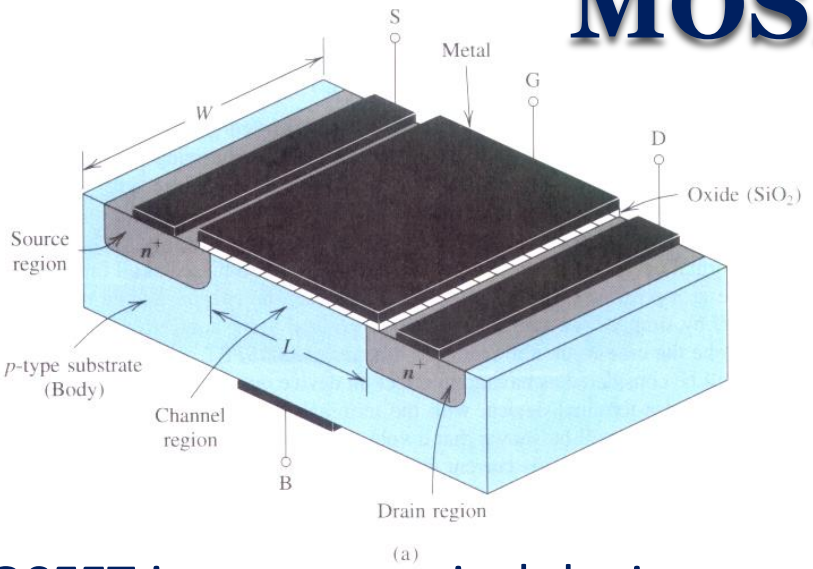
System on a Chip

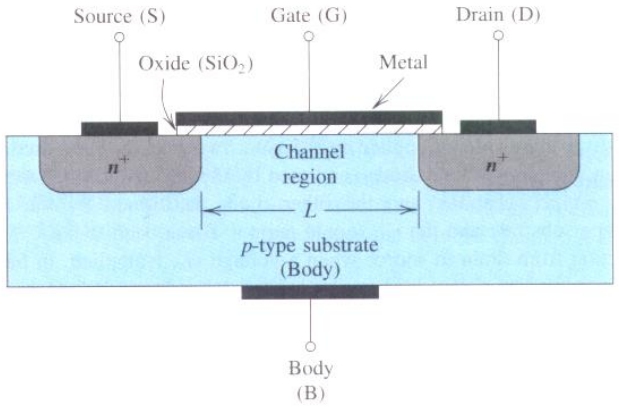
Prof. Dr. Michael Kraft

Lecture 2:
Revision: MOSFET Operation and Modelling

- Basic MOSFET Operation
- Small Signal Model
- Spice Models
- Short Channel Effects

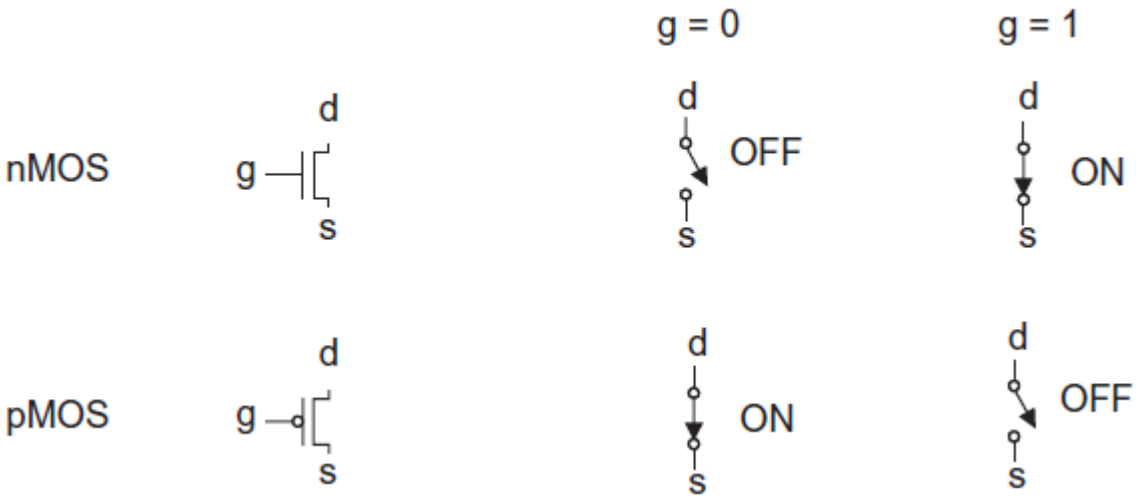
Active Components
MOSFETS





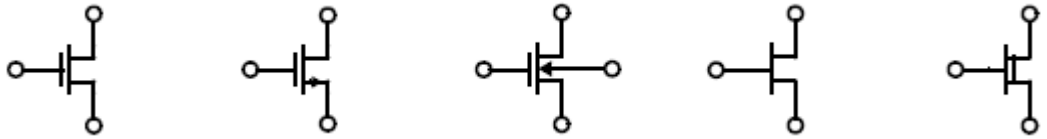
- MOSFET is a symmetrical device.
- Metal on the gate is often replaced by polysilicon.
- L is the parameter characterising a process technology. Modern processes have a gate length down to 0.065µm for digital circuits. Analogue circuits ~0.35-1µm.
- Gate oxide thickness is in the order of a few nm.
- The ITRS predicts gate lengths of 13nm by 2013 and 6nm by 2020!!!
- For $v_{GS} = 0$, the source and drain regions are separated by back-to-back *pn* junctions resulting in an extremely high resistance (about 10^{12} Ω)

MOSFET As A Switch

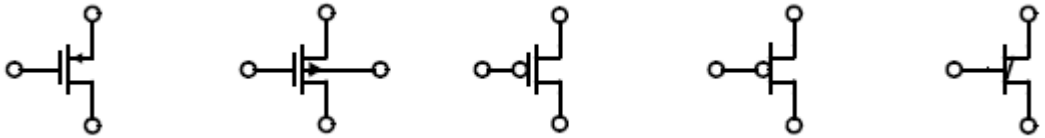


- For digital operation, the MOSFET is modelled as an ON/OFF switch
  - g=1 (High) represents the positive supply voltage (V_{DD}) applied to the gate
  - Typical values from 5V (older technologies) to 1.0V (newer technologies)

MOSFET Symbols



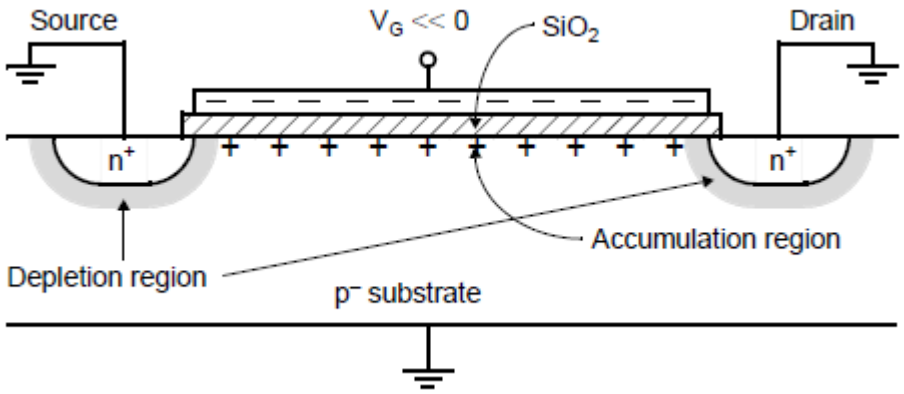
n-channel transistor symbols



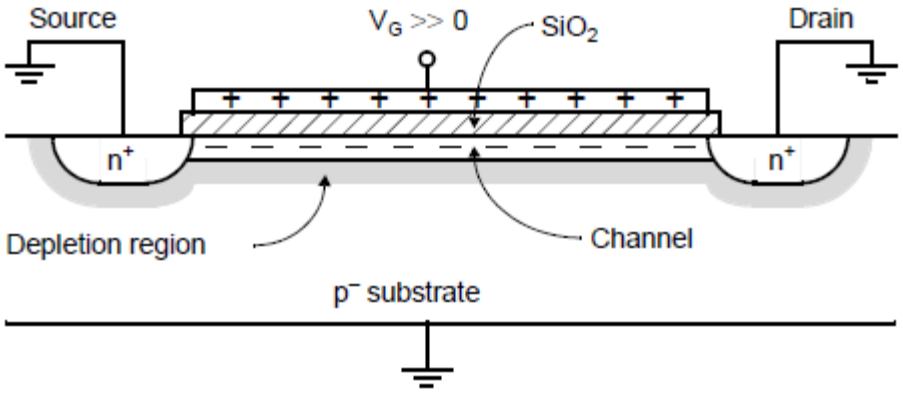
P-channel transistor symbols

- All are enhancement transistors (normally off)
  - Depletion mode transistors are not used anymore (normally on)

Basic Operation



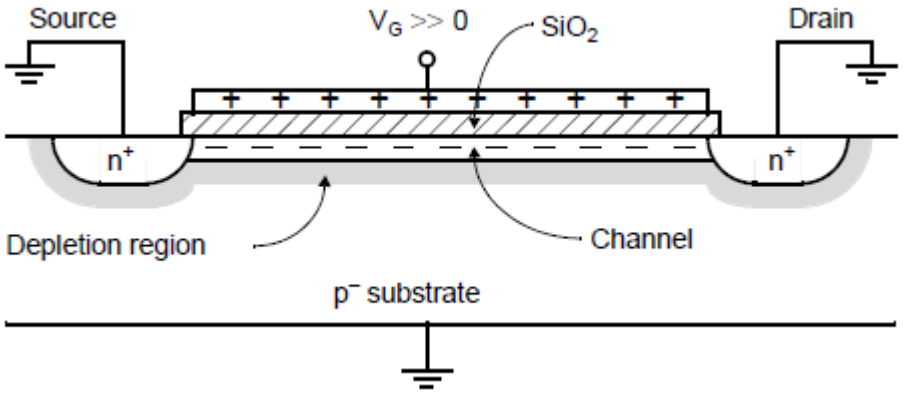
V_G<0: accumulated channel



V_G>0: inverted channel, Current flow gate drain possible

- Gate – source voltage for which the concentration of electrons is equal to concentration of holes in the substrate is called the **Threshold voltage, V_{th}**
- Charge density in channel is proportional to $V_{eff}=V_{GS}-V_{th}$

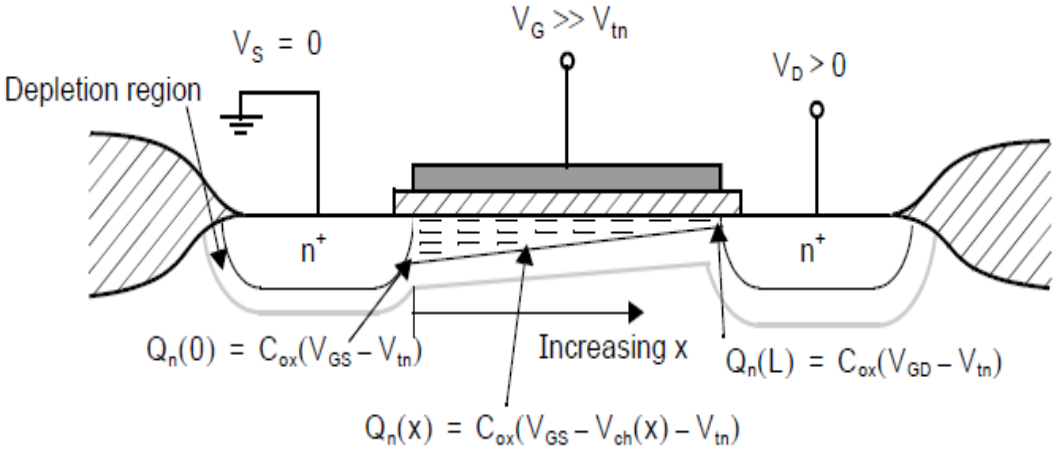
Basic Operation
Triode Region

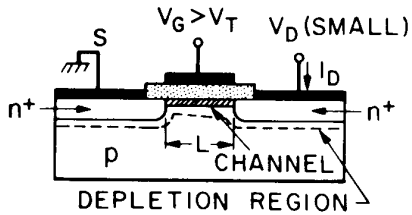


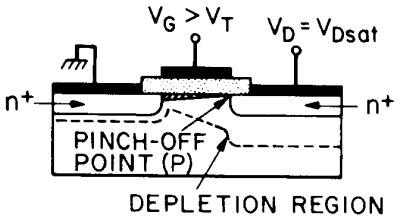
V_G>0: inverted channel, Current flow gate drain possible

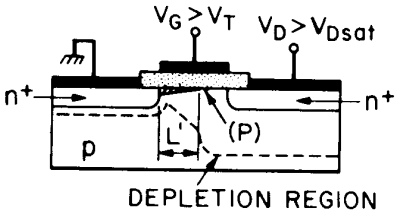
- Charge density (charge per unit area): $Q_n = C_{ox}(V_{GS}-V_{th}) = C_{ox}V_{eff}$
- C_{ox} is the **Gate Capacitance per unit area**: $C_{ox} = \frac{\varepsilon_{ox}\varepsilon_0}{t_{ox}}$
  - ε_{ox} is the relative permittivity of SiO_2
  - t_{ox} the SiO_2 thickness
- Total charge in channel: $Q_T = WLC_{ox}(V_{GS} - V_{th})$
- For $V_{DS} > 0$ but small ($V_{DS} << V_{eff}$): $I_D = \mu_n Q_n \frac{W}{L} V_{DS} = \mu_n C_{ox}(V_{GS} - V_{th}) \frac{W}{L} V_{DS}$
  - μ_n: mobility of electrons near surface (0.14m2/Vs in intrinsic Si, 0.01-0.06m^2/Vs in modern NMOS devices
  - Behaves like a resistor

Saturation Region



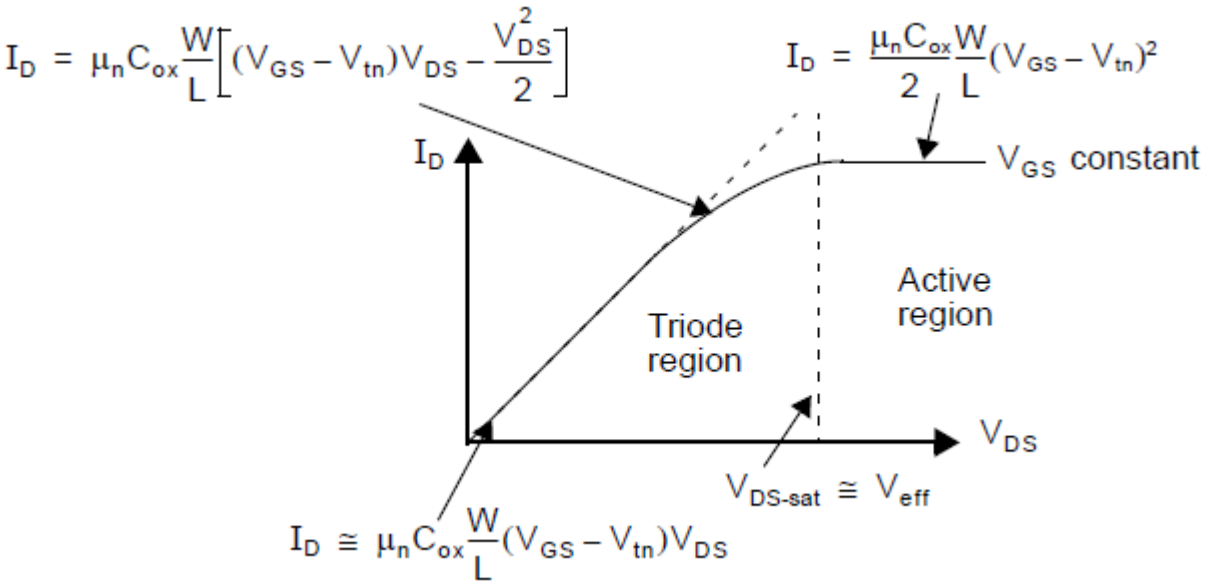






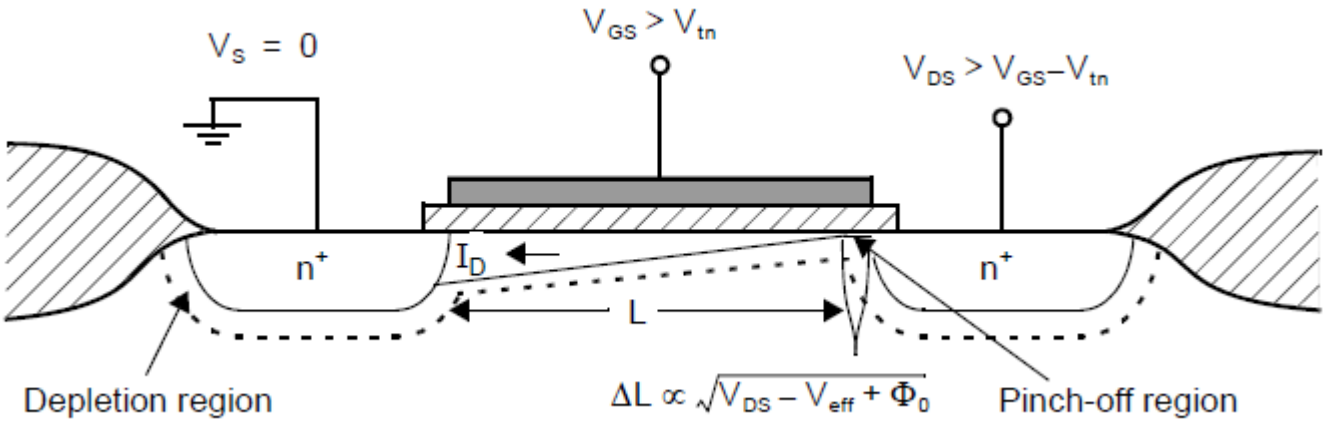
- For V_{DS} increasing, the channel charge at the drain end decreases
  - Voltage across the gate oxide is smaller at the drain end
  - Charge density has a tapered shaped
  - Charge density at x: $Q_n(x)=C_{ox}(V_{GS}-V_{ch}(x)-V_{th})$
  - At drain end: $V_G-V_{ch}(L)=V_{GD}$
- For V_{DS} further increasing, the gate to channel voltage will become smaller than the V_{th} at the drain end → channel is pinched off
  - I_D saturates
  - Pinch off occurs at $V_{DS,sat} = V_{GS}-V_{th} = V_{eff}$

Operating Regions



- Active or saturation region: mainly used for analogue circuits (amplifiers)
  - Bias transistor so that V_{GS}-V_{th}>100mV
  - Square law relationship (for BJT: exponential relationship)
  - Drain current is impendent of V_{DS}
    - This is only true as a first order approximation

Channel Length Modulation

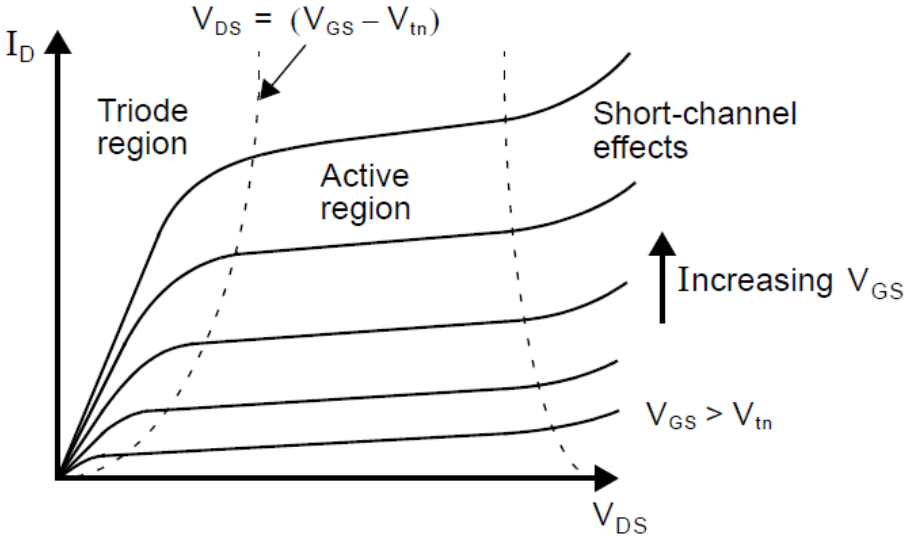


- For V_{DS} even further increasing, the point where the gate to channel voltage is equal to V_{th} moves to towards the source
  - The effective channel is shortened: **channel length modulation**
  - It can be shown that:

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

  - λ: channel length modulation parameter or output impedance constant (unit: V^{-1})

Channel Length Modulation



- Channel length modulation introduces a dependence of I_D on V_{DS} in the active region
  - Output resistance: $r_0 = \left(\frac{\partial I_D}{\partial V_{DS}} \right)^{-1} = \frac{1}{\lambda I_D}$
  - for even high values of V_{DS} other second order effects (short channel effects) dominate and cause a stronger dependence of I_D on V_{DS}

Threshold Voltage

$$V_{tno} = \Phi_{GS} - Q_{ss}/C_{OX} + 2\Phi_F + \gamma \, \mathrm{sqrt}(2\Phi_F)$$

Φ_{GS}: work function difference, gate/substrate materials
Φ_F: Fermi potential of substrate, ≈0.35V
K_S: relative permittivity of Si, 11.9

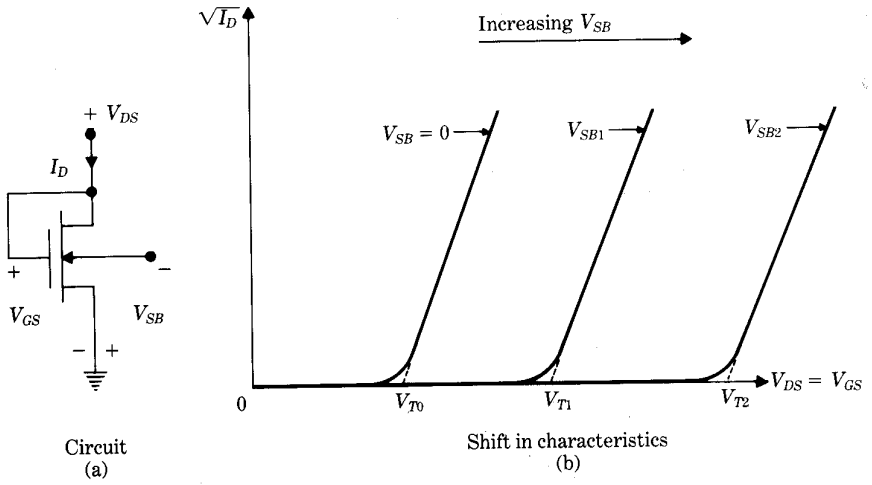
$$V_{tn} = V_{tn0} + \Delta V_{tn}$$

$$= V_{tn0} + \frac{\sqrt{2qN_AK_s\varepsilon_0}}{C_{ox}}\left[\sqrt{V_{SB} + |2\phi_F|} - \sqrt{|2\phi_F|}\right]$$

$$= V_{tn0} + \gamma\left(\sqrt{V_{SB} + |2\phi_F|} - \sqrt{|2\phi_F|}\right)$$

- The threshold voltage depends on:
  - The work-function difference between the gate material and the substrate material
  - The voltage drop between the channel and the substrate required for the channel to exist
  - The voltage drop across the thin oxide required for the depletion region, with its immobile charge, to exist
  - The voltage drop across the thin oxide due to unavoidable charge trapped in the thin oxide
  - The voltage drop across the thin oxide due to implanted charge at the surface of the silicon.

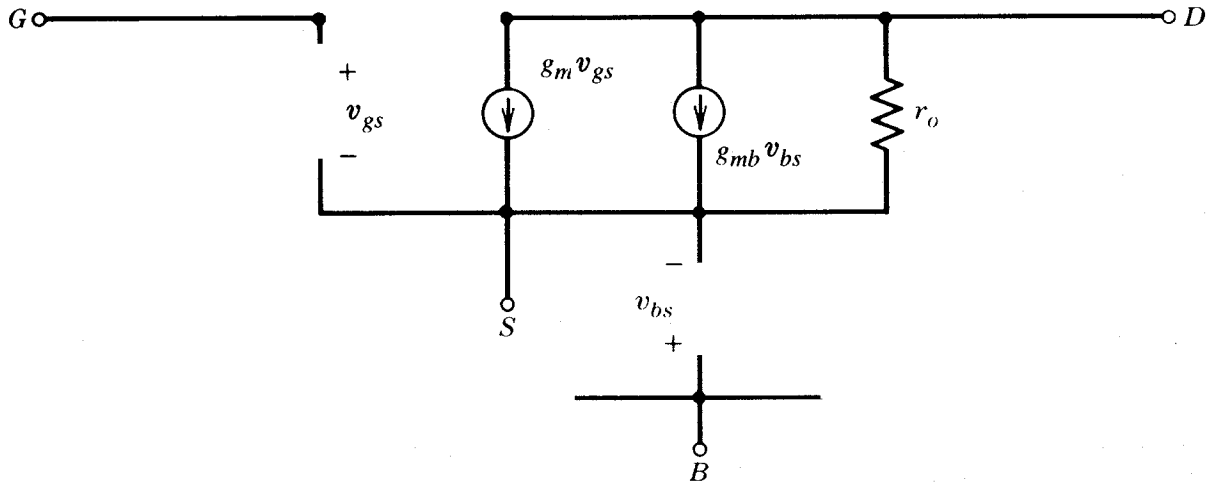
Body Effect



- The source - body voltage, V_{SB} influences the threshold voltage hence the drain current (sometimes the substrate is referred to as a second gate)
- For V_{SB} increasing, the depletion region between the channel and the substrate becomes wider, hence there is more charge
- Modelled by the Body effect constant:

$$\gamma = \frac{\sqrt{2qN_A K_s \varepsilon_0}}{C_{ox}}$$

Small Signal Model, Low Frequency



- Transconductance, g_m

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = k' \frac{W}{L}(V_{GS} - V_{th})(1 + \lambda V_{DS})$$

$$for \quad \lambda V_{DS} << 1$$

$$g_m = k' \frac{W}{L}(V_{GS} - V_{th}) = \sqrt{2k' \frac{W}{L} I_D}$$

$$with \quad k' = \mu_n C_{ox}$$

- Output resistance, r_o (or r_{ds}):

$$r_0 = \left(\frac{\partial I_D}{\partial V_{DS}} \right)^{-1} = \frac{1}{\lambda I_D}$$

Body Transconductance, g_{mb}

$$g_{mb} = \frac{\partial I_D}{\partial V_{SB}} = -k' \frac{W}{L}(V_{GS} - V_{th})(1 + \lambda V_{DS}) \frac{\partial V_T}{\partial V_{SB}}$$

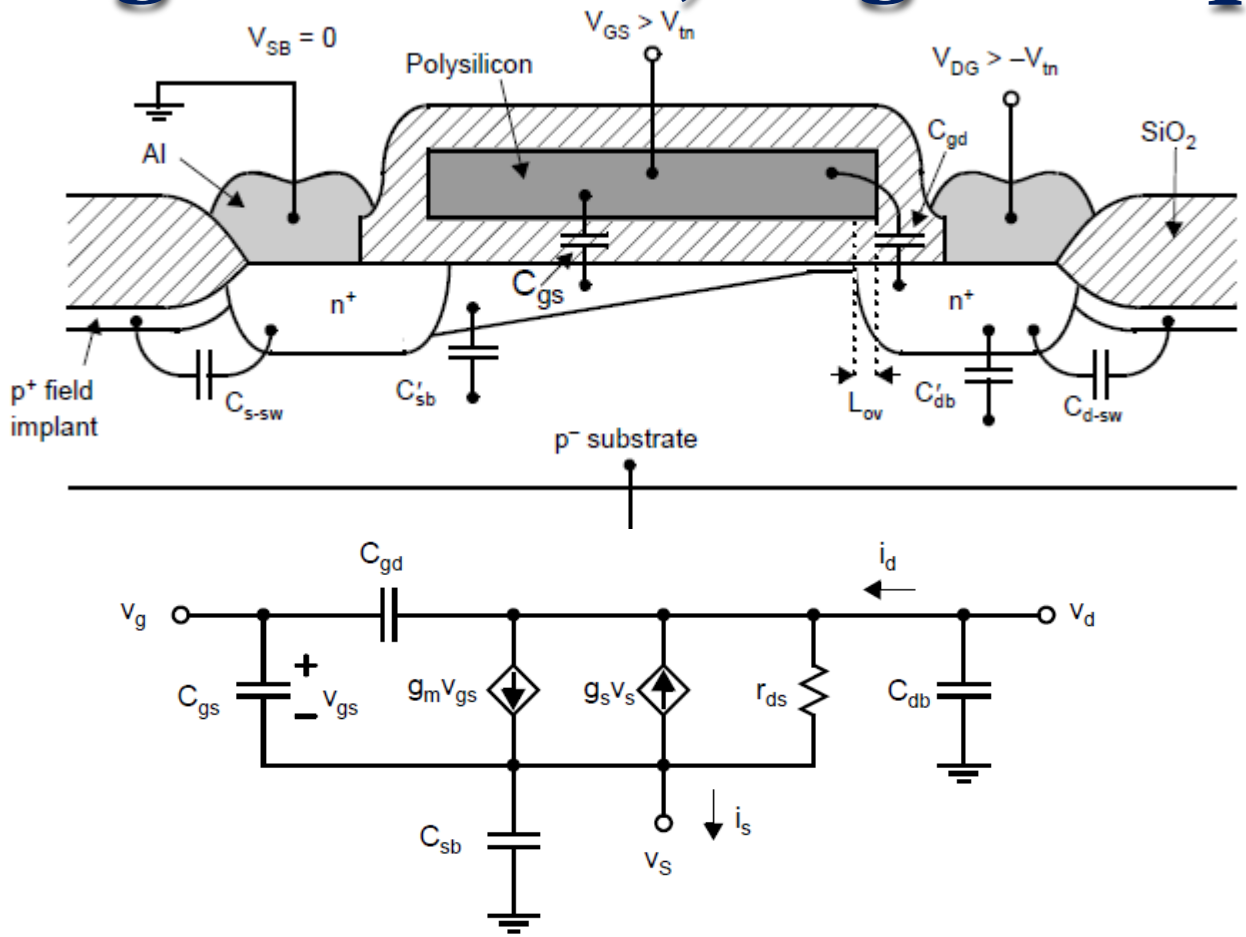
$$\frac{\partial V_T}{\partial V_{SB}} = -\frac{\gamma}{2\sqrt{2\phi_F + V_{SB}}} = -\chi$$

$$\chi = \frac{g_{mb}}{g_m} \qquad typical \ \chi = 0.1 \ldots 0.3$$

$$for \quad \lambda V_{DS} << 1$$

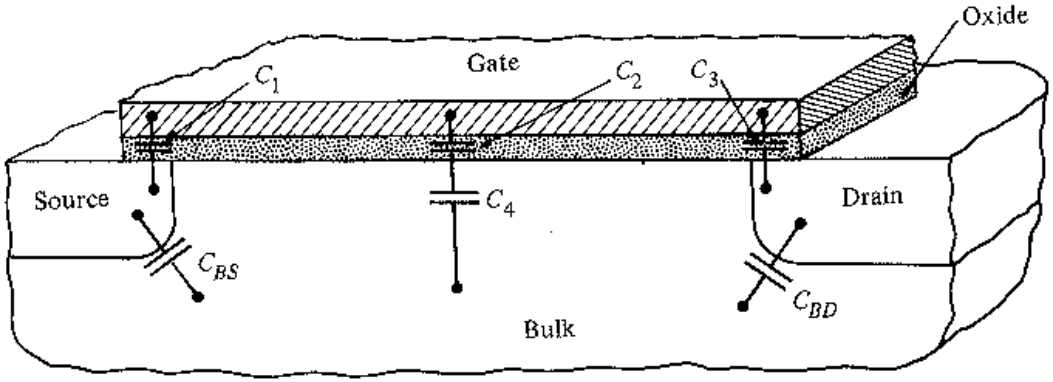
$$g_{mb} = \frac{\gamma \sqrt{k'(W/L) I_D}}{2\sqrt{2\phi_F + V_{SB}}}$$

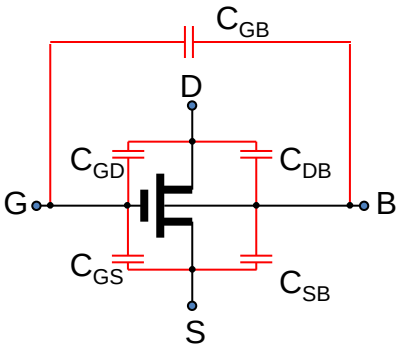
Small Signal Model, High Frequency



- For higher frequencies, capacitances need to be considered
  - The largest capacitances are $C_{gs}=2/3WLC_{ox}$ and the depletion capacitance at the source: $C'_{sb} = (A_s + A_{ch})C_{js}$ with: $$C_{js} = \frac{C_{j0}}{\sqrt{1 + \frac{V_{SB}}{\Phi_0}}}$$

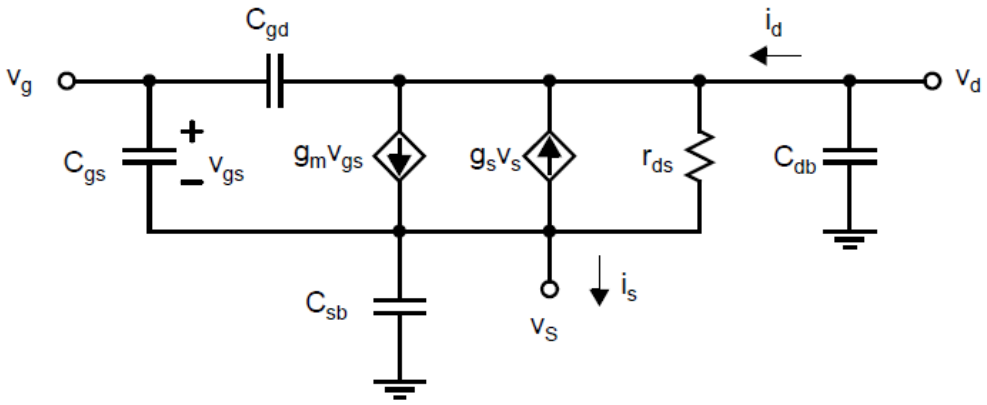
MOSFET Capacitors





- Intrinsic and extrinsic capacitances:
  - Intrinsic cap's are related to the electric field in the gate oxide which also forms the channel.
  - Extrinsic cap's are caused by parasitic effects.
- Intrinsic capacitances are associated with charge on the gate electrode and in the channel hence vary with the terminal voltages.
- Two intrinsic capacitances: gate capacitance and source-body and drain-body depletion capacitance (two reverse biased p-n junctions)

Small Signal Model, High Frequency



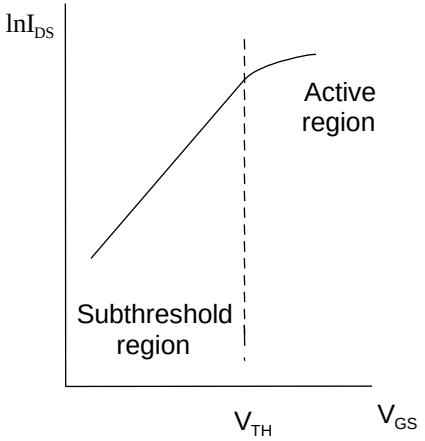
- C_{gd} is due to the overlap between the gate and the drain and fringing capacitance, $C_{gd}=C_{ox}WL_{ox}$
  - sometimes called the Miller capacitance, which is important when there is a voltage gain between gate and drain
- Figure of merit for transistor speed: *unity-gain frequency*

$$f_t \approx \frac{g_m}{2\pi(C_{gs}+C_{gd})}$$

MOSFET Capacitors

Region	C_{GS}	C_{GD}	C_{GB}
$V_{GS}<V_T$	0	0	WLC_{OX}
Triode region	$\frac{1}{2}WLC_{OX}$	$\frac{1}{2}WLC_{OX}$	0
Active region	$\frac{2}{3}WLC_{OX}$	0	0

Subthreshold (Weak Inversion)

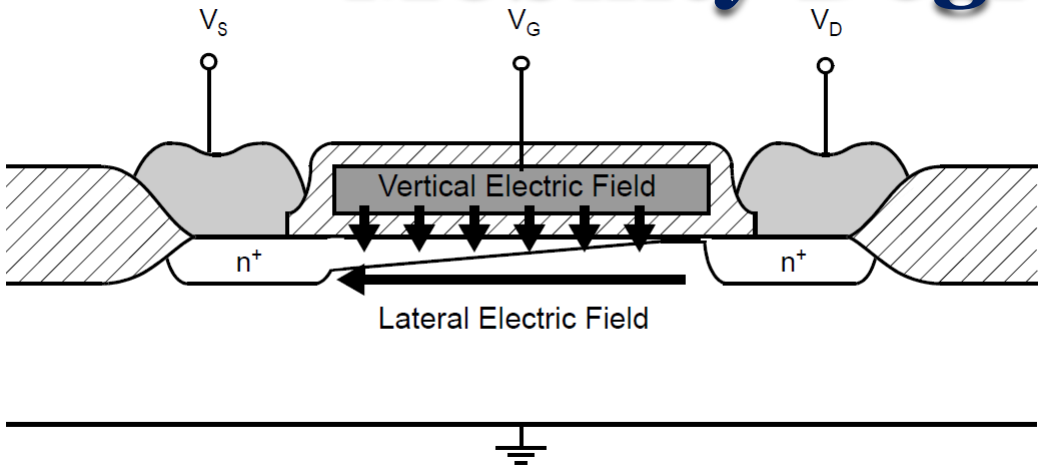


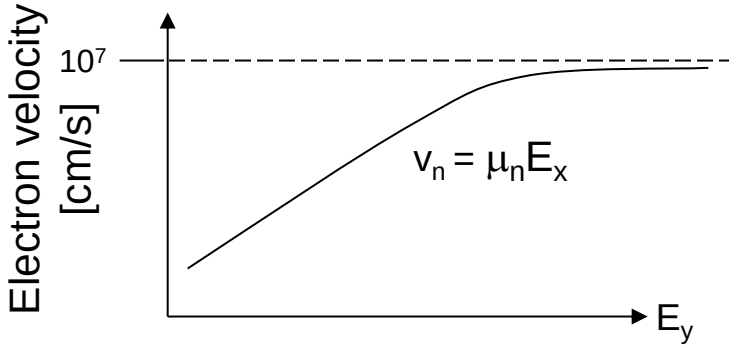
- Simple model assumes that MOST turns on suddenly for $V_{GS} > V_{th}$ and is completely off for $V_{GS} < V_{th}$
- For $0 < V_{GS} < V_{th}$ negative, there exists already a channel
  - weak inversion
  - main conduction mechanism is diffusion (not drift as in strong inversion)
  - Drain current vs gate-source voltage is exponential (like in a BJT)

$$I_{D(sub-th)} \cong I_{D0}\left(\frac{W}{L}\right) e^{(qV_{eff}/nkT)}$$

- with $n = \frac{C_{ox} + C_{j0}}{C_{ox}} \approx 1.5$ and $I_{D0} = (n-1)\mu_n C_{ox}\left(\frac{kT}{q}\right)^2$

Mobility Degradation





- The effective carrier mobility decreases under large electric fields
  - due to the vertical electric fields, electrons are pushed to the surface and scattered
- The effect causes the velocity of carriers to saturate (≈10^7 cm/s in Si)
- Effective carrier mobility: $\mu_{n,\text{eff}} \cong \dfrac{\mu_n}{([1 + (\theta V_{\text{eff}})^m])^{1/m}}$
  - θ and m are device technology parameters
- Incorporating this into the drain current equation: $I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}V_{\text{eff}}^2\left(\dfrac{1}{[1 + (\theta V_{\text{eff}})^m]^{1/m}}\right)$
- The maximum transconductance achievable with: $g_{m(\text{mob-deg})} = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\frac{1}{\theta}$
- Mobility degradation becomes a major effect for smaller technologies

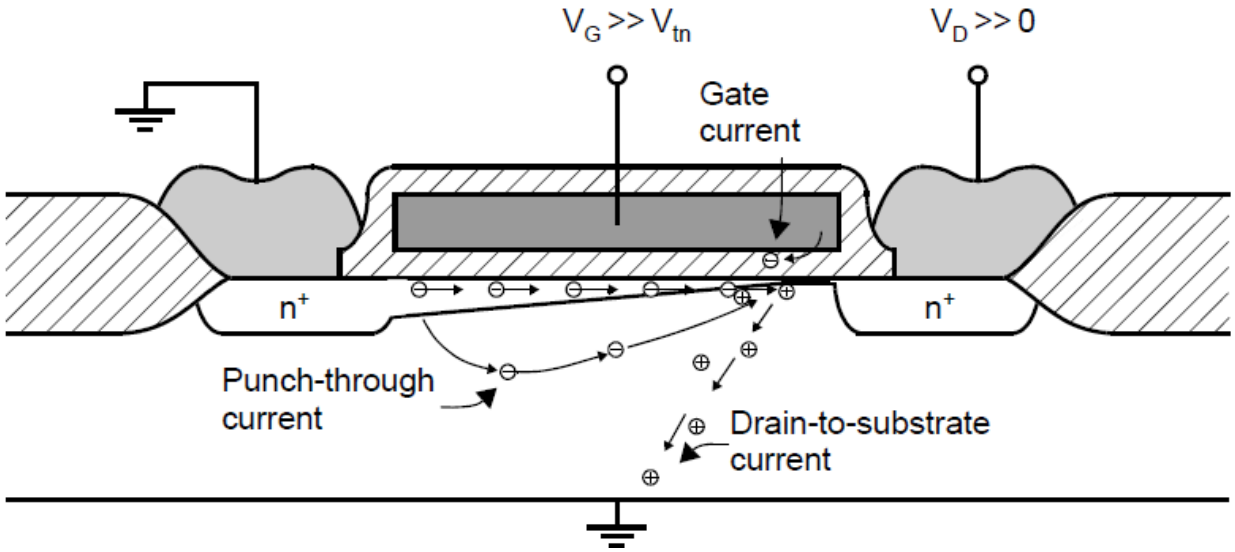
Summary Subthreshold, Mobility Degradation

	Subthreshold (exponential)	Strong Inversion (square-law)	Mobility Degradation (linear)
Region of validity	$V_{eff} \lesssim 0$	$\frac{2nkT}{q} < V_{eff} < \frac{1}{2\theta}$	$V_{eff} > \frac{1}{2\theta}$
Drain current, I_D	$I_{D0}\left(\frac{W}{L}\right) e^{(qV_{eff}/nkT)}$	$\frac{1}{2}\mu_n C_{ox}\frac{W}{L}V_{eff}^2$	$\frac{0.5\mu_n C_{ox}(W/L)V_{eff}^2}{[1+(\theta V_{eff})^m]^{1/m}}$
Small-signal transconductance, g_m	$\frac{qI_D}{nkT}$	$\frac{2I_D}{V_{eff}} = \mu_n C_{ox}\frac{W}{L}V_{eff}$	$\frac{1}{2}\mu_n C_{ox}\frac{W}{L}\frac{1}{\theta}$
Most useful	Very low-power operation	Most analog design	Very high-speed operation

Short Channel Effects

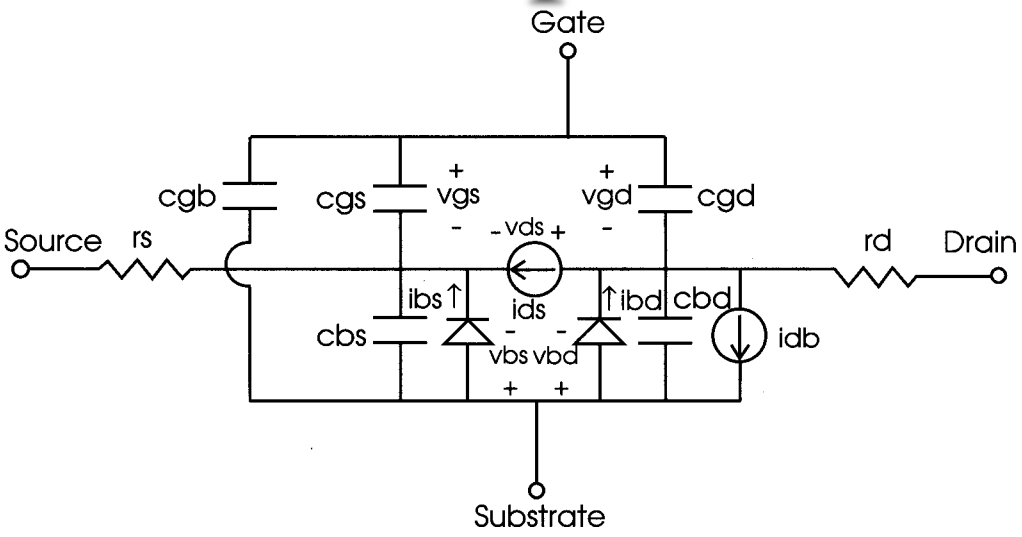
- Short Channel effects include a number of effects that become important for shrinking dimensions
  - Reduced output impedance
  - hot-carrier effects
    - oxide trapping
    - substrate currents
  - Threshold voltage dependency on device dimensions W/L
- Short channel transistors have reduced output impedance because the depletion region at drain end have an increased proportional effect on drain current
  - Additionally a phenomenon known as drain-induced barrier lowering (DIBL) effectively lowers V_{th} as V_{DS} is increased, thereby further lowering the output impedance of a short-channel device

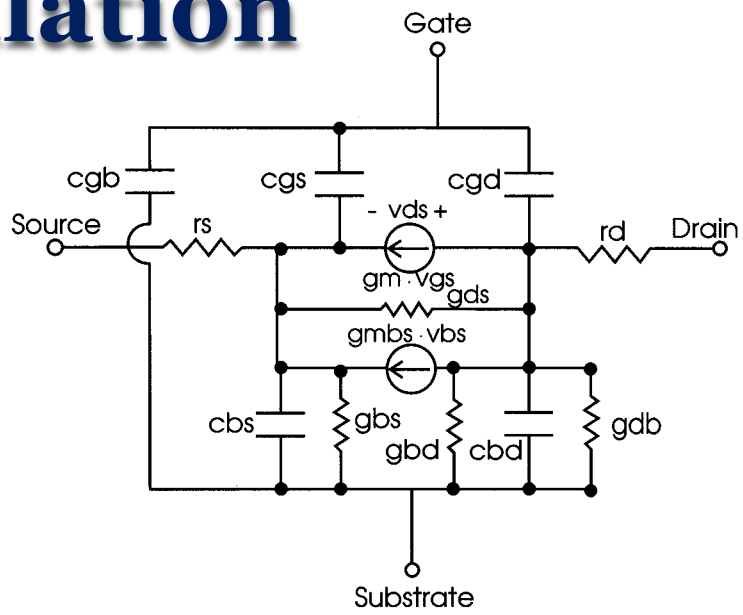
Hot Carrier Effects



- high-velocity carriers can cause harmful effects
  - generation of electron-hole pairs by impact ionization and avalanching
    - Extra electron-hole pairs are caused by impact ionization and cause currents to flow from the drain region to the substrate
    - effect can be modelled by a finite drain-to-ground impedance
    - can also cause latch-up
  - hot-carriers can also cause a tunnel current through the gate oxide
    - some get trapped and shift the threshold voltage
    - hot carriers limit the long term life time and reliability of MOST
  - hot carrier can also cause punch through from source to drain and can cause transistor breakdown

Spice Simulation





- The SPICE level 1 model (Schichman-Hodges Model*) is a simple, approximate model which is essentially the same as used for hand-calculations
- The model consists of the following components:
  - An equation for the threshold voltage
  - Equations for the drain/source current
  - Equations for the gate capacitances CGS and CGD
  - An equivalent circuit
  - Model parameters depend on the SPICE version

* [1] H. Shichman and D. A. Hodges, "Modeling and simulation of insulated-gate field-effect transistor switching circuits," IEEE Journal of Solid-State Circuits, SC-3, 285, September 1968.

SPICE Level 1 Parameters

SYMBOL	PARAMETER	DESCRIPTION
V_{TO}	VTO	Zero bias threshold voltage
K_p	KP	Transconductance parameter
λ	LAMDA	Channel length modulation
γ	GAMMA	Body effect parameter
$2\Phi_F$	PHI	Surface inversion potential
t_{ox}	TOX	Gate oxide thickness
N_A	NSUB	Substrate doping concentration
L_D	LD	Lateral diffusion
μ	UO	Surface mobility
I_S	IS	S/B & D/B diode saturation current
V_{BI}	PB	S/B & D/B diode built in voltage
C_{J0}	CJ	S/B & D/B zero bias junction cap./m^2

SPICE Level 1 Parameters

M_J	MJ	S/B & D/B doping profile grading coefficient
C_{JSW}	CJSW	S/B & D/B zero bias perimeter doping grading coefficient
M_{JSW}	MSJW	S/B & D/B perimeter doping grading coefficient
C_{GBO}	CGBO	G/B overlap capacitance/m^2
C_{GDO}	CGDO	G/D overlap capacitance/m^2
C_{GSO}	CGSO	G/S overlap capacitance/m^2
R_D	RD	Drain series resistance
R_S	RS	Source series resistance

SPICE Level 1 Equations (1/2)

From PSPICE manuals (PSPRef.pdf, pp. 198):
Drain current equations
Normal mode: Vds > 0
Case 1
for cutoff region: Vgs-V to < 0
then: Idrain = 0
Case 2
for linear region: Vds < Vgs-V to
then: Idrain = (W/L)·(**KP**/2)·(1+**LAMBDA**·Vds)·Vds·(2·(Vgs-V to)-Vds)
Case 3
for saturation region: 0 < Vgs-V to < Vds
then: Idrain = (W/L)·(**KP**/2)·(1+**LAMBDA**·Vds)·(Vgs-V to) 2
where
V to = **VTO**+**GAMMA**·((**PHI**-Vbs)1/2 -**PHI** 1/2)
Inverted mode: Vds < 0
Switch the source and drain in the normal mode equations above.

SPICE Level 1 Equations (2/2)

From PSPICE manuals (PSPRef.pdf, pp. 199):

MOSFET equations for capacitance

Cbs = bulk-source capacitance = area cap. + sidewall cap. + transit time cap.

Cbd = bulk-drain capacitance = area cap. + sidewall cap. + transit time cap.

where

if

CBS = 0 *AND* **CBD** = 0

then

Cbs = AS·**CJ**·Cbsj + PS·**CJSW**·Cbss + **TT**·Gbs

Cbd = AD·**CJ**·Cbdj + PD·**CJSW**·Cbds + **TT**·Gds

else

Cbs = **CBS**·Cbsj + PS·**CJSW**·Cbss + **TT**·Gbs

Cbd = **CBD**·Cbdj + PD·**CJSW**·Cbds + **TT**·Gds

where

Gbs = DC bulk-source conductance = *d*Ibs/*d*Vbs

Gbd = DC bulk-drain conductance = *d*Ibd/*d*Vbd

if

Vbs < **FC**·**PB**

then

Cbsj = (1-Vbs/**PB**)**-**MJ

Cbss = (1-Vbs/**PBSW**)**-**MJSW

if

Vbs > **FC**·**PB**

SPICE Level 1 Equations (3/3)

MOSFET equations for capacitance (cont.)

then

Cbsj = (1-**FC**)-(1+MJ)·(1-**FC**·(1+**MJ**)+**MJ**·Vbs/**PB**)

Cbss = (1-**FC**)-(1+MJSW)·(1-**FC**·(1+**MJSW**)+**MJSW**·Vbs/**PBSW**)

if

Vbd < **FC**·**PB**

then

Cbdj = (1-Vbd/**PB**)**-**MJ

Cbds = (1-Vbd/**PBSW**)**-**MJSW

if

Vbd > **FC**·**PB**

then

Cbdj = (1-**FC**)-(1+**MJ**)·(1-**FC**·(1+**MJ**)+**MJ**·Vbd/**PB**)

Cbds = (1-**FC**)-(1+**MJSW**)·(1-**FC**·(1+**MJSW**))

Cgs = gate-source overlap capacitance = **CGSO**·W

Cgd = gate-drain overlap capacitance = **CGDO**·W

Cgb = gate-bulk overlap capacitance = **CGBO**·L

Advanced Spice Models

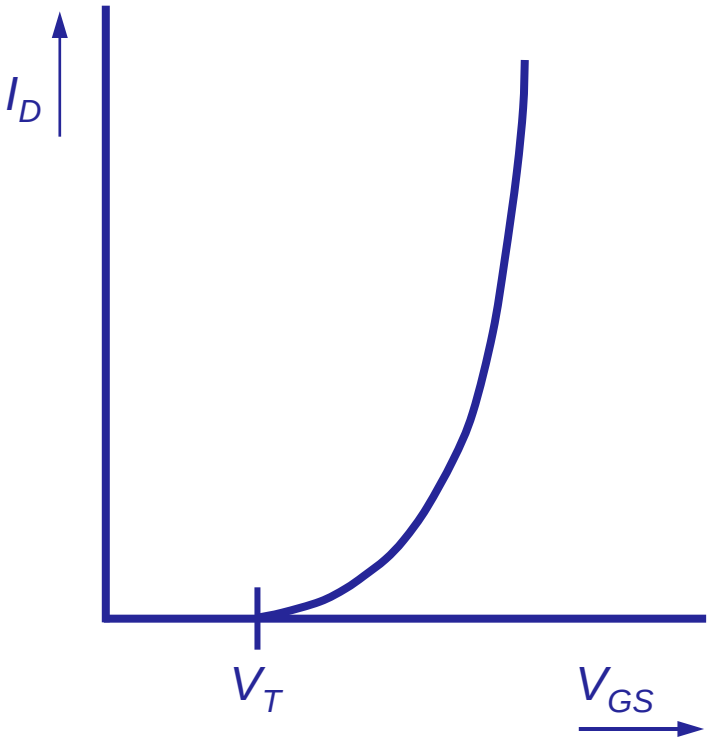
- BSIM3
  - Improved modelling of moderate inversion, and the geometry-dependence of device parameters. physics-based model
- EKV:
  - Relates terminal currents and voltages with unified equations that cover all modes of transistor operation, hence avoiding discontinuities at transitions between, for example, weak and strong inversion. Also handles geometry-dependent device parameters.
- BSIM4
  - Improved modeling of leakage currents and short-channel effects, noise, and parasitic resistance in the MOSFET terminals, as well as continued improvements in capturing the geometry dependence of device parameters.
- PSP:
  - Improved modeling of noise and the many short-channel and layout-dependent effects now dominant in nanoscale CMOS devices. Good for nonlinearities.

- sdfsd
  - werwer
  - werwer

MOSFET Equations

- n-Kanal-Enhancement-Transistor

Input characteristics



$$I_D = \frac{k'}{2} \cdot \frac{W}{L} (V_{GS} - V_T)^2 \cdot \left(1 + \frac{V_{DS}}{V_A}\right)$$

κ‘: transconductance parameter [A/V^2]

$$k' = \mu_0 C_{ox}$$

λ: Channel length modulation parameter [V^{-1}]

$$V_A = \frac{1}{\lambda}$$

g_m: Transconductance in saturation

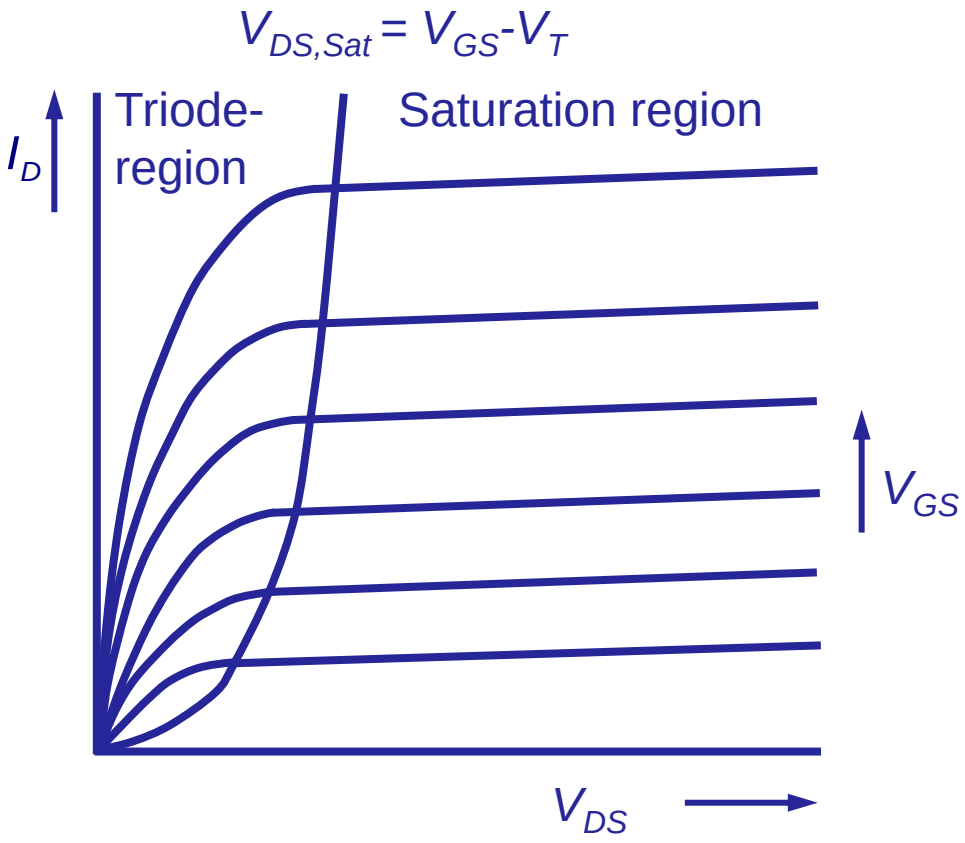
$$g_m = \frac{\partial I_D}{\partial V_{DS}}$$

$$g_m = k' \frac{W}{L} (V_{GS} - V_T) = \sqrt{2k' \frac{W}{L} I_D}$$

MOSFET Equations

- Example: n-channel enhancement transistor

Output characteristics



Triode: $V_{DS} < (V_{GS}-V_T)$
Saturation: $V_{DS} > (V_{GS}-V_T)$

Triode Region

$$I_D = k' \frac{W}{L}\left(V_{GS} - V_T - \frac{V_{DS}}{2}\right)\cdot V_{DS}$$

Saturation Region

$$I_D = \frac{k'}{2}\cdot\frac{W}{L}\left(V_{GS} - V_T\right)^2 \cdot \left(1 + \frac{V_{DS}}{V_A}\right)$$

Output resistance

$$\frac{1}{r_o} = g_{ds} = \frac{I_D}{V_A + V_{DS}}$$

Comparison BJT - MOSFET

	Field Effect Transistor (FET)	Bipolar Junction Transistor (BJT)
1	Low voltage gain	High voltage gain
2	High current gain	Low current gain
3	Very input impedance	Low input impedance
4	High output impedance	Low output impedance
5	Low noise generation	Medium noise generation
6	Fast switching time	Medium switching time
7	Easily damaged by static	Robust
8	Some require an input to turn it “OFF”	Requires zero input to turn it “OFF”
9	Voltage controlled device	Current controlled device
10	Exhibits the properties of a Resistor	
11	More expensive than bipolar	Cheap
12	Difficult to bias	Easy to bias

Comparison BJT - MOSFET

Bipolar Transistor	MOSFET
Exponential Characteristic	Quadratic Characteristic
Active: $V_{CB} > 0$	Saturation: $V_{DS} > V_{GS} - V_{TH}$
Saturation: $V_{CB} < 0$	Triode: $V_{DS} < V_{GS} - V_{TH}$
Finite Base Current	Zero Gate Current
Early Effect	Channel–Length Modulation
Diffusion Current	Drift Current
–	Voltage–Dependent Resistor

- MOSFET: controlled by an electric field
  - Very high input resistance (no current into gate)
- Bipolar devices have a higher g_m than MOSFETs for a given bias current due to its exponential IV characteristics.