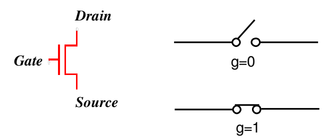
MOS Transistor Theory

- Two types of transistors
  - nMOS
  - pMOS
- Digital integrated circuits use these transistors essentially as a voltage controlled switch

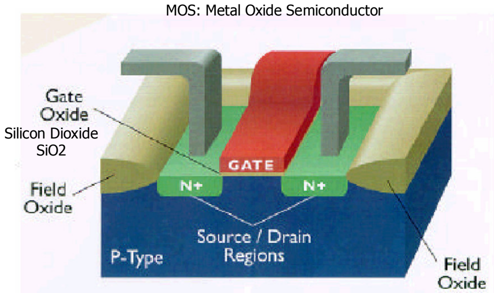
nMOS Transistor

- If the gate is "high", the switch is on
- If the gate is "low", the switch is off



nMOS Transistor

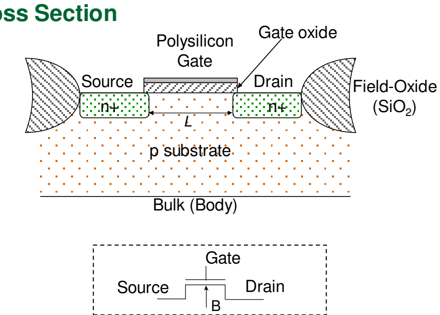
MOS: Metal Oxide Semiconductor



Cross-section of an n-type transistor

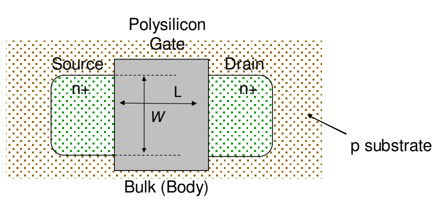
nMOS Transistor

Cross Section



nMOS Transistor

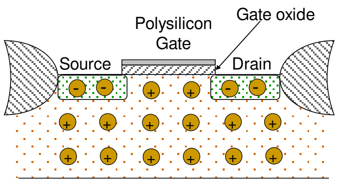
Top View



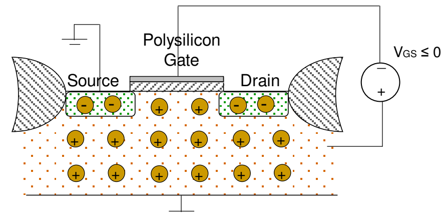
nMOS Transistor

- *n* areas have been doped with donor ions of concentration N_D - electrons are the majority carriers
- *p* areas have been doped with acceptor ions of concentration N_A - holes are the majority carriers

nMOS Transistor

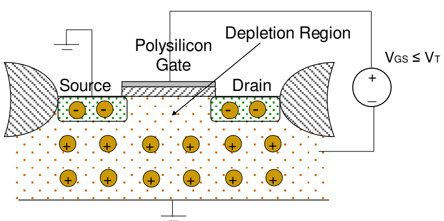


nMOS Transistor



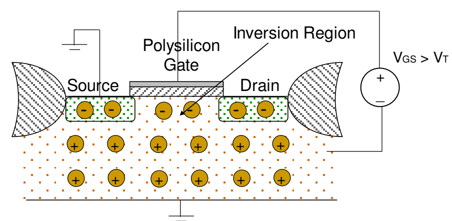
Accumulation Mode

nMOS Transistor



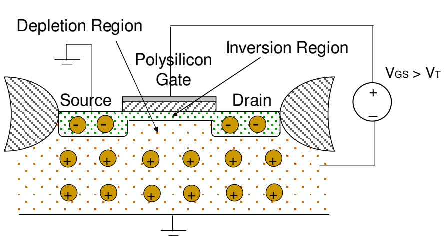
Depletion Mode

nMOS Transistor



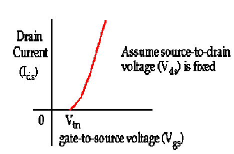
Inversion Mode

nMOS Transistor



Inversion Mode

nMOS Transistor



n-channel enhancement MOS

Threshold Voltage

- Dependent on
  - Gate conductor material
  - Gate insulator material
  - Channel Doping
  - Voltage difference between source and body

Threshold Voltage

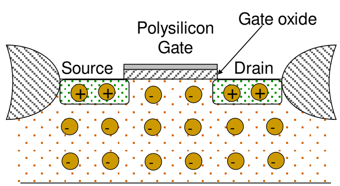
$$V_T = V_{T0} + \gamma\left(\sqrt{\left|-2\Phi_F + V_{SB}\right|} - \sqrt{\left|2\Phi_F\right|}\right)$$

- γ is the body-effect coefficient and controls the impact of the source to bulk voltage
- Φ_F is the Fermi potential and is dependent on doping levels
  - Fermi potential: potential difference between Fermi level and intrinsic Fermi level in the bulk of semiconductor.

$$\Phi_F = \frac{kT}{q}\ln\left(\frac{N_A}{n_i}\right)$$

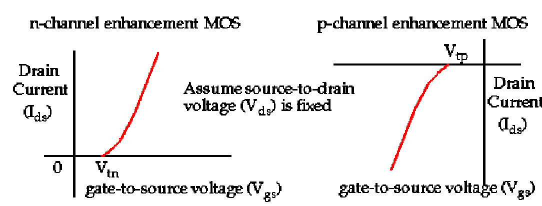
k: Boltzmann constant, T: temperature,
q: unit (electron) charge

pMOS Transistor



Accumulation Mode

pMOS Transistor

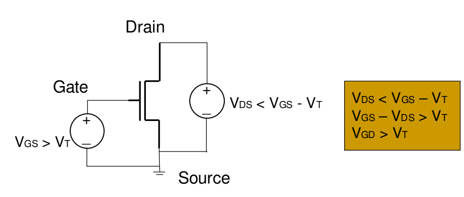


Source vs. Drain

nMOS: node with a higher voltage is drain, $V_D > V_S$

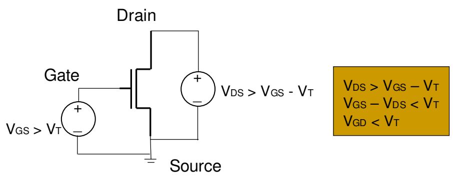
pMOS: node with a higher voltage is source, $V_S > V_D$

nMOS Transistor



Linear mode

nMOS Transistor



Saturation mode

MOS Transistor Characteristics

Linear Mode:

- $V_{GS}>V_T$ and $V_{GD}>V_T$
- Assume that V_T is constant

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

- $k_n' = (\mu_n C_{ox})$ is called the process transconductance parameter
- Gain factor of nMOS: $k_n = k_n'\ W/L$

Example

- $\mu_n = 600\ cm^2 / V\ s$
- $C_{ox} = 7 \times 10^{-8}\ F / cm^2$
- W = 20 µm
- L = 2 µm
- $K_n = \mu_n C_{ox} W/L = 0.42\ mA / V^2$

MOS Transistor Characteristics

Saturation Mode:

- $V_{GS}>V_T$ and $V_{GD}<V_T$
- Assume that V_T is constant

$$I_{DS} = k_n \frac{(V_{GS}-V_T)^2}{2}$$

In Summary

- Cutoff region ($V_{GS}<V_T$)

$$I_{DS} = 0$$

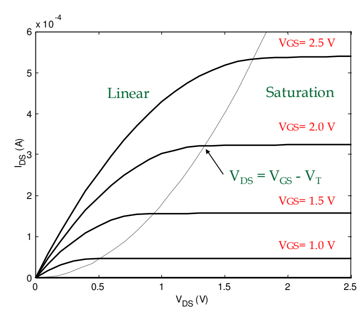
- Linear region ($V_{GS}>V_T$, $V_{DS}<V_{GS}-V_T$ or $V_{GD}>V_T$)

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

- Saturated region ($V_{GS}>V_T$, $V_{DS}>V_{GS}-V_T$ or $V_{GD}<V_T$)

$$I_{DS} = k_n \frac{(V_{GS}-V_T)^2}{2}$$

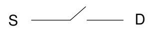
I-V Characteristics



Long channel transistor ($L = 10\mu m$)

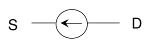
MOS Transistor

- Cutoff region ($V_{GS}<V_T$)



- Linear region ($V_{GS}>V_T$, $V_{DS}<V_{GS}-V_T$)

- Saturated region ($V_{GS}>V_T$, $V_{DS}>V_{GS}-V_T$)



Secondary Effects

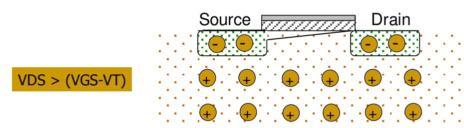
- Body effect
- Channel-length modulation
- Drain punch-through
- Short channel effect
- Velocity saturation

Body Effect

- We assumed that $V_{SB}=0$ - i.e. the source potential equals the substrate potential
- In certain situations, this assumption is not true
- Has the effect of raising the threshold voltage
  - A negative bias on the well or substrate causes the threshold to increase

Channel-Length Modulation

- We previously assumed a constant L
- In reality, when V_{DS} > ($V_{GS}-V_T$), the channel is pinched off and the effective channel length is reduced.
- Net effect is that I_{DS} is not constant in the saturated region.



MOS Transistor

- Cutoff region ($V_{GS}<V_T$)

$$I_{DS}=0$$

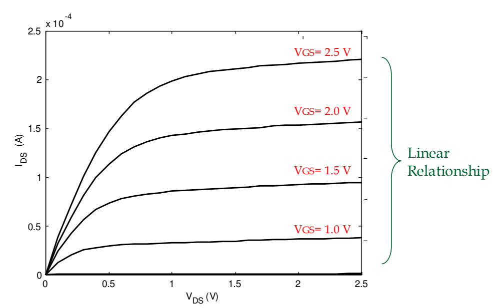
- Linear region ($V_{GS}>V_T$, $V_{DS}<V_{GS}-V_T$)

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

- Saturated region ($V_{GS}>V_T$, $V_{DS}>V_{GS}-V_T$)

$$I_{DS}=k_n\frac{(V_{GS}-V_T)^2}{2}(1+\lambda V_{DS})$$

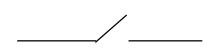
Channel-Length Modulation



Short channel transistor ($L = 0.25\mu m$)

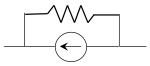
MOS Transistor

- Cutoff region ($V_{GS}<V_T$)



- Linear region ($V_{GS}>V_T$, $V_{DS}<V_{GS}-V_T$)

- Saturated region ($V_{GS}>V_T$, $V_{DS}>V_{GS}-V_T$)



Short Channel Effect

- At small gate lengths, electric field becomes more pronounced
- Electrons get excited with enough energy to cause a substrate current
- This causes change of transistor parameters - threshold voltage, current flow, etc.

Velocity Saturation

- Assumption was that carrier velocity is proportional to electric field
- When channel is small, and the voltage is large, the velocity can saturate

$$\upsilon = \begin{cases} \mu_n \xi & \xi < \xi_c \\ \mu_n \xi_c & \xi > \xi_c \end{cases}$$

ξ_c is value of electric field at which velocity saturates

MOS Transistor

- Cutoff region ($V_{GS}<V_T$)

$$I_{DS} = 0$$

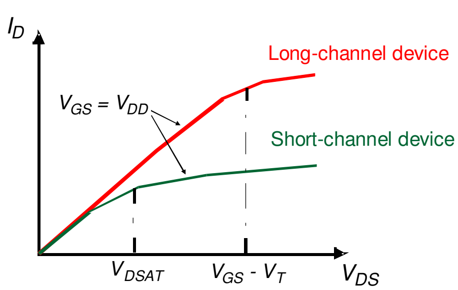
- Linear region ($V_{GS}>V_T$, $V_{DS}<V_{GS}-V_T$)

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

- Saturated region ($V_{GS}>V_T$, $V_{DS}>V_{GS}-V_T$)

$$I_{DS} = k_n\left[(V_{GS}-V_T)V_{DSAT} - \frac{V_{DSAT}^2}{2}\right]$$

Velocity Saturation



MOS Gain Characteristics

- Transconductance $g_m = \frac{dI_{DS}}{dV_{GS}}$
  - Cutoff region
    $g_m = 0$
  - Linear region
    $g_m = k_n V_{DS}$
  - Saturated region
    $g_m = k_n(V_{GS}-V_T)$

nMOS Transistor

- Cutoff region ($V_{GSn}<V_{Tn}$)

$$I_{DSn}=0$$

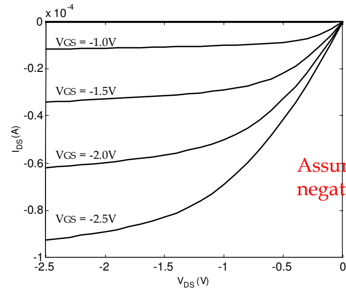
- Linear region ($V_{GSn}>V_{Tn}$, $V_{DSn}<V_{GSn}-V_{Tn}$)

$$I_{DSn}=k_n\left[(V_{GSn}-V_{Tn})V_{DSn}-\frac{V_{DSn}^2}{2}\right]$$

- Saturated region ($V_{GSn}>V_{Tn}$, $V_{DSn}>V_{GSn}-V_{Tn}$)

$$I_{DSn}=k_n\frac{(V_{GSn}-V_{Tn})^2}{2}(1+\lambda V_{DSn})$$

pMOS I-V



Assume all variables negative!

pMOS Transistor

- Cutoff region ($V_{GSp}>V_{Tp}$)

$$I_{DSp}=0$$

- Linear region ($V_{GSp}<V_{Tp}$, $V_{DSp}>V_{GSp}-V_{Tp}$)

$$I_{DSp}=-k_p\left[(V_{GSp}-V_{Tp})V_{DSp}-\frac{V_{DSp}^2}{2}\right]$$

- Saturated region ($V_{GSp}<V_{Tp}$, $V_{DSp}<V_{GSp}-V_{Tp}$)

$$I_{DSp}=-k_p\frac{(V_{GSp}-V_{Tp})^2}{2}(1+\lambda V_{DSp})$$