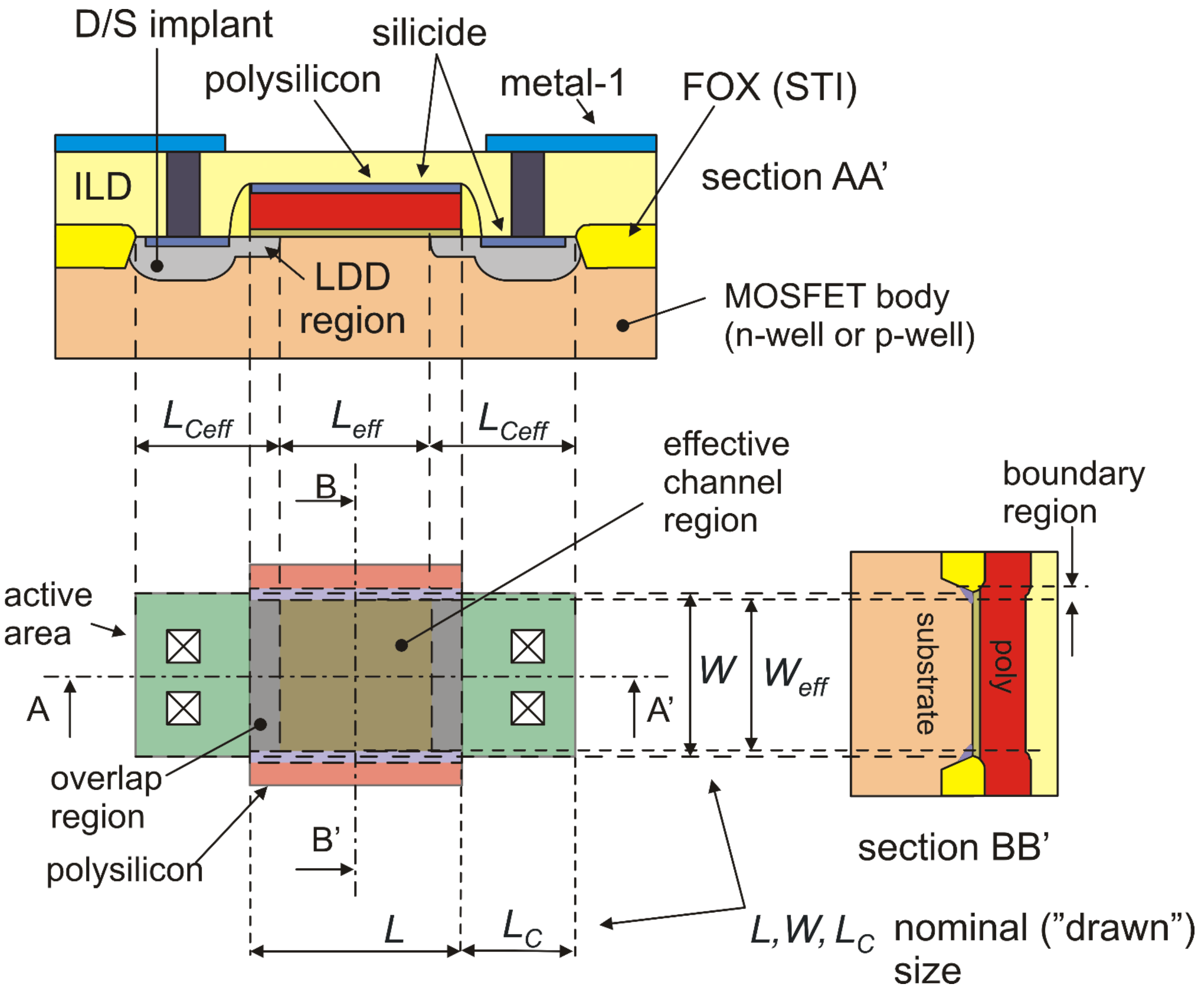
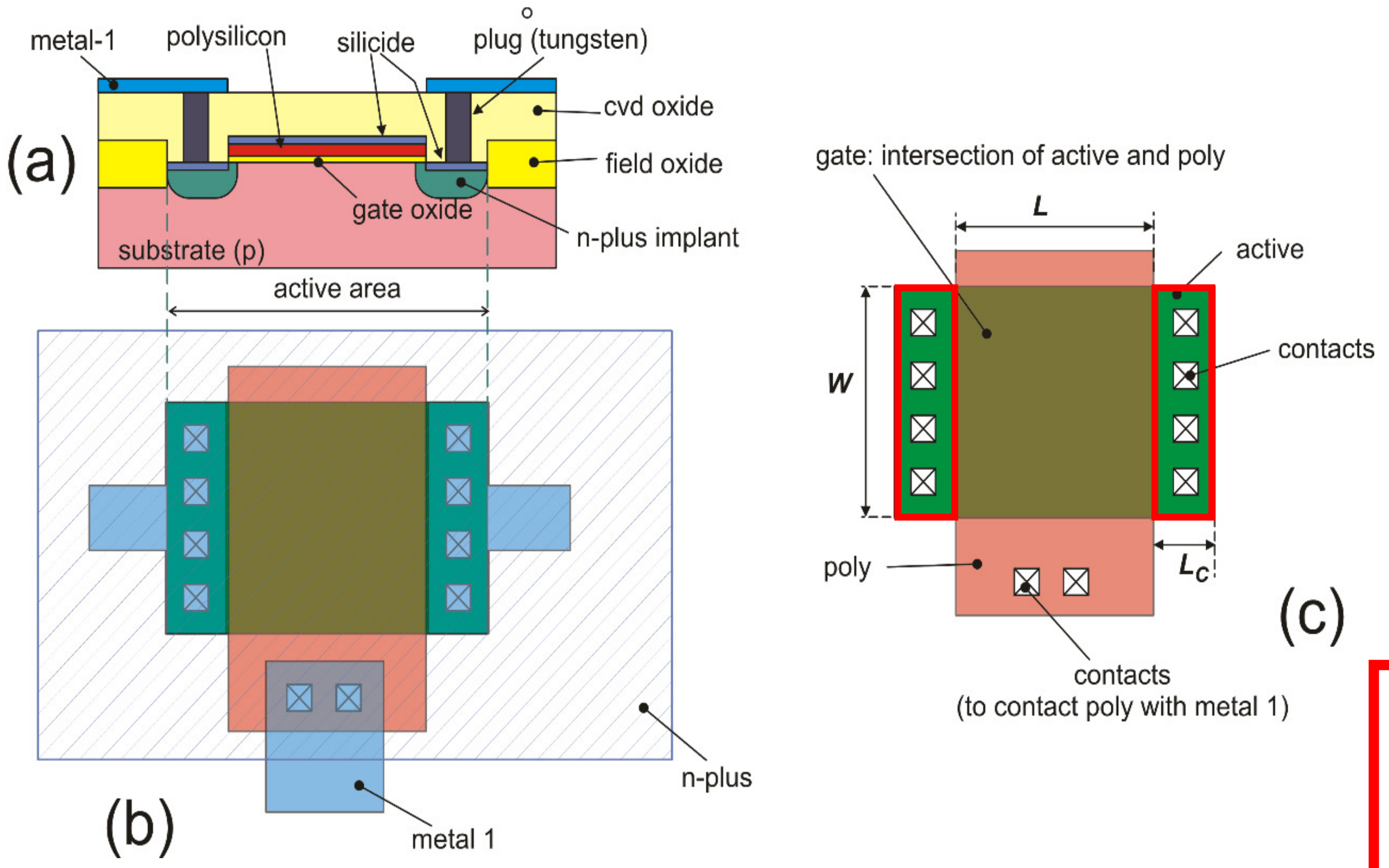
Planar n-MOSFET cross-section and layout



The designer introduces ideal geometrical values (L,W ..), while the electrical properties are determined by "effective" values:

$$\boxed{\begin{aligned} L_{eff} &= L - 2L_D \\ W_{eff} &= W - 2W_D \end{aligned}}$$

Simplified layout and cross-section ("designer view")



$$A_D = A_S = WL_C$$
$$P_D = P_S = 2L_C + 2W$$

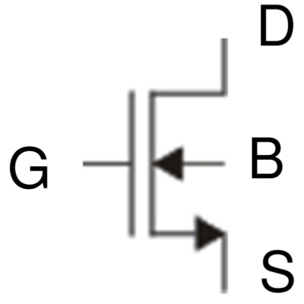
MOSFET models

- Models for accurate electrical simulations: BSIM models (Berkeley Short-channel IGFET Model), EKV (Enz, Krummenacher, Vittoz) ...
- Models for "hand calculations": square law (strong inversion)
exponential laws (weak inversion)
  - It is of primary importance to be able to manually perform first order device sizing and first order performance estimation.
  - Only very <u>simple and intuitive model</u> enable the designer to create cells that need only a final refinement and verification in the simulation phase
  - The simulator is useless if we do not know how to produce a circuit on scrap-paper. The simulator obeys to the law:

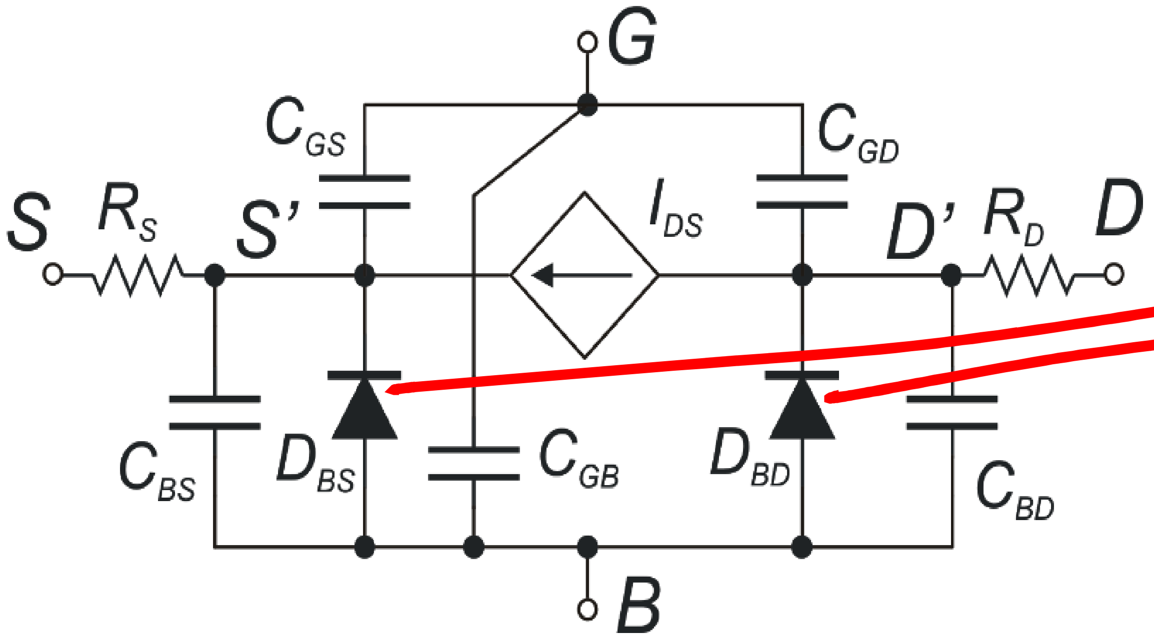
garbage in – garbage out

MOSFET models: The n-MOSFET

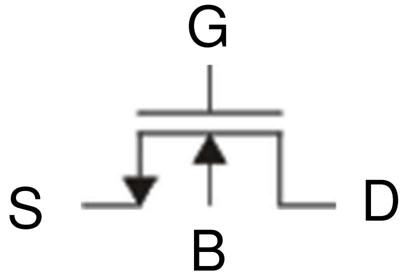
- From this point on, we will consider the behavior of the n-MOSFET, unless otherwise specified. In the end, we will suggest a simple way to transfer all the considerations made for the n-MOSFET to the p-MOSFET
- In integrated circuits, the MOSFET is a four terminal devices: Drain, Source, Gate and Body. In discrete MOSFETs, the body is generally connected to the source internally.



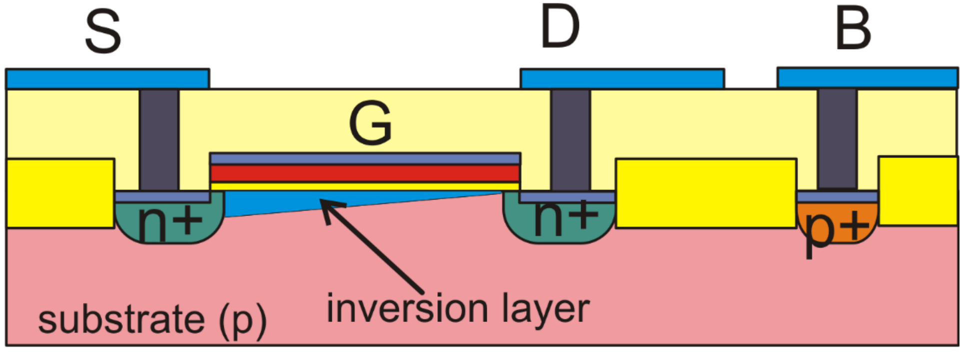
Large signal MOSFET model (n-MOSFET)



$$V_D \, , V_S \geq V_B$$



$$I_D \neq I_S \neq I_{DS}$$



In DC, we will always assume:

$$I_D \cong I_S \cong I_{DS}$$

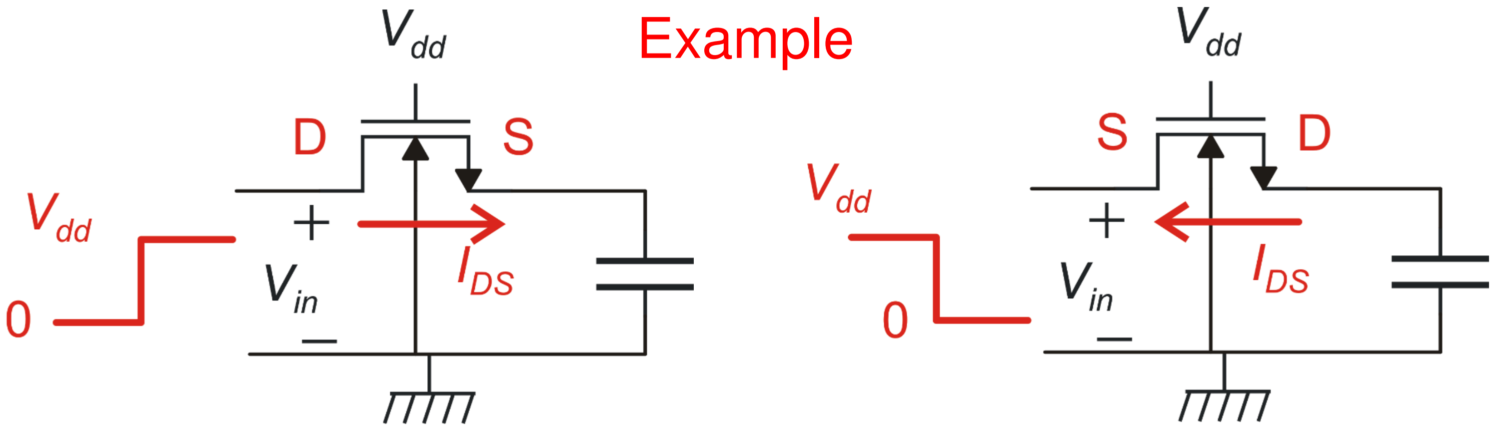
$$R_S = R_D \cong 0$$

Source and Drain Symmetry (1)

- The planar MOSFET is symmetric so that drain and source can be swapped with no consequences in the electrical characteristics.
- Equations that use the source as a reference terminal for all relevant voltages can be applied only after finding which terminal is actually playing the role of the source.
- In an **n-MOSFET**, the effective source is the terminal that has the **lower** voltage; the other one of the two, is the actual drain
- In a **p-MOSFET**, the effective source is the terminal that has the **higher** voltage; the other one of the two, is the actual drain

Source and Drain Symmetry (2)

- With this definition, it is clear that in transient situations, the effective drain and source can swap, depending on the voltage assumed by the terminals.



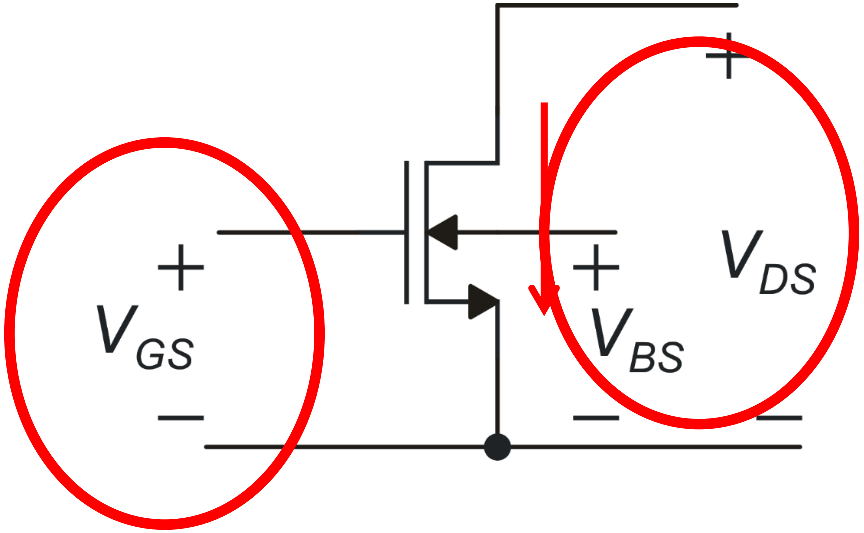
- In a schematic editor it is necessary to indicate which terminal is the drain and the source. These "conventional" terminals are used to mark all voltages for printing and plotting purposes. This choice do not affect the circuit behavior during the simulations.

Source and Drain Symmetry (3)

- If the circuit has a clear static operating point (like most analog circuits), it is convenient to mark as source the terminal that in the operating point is actually working as the source. This will facilitate reading device voltages produced as textual or graphical outputs by the simulator.
- Models like the EKV use the body as the reference for all voltages. In this way drain and sources are perfectly symmetrical also in the equations and there is no need to decide which one is actually working as the source.
- Maintaining the distinction between source and drain is more intuitive and most models oriented to hand calculations are actually based on this choice.

The $\mathbf{I_{DS}}$ model: control voltages

$$I_{DS}(V_{GS}, V_{BS}, V_{DS})$$



secondary effects:
generally they are **<u>unwanted</u>**

primary effect
it is the **<u>wanted</u>** current control

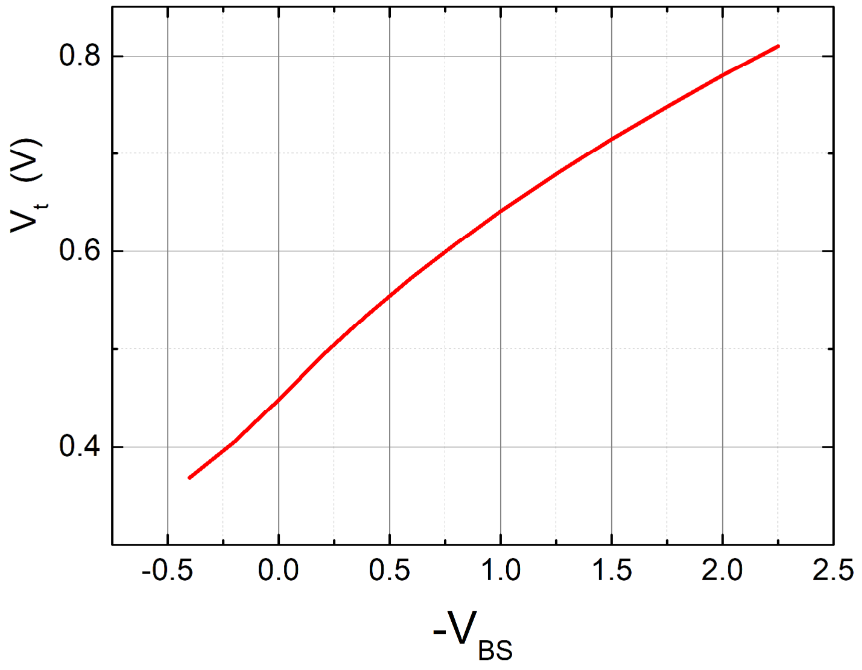
V_{GS}, V_{BS} and "overdrive voltage"

The voltage that really affects the current is the "useful" part of the V_{GS}, often called "overdrive voltage".

$$V_{OV} = V_{GT} = V_{GS} - V_t$$

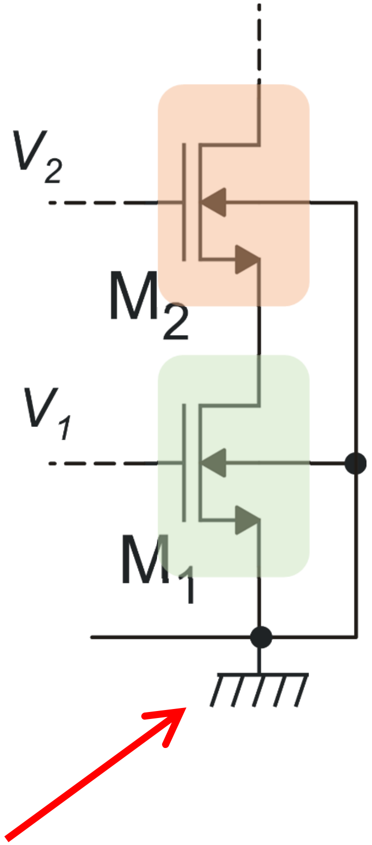
$$V_t\left(V_{BS}\right) \Rightarrow$$ Body effect

$$V_t = V_{t0} + \gamma\left(\sqrt{\phi_s - V_{BS}} - \sqrt{\phi_s}\right)$$



$$V_{t0} = V_t\left(V_{BS} = 0\right)$$ γ: body effect coefficient ϕ_S: surface potential

More on body effect: example

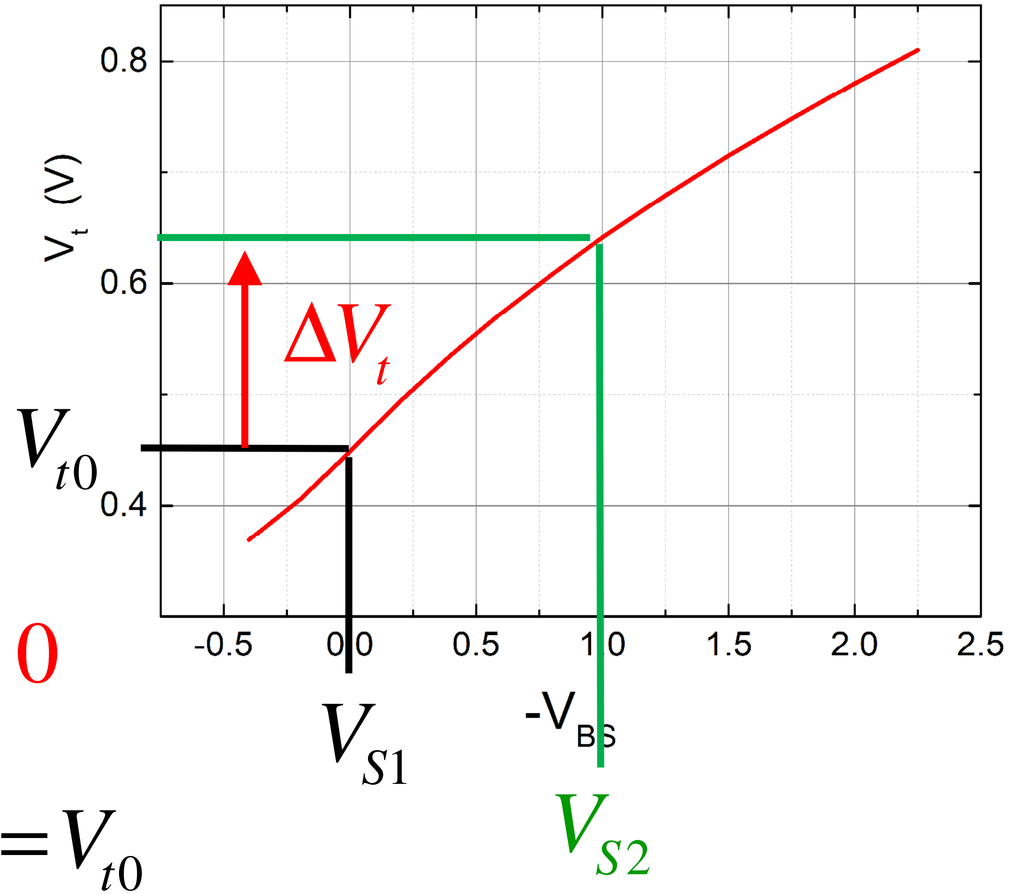


$$V_{B2} = V_{B1} = 0$$

$$V_{S1} = 0$$

$$V_{S2} = V_{D1} > 0$$

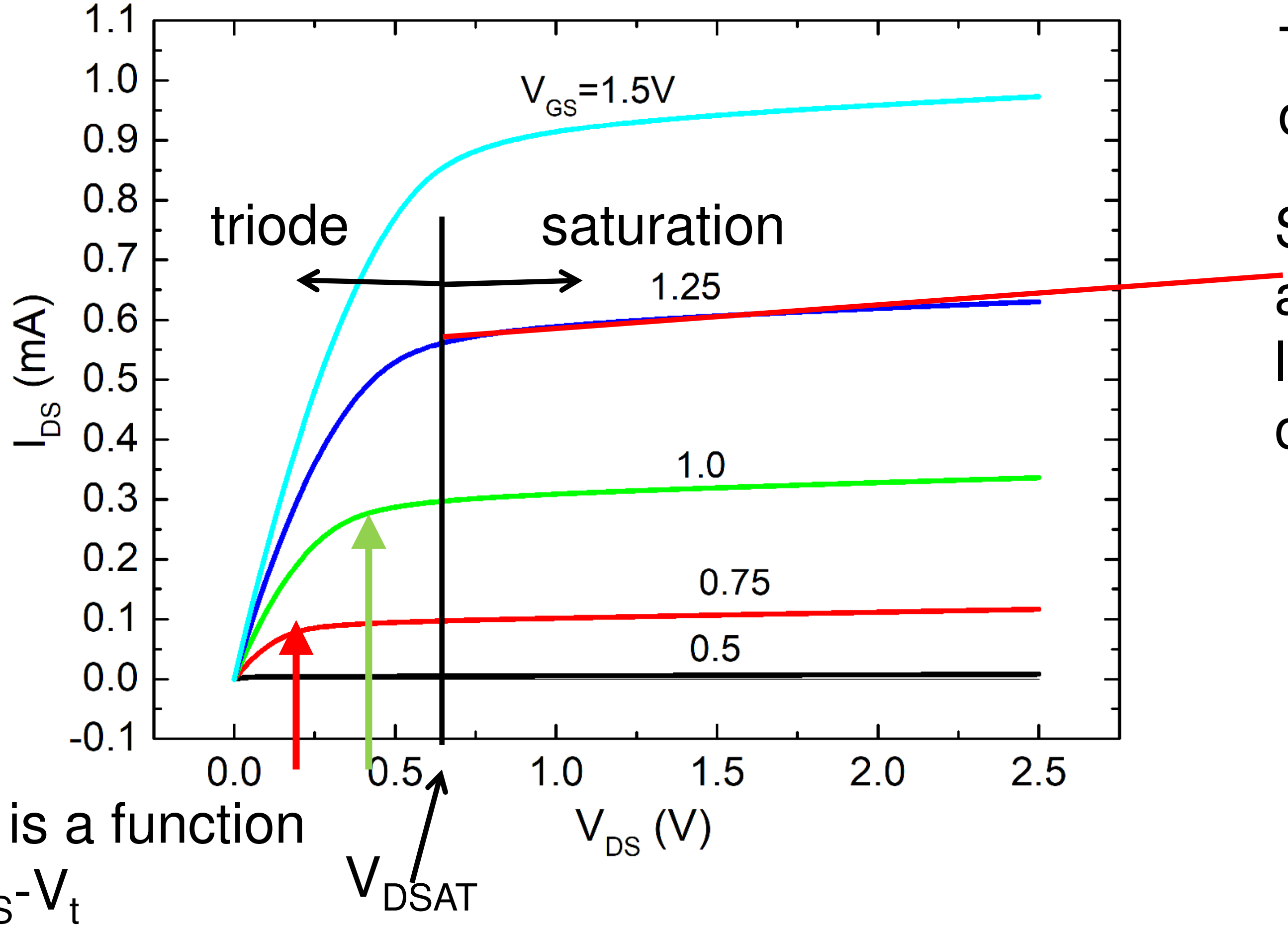
$$V_{BS2} = V_{B2} - V_{S2} < 0$$



$$V_{SS} = gnd$$

$$\begin{cases} V_{BS1} = 0 \;\Rightarrow\; V_{t1} = V_{t0} \\ -V_{BS2} = V_{S2} - V_{B2} = V_{DS1} > 0 \;\Rightarrow\; V_{t2} = V_{t0} + \Delta V_t \end{cases}$$

I_{DS} : operating zones on the basis of V_{DS}



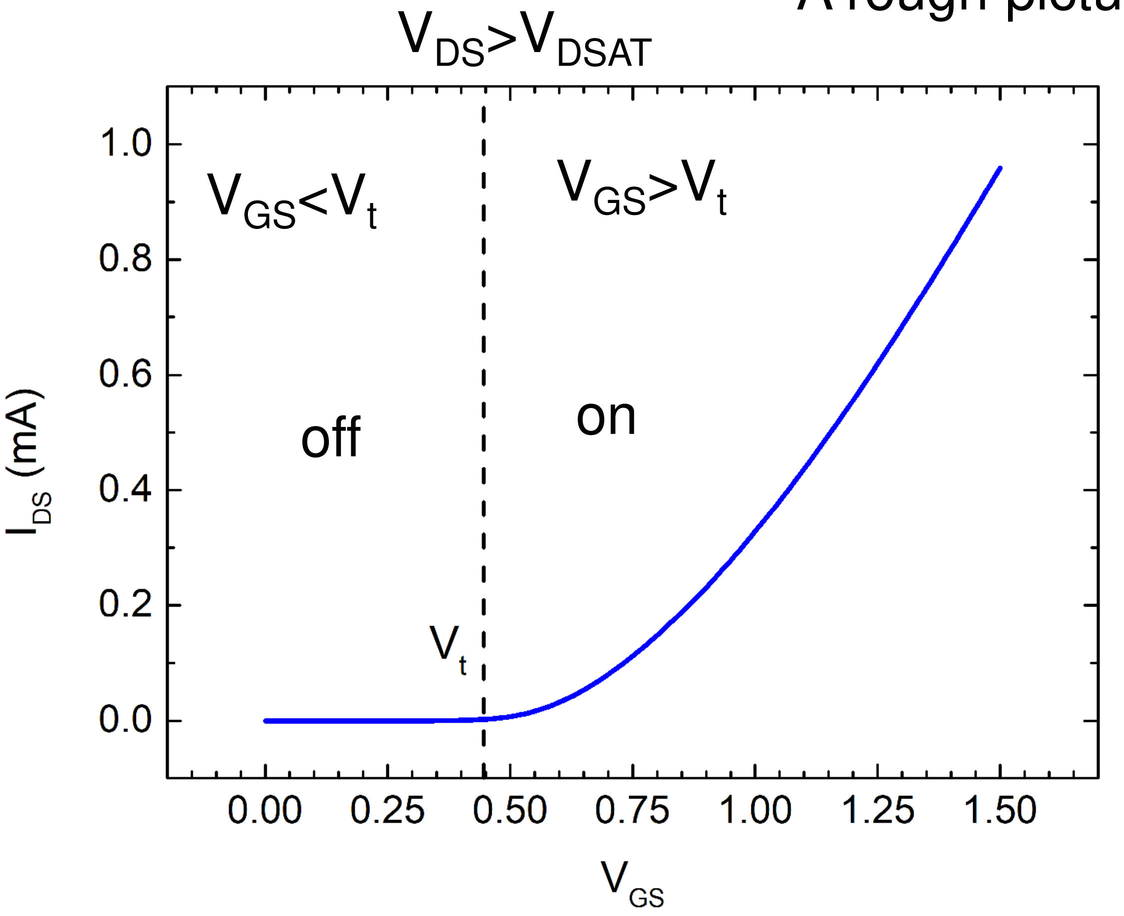
Triode: I_{DS} is strongly dependent on V_{DS}

Saturation: I_{DS} shows a weak and almost linear dependence on V_{DS}

V_{DSAT} is a function of V_{GS}-V_t

Operating zones on the basis of V_{GS}-V_t

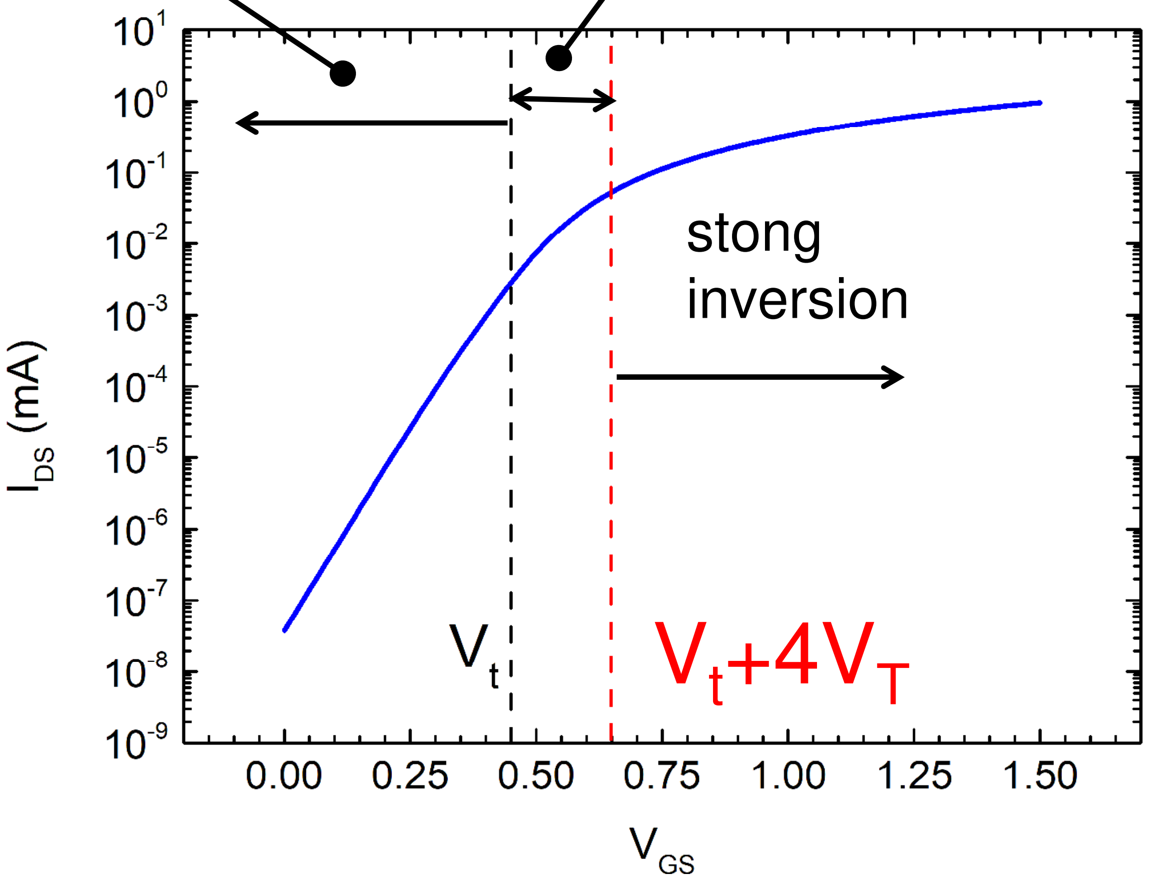
A rough picture:



A more gradual picture: same characteristic with logarithmic y-axis

weak inversion
(sub-threshold)

moderate inversion



V_T=kT/q

I_{DS}: operating zones

	$V_{GS} - V_t \le 0$	$0 \le V_{GS} - V_t \le 4V_T$	$V_{GS} - V_t \ge 4V_T$
$V_{DS} \le V_{DSAT}$	Triode – Weak Inversion	Triode – Moderate Inversion	Triode – Strong Inversion
$V_{DS} \ge V_{DSAT}$	Saturation – Weak Inversion	Saturation – Moderate Inversion	Saturation – Strong Inversion

$$V_{DSAT} \cong \begin{cases} (V_{GS} - V_t) & \text{in strong inversion} \\ 4V_T \quad (100 \;\; \text{mV}) & \text{in moderate and weak inversion} \end{cases}$$

V_{GS}-V_t>$4V_T$ Strong inversion: I_{DS} equations

$$V_{DS} \leq V_{DSAT} \text{ (Triode)} \qquad I_{DS} = \beta_n \left(V_{GS} - V_t - \frac{V_{DS}}{2} \right) V_{DS}$$

$$V_{DS} \geq V_{DSAT} \ (Saturation) \qquad I_{DS} = \beta_n \frac{\left(V_{GS} - V_t\right)^2}{2} \left[1 + \lambda \left(V_{DS} - V_{DSAT} \right) \right]$$

In some textbooks this term is omitted (V_{DSAT}) for simplicity, but this cause a discontinuity between the triode and saturation region

$$\beta_n = \mu_n C_{OX} \frac{W_{eff}}{L_{eff}} \qquad V_{DSAT} = V_{GS} - V_t$$

$$\lambda^{-1} = k_\lambda L_{eff}$$

I_{DS} simplified model in weak inversion

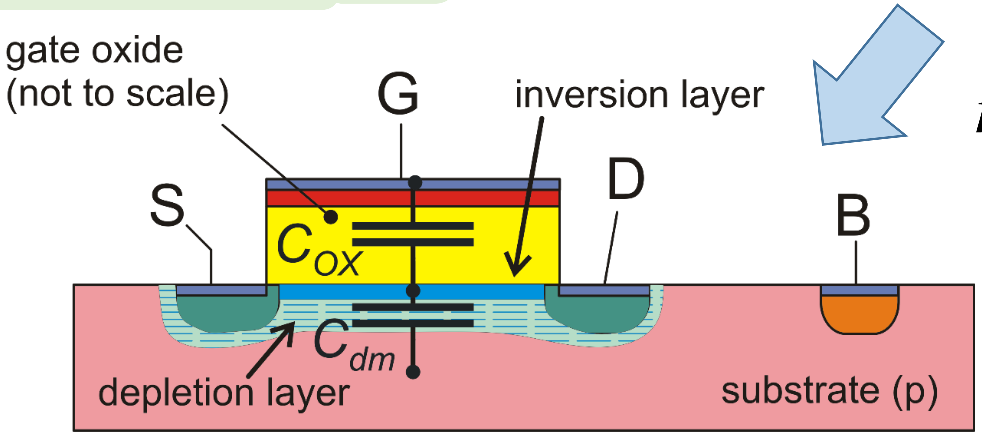
$$I_{DS} = I_{SM}\, e^{\frac{V_{GS}-V_t}{mV_T}} \left(1-e^{\frac{-V_{DS}}{V_T}}\right)\left[1+\lambda\left(V_{DS}-V_{DSAT}\right)\right]$$

m: subthreshold slope factor

$$I_{SM} = \mu_n C_{dm}\frac{W_{eff}}{L_{eff}}V_T^2 = \underbrace{\mu_n C_{ox}}_{\beta_n}(m-1)V_T^2\frac{W_{eff}}{L_{eff}}$$

$$m = 1+\frac{C_{dm}}{C_{ox}}$$

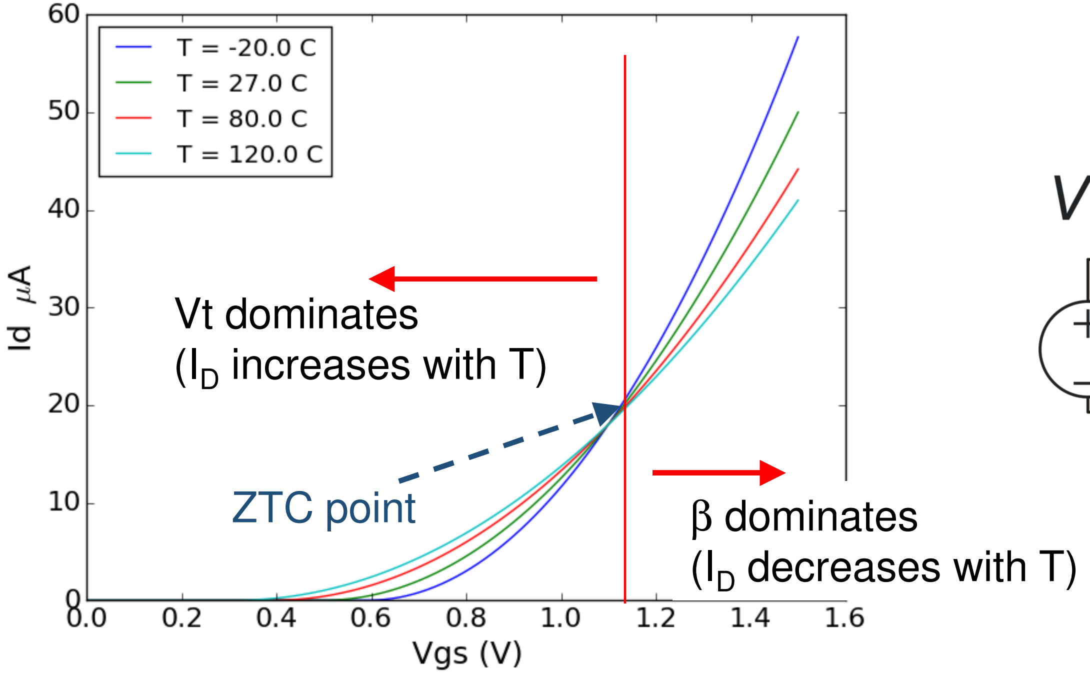
$$m \approx 1.2-1.3$$



Temperature effects on MOSFET characteristics

$$\beta_n(T) = \beta_n(T_0)\left(\frac{T}{T_0}\right)^{-\alpha_\mu} \qquad \alpha_\mu = 1.2\ - 2.4\ \text{(typical 1.5)}$$

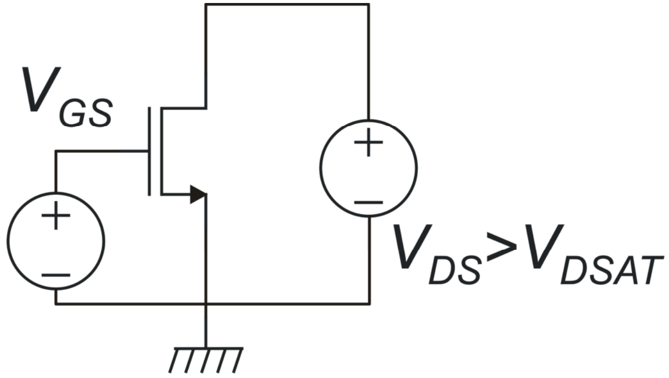
$$V_t(T) = V_t(T_0) - \alpha_{VT}\left(T - T_0\right) \qquad 1\ \text{mV/K}\ \leq\ \alpha_{VT} \leq 4\ \ \text{mV/K}$$



Vt dominates
(I_D increases with T)

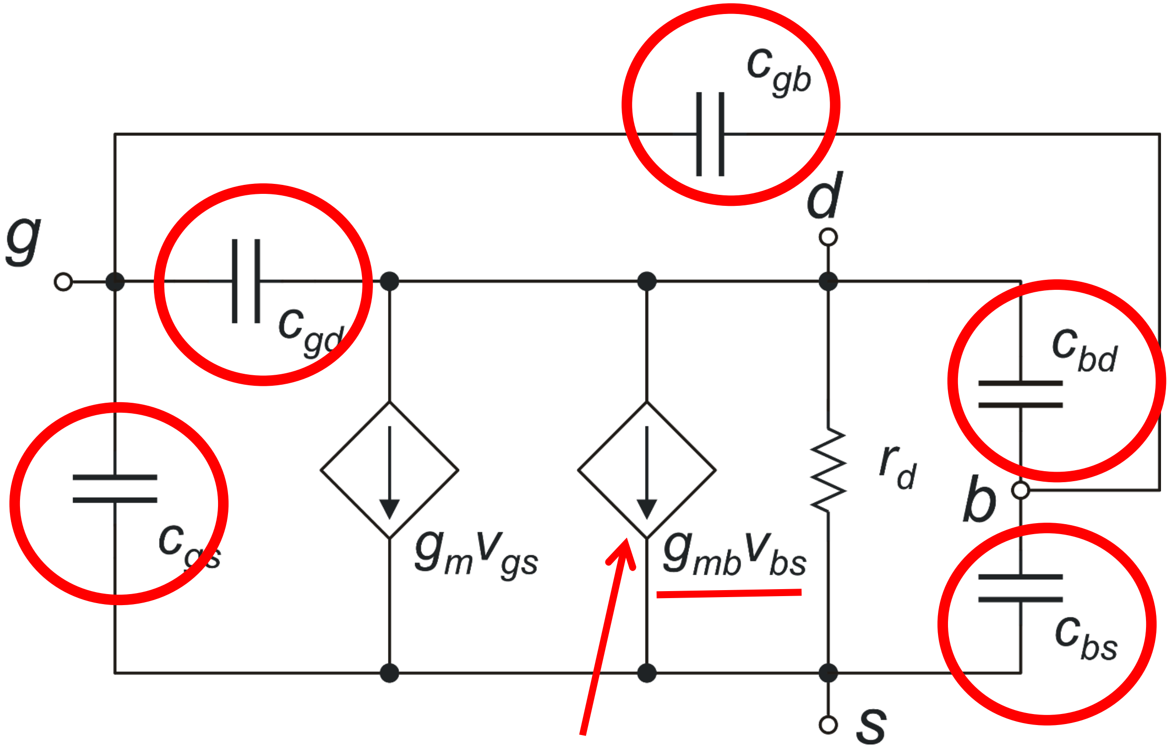
ZTC point

β dominates
(I_D decreases with T)



MOSFET Small Signal model

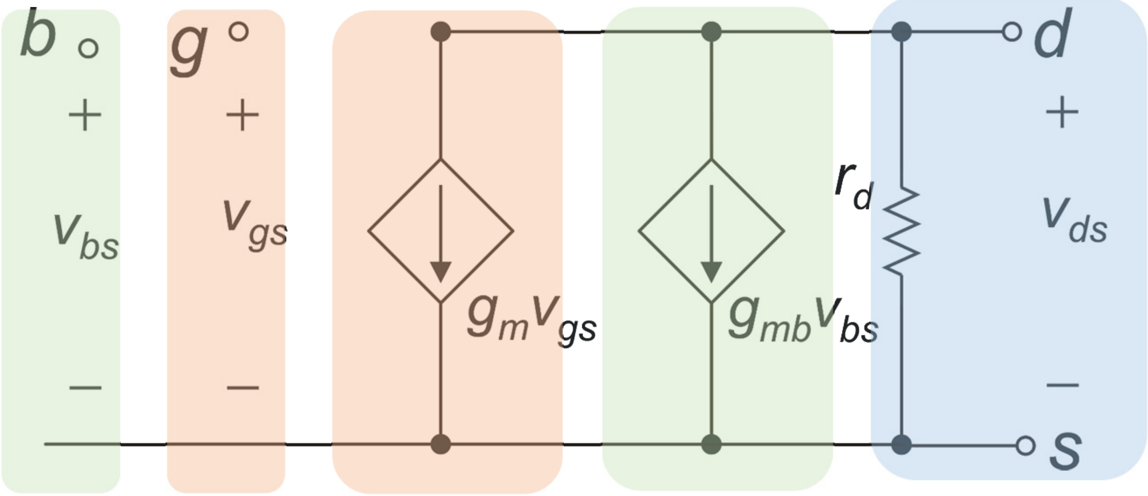
c_{gs}, c_{gd}, c_{gb}, c_{bd}, c_{bs} : small signal capacitances



effect of v_{bs}

Let's start from the dc model (capacitances are removed)

MOSFET small signal model: dc limit



$$I_D\left(V_{GS},V_{BS},V_{DS}\right)$$

small signal

$$i_d = g_m v_{gs} + g_{mb} v_{bs} + g_d v_{ds}$$

$$g_m = \left.\frac{i_d}{v_{gs}}\right|_{v_{ds},v_{bs}=0} = \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS},V_{BS}}$$

$$g_{mb} = \left.\frac{i_d}{v_{bs}}\right|_{v_{ds},v_{gs}=0} = \left(\frac{\partial I_D}{\partial V_{BS}}\right)_{V_{DS},V_{GS}}$$

$$\frac{1}{r_d} = g_d = \left.\frac{i_d}{v_{ds}}\right|_{v_{gs},v_{bs}=0} = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS},V_{BS}}$$

Body transconductance: g_{mb}

effect of V_{BS}

$$g_{mb} = \left(\frac{\partial I_D}{\partial V_{BS}}\right)_{V_{DS}=const;\ V_{GS}=const} \qquad I_D\left(V_{GS}, V_{BS}, V_{DS}\right) \cong I_D\left[\left(V_{GS} - V_t\right), V_{DS}\right]$$

let's recall the gm definition

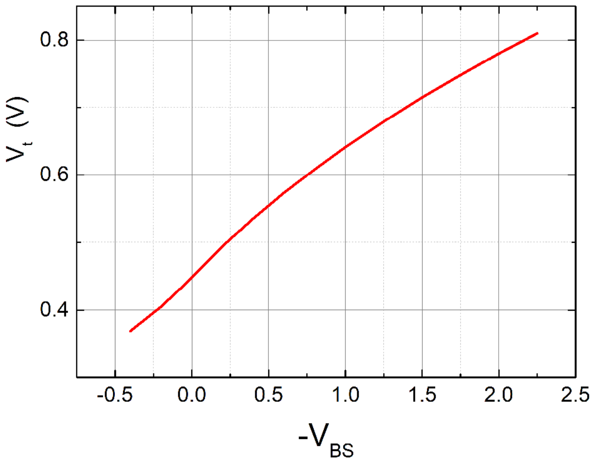
$$g_m = \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS},V_{BS}} = \left(\frac{\partial I_D}{\partial (V_{GS}-V_t)}\right)_{V_{DS}} \overset{1}{\left(\frac{\partial (V_{GS}-V_t)}{\partial V_{GS}}\right)_{V_{BS}}} = \left(\frac{\partial I_D}{\partial (V_{GS}-V_t)}\right)_{V_{DS}}$$

g_m

$$g_{mb} = \left(\frac{\partial I_D}{\partial V_{BS}}\right)_{V_{GS},V_{DS}} = \left(\frac{\partial I_D}{\partial (V_{GS}-V_t)}\right)_{V_{DS}} \left(\frac{\partial (V_{GS}-V_t)}{\partial V_{BS}}\right)_{V_{GS}} = \underline{\underline{\underline{g_m\left(-\frac{\partial V_t}{\partial V_{BS}}\right)_{V_{DS}}}}}$$

g_{mb}

Body transconductance: g_{mb}

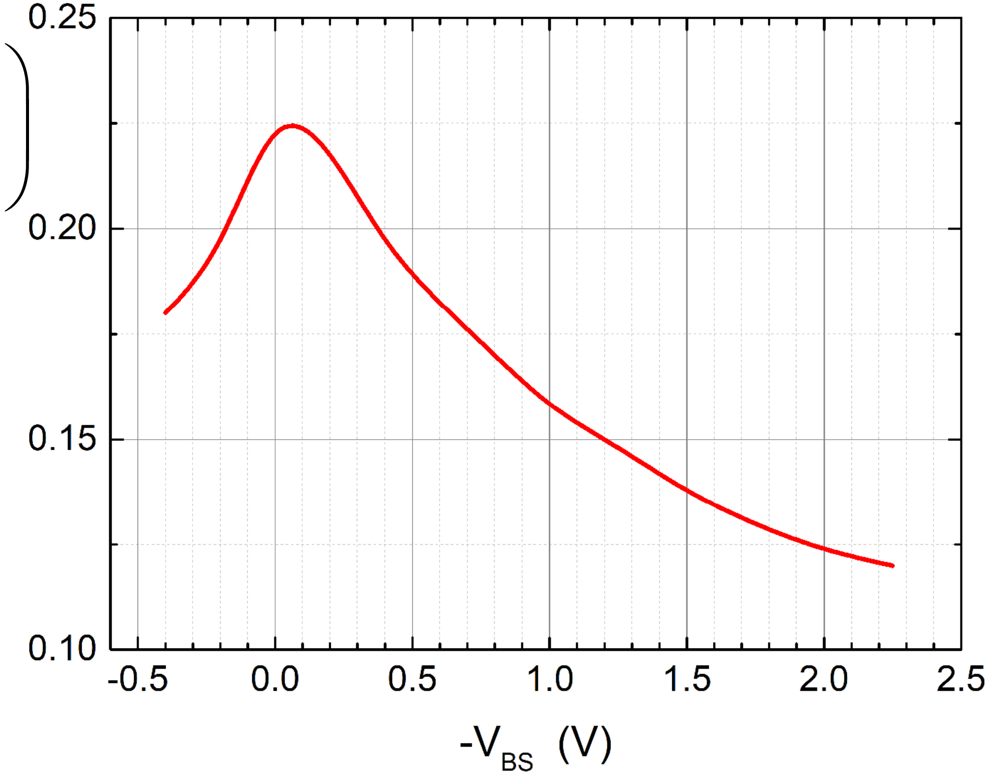


$$g_{mb} = g_m \left(-\frac{\partial V_t}{\partial V_{BS}} \right) = g_m (m-1)$$

m ~ 1.2 ➡ g_{mb} ~ 0.2g_m

Example from simulation

$$\left(-\frac{\partial V_t}{\partial V_{BS}} \right)$$



g_m, g_d in strong inversion

Triode region $$I_{DS}=\beta_n\left(V_{GS}-V_t-\frac{V_{DS}}{2}\right)V_{DS}$$

g_m $$g_m=\left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS},V_{BS}}=\underline{\beta_n V_{DS}}$$

g_d $$\frac{1}{r_d}=g_d=\left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS},V_{BS}}=\beta_n\left[\left(V_{GS}-V_t\right)-\frac{V_{DS}}{2}\right]-\beta_n\frac{V_{DS}}{2}=\underline{\beta_n\left[\left(V_{GS}-V_t\right)-V_{DS}\right]}$$

g_m, g_d in strong inversion

Saturation region

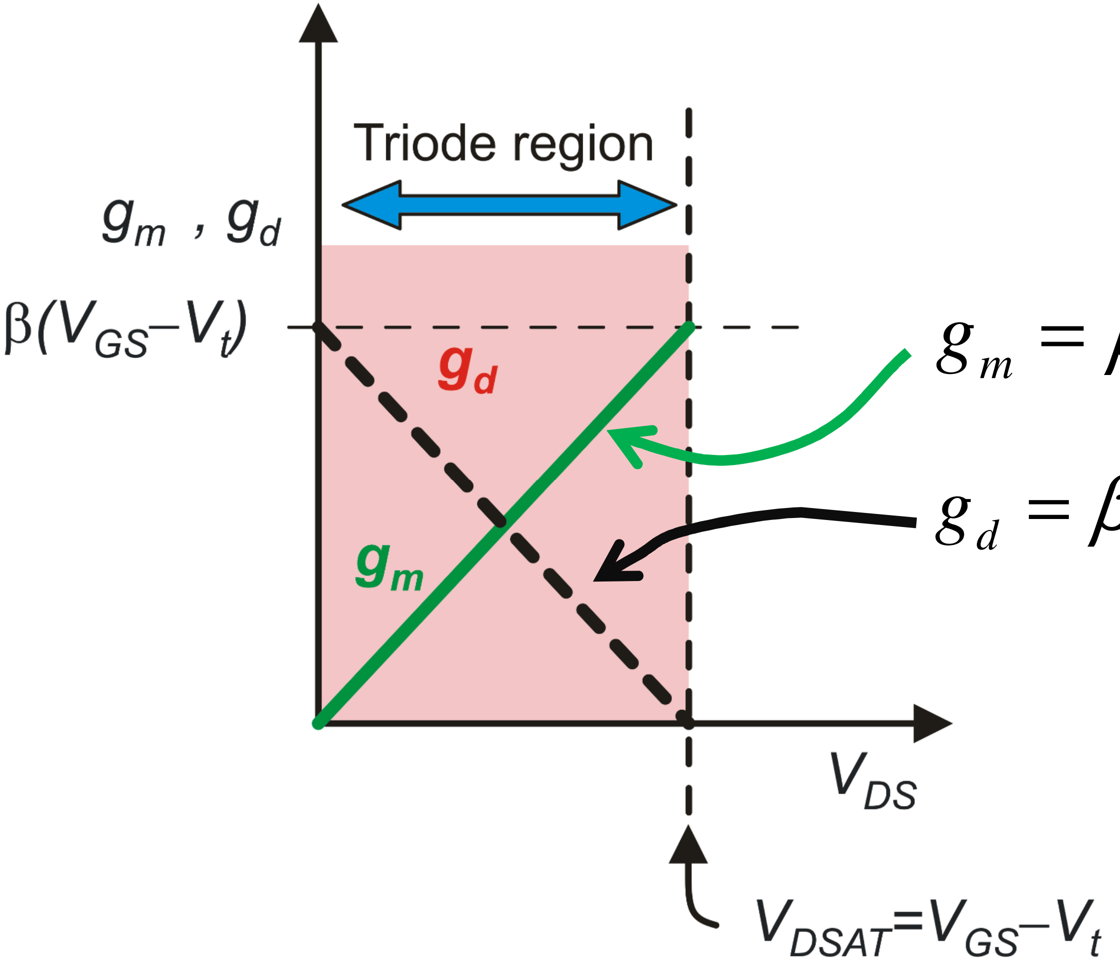
$$I_{DS} = \beta_n \frac{(V_{GS} - V_t)^2}{2}\left[1 + \lambda\left(V_{DS} - V_{DSAT}\right)\right]$$

neglecting the dependence of V_{DSAT} on V_{GS}

$$g_m \equiv \left(\frac{\partial I_{DS}}{\partial V_{GS}}\right)_{V_{DS},V_{BS}} = \beta_n (V_{GS} - V_t)\left[1 + \lambda\left(V_{DS} - V_{DSAT}\right)\right] \cong \underline{\beta_n (V_{GS} - V_t)}$$

$$\frac{1}{r_d} = g_{ds} \equiv \left(\frac{\partial I_{DS}}{\partial V_{DS}}\right)_{V_{GS},V_{BS}} = \lambda \frac{\beta_n}{2}(V_{GS} - V_t)^2 \cong \underline{\lambda I_{DS}}$$

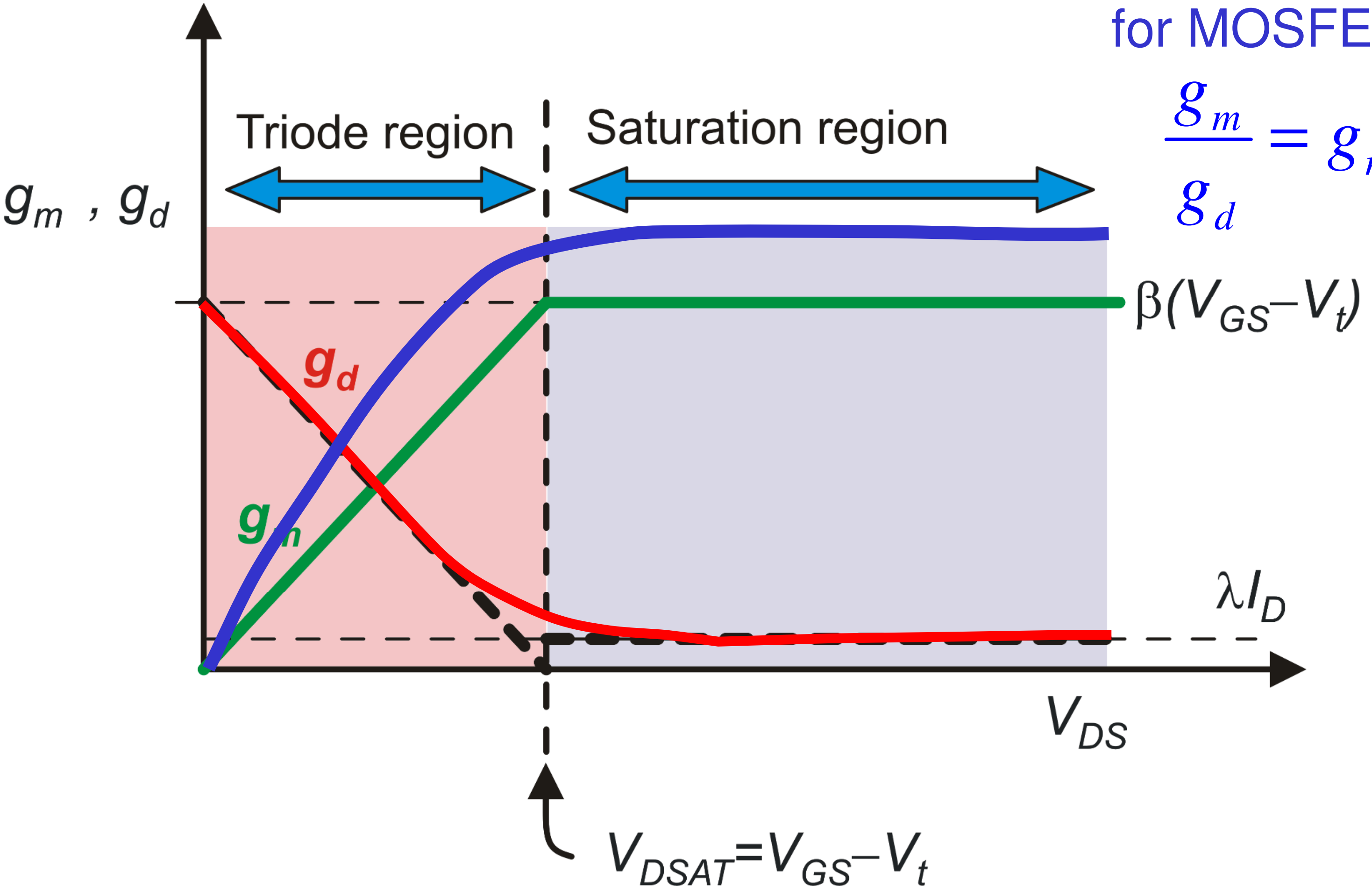
g_m, g_d in strong inversion



$$g_m = \beta_n V_{DS}$$

$$g_d = \beta_n \left[\left(V_{GS} - V_t \right) - V_{DS} \right]$$

g_m, g_d in strong inversion



Important parameter for MOSFETs:

$$\frac{g_m}{g_d} = g_m r_d$$

Transconductance models in <u>saturation</u>

In strong Inversion:

acceptable approximation, because we are studying gm, i.e the effect of V_{GS}

$$I_{DS}=\beta_n\frac{(V_{GS}-V_t)^2}{2}\left[1+\lambda\overbrace{(V_{DS}-V_{DSAT})}^{\text{only a few \%}}\right] \Rightarrow I_{DS}\cong\beta_n\frac{(V_{GS}-V_t)^2}{2}$$

$$\boxed{(V_{GS}-V_t)=\sqrt{\frac{2I_D}{\beta_n}}}$$

$$g_m=\beta_n\sqrt{\frac{2I_D}{\beta_n}}=\underline{\underline{\sqrt{2I_D\beta_n}}}$$

$$g_m=\beta_n(V_{GS}-V_t)$$

$$\beta_n=\frac{2I_{DS}}{(V_{GS}-V_t)^2}$$

$$g_m=\frac{2I_{DS}}{\underline{\underline{(V_{GS}-V_t)}}}$$

!!!

g_m,g_d in **weak inversion**

$$I_{DS}=I_{SM}e^{\frac{V_{GS}-V_t}{mV_T}}\left(1-e^{\frac{-V_{DS}}{V_T}}\right)\left[1+\lambda\left(V_{DS}-V_{DSAT}\right)\right]$$

$$g_m=\left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS},V_{BS}}=\frac{1}{mV_T}I_{SM}e^{\frac{V_{GS}-V_t}{mV_T}}\left(1-e^{\frac{-V_{DS}}{V_T}}\right)\left[1+\lambda\left(V_{DS}-V_{DSAT}\right)\right]=\frac{I_D}{mV_T}$$

Exact result

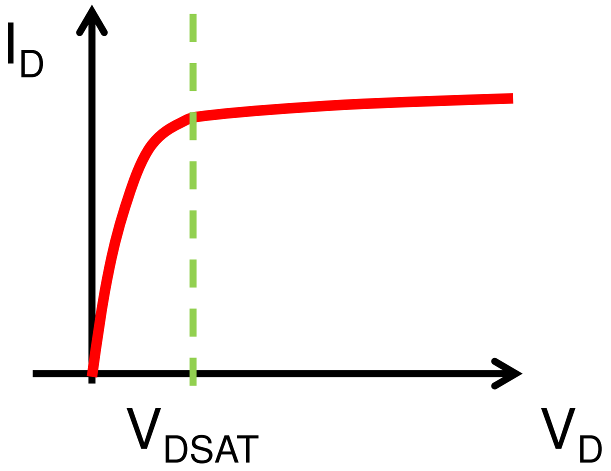
$$\frac{1}{r_d}=g_d=\left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS},V_{BS}}=$$

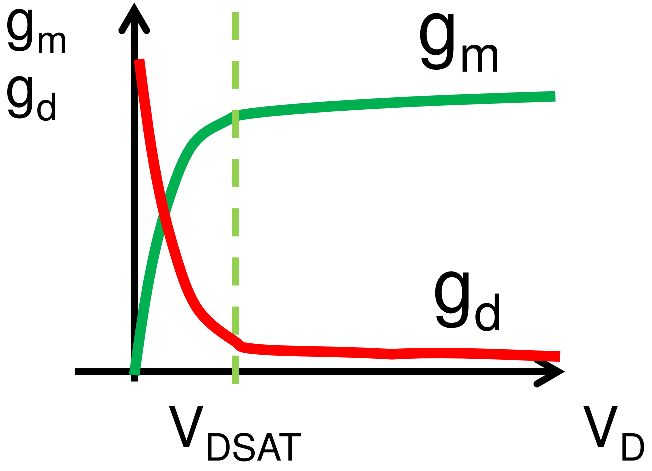
$$=\frac{I_{SM}}{V_T}e^{\frac{V_{GS}-V_t}{mV_T}}e^{\frac{-V_{DS}}{V_T}}\left[1+\lambda\left(V_{DS}-V_{DSAT}\right)\right]+\lambda I_{SM}e^{\frac{V_{GS}-V_t}{mV_T}}\left(1-e^{\frac{-V_{DS}}{V_T}}\right)$$

g_m, g_d in **weak inversion**

$$g_m = \frac{I_D}{mV_T}$$

$$g_d = \frac{I_{SM}}{V_T} e^{\frac{V_{GS}-V_t}{mV_T}} e^{\frac{-V_{DS}}{V_T}} \left[1+\lambda\left(V_{DS}-V_{DSAT}\right)\right] + \lambda I_{SM} e^{\frac{V_{GS}-V_t}{mV_T}} \left(1-e^{\frac{-V_{DS}}{V_T}}\right) \quad \Rightarrow \underline{\underline{g_d \cong \lambda I_D}}$$

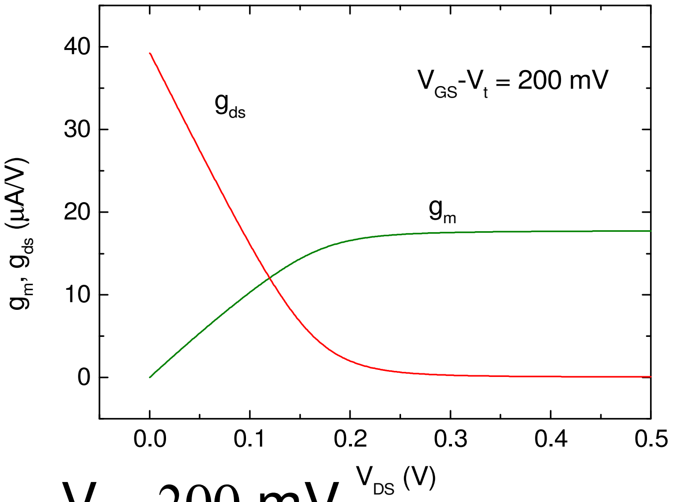




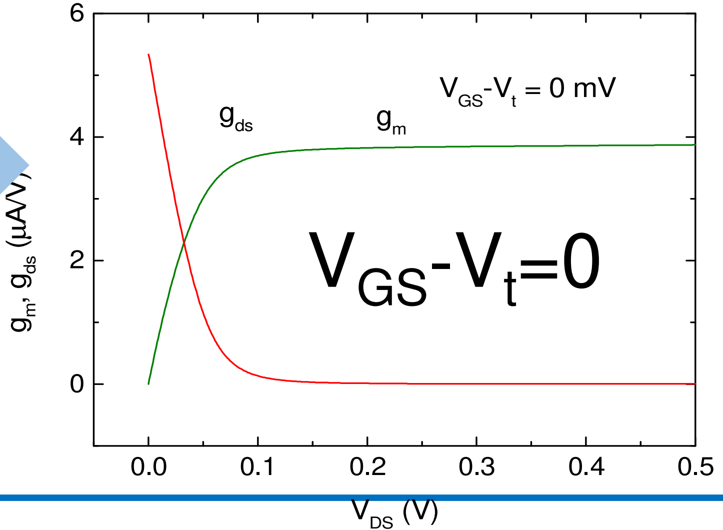
In saturation: $e^{\frac{-V_{DS}}{V_T}} \to 0$

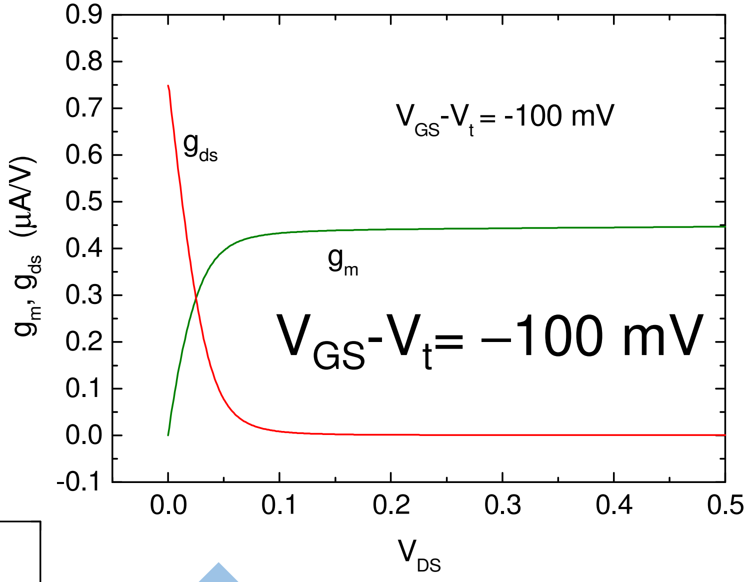
g_m, g_d everywhere: simulations

Strong inversion



V_{GS}-V_t= 200 mV





Unified model for transconductance **in saturation**

Strong Inversion

$$g_m = \frac{2I_{DS}}{\left(V_{GS}-V_t\right)} = \frac{I_{DS}}{\dfrac{\left(V_{GS}-V_t\right)}{2}}$$

$$\frac{g_m}{I_D}$$

Important parameter for analog circuit design

Weak Inversion

$$g_m = \frac{I_D}{mV_T}$$

BJT

$$g_m = \frac{I_C}{V_T}$$

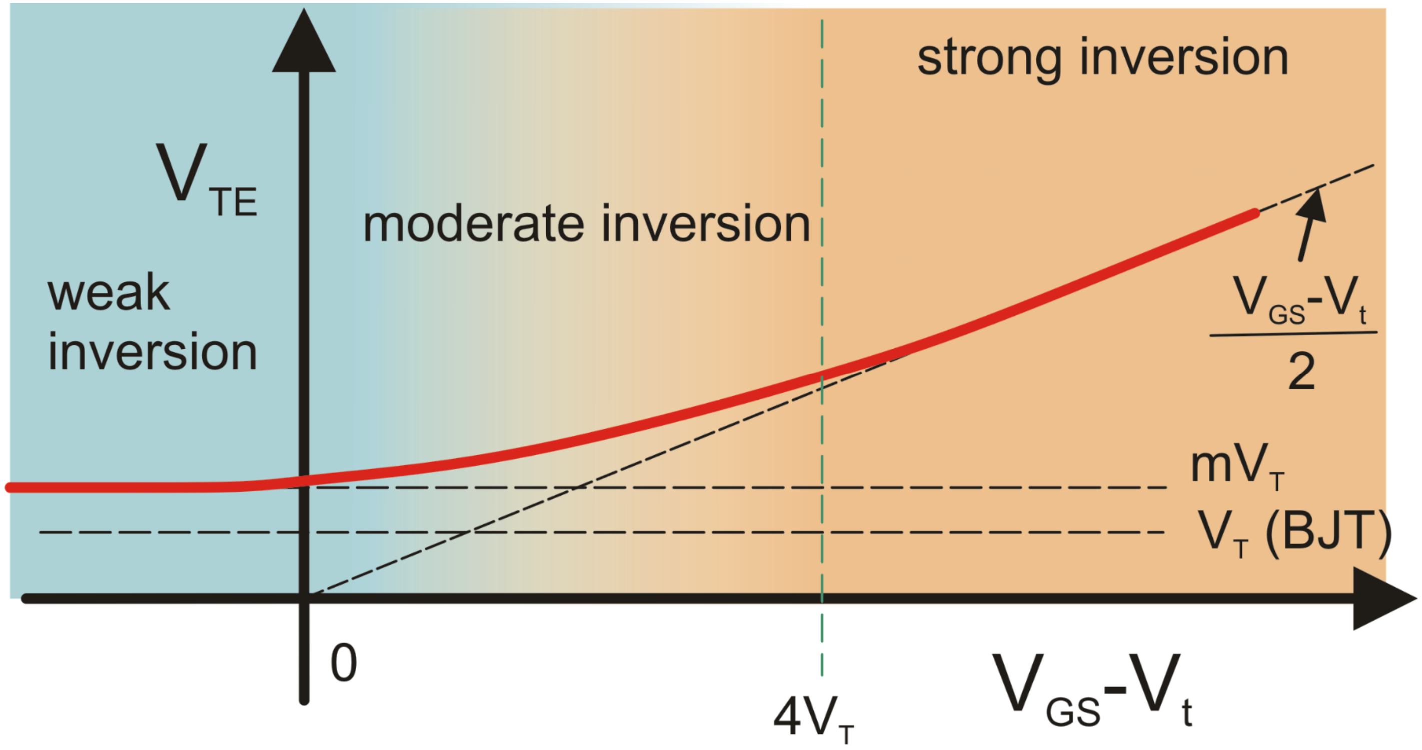
$$\boxed{g_m = \frac{I_D}{V_{TE}}}$$

with $V_{TE} = \left(\dfrac{g_m}{I_D}\right)^{-1} = \begin{cases} \dfrac{\left(V_{GS}-V_t\right)}{2} & \text{Strong inversion} \\ mV_T & \text{Weak inversion} \\ V_T & \text{BJT} \end{cases}$

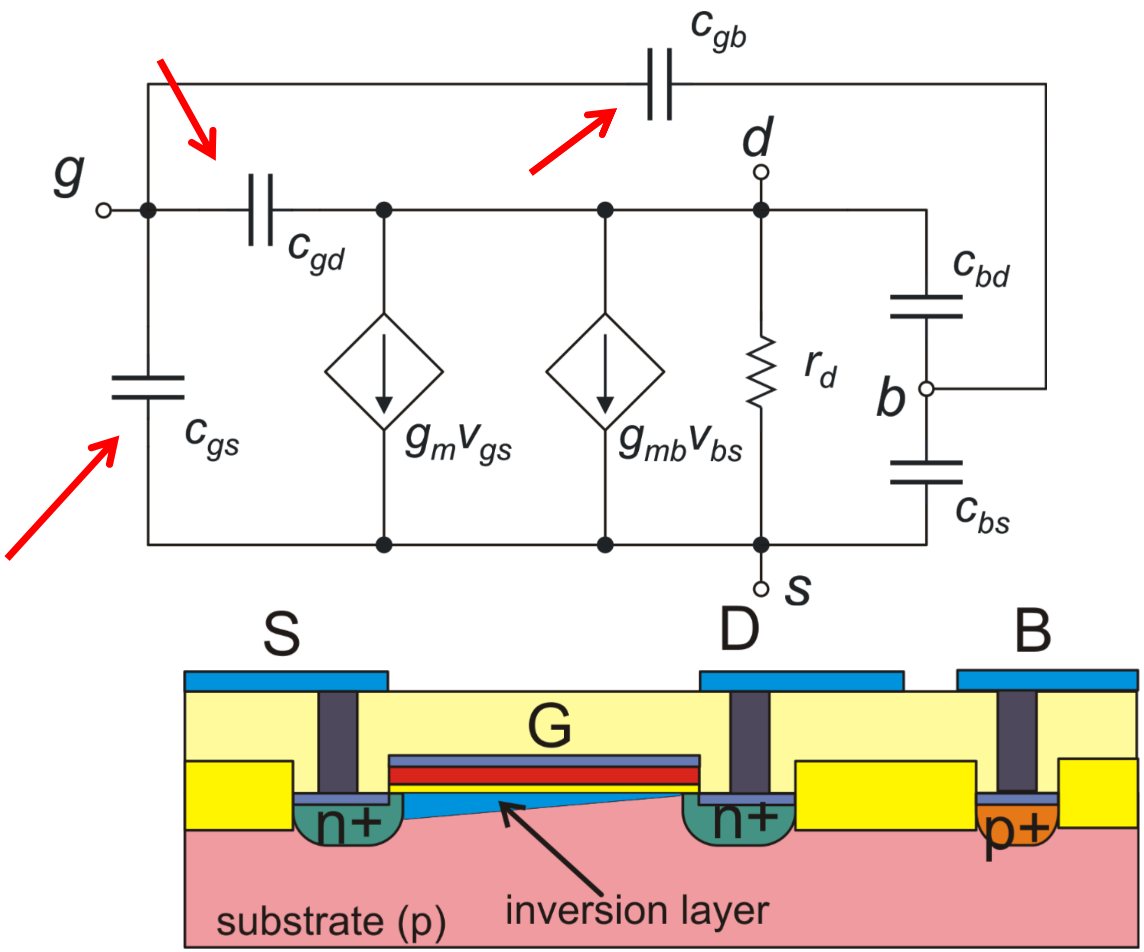
Definition of V_{TE}

Effective Thermal Voltage: V_{TE}

The smaller the V_{TE}, the higher the g_m that can be obtained with a given I_D



MOSFET Capacitance Model: gate related capacitances



estrinsic cap. intrinsic cap..

$$c_{gs} = \boxed{c_{gs}^{(ov)}} + \boxed{c_{gs}^{(i)}}$$

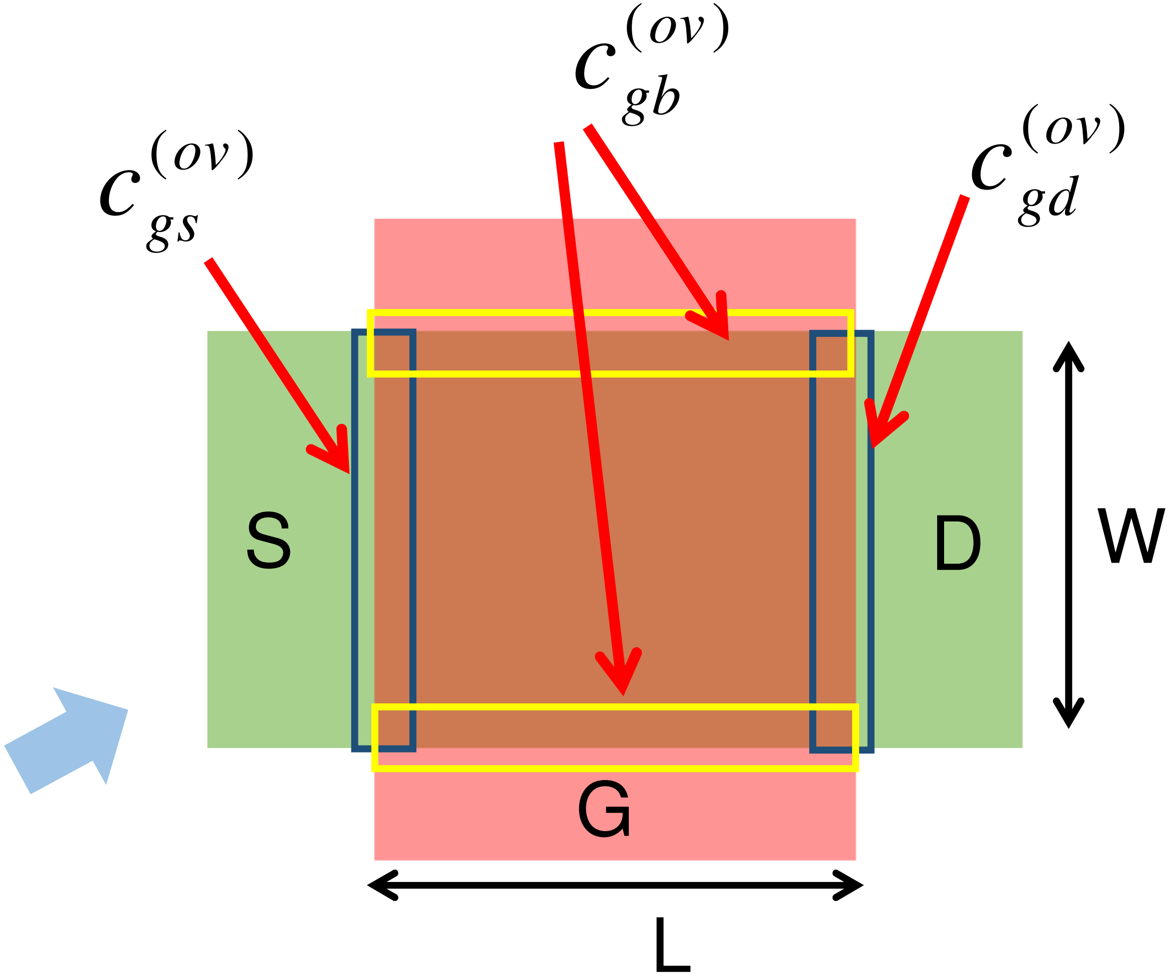
$$c_{gd} = c_{gd}^{(ov)} + c_{gd}^{(i)}$$

$$c_{gb} = c_{gb}^{(ov)} + c_{gb}^{(i)}$$

Estrinsic Capacitances

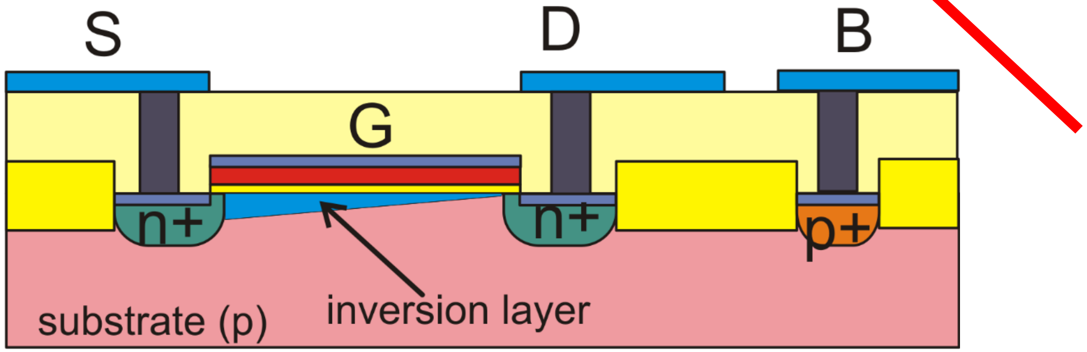
$$\begin{cases} c_{gs}^{(ov)} = c_{gso} \cdot W \\ c_{gd}^{(ov)} = c_{gdo} \cdot W \\ c_{gb}^{(ov)} = c_{gbo} \cdot L \end{cases}$$

Localization of extrinsic capacitances: along the borders of the gate



Intrinsic capacitances: The Meyer Model

	Off ($V_{GS} << V_t$)	**Triode**	**Saturation**
$C_{gs}^{(i)}$	0	$\frac{1}{2}C_{OX}WL$	$\frac{2}{3}C_{OX}WL$
$C_{gd}^{(i)}$	0	$\frac{1}{2}C_{OX}WL$	0
$C_{gb}^{(i)}$	$\left(\frac{1}{C_{OX}WL}+\frac{1}{C_{dm}}\right)^{-1}$	0	0



Series of the oxide and depletion layer capacitances. Can be approximated with only the oxide cap $C_{OX}WL$

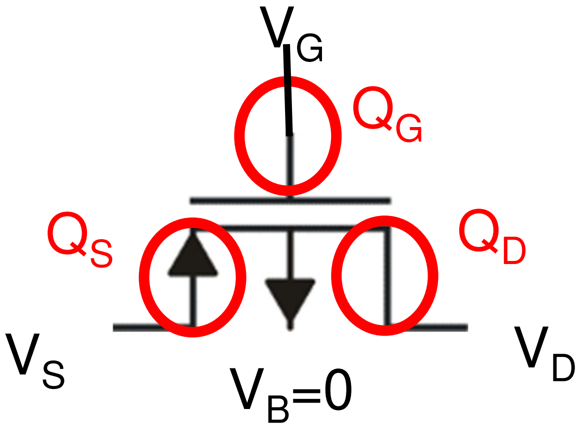
Charge oriented models (Dutton and Ward model)

Limits of the Meyer Model:

- Does not guarantee charge conservation
- Capacitances are reciprocal

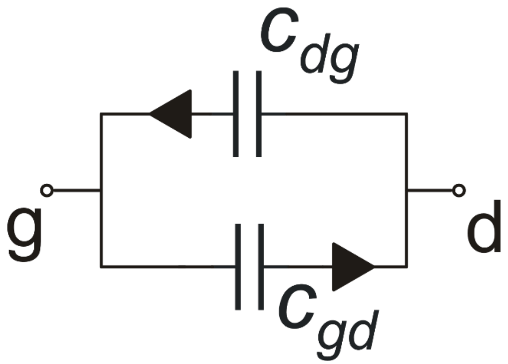
Important errors in circuits using MOSFETs as switches.

Dutt and Ward model

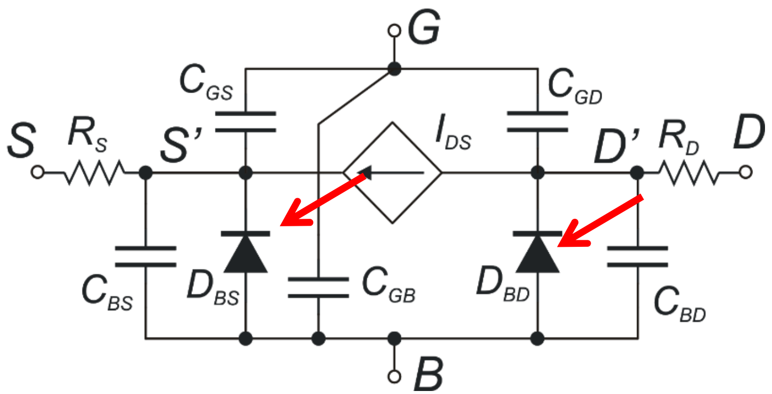


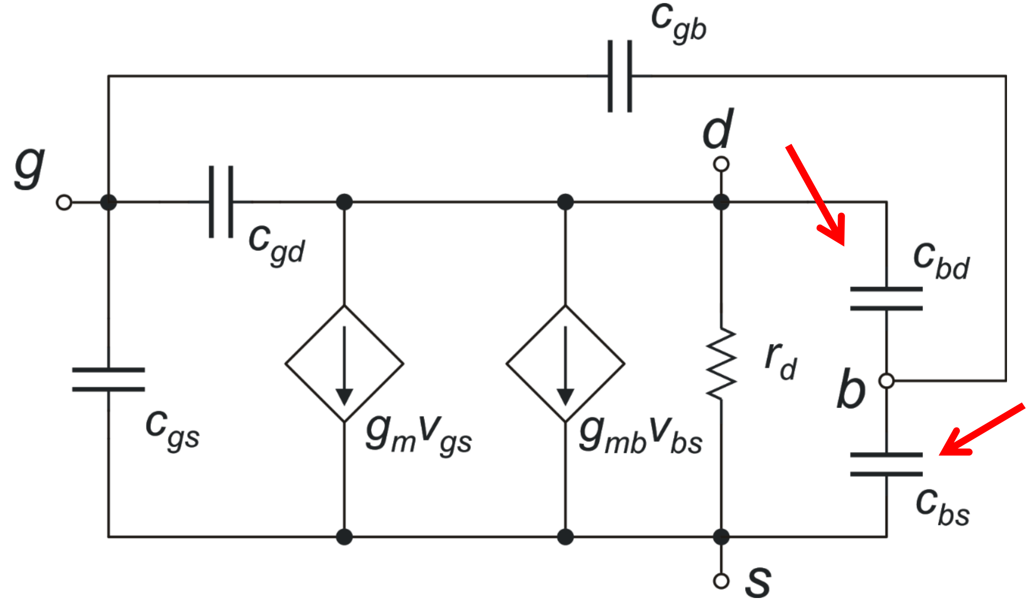
$$C_{ij} = \frac{\partial Q_i}{\partial V_j}$$

Array of 9 capacitances
Cij are < 0 for i≠j (trans-capacitances)
Cij are > 0 for i=j (self capacitances)
Generally: Cij ≠ Cji

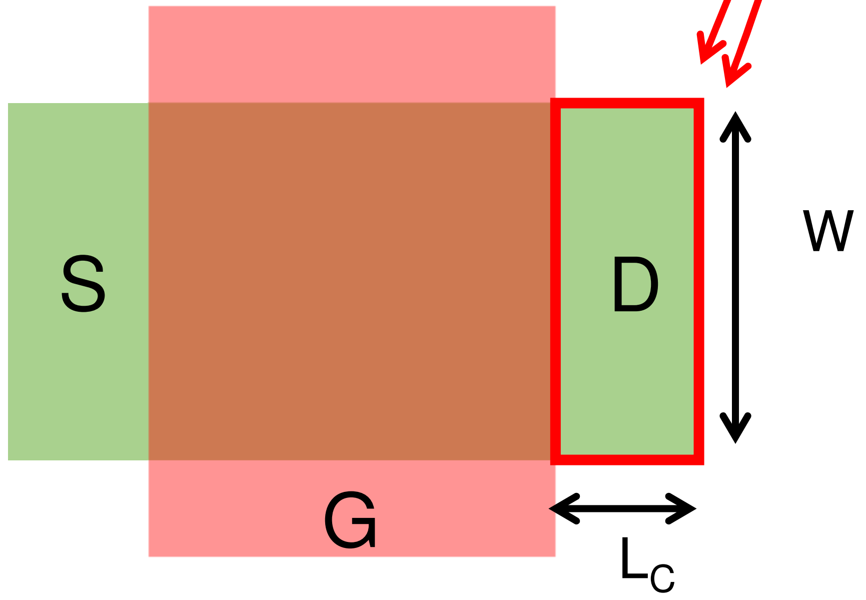


Junction capacitances





$$c_{bd} = \frac{C_J A_D}{\left(1+\frac{V_{SB}}{V_0}\right)^{m_j}} + \frac{C_{JSW} P_D}{\left(1+\frac{V_{SB}}{V_0}\right)^{m_{jsw}}}$$



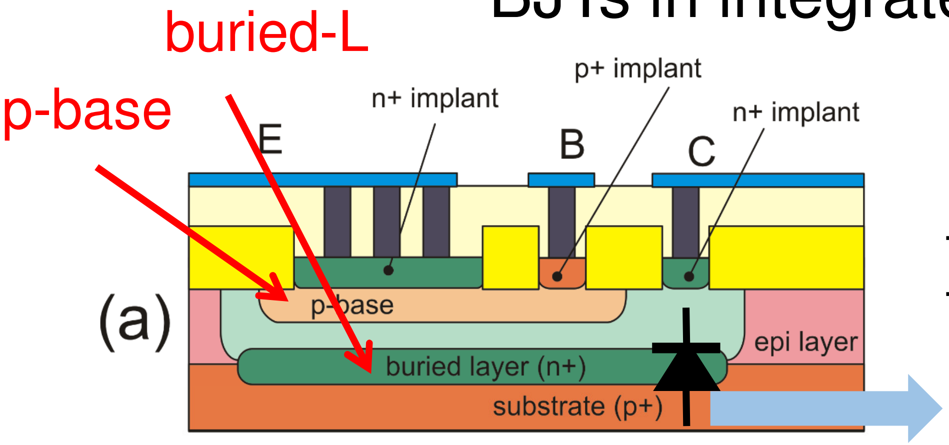
Other non-idealities of the MOSFET behaviour

Gate-bias dependent mobility → μC_{ox} depends on V_{GS} (decreases at high V_{GS}) **(all devices)**

Carrier velocity saturation → I_D dependence on V_{GS} in strong inversion tends to become linear (instead of quadratic) **(Short channel devices)**. Again, appears as a reduction of the μC_{ox} at high V_{GS}

Gate current → May be due to tunneling **(all devices)**
or hot electrons - hot holes **(Short channel devices)**

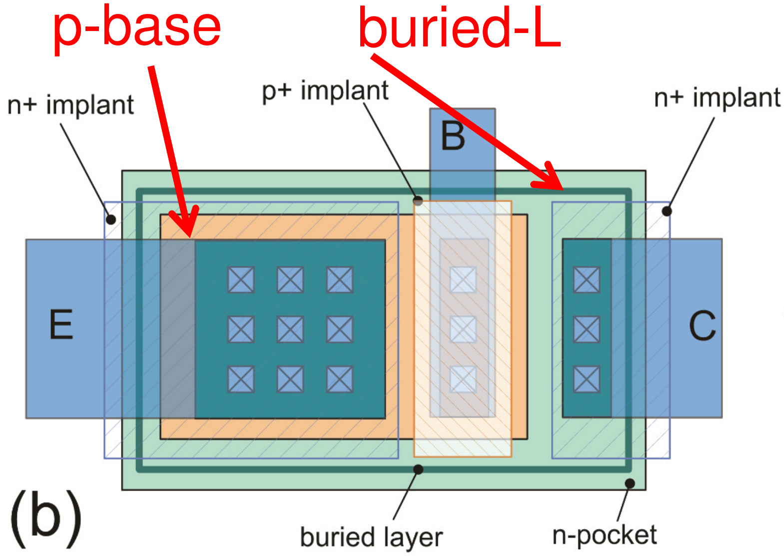
BJTs in integrated circuits: Vertical NPN

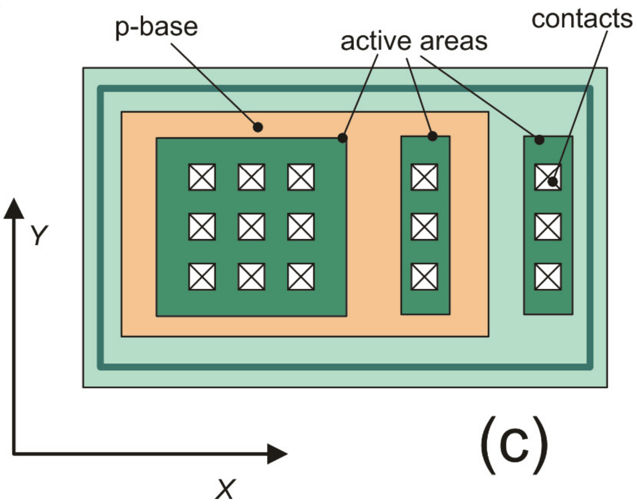


BiCMOS: CMOS + p-base and buried-layer

The **n-pocket** can be an n-well
The **buried-layer** a buried-well of a triple-well CMOS

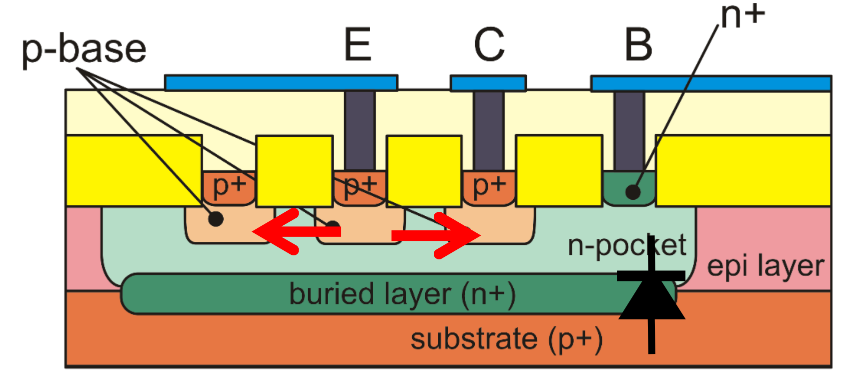
A diode between collector and substrate is present → capacitance



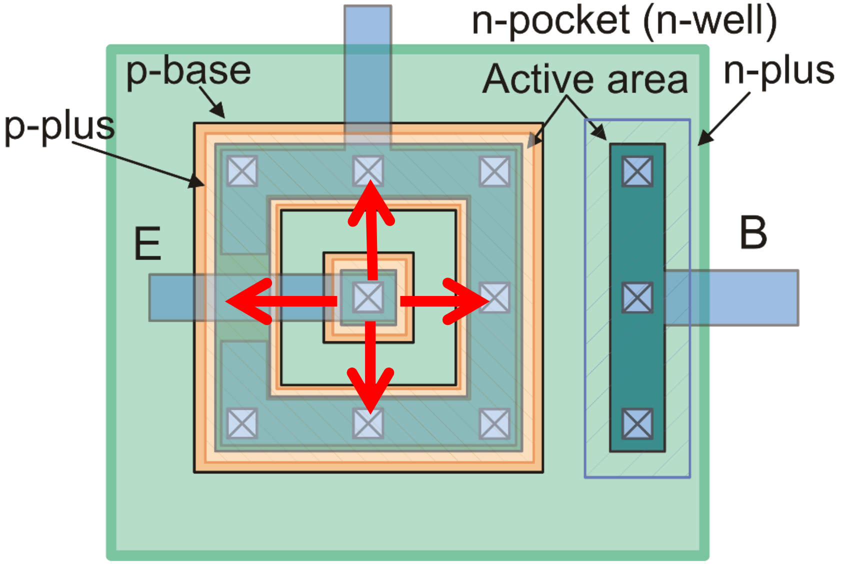


Same as (b) but with the Metal 1 removed

The lateral PNP



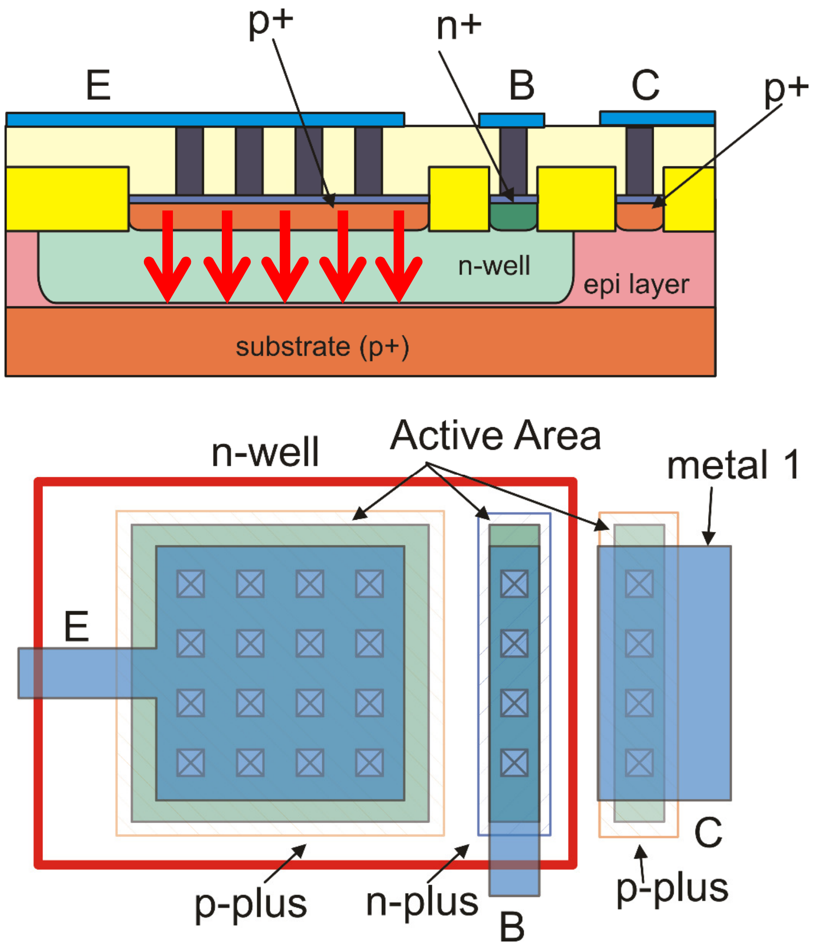
Slower than vertical devices due to large base series resistance ($r_{bb'}$) and base-to substrate capacitance



Lower early voltage (V_A), due to non-optimal collector doping.

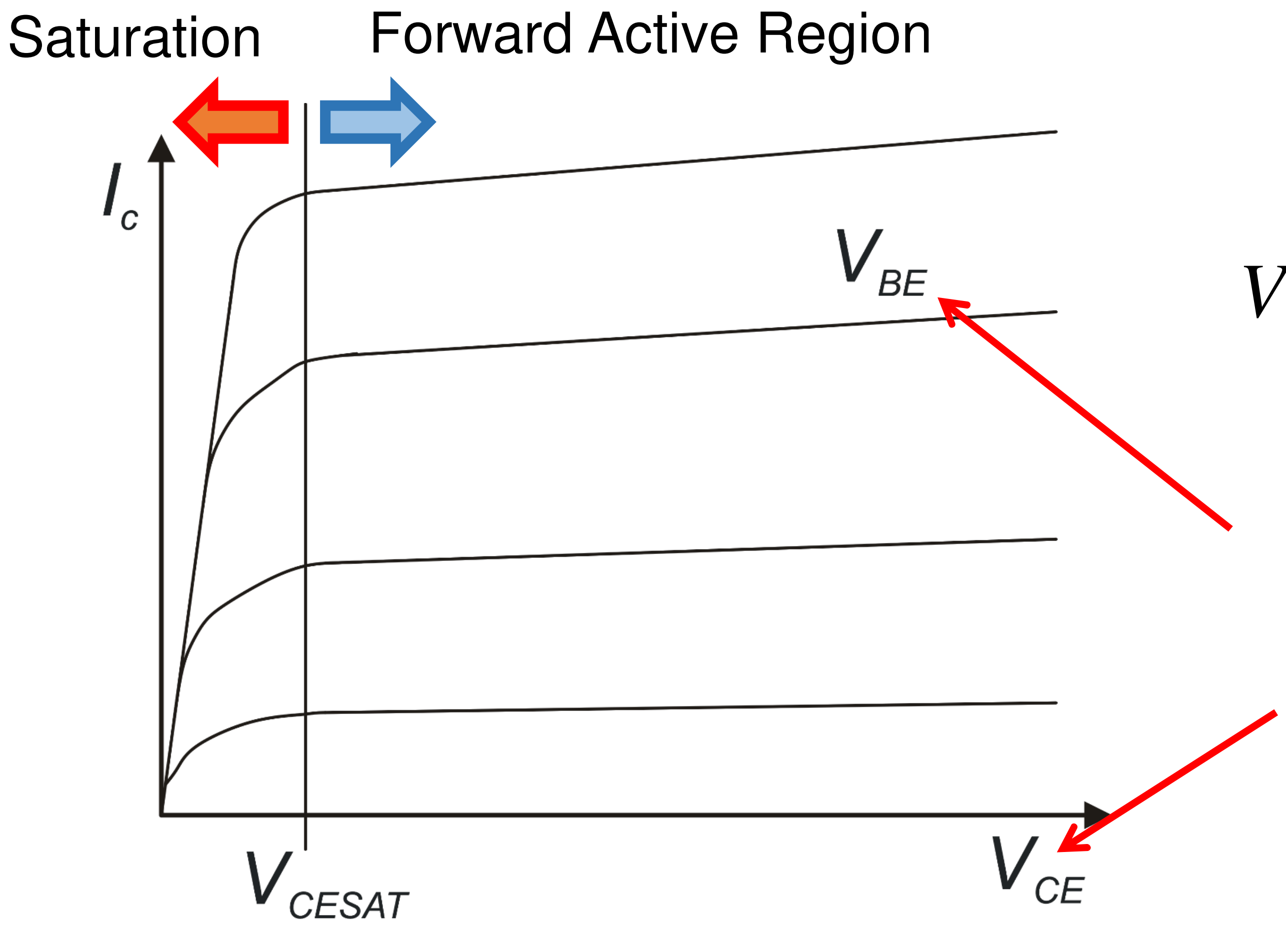
Larger than vertical devices for the same current capability

The substrate PNP: compatible with standard CMOS n-well processes



Limitation: The collector is committed to the substrate, (forced to V_{SS})

BJT output characteristics



$$V_{CESAT} \cong 0.2 \ \mathrm{V}$$

V_{BE}, V_{CE} : control voltages

I_C, I_B: dependent currents

BJT model in the forward active zone

$$I_C = I_S e^{\frac{V_{BE}}{V_T}} \left(1 + \frac{V_{CB}}{V_A} \right) \qquad V_{CB} = V_{CE} - V_{BE}$$

$$I_B = \frac{I_C}{\beta_F}$$

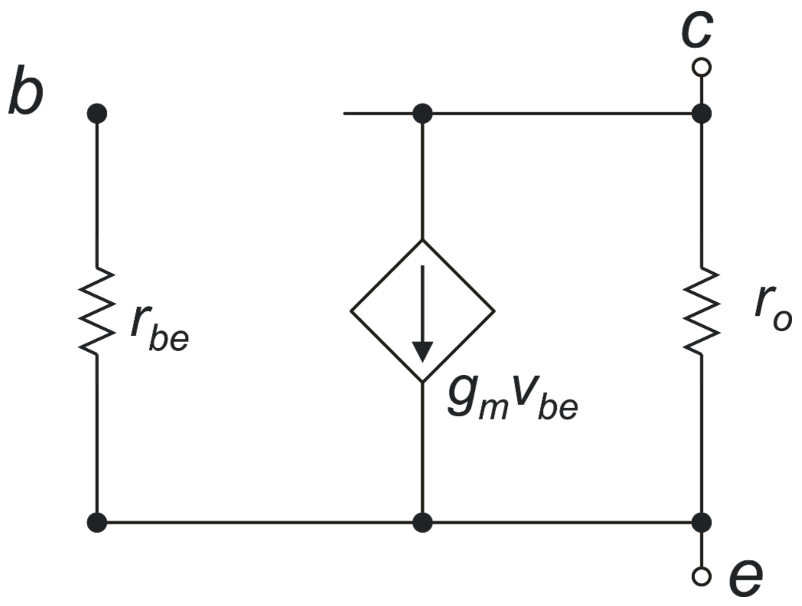
Sometimes this expression is used in order to refer to V_{BE} and V_{CE} as control voltages:

$$I_C \cong I_S e^{\frac{V_{BE}}{V_T}} \left(1 + \frac{V_{CE}}{V_A} \right)$$

For calculation of I_C and I_B in all operating zones (saturation, cut-off, forward active, reverse active) the Ebers-Moll model should be used.

BJT: small signal model

Small signal dc model



$$\boxed{g_m = \frac{I_C}{V_T} \qquad r_0 \cong \frac{V_A}{I_C}}$$

$$r_{be} = \beta_F \frac{1}{g_m}$$

Equivalence with the MOSFET parameters

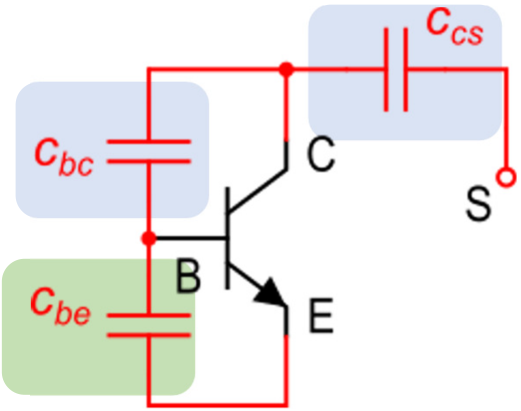
$$g_m = \frac{I_D}{V_{TE}} \qquad r_d = \frac{1}{g_d} = \frac{1}{\lambda I_D} = \frac{\lambda^{-1}}{I_D}$$

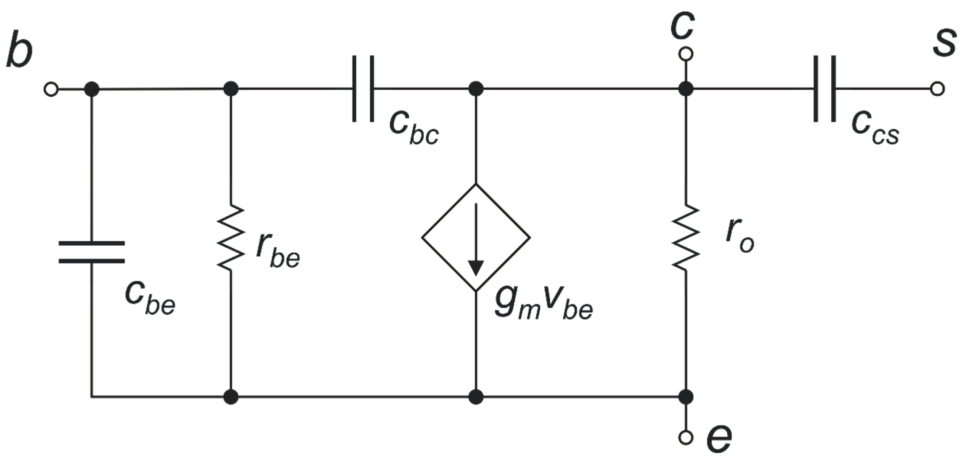
BJT		MOSFET
V_T	$\Leftrightarrow$	V_{TE}
V_A	$\Leftrightarrow$	λ^{-1}

BJT capacitances in forward active region (vertical npn)

$$c_{bc} = \frac{C_{JC}}{\left(1 + \frac{V_{CB}}{V_{JC}}\right)^{m_{jc}}}$$

$$c_{cs} = \frac{C_{JS}}{\left(1 + \frac{V_{CS}}{V_{JS}}\right)^{m_{js}}}$$



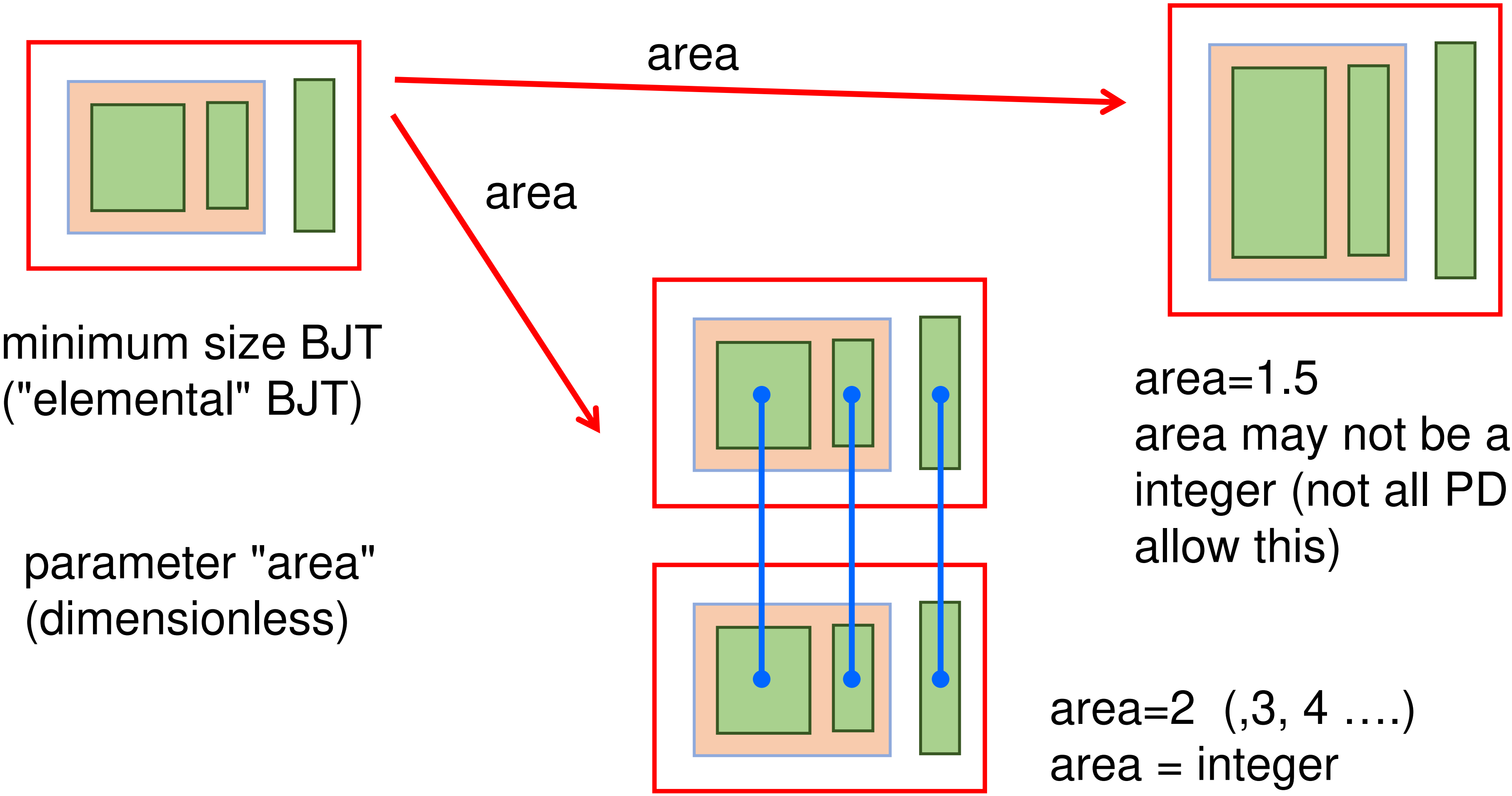


$$c_{be} = \frac{C_{JE}}{\left(1 - \frac{V_{BE}}{V_{JE}}\right)^{m_{je}}} + \underline{\underline{C_{de}}} \longrightarrow c_{de} = \tau_F g_m$$

Transition frequency

$$f_T \cong \frac{1}{2\pi\tau_F}$$

BJTs in Integrated Circuit: instance parameters



minimum size BJT
("elemental" BJT)

parameter "area"
(dimensionless)

area=1.5
area may not be an
integer (not all PDKs
allow this)

area=2 (,3, 4 ….)
area = integer

BJT sizing: Effect of the area parameter on the electrical parameters

Electrical effects of area parameter:

elemental BJT → elemental BJT with area specified as an instance parameter

$I_S \longrightarrow area \times I_S$

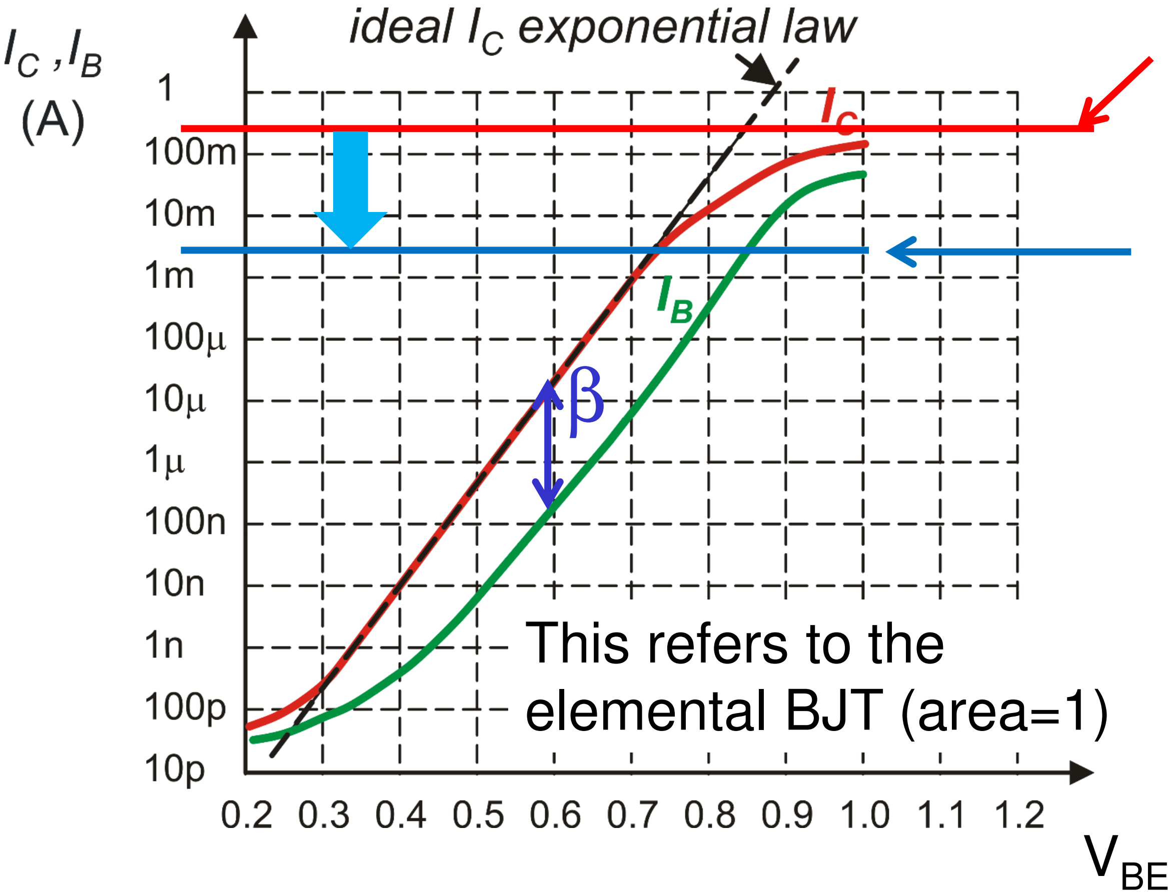
$V_A \longrightarrow V_A$

$\beta \longrightarrow \beta$

$C_{JE} \longrightarrow area \times C_{JE}$

….. → …..

BJT sizing: Gummel plot and beta plot

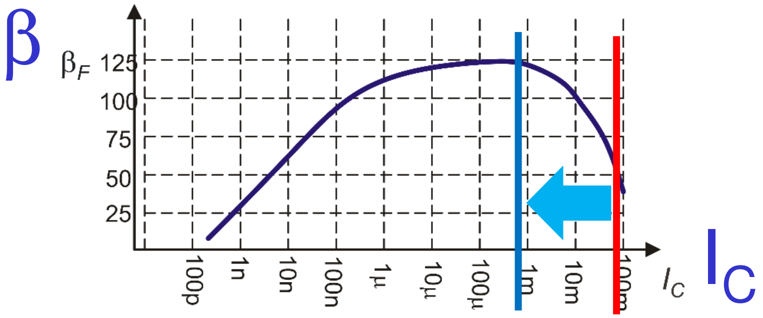


This refers to the elemental BJT (area=1)

Gummel Plot

My BJT has to carry this current (200 mA). The elemental BJT would be damaged

Using a BJT with area=100 would be equivalent to make the elemental BJT work with a current 100 times smaller. This corresponds to the operating point given by the blue line



Beta Plot