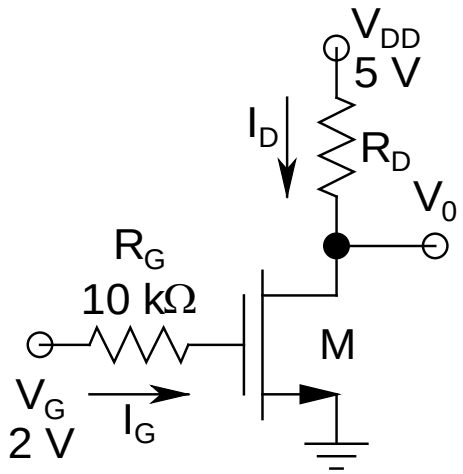
I_D-V_{DS} Characteristics in presence of CLM

DC Bias Point Calculation

- ***To find R_D for BB***
  - $V_{TN0} = 1$ V, $k'_N = 40\ \mu A/V^2$, $W/L = 10$
- ***Body terminal not shown***
  - ***Implies that it is connected to the most negative potential available in the circuit (ground in this case)***

$\Rightarrow V_{SB} = 0 \Rightarrow V_{TN} = V_{TN0}$

- $I_G = 0 \Rightarrow V_{GS} = V_G = 2$ V



- $V_{GT} = V_{GS} - V_{TN} = 1$ V
- ***Assuming saturation mode of operation and neglecting CLM***:

  $I_D = (k_N/2) V_{GT}^2 = 200$ μA
- ***For BB***, $V_{DS} = V_{DD}/2 = 2.5$ V (***2-element output branch***):

  $R_D = (V_{DD} - V_{DS})/I_D = 12.5$ kΩ
- $V_{DS} > V_{GT} \Rightarrow$ ***Assumption of saturation mode of operation validated***
- $P_D = V_{DS} \times I_D = 0.5$ mW

Small-Signal Model

- The ***electrical equivalent*** of the MOSFET at the ***DC bias point***
- ***Must be biased in saturation***
  - ***Resembles a constant current source***
- ***DC analysis must precede***, since ***need the information regarding the Q-point*** (I_D, V_{DS})
- ***This model for NMOS and PMOS is the same*** (***incremental model***)

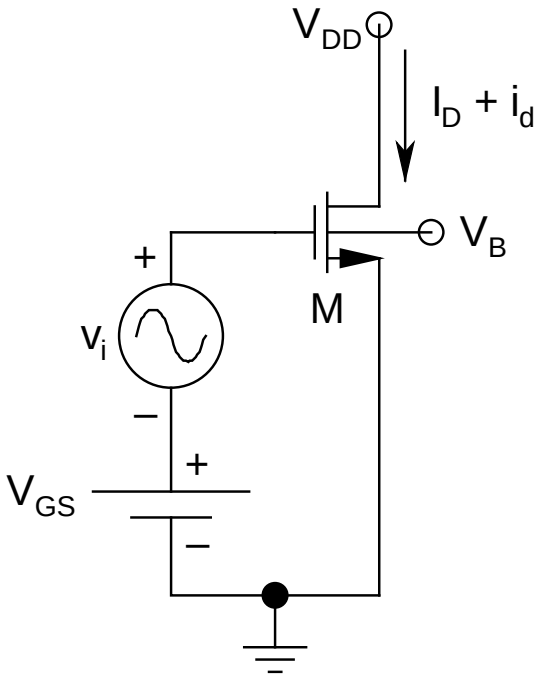
Validity of the Small-Signal Model

- The ***instantaneous current*** (***assuming $\lambda V_{DS} < 0.1$***):

$$I_d = \frac{k_N}{2}\left(V_{GT} + v_i\right)^2$$

$$= I_D + \frac{k_N}{2}\left[2V_{GT}v_i + v_i^2\right]$$

$$\Rightarrow \; i_d = k_N V_{GT} v_i \left[1 + \frac{v_i}{2V_{GT}}\right]$$



- Thus, for ***linear relationship*** between i_d and v_i, ***v_i must be << V_{GT}***
- Note that ***V_{GT} (minimum) = $3V_T$***
- Hence, ***v_i should be at least ten times less than $3V_T$***
- Recall in ***BJT***, for ***linear relationship*** between i_c and v_i, ***v_i has to be << V_T***
  - ***Three times less than that for MOSFET***
  - ***MOSFETs are inherently more linear device than BJTs*** (***compare quadratic with exponential***)

Small-Signal Model Parameters

- ***Transconductance*** (g_m):

$$g_m \triangleq \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS} \text{ and } V_{SB} \text{ constant}}$$

$$= k_N V_{GT}\left(1+\lambda V_{DS}\right) = \sqrt{2k_N I_D\left(1+\lambda V_{DS}\right)}$$

➢ ***If $\lambda V_{DS} < 0.1$***:

$$g_m \simeq k_N V_{GT} \simeq \sqrt{2k_N I_D}$$

- An important ***Figure of Merit*** is ***transconductance to current ratio***
  - For ***MOSFETs***: $g_m/I_D = 2/V_{GT}$
  - For ***BJTs***: $g_m/I_C = 1/V_T$
  - Thus, ***BJTs produce more g_m per unit current***
- As we will see later, a ***high value of g_m is highly desirable***, since it ***dictates the gain***
- ***g_m/I_D can be changed by changing the bias current and/or aspect ratio***
- ***g_m/I_C is a function only of temperature***

- ***Body Transconductance*** (g_{mb}):

$$g_{mb} \triangleq \left.\frac{\partial I_D}{\partial V_{BS}}\right|_{V_{GS} \text{ and } V_{DS} \text{ constant}} = \chi g_m$$

$$\chi = \frac{\gamma}{2\sqrt{2\phi_F + V_{SB}}} = \textbf{\textit{Body factor}} \quad (\sim 0.1\text{-}0.3)$$

  - ***Note***: ***As*** $V_{SB}\uparrow$, $V_{TN}\uparrow \Rightarrow I_D\downarrow$
  - $\partial I_D/\partial V_{SB}$ ***would have yielded negative*** g_{mb}
  - ***If both B and S are tied to fixed DC potentials (including ground),*** g_{mb} ***won't matter!***

- ***Output Conductance*** (g_0)/ ***Output Resistance*** (r_0):

$$g_0 = r_0^{-1} \triangleq \left.\frac{\partial I_D}{\partial V_{DS}}\right|_{V_{GS} \text{ and } V_{SB} \text{ constant}} = \frac{\lambda I_D}{1+\lambda V_{DS}}$$

  - ***If*** $\lambda V_{DS} < 0.1$:

    $g_0 = 1/r_0 \approx \lambda I_D$

  - ***λ has a very wide range*** ~ 0.01-0.5 V^{-1}
  - ***When*** $\lambda \rightarrow 0$, $g_0 \rightarrow 0$, ***and*** $r_0 \rightarrow \infty$
    - ***Device starts to behave like a constant current source***

- ***Gate-Source and Gate-Drain Capacitance*** (C_{gs} and C_{gd}):
  - Each has ***two components***: ***intrinsic*** (i) and ***technological*** (t)
  - ***Total intrinsic gate-body capacitance***:

    $C_{gbi} = C'_{ox} WL$
  - Using ***Meyer's model***, ***intrinsic component***:
    - In ***linear region***: $C_{gsi} = C_{gdi} = C_{gbi}/2$
    - In ***saturation region***: $C_{gsi} = (2/3)C_{gbi}$, $C_{gdi} = 0$
  - ***Technology component arises due to gate-source and gate-drain overlap*** (L_D)

- ***Technology components***:

$$C_{gst} = C_{gdt} = C'_{gs0}W = C'_{gd0}W$$

Gate-Source/Drain Overlap Capacitance per unit width: $C'_{gs0} = C'_{gd0} = C'_{ox}L_D$

- Thus, ***total capacitance in saturation***:

$$C_{gs} = (2/3)C'_{ox}WL + C'_{gs0}W$$

$$C_{gd} = C'_{gd0}W$$

- $C_{gs} >> C_{gd}$

- ***Source-Body and Drain-Body Capacitance*** (C_{sb} and C_{db}):
  - ***Both reverse-biased n^+p junctions***

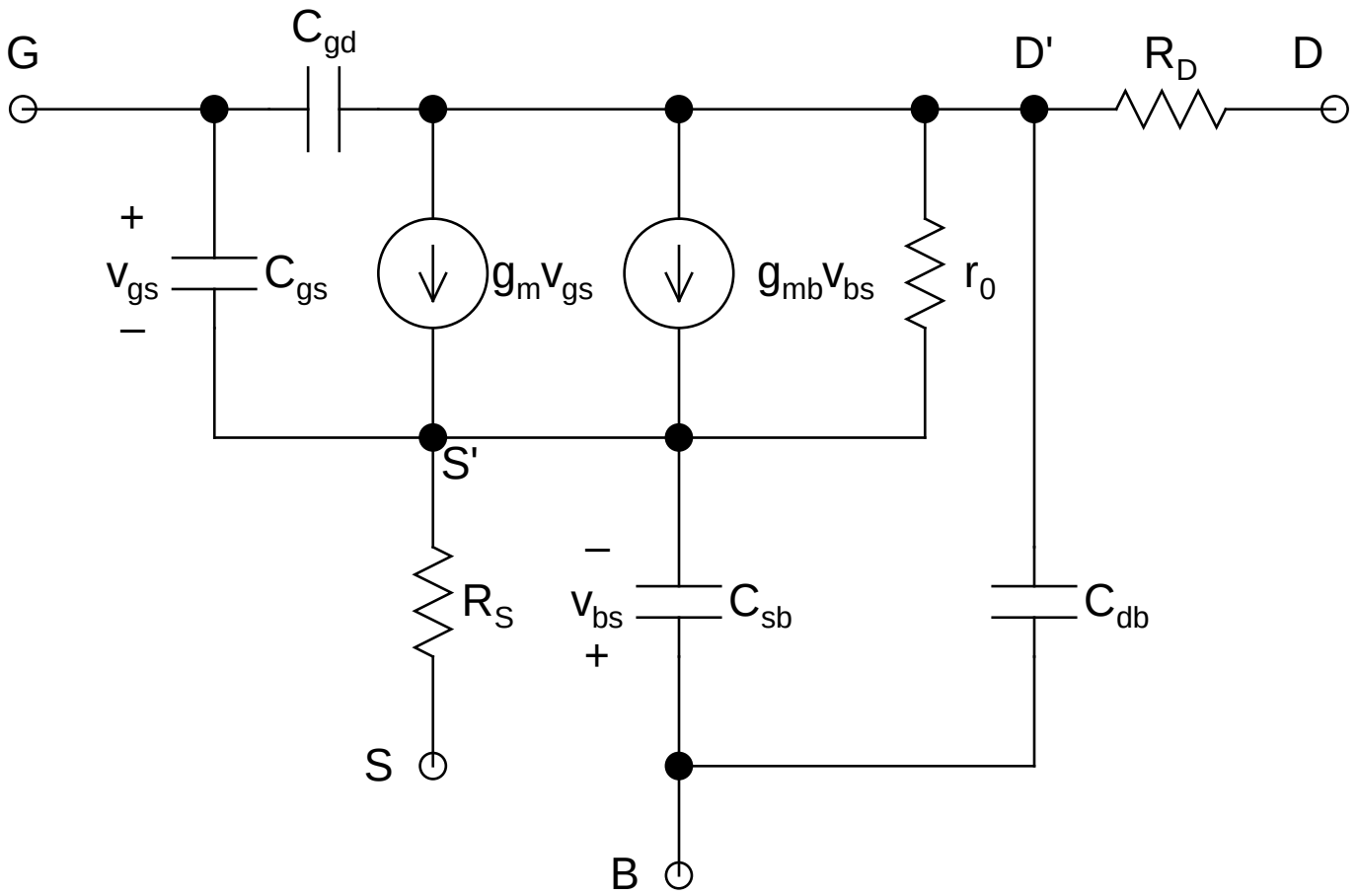
$$C_{sb} = \frac{C_{sb0}}{\left(1+V_{SB}/V_0\right)^m} \quad \text{and} \quad C_{db} = \frac{C_{db0}}{\left(1+V_{DB}/V_0\right)^m}$$

$$C_{sb0} = C_{sb}\Big|_{V_{SB}=0} \quad \text{and} \quad C_{db0} = C_{db}\Big|_{V_{DB}=0}$$

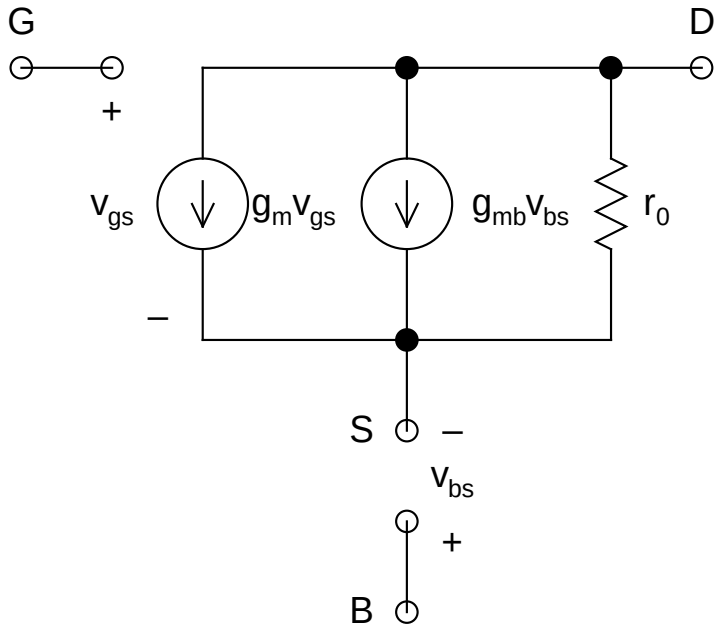
$$V_{SB} \text{ and } V_{DB} \geq 0$$

- ***Drain/Source Series Resistance*** (R_S and R_D):
  - ***Due to neutral n^+ source/drain regions***

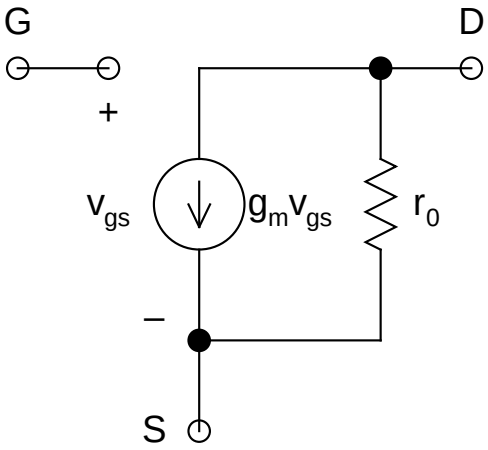
The Hybrid-π Model



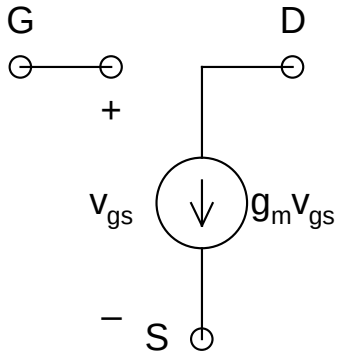
- ***Simplifications***:
  - ***R_S and R_D can be safely neglected***
  - For ***low to moderate frequencies***, the ***capacitive reactances*** of all the capacitances will be ***extremely large*** $\Rightarrow$ ***can be neglected***
  - ***If both B and S are connected to fixed DC potentials***, ***current source $g_{mb}v_{bs}$ disappears***
  - Leads to the ***Low-Frequency T-Model***, having only ***two components***: $g_m v_{gs}$ and r_0
  - ***Simplest possible equivalent results if r_0 is also neglected*** (***ideal current source***!)



Low-Frequency T-Model With Body Effect



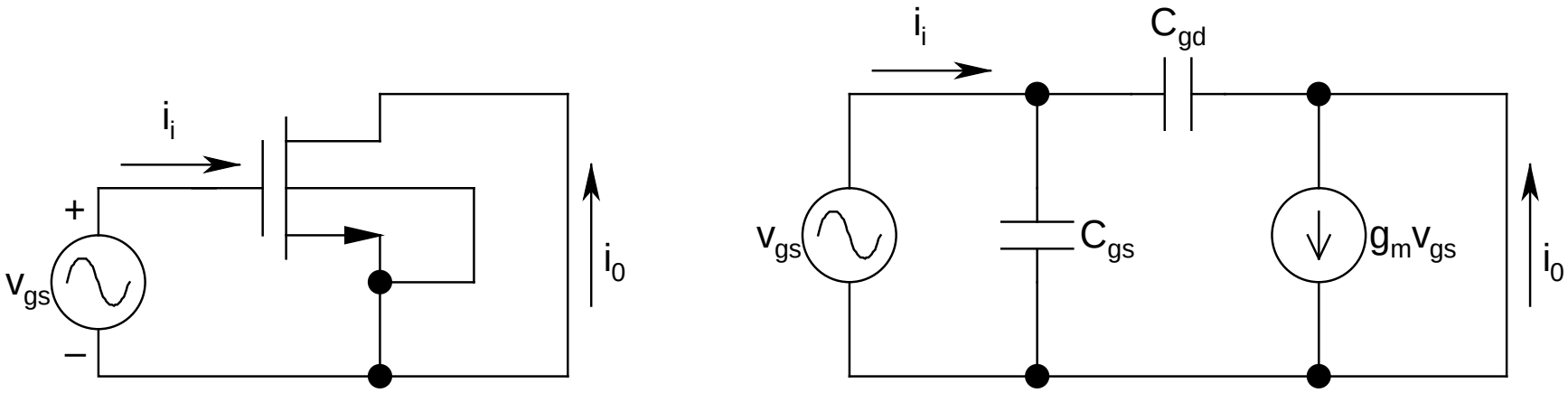
Low-Frequency T-Model Without Body Effect



Without CLM

Frequency Specification of MOSFETs

- Only ***Unity-Gain Frequency*** (f_T)



- $i_0 \approx g_m v_{gs}$ (neglecting ***reverse transmission*** through C_{gd})
- $i_i = j\omega(C_{gs} + C_{gd})$

$$\Rightarrow \frac{i_0(j\omega)}{i_i(j\omega)} = \frac{g_m}{j\omega(C_{gs} + C_{gd})}$$

$$\text{At } f = f_T,\ |i_0/i_i| = 1$$

$$\Rightarrow f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

- ***Remarkable similarity with that for BJT***

- ***Maximum Operable Frequency*** (f_{max}):
  - ***Maximum possible f_T***
  - Noting that ***$C_{gs} >> C_{gd}$***, ***neglecting C_{gst}***, and ***substituting the expressions for C_{gsi} and g_m***:

$$f_{max} = f_T\big|_{max} = \frac{3\mu_n V_{GT}}{4\pi L^2}$$

  - $f_{max} \propto 1/L^2$
    - ***Thrust towards making L as small as possible***
  - $f_{max} \propto V_{GT}$
    - ***Making V_{GT} large may be detrimental***!