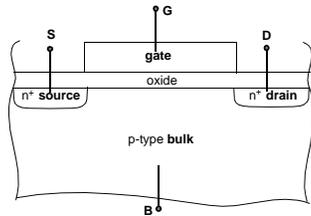


The MOS Field Effect Transistor

Cross-section:



Basic idea: add n-type regions adjacent to the MOS capacitor so that current can flow between them only when the surface is inverted

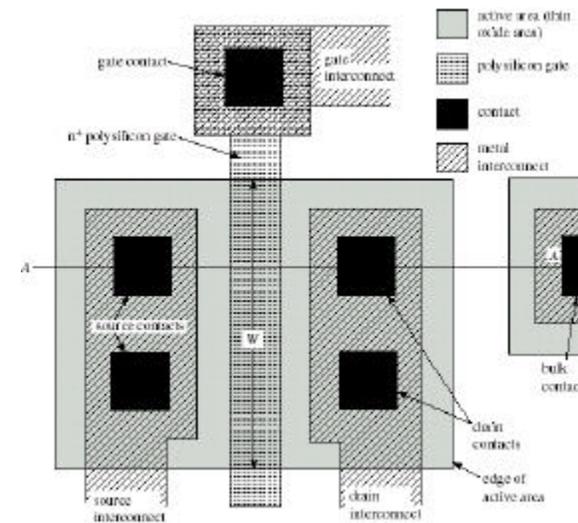
Invented 1930s, demonstrated 1960s, mass produced digital ICs 1970s

What's hard about making a MOSFET?

Oxide can contain a large amount of fixed and (worse) slightly mobile charge ... making the threshold voltage drift during use

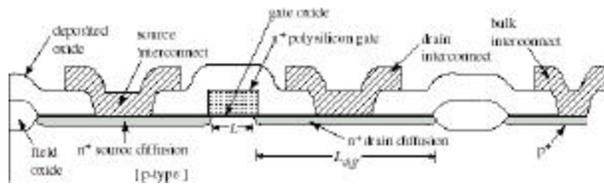
Four terminals control the electrical properties of the MOSFET

n-channel MOSFET Layout



contact to bulk (also called the "body") is made on the surface of the chip; the back of the chip is a "common" contact for all n-channel MOSFET in this process

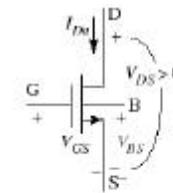
n-channel MOSFET Cross Section



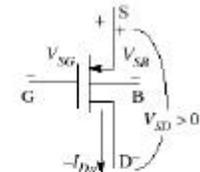
channel length is the separation of the n-type source and drain regions

the length of the source and drain regions is L_{diff}

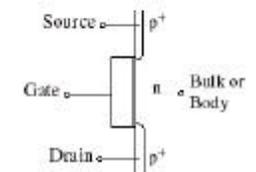
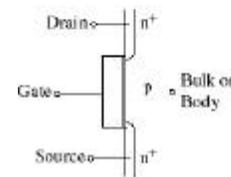
MOSFET Symbols (Analog IC Design)



n-channel MOSFET

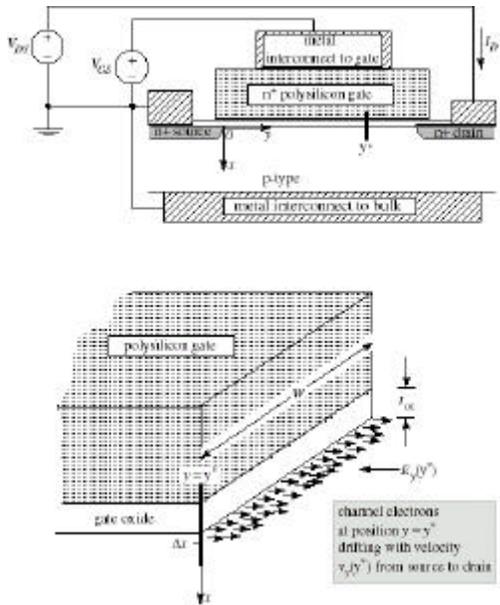


p-channel MOSFET



alternative symbols (used in digital ICs) -- not used in EE 105, but are used in Chapter 4 of textbook

Drain Current in the MOSFET



Drain current $I_D = ?$

Drift Current Equation

- Drift current for electrons in the channel:

$$J_y(x, y) = -qn(x, y)v_y(y)$$

The drain current at position y is the integral of the drift current density across the cross section. Since the conventional direction of I_D is *opposite* to the direction of the y axis, we insert a minus sign:

$$I_D = -W \int_0^{\Delta x} J_y(x, y) dx = Wv_y(y) \left(\int_0^{\Delta x} qn(x, y) dx \right)$$

- The integral is the negative of the electron charge in the channel, per unit area, at point y . The symbol for this quantity is $-Q_N(y)$:

$$I_D = -Wv_y(y)Q_N(y)$$

Note that I_D isn't a function of the position in the channel

Triode Region (Cont.)

- Next, approximate the average channel charge by averaging $Q_N(y=0)$ at the source end and $Q_N(y=L)$ at the drain end of the channel:

$$Q_N(y=0) = -C_{ox}(V_{GS} - V_{Tn})$$

At the drain end, the positive drain voltage *reduces* the magnitude of the channel charge ... why? The effect can be approximated by using V_{GD} (the drop from drain to channel, at $y = L$) --

$$Q_N(y=L) = -C_{ox}(V_{GD} - V_{Tn}) = -C_{ox}(V_{GS} - V_{DS} - V_{Tn})$$

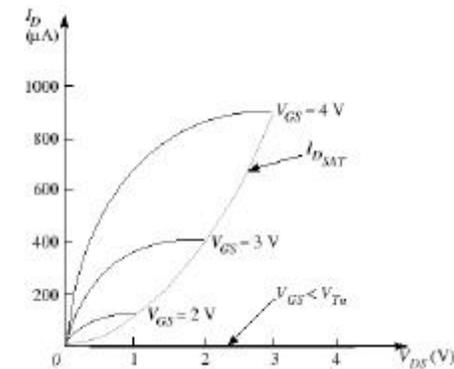
Note that $V_{GD} = V_{GS} - V_{DS} > V_{Tn}$ in order for there to be a channel left at the drain end.

- Substituting, we derive the equation for the triode region, which is defined by $V_{GS} - V_{DS} > V_{Tn}$ and $V_{GS} > V_{Tn}$.

$$I_D = \mu_n C_{ox} \left(\frac{W}{L} \right) (V_{GS} - V_{Tn} - V_{DS}/2) V_{DS}$$

Drain Characteristics

- Example: $\mu_n C_{ox} (W/L) = 50 \mu\text{A}/\text{V}^2$, $V_{Tn} = 1 \text{ V}$, and $(W/L) = 4$.



- What happens when $V_{DS} > V_{GS} - V_{Tn} = V_{DS(SAT)}$? $|Q_N(y=L)| = 0!$

Initial thought is that the lack of a channel at the drain end means that I_D must drop to zero ... WRONG!

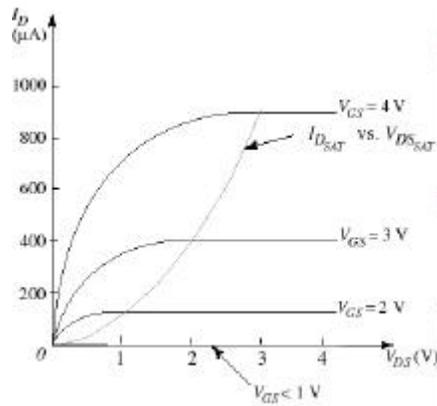
Drain terminal loses control over channel, so the drain current “saturates” and remains constant (to first approximation) at the value given by $V_{DS} = V_{DS(SAT)}$.

Saturation Region

- When $V_{GS} > V_{Tn}$ and $V_{DS} > V_{DS(SAT)} = V_{GS} - V_{Tn}$, the drain current is:

$$I_D = I_{D_{SAT}} = \mu_n C_{ox} \left(\frac{W}{2L} \right) (V_{GS} - V_{Tn})^2$$

- n-channel MOSFET drain characteristics:



MOSFET Circuit Models

- n-channel MOSFET drain current in cutoff, triode, and saturation:

$$I_D = 0 \text{ A} \quad (V_{GS} \leq V_{Tn})$$

$$I_D = \mu_n C_{ox} (W/L) [V_{GS} - V_{Tn} - (V_{DS}/2)] (1 + \lambda_n V_{DS}) V_{DS} \quad (V_{GS} \geq V_{Tn}, V_{DS} \leq V_{GS} - V_{Tn})$$

$$I_D = \mu_n C_{ox} (W/(2L)) (V_{GS} - V_{Tn})^2 (1 + \lambda_n V_{DS}) \quad (V_{GS} \geq V_{Tn}, V_{DS} \geq V_{GS} - V_{Tn})$$

Numerical values:

μ_n is a function of V_{GS} along the channel and is much less than the mobility in the bulk (typical value $200 \text{ cm}^2/(\text{Vs})$) -- therefore, we consider that $\mu_n C_{ox}$ is a measured parameter. Typical value: $\mu_n C_{ox} = 50 \mu\text{AV}^{-2}$

λ_n , sometimes called the channel length modulation parameter, increases as the channel length L is reduced:

$$\lambda_n \approx \frac{0.1 \mu\text{mV}^{-1}}{L}$$

The triode region I_D equation has $(1 + \lambda_n V_{DS})$ added in order to avoid a jump at the boundary with the saturation region. For hand calculation of DC voltages and currents, this term is usually omitted from I_D .

V_{Tn} = threshold voltage = 0.7 - 1.0 V typically for an n-channel MOSFET.

Backgate Effect

- The MOSFET has four (G, S, D, and B) electrical terminals
- If $V_{BS} \neq 0$, then X_d increases, for example X_{DMAX} :

$$\therefore \sqrt{\frac{2\epsilon(-2\phi_p)}{qNa}} \rightarrow \sqrt{\frac{2\epsilon(-2\phi_p + V_{SB})}{qNa}}$$

Remember: $V_T = V_{FB} + V_{Si} + V_{ox}$

$$\begin{array}{c} \downarrow \quad \downarrow \\ -2\phi_p \quad -qNaX_d \\ \quad \quad \quad \uparrow \\ \quad \quad \quad C_{ox} \end{array}$$

so V_T increases $\xrightarrow{\hspace{2cm}}$

$$\underbrace{V_{FB} - 2\phi_p + \frac{\sqrt{2\epsilon q Na (-2\phi_p)}}{C_{ox}}}_{V_{TO}} \rightarrow$$

Define $\gamma = \frac{\sqrt{2q\epsilon_s Na}}{C_{ox}}$

Backgate Effect

- The threshold voltage is a function of the bulk-to-source voltage V_{BS} through the *backgate effect*.

$$V_{Tn} = V_{TOn} + \gamma_n (\sqrt{-V_{BS} - 2\phi_p} - \sqrt{-2\phi_p})$$

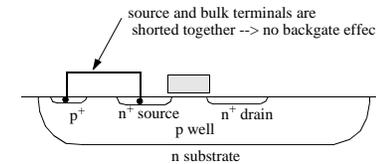
where V_{TO} is the threshold voltage with $V_{BS} = 0$ and γ is the backgate effect parameter

$$\gamma_n = (\sqrt{2q\epsilon_s Na}) / C_{ox}$$

- Physical origin: V_{BS} (a negative voltage to avoid forward biasing the bulk-to-source pn junction) increases the depletion width, which increases the bulk charge and thus, the threshold voltage.

$$I_D = I_D(V_{GS}, V_{DS}, V_{BS}) \text{ since } V_{Tn} = V_{Tn}(V_{BS})$$

Common situation is that $V_{BS} = 0$ by electrically shorting the source to the bulk (either the substrate or a deep diffused region called a *well*)



$$\text{If } V_{BS} = 0, V_{Tn} = V_{TOn}.$$

MOSFET Small-Signal Model

- Concept: find an equivalent circuit which interrelates the incremental changes in i_D , v_{GS} , v_{DS} , etc. Since the changes are small, the small-signal equivalent circuit has linear elements only (e.g., capacitors, resistors, controlled sources)
- Derivation: consider for example the relationship of the increment in drain current due to an increment in gate-source voltage when the MOSFET is saturated-- with *all other voltages held constant*.

$$v_{GS} = V_{GS} + v_{gs}, i_D = I_D + i_d \text{ -- we want to find } i_d = (?) v_{gs}$$

We have the functional dependence of the total drain current in saturation:

$$i_D = \mu_n C_{ox} (W/2L) (v_{GS} - V_{Tn})^2 (1 + \lambda_n v_{DS}) = i_D(v_{GS}, v_{DS}, v_{BS})$$

Do a Taylor expansion around the DC operating point (also called the quiescent point or Q point) defined by the DC voltages $Q(V_{GS}, V_{DS}, V_{BS})$:

$$i_D = I_D + \left. \frac{\partial i_D}{\partial v_{GS}} \right|_Q (v_{gs}) + \frac{1}{2} \left. \frac{\partial^2 i_D}{\partial v_{GS}^2} \right|_Q (v_{gs})^2 + \dots$$

If the small-signal voltage is really “small,” then we can neglect all everything past the linear term --

$$i_D = I_D + \left. \frac{\partial i_D}{\partial v_{GS}} \right|_Q (v_{gs}) = I_D + g_m v_{gs}$$

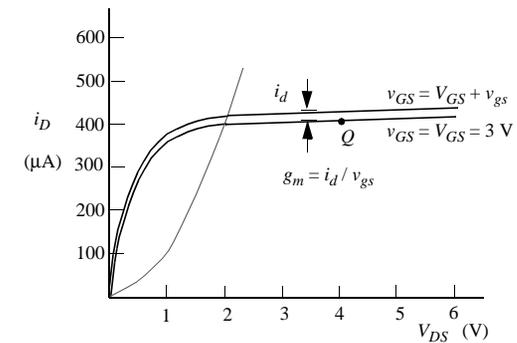
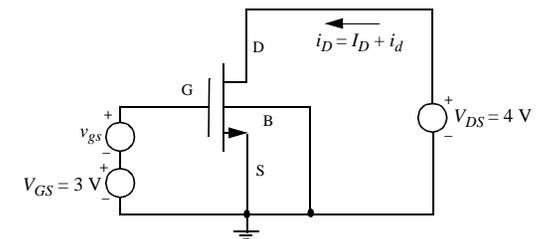
where the partial derivative is defined as the *transconductance*, g_m . That is,

$$g_{mSAT} \equiv \frac{\partial i_{DSAT}}{\partial v_{GS}}$$

Transconductance

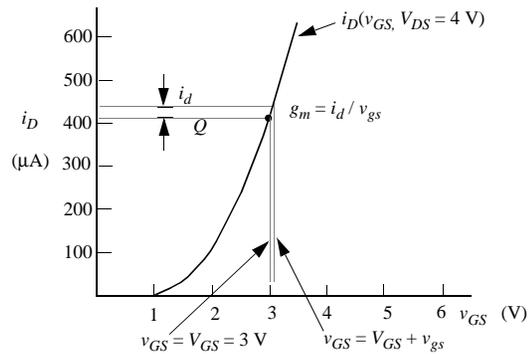
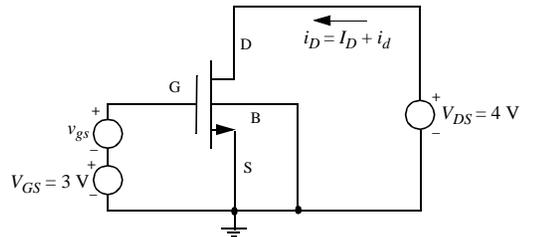
The small-signal drain current due to v_{gs} is therefore given by

$$i_d = g_m v_{gs}$$



Another View of g_m

* Plot the drain current as a function of the gate-source voltage, so that the slope can be identified with the transconductance:



Transconductance (Cont.)

- Evaluating the partial derivative:

$$g_m = \mu_n C_{ox} \left(\frac{W}{L}\right) (V_{GS} - V_{Tn})(1 + \lambda_n V_{DS})$$

Note that the transconductance is a function of the operating point, through its dependence on V_{GS} and V_{DS} -- and also the dependence of the threshold voltage on the backgate bias V_{BS} .

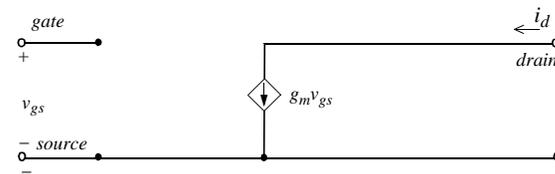
- In order to find a simple expression that highlights the dependence of g_m on the DC drain current, we neglect the (usually) small error in writing:

$$g_m = \sqrt{2\mu_n C_{ox} \left(\frac{W}{L}\right) I_D} =$$

For typical values $(W/L) = 10$, $I_D = 100 \mu\text{A}$, and $\mu_n C_{ox} = 50 \mu\text{AV}^{-2}$ we find that

$$g_m = 320 \mu\text{AV}^{-1} = 0.32 \text{ mS}$$

- How do we make a circuit which expresses $i_d = g_m v_{gs}$? We need a voltage-controlled current source:



Output Conductance

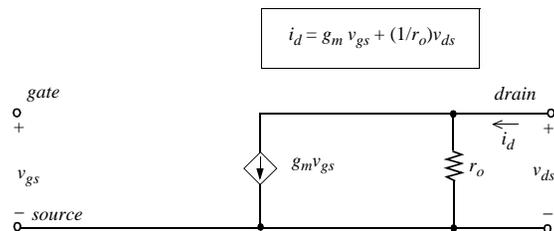
- We can also find the change in drain current due to an increment in the drain-source voltage:

$$g_o = \left. \frac{\partial i_D}{\partial v_{DS}} \right|_Q = \mu_n C_{ox} \left(\frac{W}{2L} \right) (V_{GS} - V_T)^2 \lambda_n \equiv \lambda_n I_D$$

The output resistance is the inverse of the output conductance

$$r_o = \frac{1}{g_o} = \frac{1}{\lambda_n I_D}$$

The small-signal circuit model with r_o added looks like:



Backgate Transconductance

- We can find the small-signal drain current due to a change in the backgate bias by the same technique. The chain rule comes in handy to make use of our previous result for g_m :

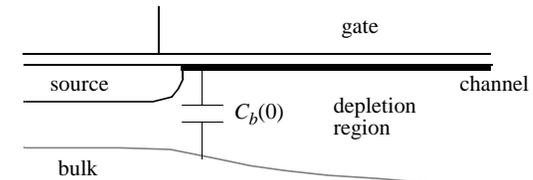
$$g_{mb} = \left. \frac{\partial i_D}{\partial v_{BS}} \right|_Q = \left. \frac{\partial i_D}{\partial V_T} \right|_Q \left. \frac{\partial V_T}{\partial v_{BS}} \right|_Q$$

$$g_{mb} = (-g_m) \left. \frac{\partial V_T}{\partial v_{BS}} \right|_Q = (-g_m) \left(\frac{-\gamma_n}{2\sqrt{-2\phi_p - V_{BS}}} \right) = \frac{\gamma_n g_m}{2\sqrt{-2\phi_p - V_{BS}}}$$

The ratio of the “front-gate” transconductance g_m to the backgate transconductance g_{mb} is:

$$\frac{g_{mb}}{g_m} = \frac{\sqrt{2q\epsilon_s N_a}}{2C_{ox}\sqrt{-2\phi_p - V_{BS}}} = \frac{1}{C_{ox}} \sqrt{\frac{q\epsilon_s N_a}{2(-2\phi_p - V_{BS})}} = \frac{C_b(y=0)}{C_{ox}}$$

where $C_b(y=0)$ is the depletion capacitance at the source end of the channel --



MOSFET Capacitances in Saturation

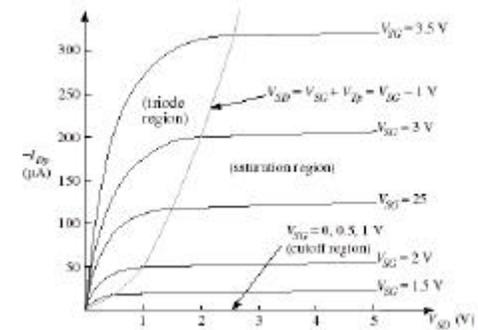
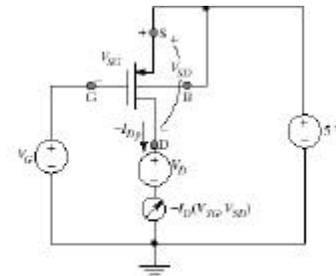
- There are *lots* of capacitances built into the MOSFET
- Gate-source capacitance:
the “wedge” channel charge in saturation: less charge than for a MOS capacitor
see text: $C_{gs} = (2/3) C_{ox} WL +$ “overlap” capacitance

$$C_{ov} = C_{ox} (W\Delta)$$

- Gate-drain capacitance:
 $C_{gd} =$ overlap capacitance
- Parasitic depletion capacitances: drain-to-bulk, source-to-bulk
- “Complete” model:

p-channel MOSFET

Test circuit for measuring drain characteristics



p-channel MOSFET Models

- DC drain current in the three operating regions: $-I_{Dp} > 0$

$$\begin{aligned}
 -I_{Dp} &= 0 \text{ A} && (V_{SG} \leq -V_T) \\
 -I_{Dp} &= \mu_p C_{ox} (W/L) [V_{SG} + V_{Tp} - (V_{SD}/2)] (1 + \lambda_p V_{SD}) V_{SD} && (V_{SG} \geq -V_{Tp}, V_{SD} \leq V_{SG} + V_{Tp}) \\
 -I_{Dp} &= \mu_p C_{ox} (W/(2L)) (V_{SG} + V_{Tp})^2 (1 + \lambda_p V_{SD}) && (V_{SG} \geq -V_{Tp}, V_{SD} \geq V_{SG} + V_{Tp})
 \end{aligned}$$

- The threshold voltage with backgate effect is given by:

$$V_{Tp} = V_{TOP} - \gamma_p (\sqrt{-V_{SB} + 2\phi_n} - \sqrt{2\phi_n})$$

Numerical values:

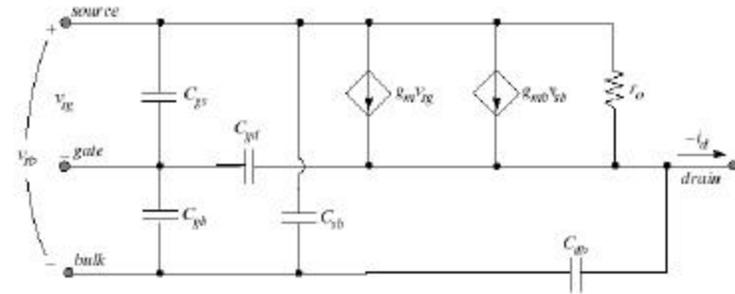
$\mu_p C_{ox}$ is a measured parameter. Typical value: $\mu_p C_{ox} = 25 \mu\text{A V}^{-2}$

$$\lambda_p \approx \frac{0.1 \mu\text{m V}^{-1}}{L}$$

$V_{Tp} = -0.7$ to -1.0 V, which should be approximately $-V_{Tn}$ for a well-controlled CMOS process

PMOS Small-Signal Model

Same functional forms as for NMOS ... control voltages are v_{sg} , v_{sd} , v_{sb}



Source is located at the top (corresponding to its position on the schematic as the higher potential compared to the drain); however, the source is often connected to a DC voltage --> a short for small signals.

Circuit Simulation

Objectives:

- fabricating an IC costs \$1000 ... \$100,000 per run
--> nice to get it "right" the first time
- check results from hand-analysis
(e.g., validity of assumptions)
- evaluate functionality, speed, accuracy, ... of large circuit blocks or entire chips

Simulators:

- SPICE**: invented at UC Berkeley circa 1970-1975
commercial versions: HSPICE, PSPICE, I-SPICE, ... (same core as Berkeley SPICE, but add functionality, improved user interface, ...)
EE 105: student version of PSPICE on PC, limited to about 15 transistors
- other simulators for higher speed, special needs (e.g. SPLICE, RSIM)

Limitations:

- simulation results provide no insight (e.g. how to increase speed of circuit)
- results sometimes wrong (errors in input, effect not modeled in SPICE)

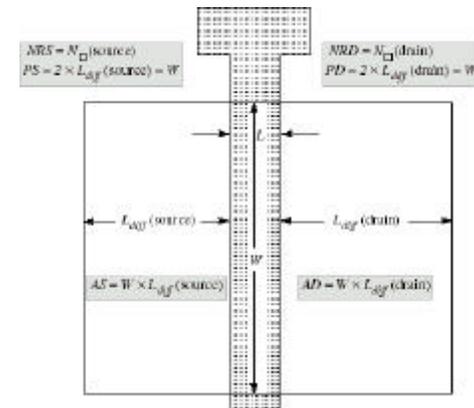
====> always do hand-analysis first and COMPARE RESULTS

MOSFET Geometry in SPICE

- Statement for MOSFET ... D, G, S, B are node numbers for drain, gate, source, and bulk terminals

$Mname\ D\ G\ S\ B\ MODname\ L=_\ W=_\ AD=_\ AS=_\ PD=_\ PS=_$

$MODname$ specifies the model name for the MOSFET



MOSFET Model Statement

.MODEL *MODname* {NMOS or PMOS} *VTO*=_ *KP*=_ *GAMMA*=_ *PHI*=_
LAMBDA=_ *RD*=_ *RS*=_ *RSH*=_ *CBD*=_ *CBS*=_ *CJ*=_ *MJ*=_ *CJSW*=_
MJSW=_ *PB*=_ *IS*=_ *CGDO*=_ *CGSO*=_ *CGBO*=_ *TOX*=_ *LD*=_

Parameter name (SPICE / this text)	SPICE symbol	Analytical symbol	Units
channel length	<i>L_{eff}</i>	<i>L</i>	m
polysilicon gate length	<i>L</i>	<i>L_{gate}</i>	m
lateral diffusion/ gate-source overlap	<i>LD</i>	<i>L_D</i>	m
transconductance parameter	<i>KP</i>	$\mu_n C_{ox}$	A/V ²
threshold voltage / zero-bias threshold	<i>VTO</i>	<i>V_{Th0}</i>	V
channel-length modulation parameter	<i>LAMBDA</i>	λ_n	V ⁻¹
bulk threshold / backgate effect parameter	<i>GAMMA</i>	γ_n	V ^{1/2}
surface potential / depletion drop in inversion	<i>PHI</i>	$-\phi_p$	V

NMOS SPICE DC Drain Current Equations:

$$I_{DS} = 0 \quad (V_{GS} \leq -V_{TH})$$

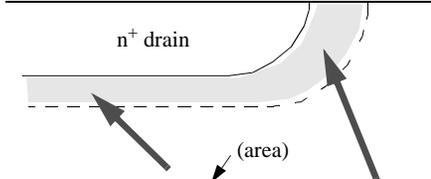
$$I_{DS} = \frac{KP}{2} (W/L_{eff}) V_{DS} [2(V_{GS} - V_{TH}) - V_{DS}] (1 + LAMBDA \cdot V_{DS}) \quad (0 \leq V_{DS} \leq V_{GS} - V_{TH})$$

$$I_{DS} = \frac{KP}{2} (W/L_{eff}) (V_{GS} - V_{TH})^2 (1 + LAMBDA \cdot V_{DS}) \quad (0 \leq V_{GS} - V_{TH} \leq V_{DS})$$

$$V_{TH} = V_{TO} + GAMMA (\sqrt{2 \cdot PHI - V_{BS}} - \sqrt{2 \cdot PHI})$$

Capacitances

SPICE includes the “sidewall” capacitance due to the perimeter of the source and drain junctions --



$$C_{BD}(V_{BD}) = \frac{CJ \cdot AD}{(1 - V_{BD}/PB)^{MJ}} + \frac{CJSW \cdot PD}{(1 - V_{BD}/PB)^{MJSW}}$$

Gate-source and gate-bulk overlap capacitance are specified by *CGDO* and *CGSO* (units: F/m).

Level 1 MOSFET model (for 3 μm channel length NMOS device)

.MODEL MODN NMOS LEVEL=1 *VTO*=1 *KP*=50U *LAMBDA*=.033 *GAMMA*=.6
+ *PHI*=0.8 *TOX*=1.5E-10 *CGDO*=5E-10 *CGSO*= 5e-10 *CJ*=1E-4 *CJSW*=5E-10
+ *MJ*=0.5 *PB*=0.95

The Level 1 model is adequate for channel lengths longer than about 1.5 μm -- see the MOSFET experiment for an example of its limitations

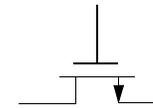
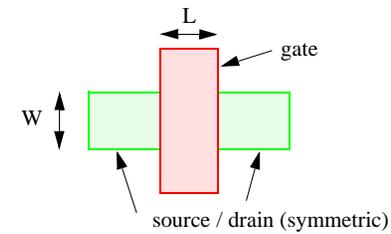
For sub-μm MOSFETs, BSIM = “Berkeley Short-Channel IGFET Model” is the industry-standard SPICE model.

CMOS Layers

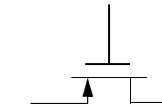
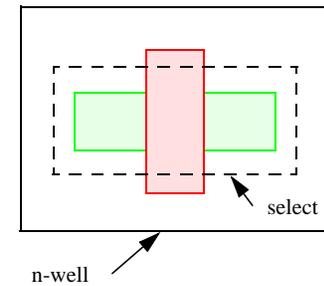
Layer	Representation	Color Convention (EE 105)
n-well		purple
active		green
select p+		brown
polysilicon		red
metal		blue
contact		black

Active Devices

- polysilicon crossing active results in an NMOS device:



- PMOS devices are placed in n-wells:



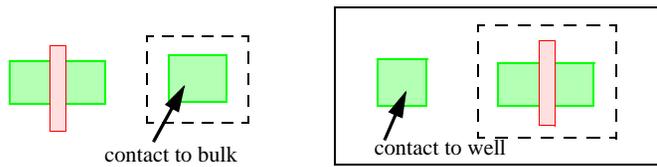
The select mask is used twice:

clear field: photoresist masks the heavy n-type source/drain implant for the NMOS transistors

dark field: photoresist masks the heavy p-type source/drain implant for the PMOS transistors

Bulk and Well Contacts

- use p-doped *active* (*select* mask) as contact to the bulk
- use n-doped *active* (no *select* mask) as a contact to *n*-wells



Layout Rules (EE 105 n-well CMOS Technology)

- minimum dimensions and separations (in μm , not to scale):

