I] MOSFET

Metal-Oxide-Semiconductor-Field-Effect-Transistor

- most ~~usually~~ widely used semiconductor device

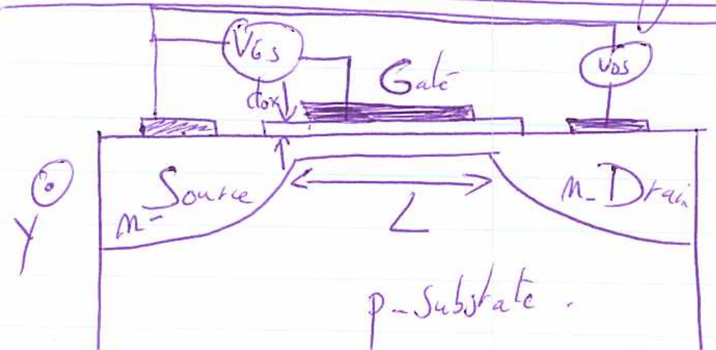
[Figure 7.1 of textbook → Moore's law; exponential growth of integration density with time.

- 2 types of MOSFET → n-channel MOSFET or nFET
  → p-channel MOSFET or pFET

[the most commonly used technology is CMOS (Complementary MOS) in which both n-channel and p-channel transistor are fabricated

- we limit here our analysis to nFET where ~~e~~ ~~are responsible~~ the current flow is due to e⁻

① Introduction to MOSFET operation



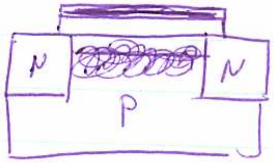
L = channel length

dox = oxide thickness

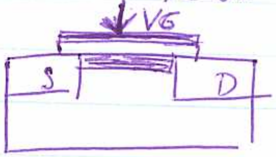
W = channel width. in Y direction

operation principles

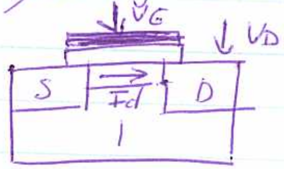
(i) * Transistor is "off" state



* Applying a voltage V_T to the gate "inverts" the channel region creating an electrical path between the source and drain.



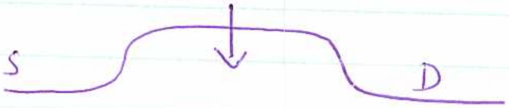
* Applying a voltage to the Drain pulls current-carriers across the channel, creating the drive current (I_D)



* The role of gate electrode for n-channel MOSFET

Positive gate voltage does two things-

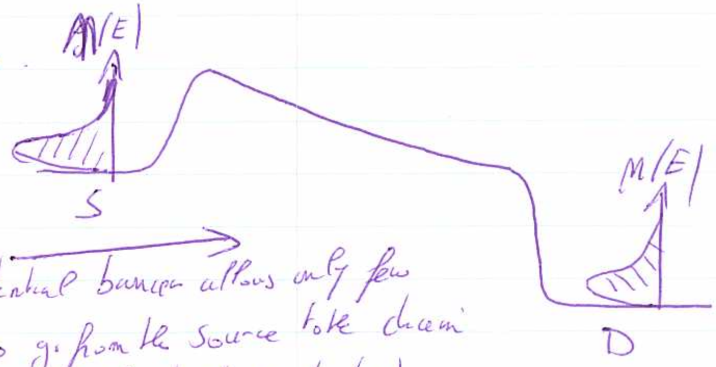
(1) Reduces the potential energy barrier seen by the e^- from the source and the drain regions



(2) Inverts the surface, and increases the conductivity of the channel.

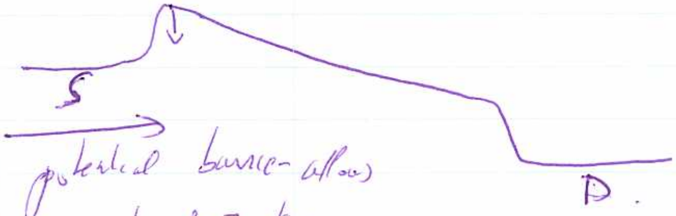
* The role of the Drain electrode for n-channel MOSFET

$V_G = 0$, $V_D > 0$.



Large potential barrier allows only few e^- to go from the source to the drain (subthreshold conduction).

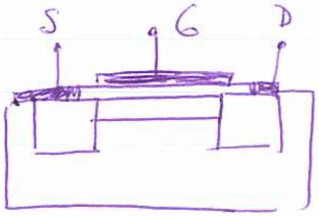
$V_G > 0$, $V_D > 0$.



Smaller potential barrier allows a large number of e^- to go from the source to the drain.

* Qualitative description of MOSFET operation

ⓐ $V_G > V_T$, $V_D > 0$ (small)

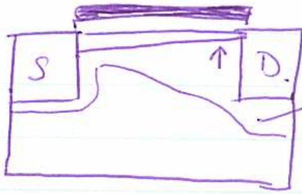


Variation of e^- density along the channel is small

$I_D \sim V_D \Rightarrow$ linear regime

ⓑ $V_G > V_T$, $V_D > 0$ (large).

Increase in the drain current reduce due to the reduced conductivity of the channel at the drain end



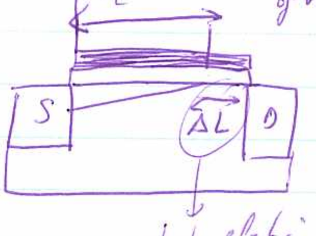
ⓒ $V_G > V_T$, $V_D =$ ~~$V_G - V_T$~~ V_{Dsat}
Drain saturation voltage.

Pinch-off point: e^- density at the end channel is identically zero.

ⓓ $V_G > V_T$; $V_D > V_{Dsat}$

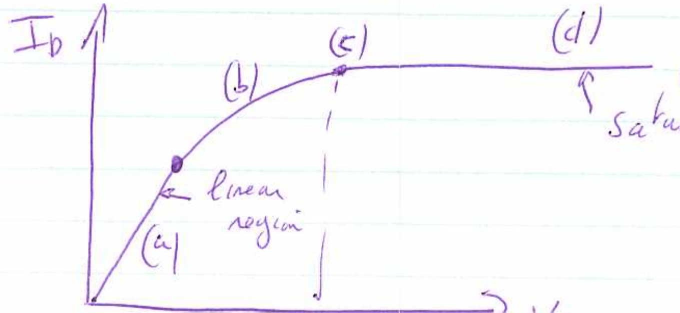
Post pinch-off characteristic.
the excess drain voltage is dropped across the highly resistive pinch-off region denoted by ΔL

potential across the channel $[V_{Dsat} - V_S]$ independent of VDS



high-electric field ⇒ drift of carriers.
⇒ current saturates since it is fixed by the potential dropped across the channel.

⇒



saturation region.

② MOSFET analysis

ⓐ The linear Model

it describes the behavior of a MOSFET biased with a small drain-to-source voltage.

* MOSFET acts as a linear device ⇒ can be modeled as a linear resistor whose resistance is modulated by the gate voltage.

General expression for the Drain current.

$$I_D = -\frac{Q_{inv} W L}{t_r}$$

Q_{inv} = inversion layer charge per unit of area.

W = gate width

L = gate length.

t_r = transit time $\boxed{t_r = \frac{L}{v}}$ v = velocity

$$\boxed{v = \underset{\text{mobility}}{\mu}\,\underset{\text{electric field}}{\mathcal{E}} = \mu \frac{V_{DS}}{L}}$$

$$\Rightarrow \boxed{I_D = -\mu Q_{inv} \frac{W}{L} V_{DS}.}$$

we saw that (MOS capacitor) $V_G = V_T - \frac{Q_{inv}}{C_{ox}}$

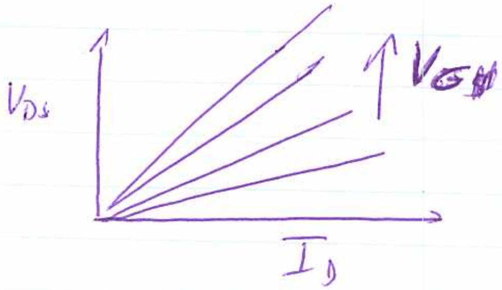
$$\Rightarrow \boxed{Q_{inv} = -C_{ox}(V_G - V_T)} \text{ if } V_G > V_T$$

with $\boxed{V_T = V_{FB} + V_S + \frac{\sqrt{2\epsilon q N_A V_S}}{C_{ox}}}$ $V_S = 2V_F$

$$\text{so } \boxed{I_D = \mu C_{ox} \frac{W}{L}(V_G - V_T)V_{DS}}$$

for $|V_{DS}| \ll (V_G - V_T)$

- Also $I_D = 0$ if $V_G < V_T$.
- $|V_{DS}| \ll (V_G - V_T) \Rightarrow$ means that n, E, Q_{inv} is constant between source and drain.



ⓑ The quadratic model

The model for the charge includes the voltage variation between source and drain. The variation of the depletion layer charge is ignored, not the variation of the inversion layer charge. it is the most commonly used model.

⑦

[the charge on the gate is completely balanced by $Q_N(x)$
$Q_N(x) \simeq -C_{ox}(V_G - V_T - V_c(x))$]

We consider a small section with width dy, and channel voltage V_c~~$+V_S$~~.

~~$V_{GS} \Rightarrow V_G - V_c$ inside the channel. since $V_S \Rightarrow$~~ [illegible]

so the previous expression from the linear model become

$$\boxed{I_D = \mu C_{ox} \frac{W}{dy} (V_G - V_c - V_T)\, dV_c}$$

where $\left[\begin{array}{l} L \rightarrow dy \quad \text{locally} \\ V_{DS} \rightarrow dV_c \end{array}\right.$

integrating from source to drain.

$$\int_0^L I_D\, dy = \mu C_{ox} W \int_0^{V_{DS}} (V_G - V_c - V_T)\, dV_c$$

I_D is constant along the channel so $\Longrightarrow$

$$\boxed{I_D = \mu C_{ox} \frac{W}{L} \left[(V_G - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right]} \quad (*)$$

$V_{DS} < V_{GS} - V_T$

the Drain current first increases linearly with the applied drain to source voltage, but then reaches a maximum value.

if $\boxed{V_{DS} = V_G - V_T} \equiv V_{DS,sat} \left[\frac{dI_D}{dV_{DS}} = 0 \right]$

Drain saturation voltage, $V_{DS,sat}$

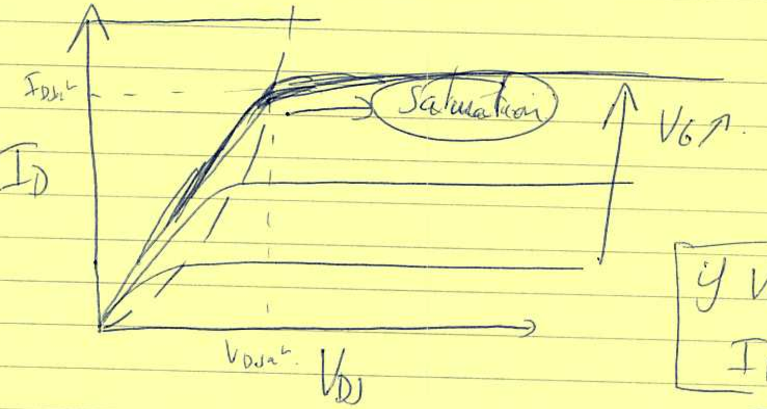
We call $I_{D,satt}$, the value of the saturated drain current. ⇒ pinch-off point.

$Q_{inv}=0$ at the Drain interface

$$\Rightarrow \boxed{I_{D,SAT} = \mu C_{ox} \frac{W}{L}\left(\frac{V_{G}-V_T}{2}\right)^2}$$

The equation (*) is not valid beyond the pinch-off point if $\boxed{V_{DS} > V_{G}-V_T}$ ⇒ the current maintains the value $I_{D,SAT}$.

The quadratic model explains the typical I-V characteristics of a MOSFET, which are normally plotted for different V_G.



$\boxed{\text{if } V_{GS} < V_T \quad I_D = 0}$

* We would like to calculate the transconductance gm and

(i) the ~~output~~ transconductance gm is given by

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS}\ cte}$$ (if gm ↗ speed of the device ↗

Here we get in the quadratic region $\boxed{g_{m,quad} = \mu C_{ox} \frac{W}{L} V_{DS}}$ if $V_{DS} < V_{GS} - V$

In saturation $\boxed{g_{m,sat} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T)}$

(ii) the output conductance gd is given by

$$g_d = \left.\frac{\partial I_D}{\partial V_{DS}}\right|_{V_{GS}}$$ [" ease of with which an electric current flows through a material" → reciprocal of resistance.]

$\boxed{g_d = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T - V_{DS})}$ [in the quadratic region.]

gd ↘ if V_{DS} ↗.

and $g_d \longrightarrow 0$ as the device operated in the saturated region $\boxed{g_{d,sat} = 0.}$

ⓒ The variable depletion layer model

include now the variation of the charge in the depletion layer between Source and Drain.

$$Q_{inv} = -C_{ox}(V_{GS}-V_T) \quad V_{GS}>V_T$$

however now ~~the~~ we include the ~~explicit~~ implicit dependence of V_T on the charge ~~of~~ in the depletion layer.

$$V_T = V_{FB} + V_S + \frac{\sqrt{2\epsilon q N_A V_S}}{C_{ox}} \qquad \boxed{V_S = 2V_F + V_C}$$

$$\Rightarrow \boxed{V_T = V_{FB} + 2V_F + V_C + \frac{\sqrt{2\epsilon q N_A(2V_F+V_C)}}{C_{ox}}}$$

the voltage ~~[illegible]~~ V_C is the ~~difference~~ channel voltage $V_C(y)$

$$\Rightarrow I_D = \mu C_{ox}\frac{W}{dy}\left(V_{GS} - V_{FB} - 2V_F - V_C(y) - \frac{\sqrt{2\epsilon q N_A(2V_F+V_C(y))}}{C_{ox}}\right)dV$$

$\downarrow$ y varies from 0 to L. $V_C(y)$ varies from 0 to V_{DS}

$$\left(\begin{aligned} I_D &= \mu C_{ox}\frac{W}{L}\left(V_G - V_{FB} - 2V_F - \frac{V_{DS}}{2}\right)V_{DS} \\ &- \frac{2}{3}\mu\frac{W}{L}\sqrt{2\epsilon q N_a}\left[(2V_F + V_{DS})^{3/2} - (2V_F)^{3/2}\right]\end{aligned}\right.$$

on integration from V_S to V_D

$$\boxed{I_D = \mu C_{ox}\frac{W}{L}\left\{\left(V_G - V_{FB} - 2V_F - V_S - \frac{V_{DS}}{2}\right)V_{DS} - \frac{2}{3}\gamma\left[\left(2V_F + V_D\right)^{3/2} - \left(2V_F + V_S\right)^{3/2}\right]\right\}}$$

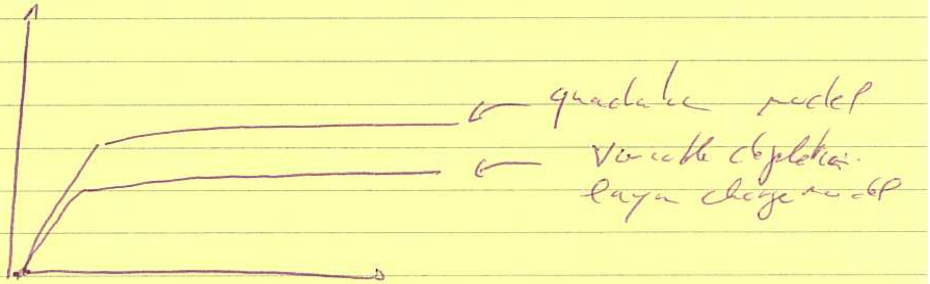
$$\gamma = \frac{\sqrt{2q\epsilon_s N_a}}{C_{ox}}$$

Again this expression is not valid after the pinch-off point for describing the saturation regime.

the drain voltage at which saturation occurs can be ~~$V_{D,sat} = V_G$~~ found setting $\frac{dI_D}{dV_D} = 0$.

$$\Rightarrow \boxed{V_{Dsat} = V_G - V_{FB} - 2V_F + \frac{\gamma^2}{2} - \gamma\sqrt{V_G - V_{FB} + \frac{\gamma^2}{4}}}$$

Actually the quadratic model yields to larger current compared to this more accurate derivation.



it comes for the transconductance.

~~$g_{m,sat} = \mu_n C_{ox} \frac{W}{L} \frac{W}{L} [V_G - V_{FB} - 2V_F$~~

$$g_{m,\cancel{sat}} = \mu C_{ox} \frac{W}{L} V_{DS} \quad \text{(still)}$$

and
$$\boxed{g_{m,sat} = \mu C_{ox} \frac{W}{L} V_{Dsat}.} \Rightarrow \text{almost linear with } V_G$$

So it can be also written

$$\boxed{g_{m,sat} = \mu^* C_{ox} \frac{W}{L} (V_G - V_T)} \quad \text{with modified } \mu^*$$

$$\mu^* = \mu \left[1 - \frac{1}{\sqrt{1 + \frac{2(2V_F + V_S) C_{ox}^2}{q N_a \epsilon_{si}}}} \right]$$

ⓓ Influence of substrate bias on threshold voltage

we saw that $\boxed{V_T = V_{FB} + 2V_F + \gamma\sqrt{2V_F}}$

$$\gamma = \frac{\sqrt{2q\epsilon_{si}N_a}}{C_{ox}}$$

where we supposed that both the source and the substrate were grounded. $V_s - V_{sub} = 0$.

In many applications, the source and the substrate may be at different potential.

the depletion charge under the channel is obtained by.

$$Q_d(y) = -\sqrt{2q\epsilon_{si}N_a(2V_F + V_c(y) - V_{sub})} \qquad \underline{V_{sub} < 0}$$

$$\left[V_T = V_{FB} + 2V_F + V_c(y) + \gamma\sqrt{2V_F + V_c(y) - V_{sub}}\right]$$

at $y=0 \Rightarrow V_c(y) = V_s$

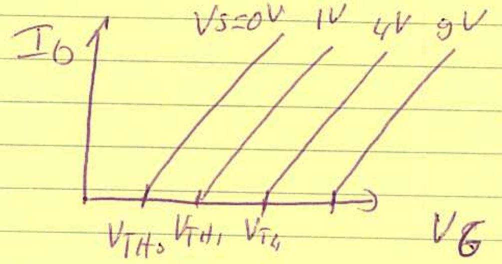
if (i) $V_{sub} = 0$; $V_s > 0 \Rightarrow$

$$V_{TH} = V_s + V_{FB} + 2V_F + \gamma\sqrt{2V_F + V}$$

(ii) $V_{sub} < 0$; $V_s = 0$

$$V_{TH} = V_{FB} + 2V_F + \gamma\sqrt{2V_F - V_{sub}}$$

In conclusion, V_T increases as a function of the potential difference between source and substrate.



V_{DS} is small and $V_{sub} = 0$ (grounded).

③ Surface mobility

The mobility μ_n that we used is the one associated with the "surface mobility". Surface mobility is lower than bulk mobility because of scattering at the silicon-oxide interface.
The surface mobility depends on how much e^- interact at the interface, and therefore it depends on the vertical electrical field which 'pushes' e^- against the interface.
if $\vec{E} = 0$ we denote $\mu_n \equiv \mu_{n0}$.

The current in the transistor is given by

$$\boxed{I_D = \frac{W}{L}\int_{V_S}^{V_D} \mu_n(-Q_{inv}(y))\, d(V(y))}$$

[illegible] [illegible] the electric

There exists an empirical relationship that describes the dependence of surface mobility on vertical electric field in the channel, $\mathcal{E}_x$.

$$\mu_n(y) = \frac{\mu_{no}}{1 + \theta\, \mathcal{E}_x(y)}$$

θ = mobility reduction factor.

The average electric field in the channel is given by

$$\mathcal{E}_x(y) = \frac{\mathcal{E}_{so}(y) + \mathcal{E}_{si}(y)}{2}$$

$\mathcal{E}_{so}$ = electric field at the silicon-oxide interface

$\mathcal{E}_{si}$ = electric field at the boundary between inversion layer and depletion region

According to Gauss' Law

(i) $\epsilon_{si}\, \mathcal{E}_{so}(y) = \epsilon_{ox}\, \mathcal{E}_{ox}(y)$. $\quad \mathcal{E}_{ox} = \dfrac{\left(V_G - V_{FB} - V_{s}(y)\right)}{d_{ox}}$ $\quad$ surface $\cong 2\phi_F + V_{CH}$

$$\Rightarrow \epsilon_{si}\, \mathcal{E}_{so}(y) = C_{ox}\left[V_G - V_{FB} - V_{surf}(y)\right] = -\left[Q_{inv}(y) + Q_d(y)\right]$$

$$= Q_G(y)$$

(ii) $\epsilon_{si}\, \mathcal{E}_{si}(y) = -Q_d(y) = \gamma\, C_{ox} \sqrt{V_{surf}(y)}$

$$\Rightarrow \boxed{\mathcal{E}_x(y) = \frac{C_{ox}}{2\epsilon_{si}}\left(V_G - V_{FB} - V_{surf}(y) + \gamma\sqrt{V_{surf}(y)}\right)}$$

$$\Rightarrow \mu_n(y) = \frac{\mu_{n0}}{1+\theta\frac{C_{ox}}{2\epsilon_{si}}\left(V_G-V_{FB}-V_{surf}(y)+\gamma\sqrt{V_{surf}(y)}\right)}$$

We want to simplify the calculation of the Drain current.

$$\boxed{I_D \text{ calculated with } \mu_n(y) \text{ depending on } E_x(y) \equiv I_D \text{ calculated with a cte } \mu_{eff}}$$

- Let us consider a small element dy.

$$\Rightarrow I_D\,dy = W\mu_{eff}\left(-Q_{in}(y)\right)dV(y)$$

$$\Rightarrow I_D\cdot dy = \frac{W\mu_{n0}\left(-Q_{in}(y)\right)dV(y)}{1+\theta\frac{C_{ox}}{2\epsilon_{si}}\left(V_G-V_{FB}-V_{surf}(y)+\gamma\sqrt{V_{surf}(y)}\right)}$$

we substitute and integrate from source to drain.

$$\left[\int_0^L\left\{1+\theta\frac{C_{ox}}{2\epsilon_{si}}\left(V_G-V_{FB}-V_{surf}(y)+\gamma\sqrt{V_{surf}(y)}\right)\right\}dy = L\frac{\mu_{n0}}{\mu_{eff}}\right] \quad (*)$$

<u>Remark</u> $V_{surf}(y) = 2V_F + V_c(y)$

$$\frac{dV}{dy} \simeq \frac{V_D-V_S}{L} \quad \text{(linear approximation).}$$

we can write the following identity:

$$\int_0^L \left[V_C(y) - \gamma\sqrt{2V_F + V_C(y)}\right] dy = \frac{L}{V_D - V_S} \int_{V_S}^{V_D} \left[V_C - \gamma\sqrt{2V_F + V_C}\right] dV$$

$$= \frac{L}{V_D - V_S}\left[\frac{V_D^2}{2} - \frac{V_S^2}{2} - \frac{2}{3}\gamma\left\{(2V_F + V_D)^{3/2} - (2V_F + V_S)^{3/2}\right\}\right]$$

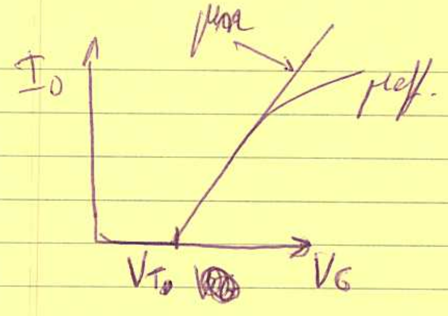
introducing in (*), we get (after calculation).

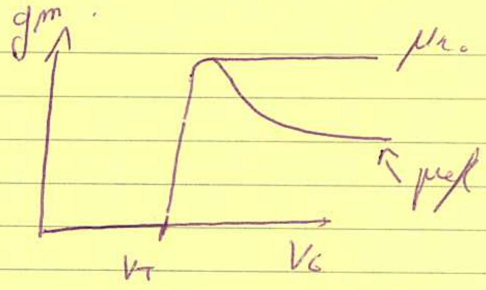
$$\frac{\mu_{no}}{\mu_{eff}} = 1 + \frac{\theta C_{ox}}{2\epsilon_{si}}\left[V_G - V_T + \gamma\sqrt{2V_F} - \frac{V_S + V_D}{2} + \frac{2}{3}\gamma\frac{(2V_F + V_D)^{3/2} - (2V_F + V_S)^{3/2}}{V_{DS}}\right.$$

$$\Rightarrow \boxed{\mu_{eff} = \frac{\mu_{no}}{1 + \frac{\theta C_{ox}}{2\epsilon_{si}}\left(V_G - V_T - V_S + 2\gamma\sqrt{2V_F + V_S}\right)}}$$

~~Assuming~~ Neglecting the influence of depletion charge near the source; after simplification and also with $\alpha = \frac{\theta C_{ox}}{2\epsilon_{si}}$

$$\Rightarrow \boxed{\mu_{eff} = \frac{\mu_{no}}{1 + \theta(V_G - V_{TH} - V_S)}}$$





④ Carrier velocity saturation

So far we consider a linear dependency of the drift current on the lateral electric field. ——→ y

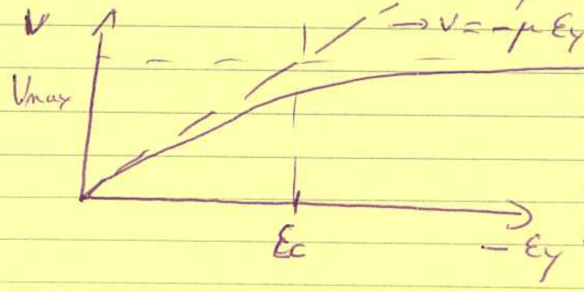
$$\boxed{q \mu_n n E_y = -q v n}$$

v = velocity of carrier in the inversion layer

Actually, this is working for low electric field only. At high electric field, and above a critical value, the carrier velocity saturates.

ex = e^- in silicon, max velocity is 10^7 cm/s

$$\boxed{E_c = \frac{v_{max}}{\mu_n}}$$



if $E_y \to +\infty$

v can be expressed as follows

$$\boxed{v = v_{max} \frac{E_y/E_c}{E_y/E_c - 1}}$$

what is the impact of the velocity saturation effect on the expression of the drain current of a MOSFET_

if $V_D \geq V_{Dsat}$ $\quad \mathcal{E}_y$ reaches high value near the drain junction_

the drain current is:

$$\boxed{I_D = -W\, Q_{in}(y)\, \mu_n \frac{dV(y)}{dy} = W\, Q_{in}\, \mu_n\, \mathcal{E}_y}$$

if $\mathcal{E}_y \to +\infty$ then $Q_{in}(y) \to 0$.

from $v = v_{max} \dfrac{\mathcal{E}_y/\mathcal{E}_c}{\mathcal{E}_y/\mathcal{E}_c - 1}$ and using $\left[\begin{array}{l} \mathcal{E}_c = v_{max}/\mu_n \\ \mathcal{E}_y = -dV(y)/dy \end{array} \right.$

$$\Longrightarrow \quad V(y) = V_{max} \frac{-\frac{1}{\mathcal{E}_c}\frac{dV(y)}{dy}}{-\frac{1}{\mathcal{E}_c}\frac{dV(y)}{dy} - 1} = \mu_n \frac{\frac{dV(y)}{dy}}{1 + \frac{1}{\mathcal{E}_c}\frac{dV(y)}{dy}}$$

the expression of the current becomes.

$$\boxed{I_D \doteq -W \left(\frac{\mu_n}{1 + \frac{1}{\mathcal{E}_c}\frac{dV(y)}{dy}} \right) Q_{in}(y) \frac{dV(y)}{dy}}$$

just replacing $\mu_n \to \dfrac{\mu_n}{1 + \frac{1}{\mathcal{E}_c}\frac{dV(y)}{dy}}$ to account for velocity saturation integrating from source to Drain yields.

$$I_D = -\frac{W}{L}\left(\frac{\mu_n}{1+\frac{V_{DS}}{L\mathcal{E}_c}}\right)\int_{V_S}^{V_D} Q_{inv}(y)\,dV(y).$$

using quadratic model

$$\Rightarrow \boxed{I_D = \frac{W}{L}\left(\frac{\mu_n}{1+\frac{V_{DS}}{L\mathcal{E}_c}}\right) C_{ox}\left[(V_G-V_T)V_{DS} - \frac{V_{DS}^2}{2}\right]}$$

by imposing $\frac{dI_D}{dV_{DS}} = 0 \Rightarrow$ drain saturation voltage is

$$\boxed{V'_{D\,sat} = L\mathcal{E}_c\left[\sqrt{1+\frac{2(V_G-V_T)}{L\mathcal{E}_c}} - 1\right]}$$

when velocity saturation is taken into account this is equivalent to making the channel longer L is multiply by $\left(1+\frac{V_{DS}}{L\mathcal{E}_c}\right)$.

Therefore $\Rightarrow$ the Drain saturation voltage and the drain saturation current are reduced.

21

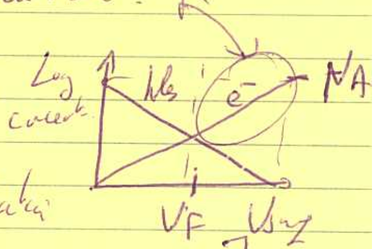
(5) Subthreshold ~~current / Subthreshold slope~~ Characteristics

(a) Subthreshold current

So far we assumed that $I_D = 0$ if $V_G \leq V_T$

Actually there is a significant amount of e^- near the surface if the device operates ~~at~~ below strong inversion.

$V_F < V_{surf} < 2V_F$ Remember



[the actual dependence of the e^- concentration at the surface is an exponential function of V_{surf}]

⇒ Experimentally we prove that the Drain current below threshold "subthreshold current" is independent of the drain voltage. ⇒ this means that the subthreshold current is caused by diffusion rather than drift mechanism.

current density from Source to Drain can be written

$$J_{ny} = -qD\frac{dn}{dy} \Rightarrow \boxed{I_D = qAD_n\frac{n(0)-n(L}{L}}$$

A is the cross-sectional area of the channel region.

e^- density at the edge of source and drain =

$$n(0) = n_{p0} \exp\left(\frac{q V_{surf}}{k_B T}\right) \qquad n(L) = n_{p0} \exp\left(\frac{q (V_{surf} - V_D)}{k_B T}\right)$$

$$n_{p0} = \frac{n_i^2}{N_a}$$

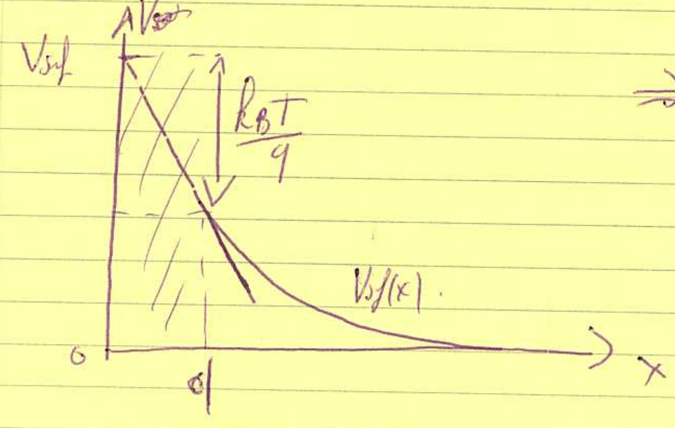
(→ Source is considered here at ground).

what is the area A?

we know that the e^- density varies as $\exp\left(\frac{q V_{surf}}{k_B T}\right)$ near the surface. we however going to

approximate the exp-profile by a constant e^- density extending to a depth d below the surface.

the depth d is defined as the depth a which the potential has decreased by k_BT/q below the surface potential value



⇒ so $A = W \times d$

W (width of the transistor)

$$d = \frac{k_BT/q}{\mathcal{E}_s} \qquad \mathcal{E}_s = -\left.\frac{dV_{sf}}{dx}\right|_0$$

Using the ~~[illegible]~~ Einstein relationship $\mu_n \frac{k_BT}{q} = D_n$

we get the subthreshold current

$$\boxed{I_D = \mu_n \frac{W}{L} q \left(\frac{k_BT}{q}\right)^2 \frac{n_i^2}{N_A} \left[1 - \exp\left(-qV_D/k_BT\right)\right] \times \frac{\exp\left(q V_{surf}/k_BT\right)}{-\frac{dV_{surf}}{dx}}}$$

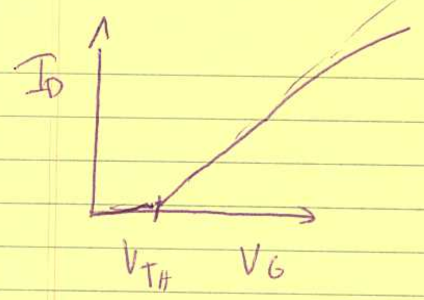
the electric field at the surface $\mathcal{E}_s$ is given by

(MOS capacitor section) $-\frac{dV_{surf}}{dx}\Big|_{x=0} = \mathcal{E}_s = \sqrt{\frac{2q N_A V_{surf}}{\epsilon_{si}}}$

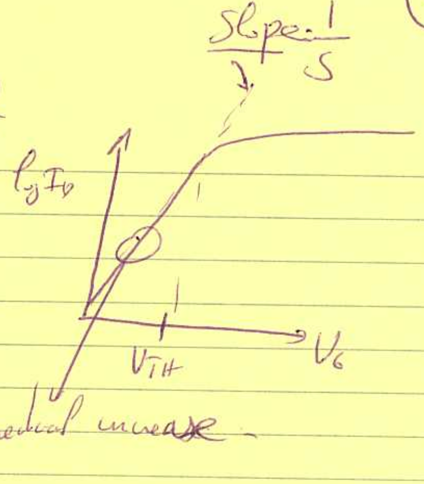
$$\Rightarrow \boxed{\mathcal{E}_s = q\frac{N_A}{C_D}}$$

⇒ I_D is independent of V_D. as long as V_D is larger than a few $\frac{k_BT}{q}$. (exponential term → 0)

⇒ I_D also increases exp. with the surface potential.



$\xrightarrow{\text{in log scale}}$



(b) Subthreshold slope

it is $\frac{1}{S}$ (inverse of the slope) $\Longrightarrow$ subthreshold slope $\boxed{S = \frac{dV_G}{d\log(I_D)}}$

⇒ represents how many millivolts per decade

⇒ "how many millivolts V_G be increased to increase I_D by a factor 10"

if S ↓ efficiency and switching speed of the device from off-state to on-state ↑

$$S = \frac{\ln(10)}{\frac{d(\ln(I_D))}{dV_G}}$$

$\log_{10} \rightarrow \log_e \equiv \ln$.

Also $$\frac{d\ln(I_D)}{dV_G} = \frac{d(\ln I_D)}{dV_{surf}} \frac{dV_{surf}}{dV_G}$$

$$\frac{d\ln(I_D)}{dV_{surf}} = \frac{d}{dV_{surf}}\left[\ln\left(e^{+\frac{qV_{surf}}{k_BT}}\right) - \ln\left(\frac{dV_{surf}}{dx}\right)\right]$$

$$= \left[\frac{q}{k_BT} - \frac{\frac{d}{dV_{surf}}\left(-\frac{dV_{surf}}{dx}\right)}{-\frac{dV_{surf}}{dx}}\right] \quad (*)$$

$$\frac{d}{dV_{surf}}\left(-\frac{dV_{surf}}{dx}\right) = \frac{C_D}{\epsilon_{si}} \leftarrow \text{(from Poisson)}$$

C_D is the depletion capacitance $\boxed{C_D = -\frac{dQ_d}{dV_{surf}}} = \frac{\epsilon_{si}}{x_d}$

$$\Rightarrow \frac{\frac{d}{dV_{surf}}\left(-\frac{dV_{surf}}{dx}\right)}{-\frac{dV_{surf}}{dx}} = \frac{C_D^2}{qN_a\epsilon_{si}} = \frac{1}{2V_{surf}}$$

in weak inversion regime $V_F < V_{surf} < 2V_F$,

$\frac{1}{2V_{surf}}$ is small compared to $\frac{q}{k_BT}$ ⇒ it comes

(second term in * is neglected).

$$\boxed{\frac{d(\ln(I_D))}{dV_G} \simeq \frac{q}{k_BT}\frac{dV_{surf}}{dV_G}}$$

Also $\left[V_G = V_{FB} + V_{surf} - \frac{Q_d}{C_{ox}}\right]$

$$\Rightarrow \frac{dV_{surf}}{dV_G} = \left(1 + \frac{C_D}{C_{ox}}\right)^{-1} \quad \left[\text{do } \frac{dV_G}{dV_{surf}} \text{ first}\right]$$

⇒ finally

$$\boxed{S = \frac{k_B T}{q} \ln(10) \left(1 + \frac{C_D}{C_{ox}}\right)} \qquad \boxed{\left(1 + \frac{C_D}{C_{ox}}\right) = \underset{\uparrow}{n}}$$

body factor

Also since $\frac{d(\ln I_D)}{dV_G} = \frac{q}{k_B T n}$

$$\Rightarrow \boxed{I_D \sim \exp\left(\frac{qV_G}{n k_B T}\right)}$$

the subthreshold current varies exponentially as a function of the gate voltage.

* In case where we consider interface states (or traps) in the Si energy bandgap at the silicon/oxide interface. One can ~~define~~ associate a charge of these interfaces state Q_{it}.

$$\Rightarrow C_{it} = \frac{-dQ_{it}}{dV_{surf}}.$$

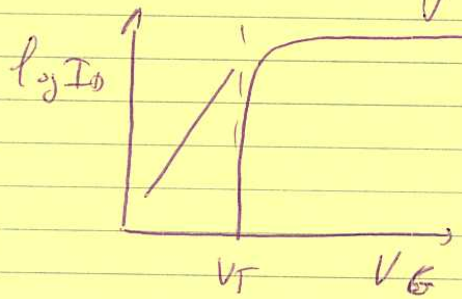
we get after calculations.

$$\boxed{S = \frac{k_B T}{q} \ln(10) \left(1 + \frac{C_D + C_{it}}{C_{ox}}\right)}$$

$$= 60 \left(1 + \frac{C_D + C_{it}}{C_{ox}}\right) \text{ mV/decade}$$ at room temperature

⑥ Continuous model

The model developed so far for $V_G > V_T$ or for $V_G < V_T$ do not connect well around $V_G = V_T$ (discontinuity). Since often V_G is close to V_T, this is a problem.



A model which is valid everywhere can be derived.

Indeed, $I_D = \mu_n C_{ox} \frac{W}{L} \int_{V_S}^{V_D} -\frac{Q_{inv}(y)}{C_{ox}}\, dV(y)$.

can be rewritten

$$I_D = \mu_n C_{ox} \frac{W}{L} \left\{ \int_{V_S}^{V_P} -\frac{Q_{inv}(y)}{C_{ox}}\, dV(y) - \int_{V_D}^{V_P} -\frac{Q_n(y)}{C_{ox}}\, dV(y) \right\}$$

$V_P = (V_G - V_T) \equiv$ drain saturation voltage if source is grounded and $V_G > V_T$

$$I_D = \underset{\substack{\uparrow \\ \text{forward current}}}{I_F} - \underset{\substack{\uparrow \\ \text{Reverse current}}}{I_R}$$

It is possible to find a mathematical function which describes the evolution of the current as a function of gate and drain voltage for all regime of operation (depletion, weak and strong inversion)

⟹ Enz, Krummenacher, Vittoz

⟹ EKV Model

$$I_D = 2\mu_n C_{ox} \frac{W}{L} \left(\frac{k_B T}{q}\right)^2 \left\{ \cancel{\tfrac{1}{4}} \left[\ln\left(1 + \exp\left(\frac{V_P - V_S}{2 k_B T/q}\right)\right)\right]^2 - \left[\ln\left(1 + \exp\left(\frac{V_P - V_D}{2 k_B T/q}\right)\right)\right]^2 \right\}$$

For example:

if $V_S = 0$, $V_D < V_P$ $V_G > V_T$ ⇒ the transistor operates in non-saturated regime.

⇒ exponential terms ≫ 1

$$I_D = 2\mu_n C_{ox}\frac{W}{L}\left(\frac{k_BT}{q}\right)^2\left[\left(\frac{V_P}{2k_BT/q}\right)^2-\left(\frac{V_P-V_D}{2k_BT/q}\right)^2\right]$$

$$= \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left[2V_DV_P - V_D^2\right].$$

$$\Rightarrow \boxed{I_D = \cancel{\tfrac{1}{2}}\,\mu_n C_{ox}\frac{W}{L}\left[(V_G - V_T)V_D - \frac{V_D^2}{2}\right]}$$

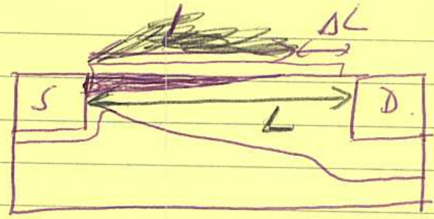
quadratic model

⑦ Channel length modulation

Previously we assumed that when $V_D > V_{Dsat}$, the Drain current of a MOSFET is constant and equal to I_{Dsat}.

Actually if V_D ↗ beyond V_{Dsat}, the depletion region and local threshold voltage near the Drain are increased the effective length of the channel is $L-\Delta L$

~~(ΔL [illegible] region)~~



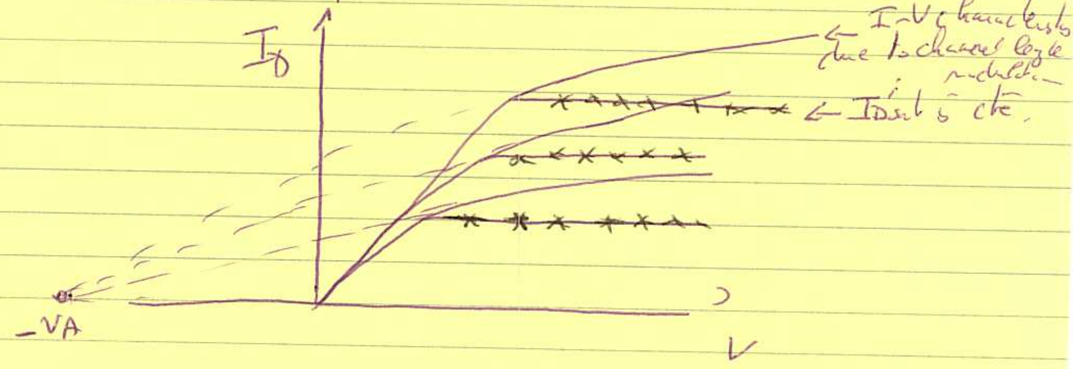
and shorter channel length ⟹ increase in Drain current.

if I_{Dsat} ~~[illegible]~~ is obtained if $(V_D = V_{Dsat})$

if $V_D > V_{Dsat}$

$$I'_{Dsat} = I_{Dsat}\left(1 + \frac{V_D - V_{Dsat}}{V_A + V_{Dsat}}\right)$$

where V_A is a positive voltage value that can be obtained experimentally. V_A is called "early voltage"



in term of channel length modulation, we get

$$\frac{I'_{Dsat}}{L} = \frac{I_{Dsat}}{L-\Delta L}$$

$$\Rightarrow \Delta L = L\left[1 - \frac{1}{1+\frac{V_D - V_{Dsat}}{V_A + V_{Dsat}}}\right]$$

the saturation output conductance, which was previously considered equal to zero, is now given by:

$$g_{Dsat} = \frac{dI'_{Dsat}}{dV_D} = \frac{I_{Dsat}}{V_A + V_{Dsat}} \simeq \frac{I_{Dsat}}{V_A}$$

* the channel length modulation effect typically increases in small devices with low-doped substrates.

* Proper scaling can reduce the channel length modulation by increasing the doping density (as the gate length is reduced)

⑧ Short-channel effect

evolution of semiconductor processing technology enables. Scaling of transistors.

Ex gate length of MOSFET used in 256K DRAM in 1984 was ~ 1.2 μm.

in 1994, 64 M DRAM were produced using 0.4 μm

prediction 50 nm in 2009
35 nm in 2012.

⇒ aggressive scaling ⇒ undesirable effects including the short-channel effect.

[here the depletion layer widths of the source and drain junction become comparable to the channel length.

Let us look at the threshold $V_T = V_{FB} + 2V_F - \frac{Q_d}{C_{ox}}$

where Q_d can be represented by a trapezoid area.

Area of the trapezoid is $(L+L_1)\dfrac{x_{dmax}}{2}$.

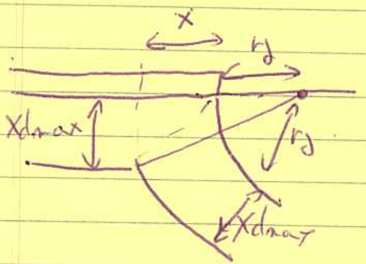
- if the channel is long $L \simeq L_1 \Rightarrow$ area $\simeq L\,x_{dmax}$ (rectangular area)

  V_T is then accurately described by the equation derived previously.

- if the channel (gate) is short $L_1 < L$, then the depletion charge due to V_G under the gate electrode is reduced.

Consider that $V_G = V_T$ and $V_{DS} \simeq 0$.

Based on geometrical consideration we get:



$$X_{dmax}^2 + (x+r_j)^2 = (r_j + X_{dmax})^2$$

Here we suppose also that the built-in potential of the Source and Drain junction is approximately equal to $2V_F$ $\Rightarrow$ width of the depletion region around source and drain is then $\simeq X_{dmax}$

$$\Rightarrow \quad x^2 + 2r_j x - 2r_j X_{dmax} = 0$$

where

$$x = -r_j \pm \sqrt{r_j^2 + 2r_j X_{dmax}} = r_j\left(-1 \oplus \sqrt{1 + \frac{2X_{dmax}}{r_j}}\right)$$

$\Rightarrow$ x must be positive. So

L_1 then is given by

$$L_1 = L - 2x = L - 2r_j\left(\sqrt{1+\frac{2X_{dmax}}{r_j}} - 1\right)$$

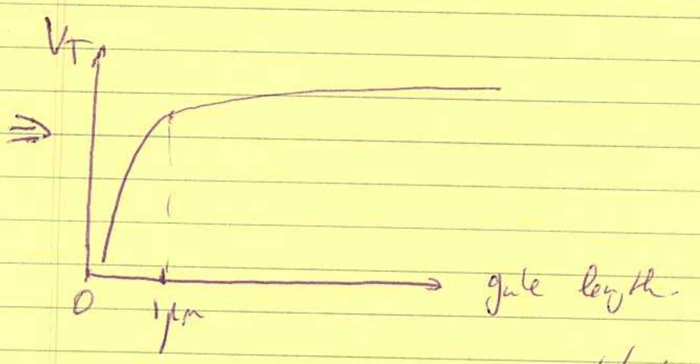
the depletion charge controlled by the gate voltage is

$$Q_d = -q\, X_{dmax}\, N_A \left(\frac{L+L_1}{2}\right)$$

$$= \underbrace{-q\, X_{dmax}\, N_A}_{\downarrow} L \left(1 - \frac{r_j}{L}\left(\sqrt{1+\frac{2X_{dmax}}{r_j}} - 1\right)\right)$$

Q_{do} which is the depletion charge if the depletion region would be considered rectangular.

$$V_T = V_{FB} + 2V_F - \frac{Q_d}{C_{ox}} = V_{FB} + 2V_F + q\,\frac{X_{dmax} N_a}{C_{ox}} \times \left(1 - \frac{r_j}{L}\left(\sqrt{1+\frac{2X_{dmax}}{r_j}} - 1\right)\right)$$



in short devices. Small statistical variations in gate length give rise to large statistical variations of V_T (problem of reproducibility in integrated circuits)

Remarks

- The short channel effect can be reduced by using shallower junctions and higher substrate doping ⇒ to reduce the extension of the Source & Drain depletion regions in the channel.

- if the gate length is very small and V_{DS} is high enough, the source and Drain depletion region can touch one another $L_1 < 0$. ⇒ the potential in the channel region is no longer controlled by the gate and a large current can flows from Source to Drain.

⇒ this is called "punchthrough"