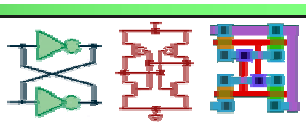
MOS Transistor



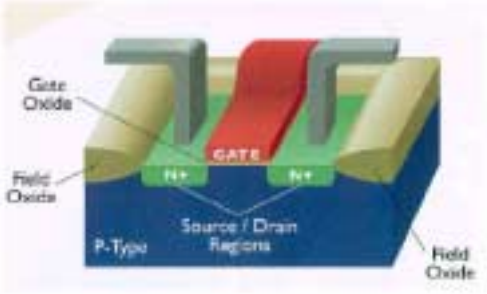
Spring 2003

Poras T. Balsara
Center for Integrated Circuits and Systems
Department of Electrical Engineering
University of Texas at Dallas

Outline

- MOSFET structure
- Threshold voltage concept
- I-V characteristics and different regions of operation
- Secondary effects
- Short-channel devices
- SPICE model

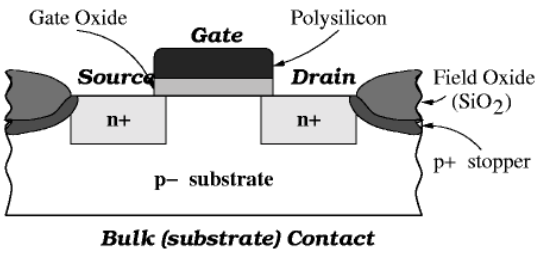
The MOSFET

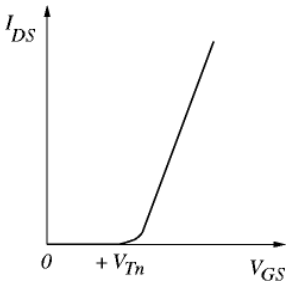


Cross-section of an NMOSFET

- N areas have been doped with donor ions (arsenic) with concentration N_D -- *electrons* are majority carriers.
- P areas have bee doped with acceptor ions (boron) with concentration N_A -- *holes* are majority carriers.

N-Channel MOSFET (NMOSFET)



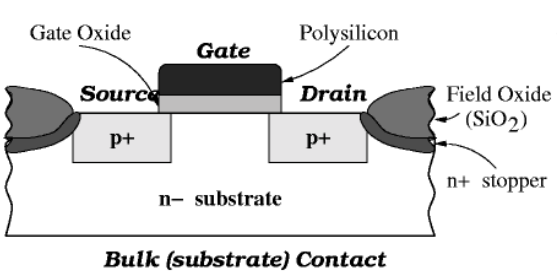


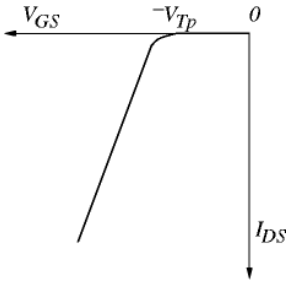
- Conducting channel is formed between drain and source when the gate-source voltage,

$$V_{GS} \geq V_{Tn}$$

- V_{Tn} is the NMOSFET *Threshold Voltage*

P-Channel MOSFET (PMOSFET)





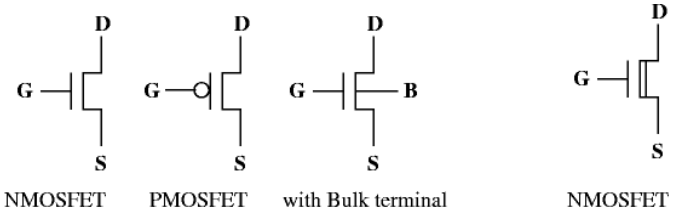
- Conducting channel is formed between drain and source when the gate-source voltage,

$$V_{GS} \leq V_{Tp}$$

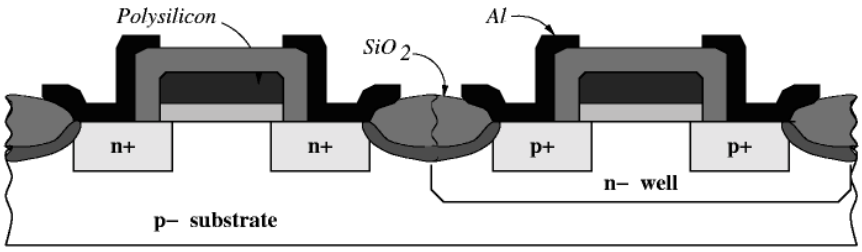
- V_{Tp} is the PMOSFET *Threshold Voltage*

Enhancement vs. Depletion Mode

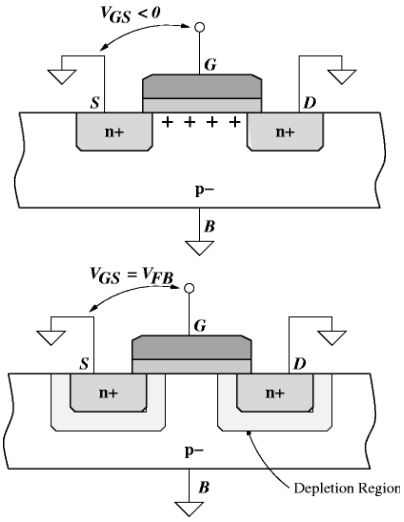
- *Enhancement Mode Devices*: In these devices, application of V_{GS} forms a conducting channel between drain and source. Channel does not exist under zero gate bias.
- *Depletion Mode Devices*: In these devices, channel exists under zero gate bias. Application of V_{GS} turns these devices off.
- MOSFET Symbols:



Cross-section of CMOS Technology

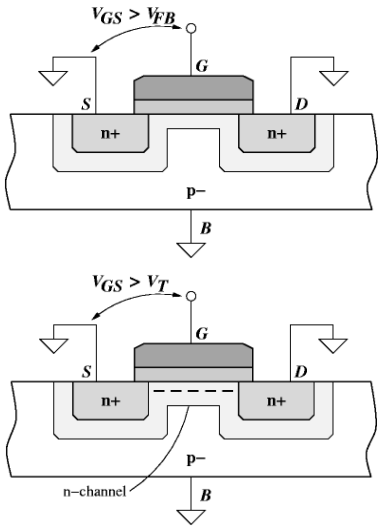


NMOSFET – Threshold Voltage



- Accumulation Region:
  $V_{GS} < 0$
- Flat-band Condition:
  $V_{GS} = V_{FB}$

NMOSFET – Threshold Voltage



- Depletion Region:
  $V_{GS} > V_{FB}$
- Surface Inversion:
  $V_{GS} \geq V_{Tn}$

Threshold Voltage

$$V_T = V_{T0} + \gamma\left(\sqrt{\left|-2\phi_F + V_{SB}\right|} - \sqrt{\left|-2\phi_F\right|}\right)$$

Threshold voltage at V_{SB}=0 and is mostly a function of manufacturing process.

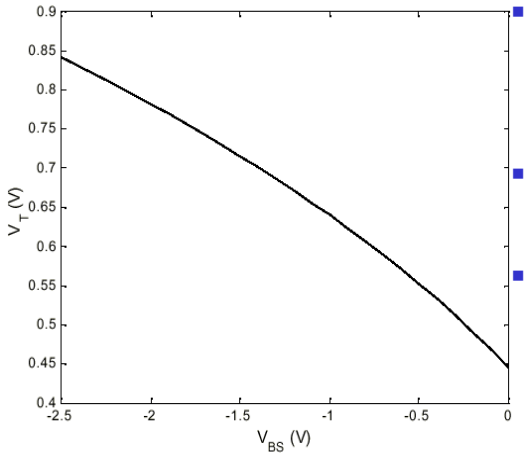
The ***body-effect coefficient***: it expresses impact of changes in V_{SB}

$$= \sqrt{\left(2q\varepsilon_{si} N_A\right)/C_{ox}}$$

Substrate Bias (source bulk voltage)

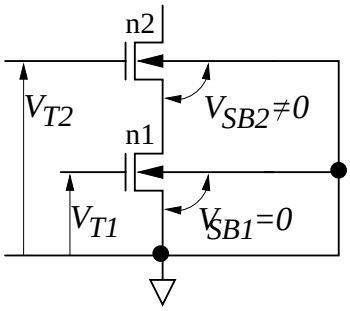
Note: The *threshold voltage* has a **positive** value for a typical **NMOSFET**, while it is **negative** for a normal **PMOSFET**.

The Body Effect



- Effect of well bias on an NMOSFET is plotter here. (V_{Tn} *is positive for a normal NMOSFET with body tied to ground*).
- Negative bias on well or substrate causes V_{Tn} to increase from 0.45V to 0.85V
- If V_{SB} is increased, the two depletion regions under the n+ regions will become deeper ⇨ larger V_{GS} is needed to attain similar depth of depletion under the channel.

Body Effect in a Logic Gate

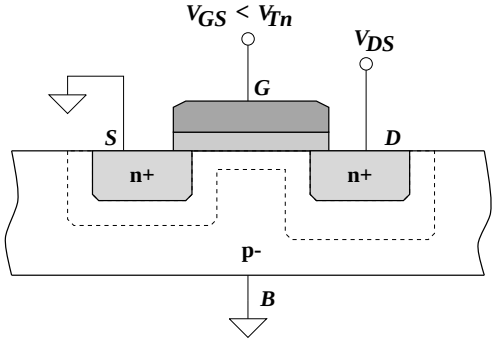


strong inversion: $V_{GS} > V_T$

$V_G - V_S > V_T$

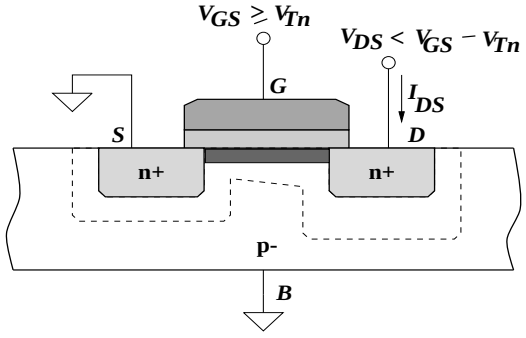
body effect: $V_{T2} > V_{T1}$

NMOSFET in Cut-off Region



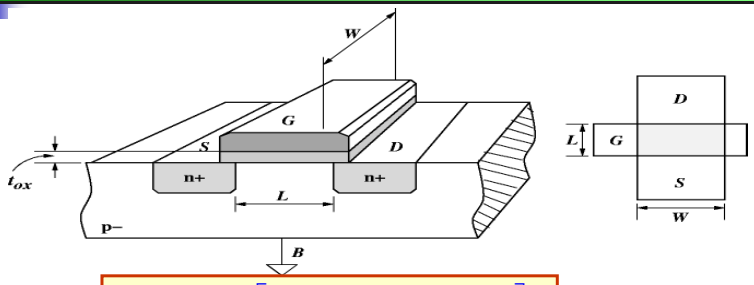
- $V_{GS} < V_{Tn}$
- $I_{DS} \approx 0$
- I_{DS} is due to diode reverse-bias leakage current and/or sub-threshold conduction.

NMOSFET in Linear (Resistive) Region



- $V_{GS} \geq V_{Tn}$ and $V_{DS} < V_{GS} - V_{Tn}$
- $I_{DS} = k_n' W/L\left[(V_{GS} - V_{Tn})V_{DS} - V_{DS}^2/2\right]$

NMOSFET V-I: Linear Mode



$$I_{DS} = k_n' \frac{W}{L}\left[(V_{GS} - V_{Tn})V_{DS} - \frac{V_{DS}^2}{2}\right]$$

Process Transconductance Parameter

$k_n' = \mu_n C_{ox} = \mu_n \varepsilon_{ox} / t_{ox}$ in $\left(A/V^2\right)$

Oxide Permittivity:

$\varepsilon_{ox} = 3.97 \times \varepsilon_0 = 3.5 \times 10^{-11}\ F/m$

MOSFET Gain Factor

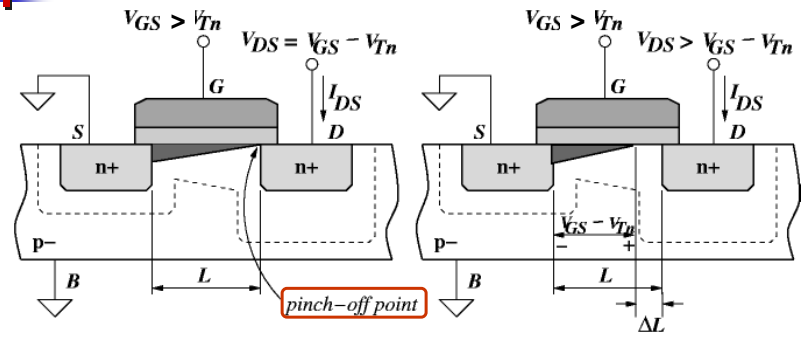
$k_n = k_n' W/L$

Carrier Mobility

$\mu_n \approx 500 cm^2/V\sec$

$\mu_p \approx 180 cm^2/V\sec$

NMOSFET in Saturation Region



- $V_{GS} > V_{Tn}$ and $V_{DS} \geq V_{GS} - V_{Tn}$
- $$I_{DS} = \frac{k_n'}{2}\frac{W}{L}(V_{GS} - V_{Tn})^2$$

Why is this equation not entirely correct?

Channel Length Modulation

- The position of the *pinch-off point* shifts as V_{DS} is increased. Consequently, the *effective channel length* (i.e., $L' = L - \Delta L$) is modulated by V_{DS}.
- I_{DS} is increased when the effective length is decreased.

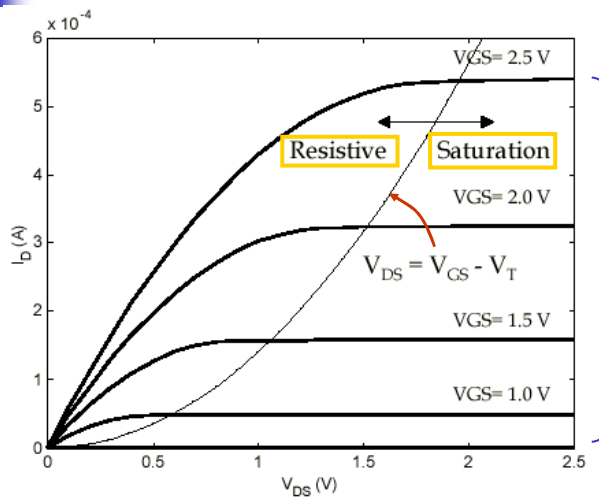
$$I_{DS} = \frac{k_n'}{2}\frac{W}{L}\left(V_{GS} - V_{Tn}\right)^2 \left(1+\lambda V_{DS}\right)$$

- λ is an empirical constant parameter, called the *channel-length modulation coefficient* (varies roughly with the inverse of the channel length)

Current Determinates

- For a fixed V_{DS} and V_{GS} ($\geq V_{Tn}$), I_{DS} is a function of
  - the distance between the source and drain - ***L***
  - the channel width - ***W***
  - the threshold voltage - V_T
  - the thickness of the oxide - t_{ox}
  - the dielectric of the gate insulator (SiO_2) - ε_{ox}
  - the carrier mobility:
    - for NMOSFETs: $\mu_n \approx 500\ cm^2/V\text{-}sec$
    - for PMOSFETs: $\mu_p \approx 180\ cm^2/V\text{-}sec$

Long Channel NMOS I-V Curves



Quadratic dependence

NMOSFET: 0.25μm CMOS, V_{DD}=2.5V, V_{Tn}=0.5V L_d=10μm W/L=1.5

I-V Characteristics: Summary

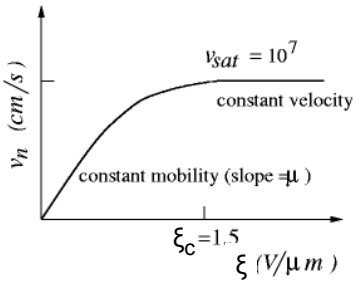
Region	NMOSFET	PMOSFET
Cutoff	$V_{GS} < V_{Tn}$ $I_{DS} \approx 0$	$V_{GS} > V_{Tp}$ $I_{DS} \approx 0$
Linear	$V_{GS} \geq V_{Tn}$ $V_{DS} < V_{GS} - V_{Tn}$ $I_{DS} = k_n' \frac{W}{L}\left[(V_{GS} - V_{Tn})V_{DS} - \frac{V_{DS}^2}{2}\right]$	$V_{GS} \leq V_{Tp}$ $V_{DS} > V_{GS} - V_{Tp}$ $I_{DS} = k_p' \frac{W}{L}\left[(V_{GS} - V_{Tp})V_{DS} - \frac{V_{DS}^2}{2}\right]$
Saturation	$V_{GS} \geq V_{Tn}$ $V_{DS} \geq V_{GS} - V_{Tn}$ $I_{DS} = \frac{k_n'}{2}\frac{W}{L}(V_{GS} - V_{Tn})^2(1+\lambda V_{DS})$	$V_{GS} \leq V_{Tp}$ $V_{DS} \leq V_{GS} - V_{Tp}$ $I_{DS} = \frac{k_p'}{2}\frac{W}{L}(V_{GS} - V_{Tp})^2(1+\lambda V_{DS})$

Short Channel Effects

- The behavior of MOSFET with sub-micron channel length deviates considerably from the first-order long channel models discussed earlier.
- Short channel effects for sub-micron transistors:
  - Velocity saturation and mobility degradation
  - Sub-threshold conduction
  - Threshold voltage variations
  - Parasitic resistance
- Secondary effects:
  - Temperature effects
  - Hot carrier effects
  - Latchup

Velocity Saturation

- Carrier velocity saturates when the elect. field along the channel reaches a critical value ξ_c. This is due to scattering (collisions suffered by the carriers).



	ξ_c	v_{sat}
NMOS	1.5×10^4 V/cm (or 1.5 V/μm)	10^7 cm/s
PMOS	≥ 10^5 V/cm	10^7 cm/s

For an NMOSFET with L=1μm, only a couple of volts difference between D and S are needed to reach velocity saturation.

V-I Relation: Velocity Saturation

- NMOS Linear Region: $V_{DS} < V_{GS} - V_{Tn}$

$$I_{DS} = \kappa(V_{DS})\, k_n' \frac{W}{L}\left[(V_{GS} - V_{Tn})V_{DS} - \frac{V_{DS}^2}{2}\right]$$

where,

$\kappa(V) = 1/(1 + (V/\xi_c L))$ is a measure of the degree of velocity saturation.

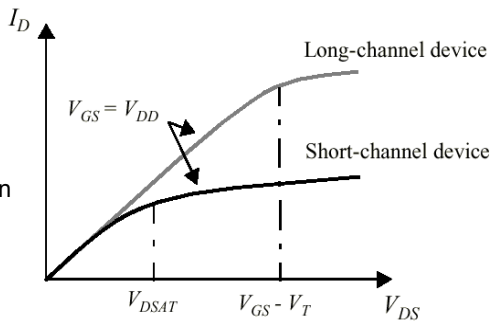
- NMOS Saturation Region: $V_{DS} = V_{DSAT} \geq V_{GS} - V_{Tn}$

$$I_{DSAT} = \kappa(V_{DSAT})\, k_n' \frac{W}{L}\left[(V_{GS} - V_{Tn})V_{DSAT} - \frac{V_{DSAT}^2}{2}\right]$$

$$V_{DSAT} = \kappa(V_{GS} - V_{Tn})(V_{GS} - V_{Tn})$$

Effects of Velocity Saturation

- For short-channel devices with large enough $(V_{GS}-V_T)$, $\kappa \ll 1$
  ⇨ $V_{DSAT} < V_{GS}-V_T$
  - Device enters saturation before V_{DS} reaches $V_{GS}-V_T$
    ⇨ extended saturation
    ⇨ device operates more often in saturation.
  - I_{DSAT} has a linear dependence with respect to V_{GS}
    ⇨ reduction in the amount of current for a given control voltage.



I_{DSAT} Approximation

- For first-order analysis the current equations for velocity saturated devices can be approximated by assuming,

$$v = \mu_n \xi \quad \text{for } \xi < \xi_c$$
$$= v_{sat} = \mu_n \xi_c \quad \text{for } \xi \geq \xi_c$$

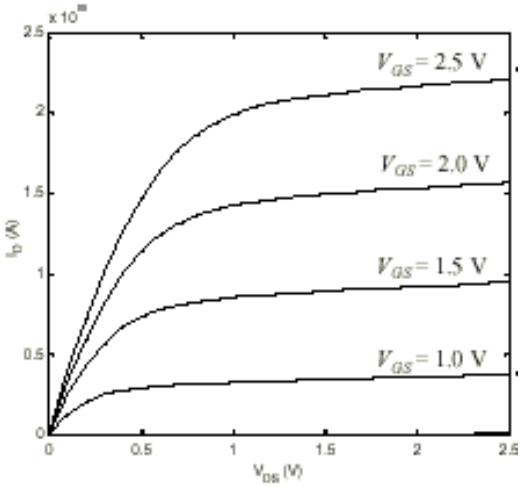
and

$$V_{DSAT} = L\xi_c = \frac{L v_{sat}}{\mu_n}$$

- I_{DSAT} can be obtained by plugging in the saturation voltage.

$$I_{DSAT} = I_{DS}\left(V_{DS} = V_{DSAT}\right)$$
$$= \mu_n C_{ox} \frac{W}{L}\left[\left(V_{GS} - V_{Tn}\right)V_{DSAT} - \frac{V_{DSAT}^2}{2}\right]$$
$$= v_{sat} C_{ox} W \left(V_{GS} - V_{Tn} - \frac{V_{DSAT}}{2}\right)$$

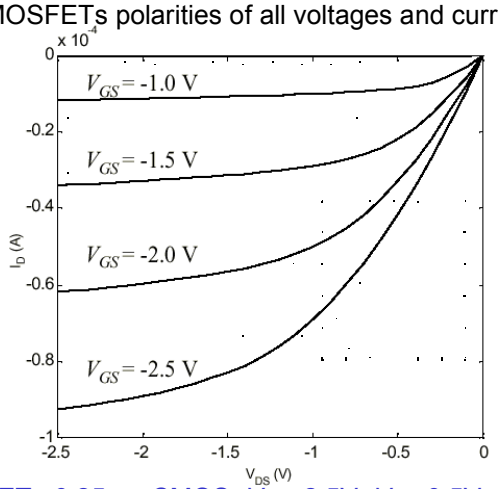
Short Channel NMOS I-V Curves



NMOSFET: 0.25μm CMOS, V_{DD}=2.5V, V_{Tn}=0.5V L_d=0.25μm W/L=1.5

Short Channel PMOS I-V Curves

For PMOSFETs polarities of all voltages and currents are reversed

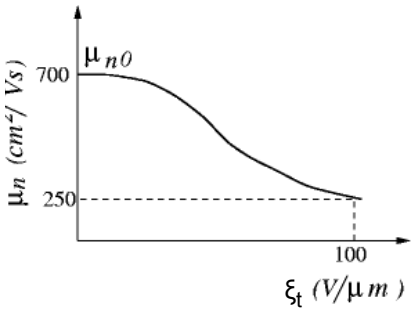


NMOSFET: 0.25μm CMOS, V_{DD}=2.5V, V_{Tn}=0.5V L_d=0.25μm W/L=1.5

Mobility Degradation

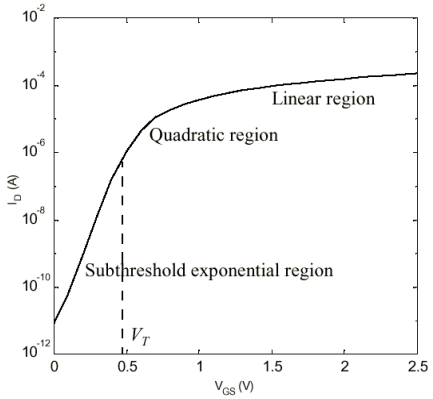
- In short channel devices, the vertical field originating from the gate voltage reduces surface carrier mobility with respect to the bulk mobility ⇨ *mobility degradation*

$$\mu_{n,eff} = \frac{\mu_{n0}}{1+\eta\left(V_{GS}-V_T\right)}$$



where, μ_{n0} is the bulk mobility and
η is an empirical parameter

Subthreshold Conduction



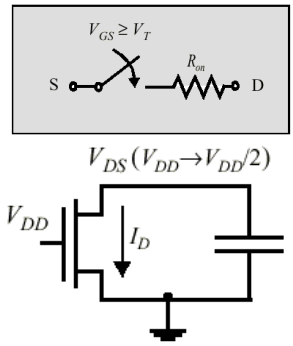
- For $V_{GS} < V_{Tn}$ NMOSFET conducts partially and the current decays exponentially (does not switch abruptly) – *subthreshold conduction* or *weak inversion*
- For ideal device ($n = 1$), the *subthreshold slope factor, S* at room temperature is 60 mV/decade
- For actual devices $n > 1$, typically around 1.5 ⇨ relatively larger voltage is required to drop the current, i.e., higher S.
- Current roll-off is adversely affected by increase in temp.

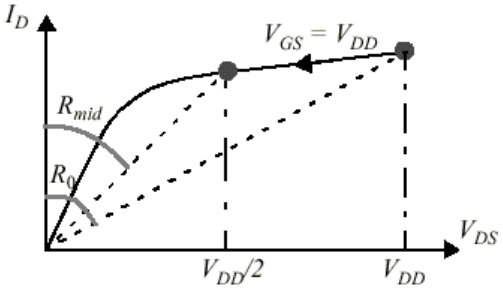
Subthreshold Slope Factor:

$$S = \left[\frac{d}{dV_{GS}}\ln\left(I_{DS}\right)\right]^{-1} = n\left(\frac{kT}{q}\right)\ln(10)$$

MOSFET Modeled as a Switch

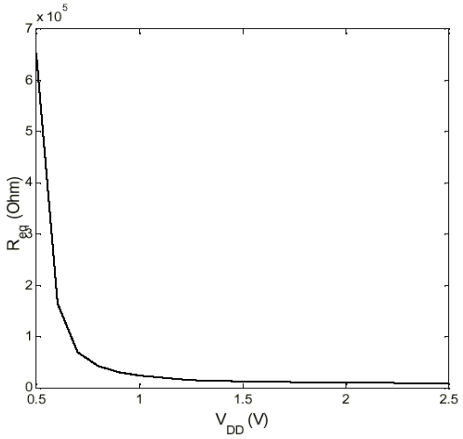
- Modeled as a switch with infinite off resistance and a finite on resistance, R_{on}.
- A simple model: use average value of resistance over operation region of interest or even simpler, use average of resistance at the end-points of the transition.





R_{on} for a MOSFET

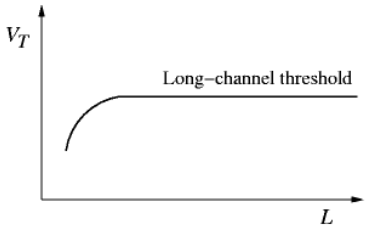
- Plot for average R_{on} for $V_{GS} = V_{DD}$ and $V_{DS} = V_{DD} \rightarrow V_{DD}/2$
- R_{on} is inversely proportional to the ratio, W/L
- For $V_{DD} >> V_T + V_{DSAT}/2$, R_{on} becomes independent of V_{DD}
- Once V_{DD} approaches V_T, R_{on} increases dramatically.
- R_{on} for 0.25µm CMOS, with W/L=1 and $L=L_{min}$



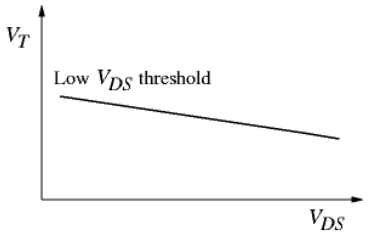
V_{DD}(V)	1	1.5	2	2.5
NMOS (kΩ)	35	19	15	13
PMOS (kΩ)	115	55	38	31

Threshold Voltage Variation

- V_T is a function of technology and the substrate bias.
- For short channel devices it is also a function of L and V_{DS}
  - *V_T decreases with L*
  - *V_T decreases with increasing V_{DS}* (DIBL). *For high enough V_{DS} the depletion region around drain may extend to the source* ⇨ *I_{DS} flows regardless of V_{GS}.*



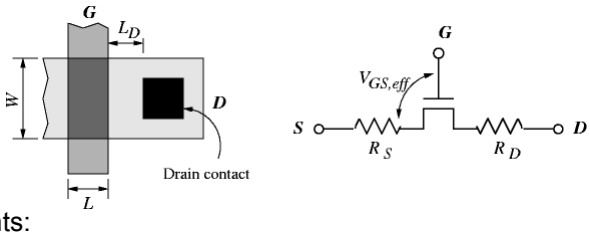
V_T as a function of L (for low V_{DS})



Drain–induced Barrier Lowering (for low L)

Source Drain Resistance

- For scaled MOSFETs, junctions are shallower, and contact openings become smaller ⇨ parasitic resistance in series with the drain and source regions increases.
  - Deterioration in performance since I_{DS} for a given voltage is reduced.



- Improvements:
  - cover drain source regions with low-ρ material (e.g., titanium or tungsten) to effectively reduce parasitic resistance – *silicidation*.
  - make the device wider than needed.

Temperature Effects

- Absolute value of V_T decreases with an increase in temperature. Variation is approx. -4mV/°C for high substrate doping levels and -2mV/°C for low doping levels ⇨ I_{DS} should increase with increase in temperature, T

- However, increase in I_{DS} is overwhelmingly offset by the degradation of mobility, μ

$$\mu(T) = \mu(295^{\circ}\text{K})\left[\frac{295}{T}\right]^{\alpha}$$

where, $\alpha \approx 1.5$ for electrons and 1 for holes.

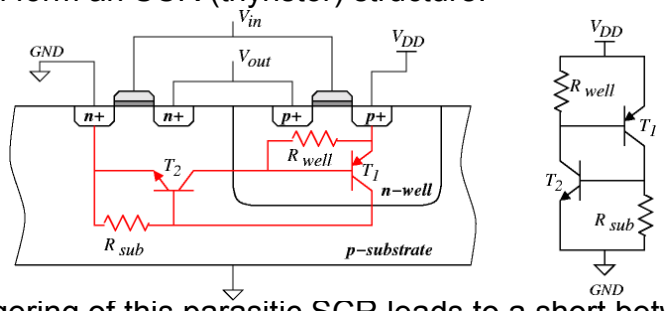
$\Rightarrow I_{DS} \propto T^{-\alpha}$

Hot Carrier Effects

- As *L* is reduced, the electric field at the drain of a MOSFET in saturation increases (for a fixed V_{DS}) ⇨ electrons are imparted high enough energy to become "hot".
- Hot electrons can leave silicon and tunnel into the gate oxide and get trapped therein, increasing V_{Tn} for NMOSFET and decreasing V_{Tp} for PMOSFETs
  - Electric fields of at least 10^4V/cm is needed for electrons to become hot.
  - Long-term reliability problems leading to circuit failure due to the degradation of device parameters.
  - Improvements done in present day devices:
    - specially engineered source drain regions to bound peaks in elect. fields.
    - reduced supply voltages.

Latchup in CMOS

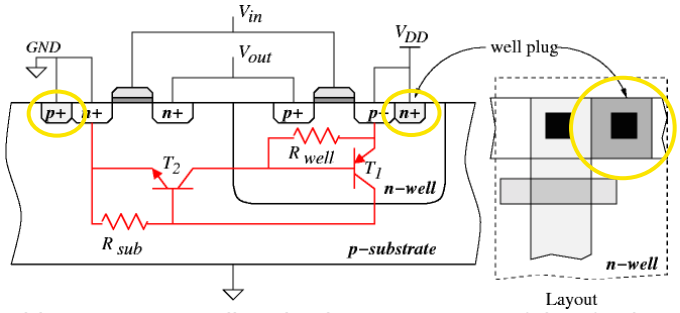
- Parasitic bipolar transistors exist in CMOS – a couple of them form an SCR (thyristor) structure.



- Triggering of this parasitic SCR leads to a short between V_{DD} and *GND*, usually resulting in a destruction of the chip, or a system failure that can only be resolved by power-down.

Prevention of Latchup

- To avoid latchup, R_{well} and R_{sub} should be minimized.

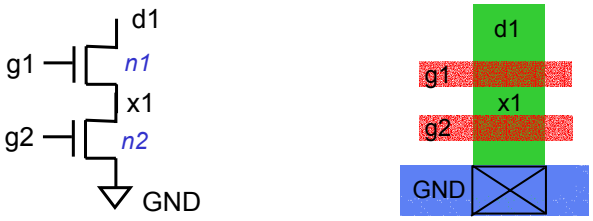


- Provide numerous well and substrate contacts (plugs), placed close to the source connections of the devices. *Minimum of one plug per well. One plug for every 5-8 devices in a well.*
- High current or highly susceptible devices like I/O drivers should be surrounded by *guard rings.*

MOSFET Models in SPICE3

- Level 1: *Shichman-Hodges* model – long-channel equations.
- Level 2: Geometry-based model; uses detailed device physics to define its equations; handles velocity saturation, DIBL, threshold voltage variations.
- Level 3: Semi-empirical model. Relies on measured device data to determine its parameters.
- Level 4: *BSIM model*. Analytically simple, based on parameters normally extracted from experimental data. Its accuracy and efficiency make it one of the most popular SPICE MOSFET models.
- **Level 8 (Star-HSPICE Level 49): *BSIM3 ver. 3.1***

MOSFET Description in SPICE3




```
Mn1  d1  g1  x1   GND  nenh   W=3.6u  L=0.6u
+  ad=6.5p    pd=7.2u
+  as=1.1p    ps=0.6u
Mn2  x1  g2   GND   GND   nenh   W=3.6u  L=0.6u
+  ad=1.1p    pd=0.6u
+  as=6.5p    ps=7.2u
```


Bibliography

1. Weste & Eshraghian: "*Principles of CMOS VLSI Design*", Addison Wesley.
2. J. Rabaey: "*Digital Integrated Circuits*", Prentice Hall
3. S-M. Kang & Y. Leblebici: "*CMOS Integrated Circuits: Analysis and Design*", McGraw Hill.
4. Y. Tsividis: "*Mixed Analog-Digital VLSI Devices and Technology*", McGraw Hill.