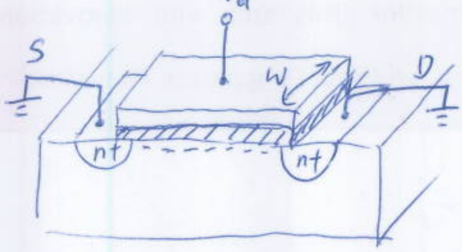
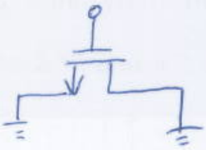
Chapter 1 - Basic MOS Device Physics and Models

charge density

velocity of the charge

$$I = Q_d \, v$$





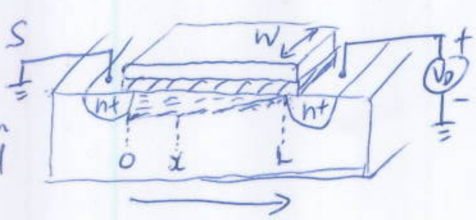
$V_{GS} = V_{TH}$, inversion layer occurs

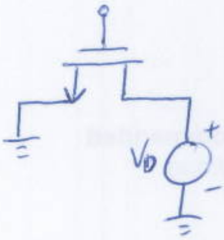
→ charge density produced by gate oxide the gate oxide capacitance

→ proportional to $V_{GS} - V_{TH}$

$$Q_d = \underbrace{W C_{ox}}_{\text{total capacitance per unit length}} (V_{GS} - V_{TH})$$

local voltage difference between gate and channel V_G to $V_G - V_D$





$\cancel{V_{GS} = V_{TH}}$

At point x, $Q_d(x) = W C_{ox} [V_{GS} - V(x) - V_{TH}]$

∴ The current is given by $I_D = - W C_{ox} [V_{GS} - V(x) - V_{TH}] \, v$

negative sign because charge carriers are negative.

For semiconductors, $v = \mu E$ where μ is mobility [μ_n for electron]

E is electric field. $\left[E(x) = - \frac{dV}{dx} \right]$

$$\therefore \; I_D = W C_{ox} [V_{GS} - V(x) - V_{TH}] \, \mu_n \frac{dV(x)}{dx}$$

→ boundry conditions $V(0) = 0$ and $V(L) = V_{DS}$

→ Multiply both sides by dV and perform integration

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

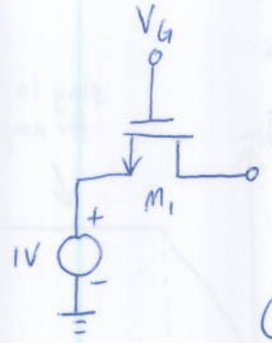
$$I_D \cdot L = W C_{ox} \mu_n \left[V_{GS} V - \frac{V^2}{2} - V_{TH} V \right]_0^{V_{DS}}$$

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

$$I_{D,max} = \cancel{\frac{\partial I_D}{\partial V_{DS}}} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2$$

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

E.g.



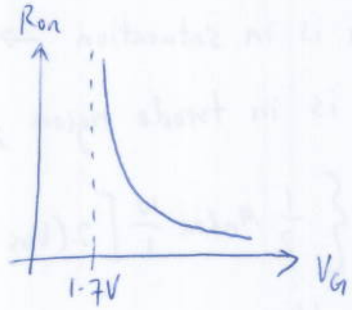
Assume $V_{TH} = 0.7V$, plot the on-resistance of M_1 as a function of V_G

Ans: $I_D = 0$, $V_{DS} = 0$ (drain terminal is open)

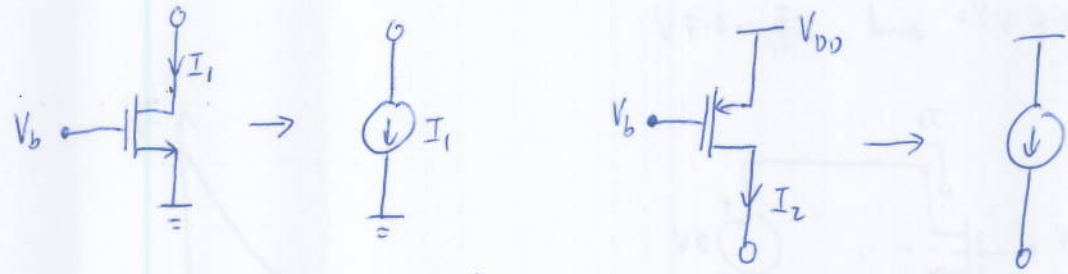
If device is on, it operates in the deep triode region.

① For $V_G < 1V + V_{TH}$, M_1 is off $\rightarrow R_{on} = \infty$

② For $V_G > 1V + V_{TH}$, $R_{on} = \dfrac{1}{\mu_n C_{ox} \frac{W}{L}(V_G - 1V - 0.7V)}$



- Saturated MOSFETs operating as current sources



inject current into ground or draw current from V_{DD}

- To ~~see~~ define how well a device converts a voltage to a current, we look at the change in the drain current divided by the change in the gate-source voltage (This is known as transconductance, g_m)

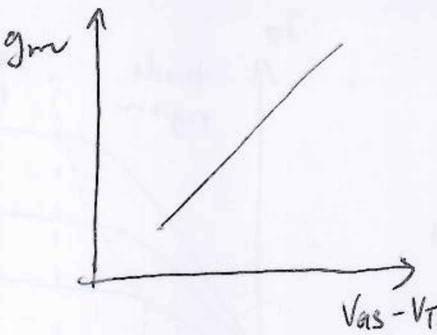
$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS},\,const.}$$

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

→ The higher the g_m the more sensitive the device is.

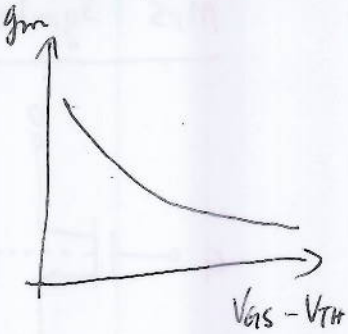
→ See the three graphs

W/L constant



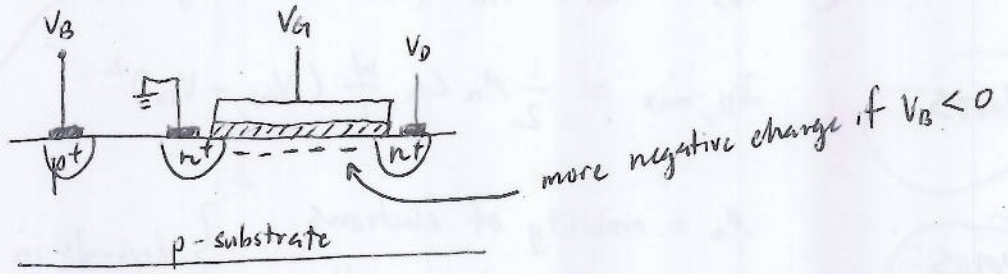
W/L constant

I_D Constant



Second-order Effects

① – body effect / backgate effect when $V_B < 0$



As V_B becomes more negative, more holes are attracted to the substrate connection, leaving larger negative charge.

$V_{THO} \approx \phi$

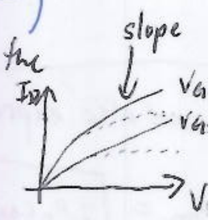
$$V_{TH} = V_{THO} + \gamma\left(\sqrt{|2\phi_F + V_{SB}|} - \sqrt{|2\phi_F|}\right) \qquad \gamma = \frac{\sqrt{2q\varepsilon_{Si}N_{sub}}}{C_{ox}}$$

where V_{THO} will be given, γ = body effect coefficient,
V_{SB} = source-bulk potential difference.

γ typically lies in the range of 0.3 to 0.4 $V^{1/2}$

② Channel-Length Modulation $\left(L' = L - \Delta L,\ \frac{1}{L - \Delta L} \approx \frac{1 + \Delta L/L}{L} = \frac{1 + \lambda V_{DS}}{L}\right)$ first order relationship

– actual length of the inverted channel gradually decreases as the potential difference between the gate and the drain increases.



longer channel
$\lambda \downarrow$ smaller

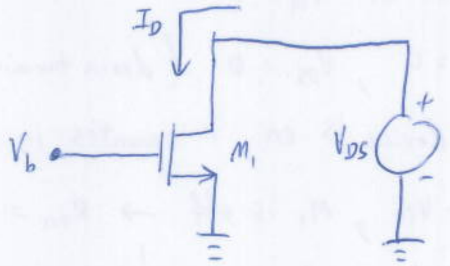
In saturation, $I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})^2(1 + \lambda V_{DS})$ ← channel-length modulation coefficient.

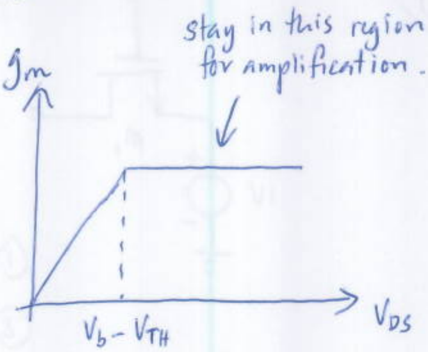
$$g_m = \mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})(1 + \lambda V_{DS})$$
$$= \sqrt{\frac{2\mu_n C_{ox}(W/L) I_D}{1 + \lambda V_{DS}}}$$

③ Subthreshold Conduction (operates in weak inversion) $\rightsquigarrow I_D = I_0 \exp\frac{V_{GS}}{\delta V_T}$ $\quad \delta > 1$ nonideality factor $\quad V_T = \frac{kT}{q}$

④ Voltage Limitations (punchthrough) ⟶ breakdown effect if their terminal voltage difference exceeds certain limits

(E.g.) Plot the transconductance as a function of V_{DS}





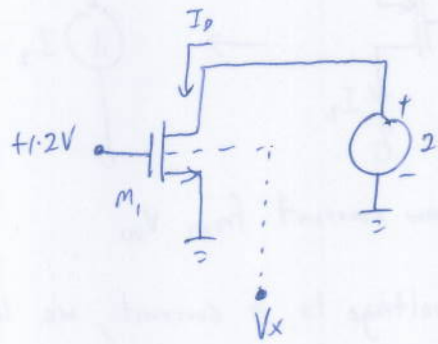
For $V_{DS} \geq V_b - V_{TH}$, M_1 is in saturation $\rightarrow$ I_D relatively constant. $\rightarrow$ g_m cons

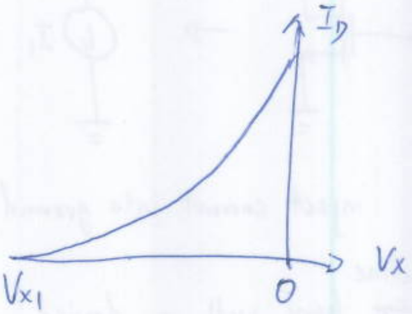
For $V_{DS} < V_b - V_{TH}$, M_1 is in triode region,

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

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

(E.g.) Plot the drain current if V_X varies from $-\infty$ to 0. Assume $V_{TH0} = 0.6V$, $\gamma = 0.4\ V^{1/2}$ and $2\Phi_F = 0.7V$





V_X is sufficiently negative, the threshold voltage of M_1 exceeds 1.2V and the device is off. That is

$$1.2V = 0.6 + 0.4\left(\sqrt{0.7 - V_{X1}} - \sqrt{0.7}\right)$$

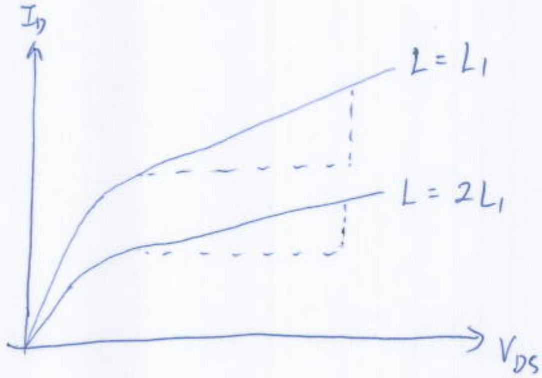
and hence $V_{X1} = -4.76V$. For $V_{X1} < V_X < 0$, I_D increases according to

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

(E.g.) Keeping all other parameters constant, plot I_D/V_{DS} characteristic of a MOSFET for $L = L_1$ and $L = 2L_1$.

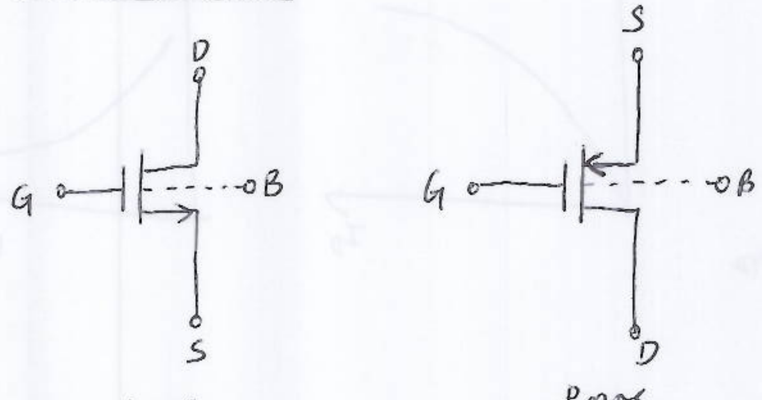
$$I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH})^2 (1 + \lambda V_{DS})$$

and $\lambda \propto 1/L$, we note that if the length is doubled, the slope of I_D vs V_{DS} is divided by four because $\partial I_D / \partial V_{DS} \propto \lambda / L \propto 1/L^2$.

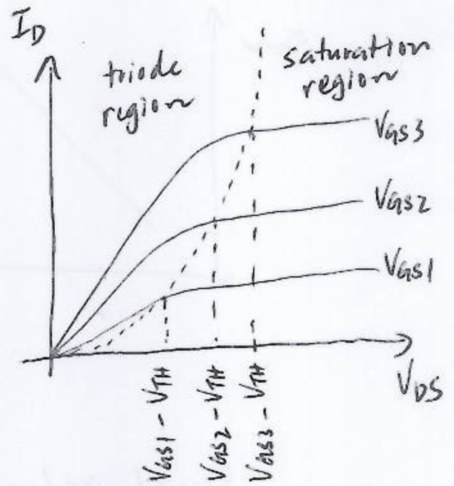


For a given gate-source overdrive, a larger L gives a more ideal current source while degrading the current capability of the device. Thus, W may need to be increased proportionally.

MOS Symbols



NMOS PMOS



NMOS

$\boxed{V = \mu E}$

① $V_{DS} = V_{GS} - V_{TH}$ (saturation region)

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

$$I_{D,max} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS}-V_{TH})^2$$

μ_n = mobility of electrons
C_{ox} = gate oxide capacitance } depends on technology

W = width
L = length } W/L : aspect ratio

PMOS put minus sign in front

$V_{GS} - V_{TH}$ = overdrive voltage / effective voltage

② If $V_{DS} \leq V_{GS} - V_{TH}$ → (triode region / linear region)

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

Later

$\sim\!\!\to g_m = \mu_n C_{ox}\frac{W}{L}V_{DS}$

③ If $V_{DS} \ll 2(V_{GS}-V_{TH})$, → (in deep triode region)

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

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

can also express

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

Later

$\boxed{g_m = \frac{1}{R_{on}}}$ transconductance $g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS},\text{const}}$

saturation region

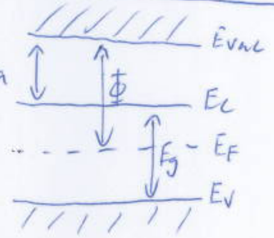
④ If $V_{DS} > V_{GS} - V_{TH}$,

same as $V_{DS} = V_{GS} - V_{TH}$

minimum energy needed to remove an electron from a solid

For second-order effects:

electron affinity $= E_{ea}$



$$V_{TH0} = \phi_{ms} + 2\phi_F + \frac{Q_{dep}}{C_{ox}}$$

- ϕ_{ms} is the difference between the work functions of the polysilicon gate and the silicon substrate
- $\phi_F = \frac{kT}{q} \ln\left(\frac{N_{sub}}{n_i}\right)$, q is electron charge, N_{sub} is the doping concentration of the substrate.
- Q_{dep} is the charge in the depletion region
- C_{ox} is the gate oxide capacitance per unit area.

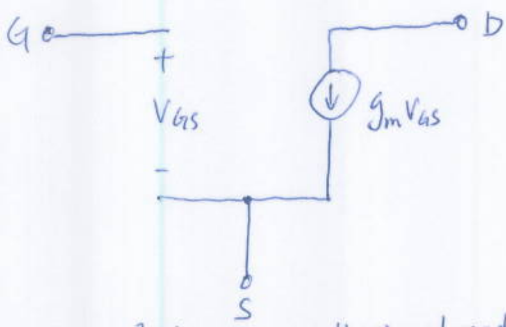
From PN junction theory: $\simeq \sqrt{4q\varepsilon_{si}|\phi_F|N_{sub}}$, ε_{si} is the dielectric constant of silicon.

Typical $t_{ox} \approx 50\,\text{Å}$, $C_{ox} \approx 6.9\ fF/\mu m^2$

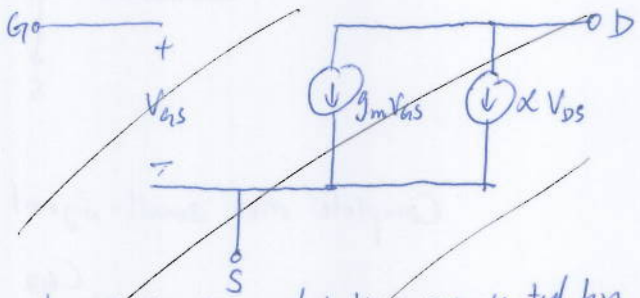
MOS Small-Signal Model and capacitance

Small-signal vs large signal

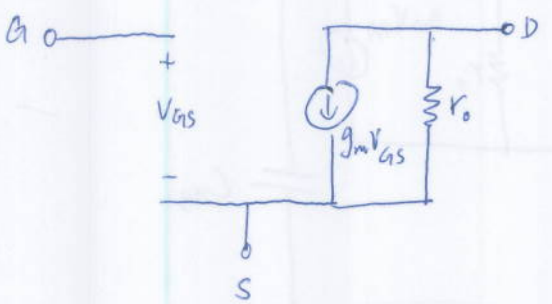
- Small signal model takes a circuit and based on an operating point (bias) and linearizes all the components. Nothing changes because the assumption is that the signal is so small that the operating point (gain, capacitance, etc.) doesn't change.
- A Large signal takes into account the fact that the large signal actually affects the operating point and takes into account that elements are non-linear and that circuits can be limited by power supply values.



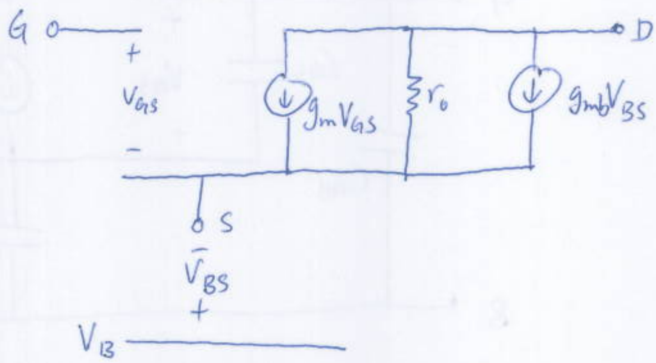
Basic MOS small-signal model



~~channel-length modulation represented by a dependent current source.~~



channel-length modulation represented by a resistor



body effect represented by a dependent current source

$$r_o = \frac{\partial V_{DS}}{\partial I_D}$$

$$= \frac{1}{\partial I_D / \partial V_{DS}}$$

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

$$\approx \frac{1}{\lambda I_D}$$

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}}$$

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

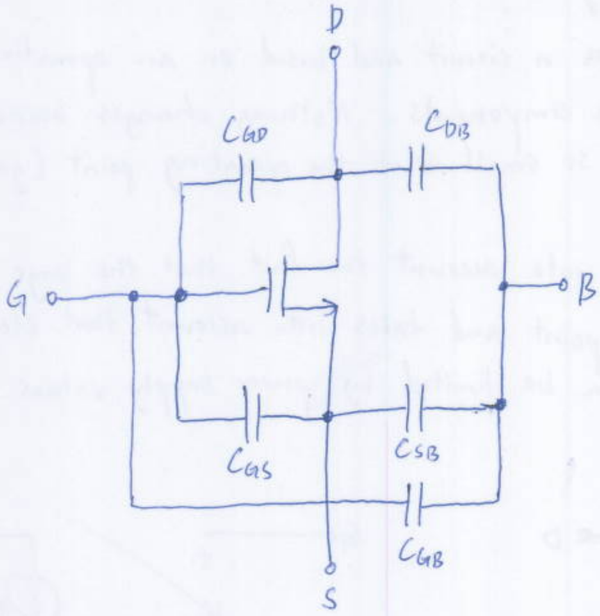
$$\frac{\partial V_{TH}}{\partial V_{BS}} = -\frac{\partial V_{TH}}{\partial V_{SB}}$$

$$= -\frac{\gamma}{2}(2\phi_F + V_{SB})^{-1/2}$$

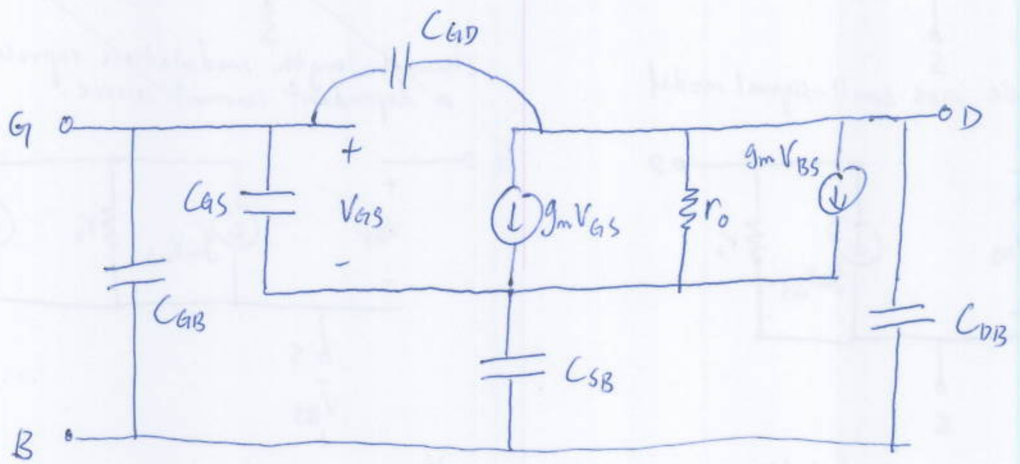
$$g_{mb} = g_m \frac{\gamma}{2\sqrt{2\phi_F + V_{SB}}}$$

$$= \eta\, g_m$$

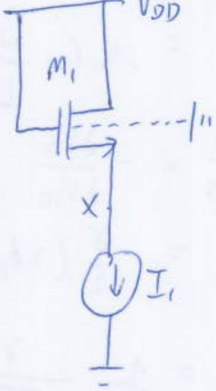
Mos capacitance . . .

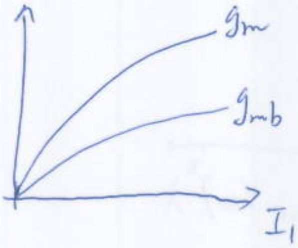


Complete mos small-signal model.



Eg. Sketch g_m and g_{mb} of M_1 in Fig. below as a function of the bias current I_1.





$g_m = \sqrt{2\mu_n C_{ox}(W/L) I_D} \Rightarrow g_m \propto \sqrt{I_1}$

For g_{mb}, less straight forward

As I_1 increases, V_X decreases and so does V_{SB}