CCCD Workshop 2003, Lund, Oct. 2-3

WEAK INVERSION IN ANALOG AND DIGITAL CIRCUITS

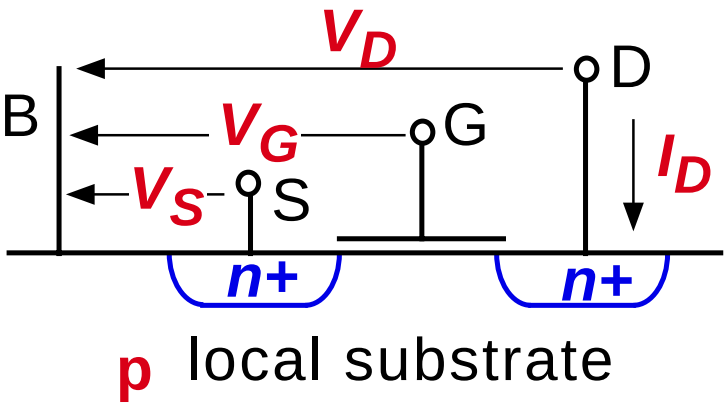
Eric A.Vittoz

CSEM, Centre Suisse d'Electronique et de Microtechnique SA
Jaquet-Droz 1, CH 2007 Neuchâtel, Switzerland
eric.vittoz@csem.ch

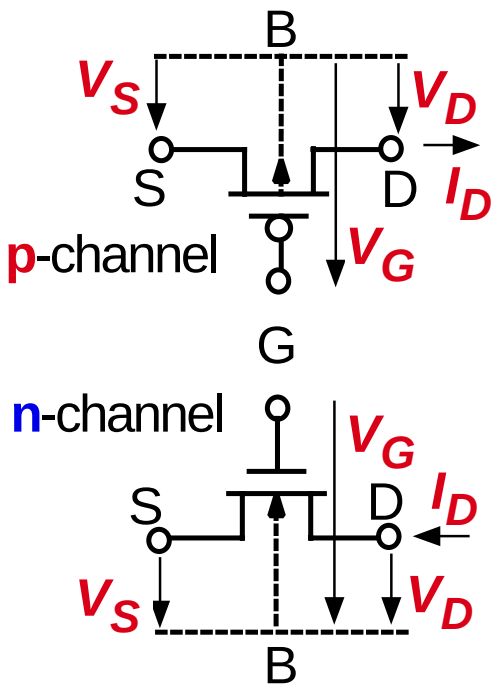
- Behaviour and model of MOS transistors in weak inversion **[1,2,3]**.
- Examples of analog circuits.
- Exploratory analysis of weak inversion logic **[4,5]**.

MOS TRANSISTOR : DEFINITIONS

n-channel



symbols:



W,L width, length of the channel

C_{ox} gate capacitance per unit area

$U_T = kT/q$ (= 26 mV at 300°K)

V = local non-equilibrium voltage in channel : **channel voltage**
(quasi-Fermi potential of electrons)

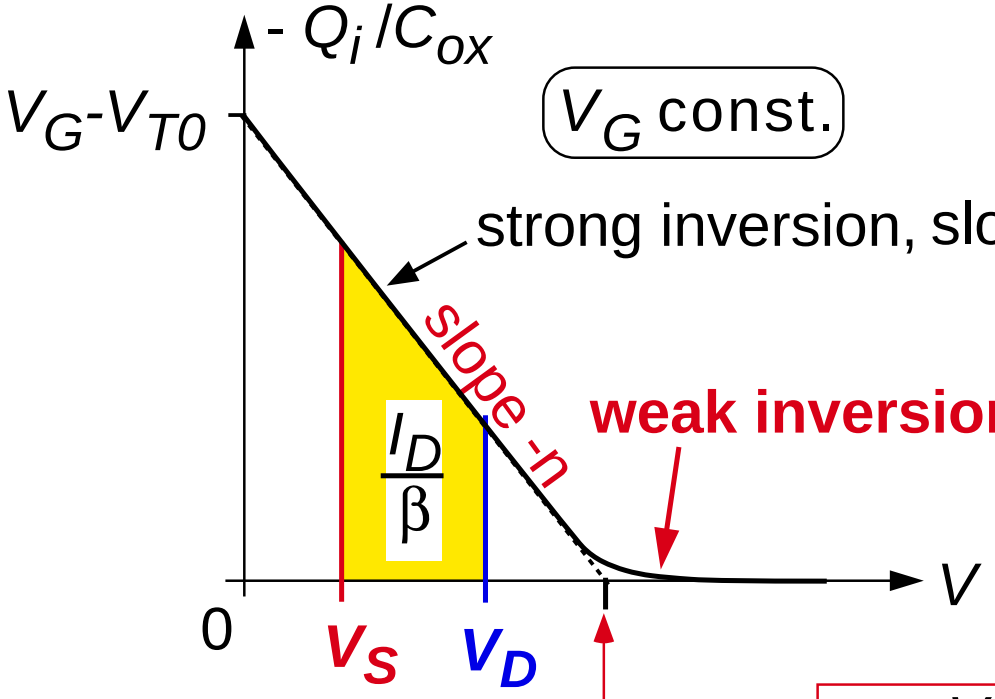
- at source end of channel: $V = V_S$
- at drain end of channel: $V = V_D$

Q_i local mobile inversion charge in channel (electrons)

V_{T0} gate threshold voltage for V=0.

DRAIN CURRENT

- Given by: $I_D = \beta \int_{V_S}^{V_D} -\frac{Q_i}{C_{ox}} dV$ with $\beta = \mu C_{ox} \frac{W}{L}$ (μ=mobility)



strong inversion, slope factor ***n*** =1.2 to 1.6

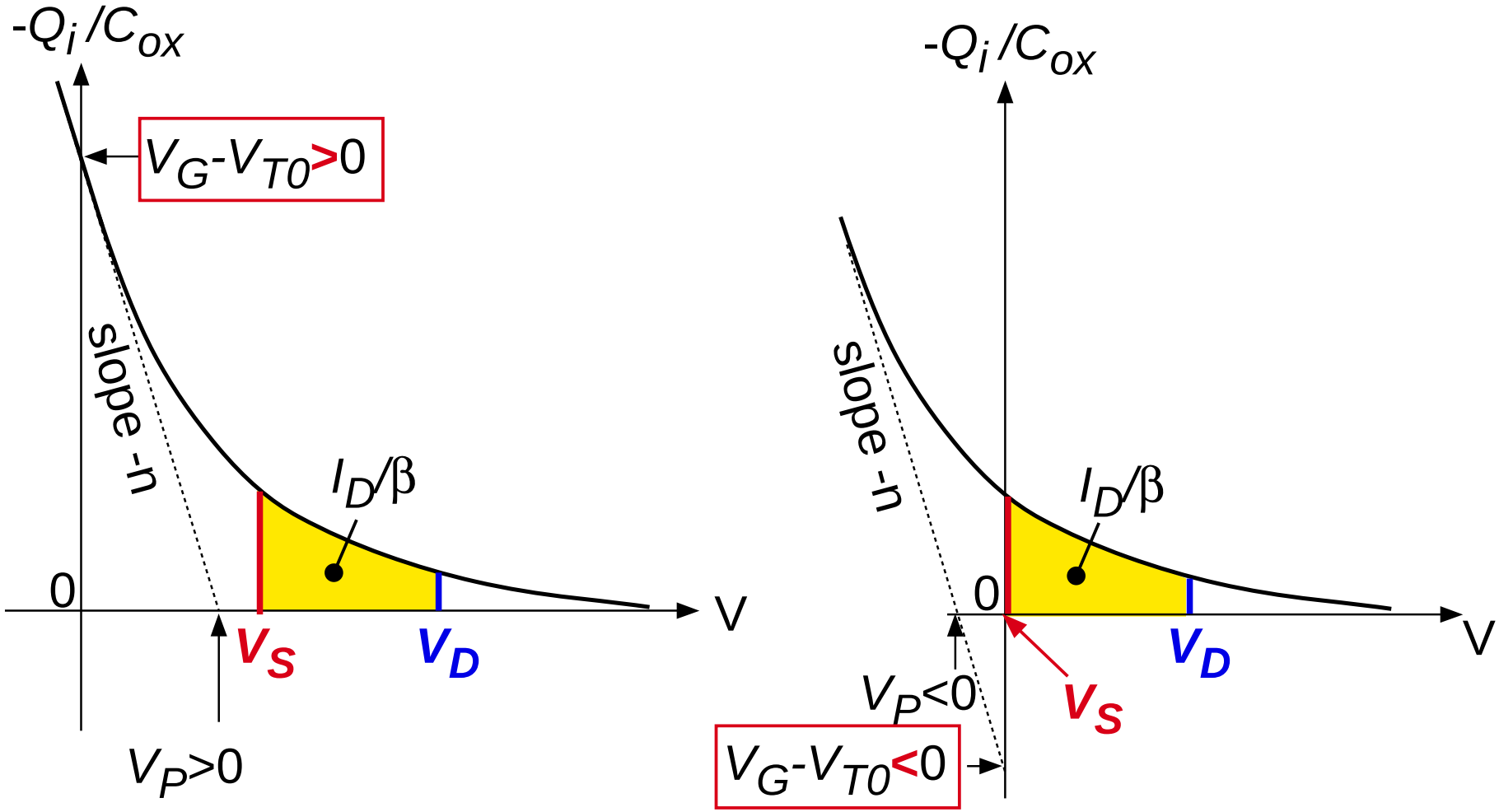
$$\frac{-Q_i}{C_{ox}} = 2nU_T \exp\frac{V_P - V}{U_T}$$ exponential

$$V_P \cong \frac{V_G - V_{T0}}{n}$$

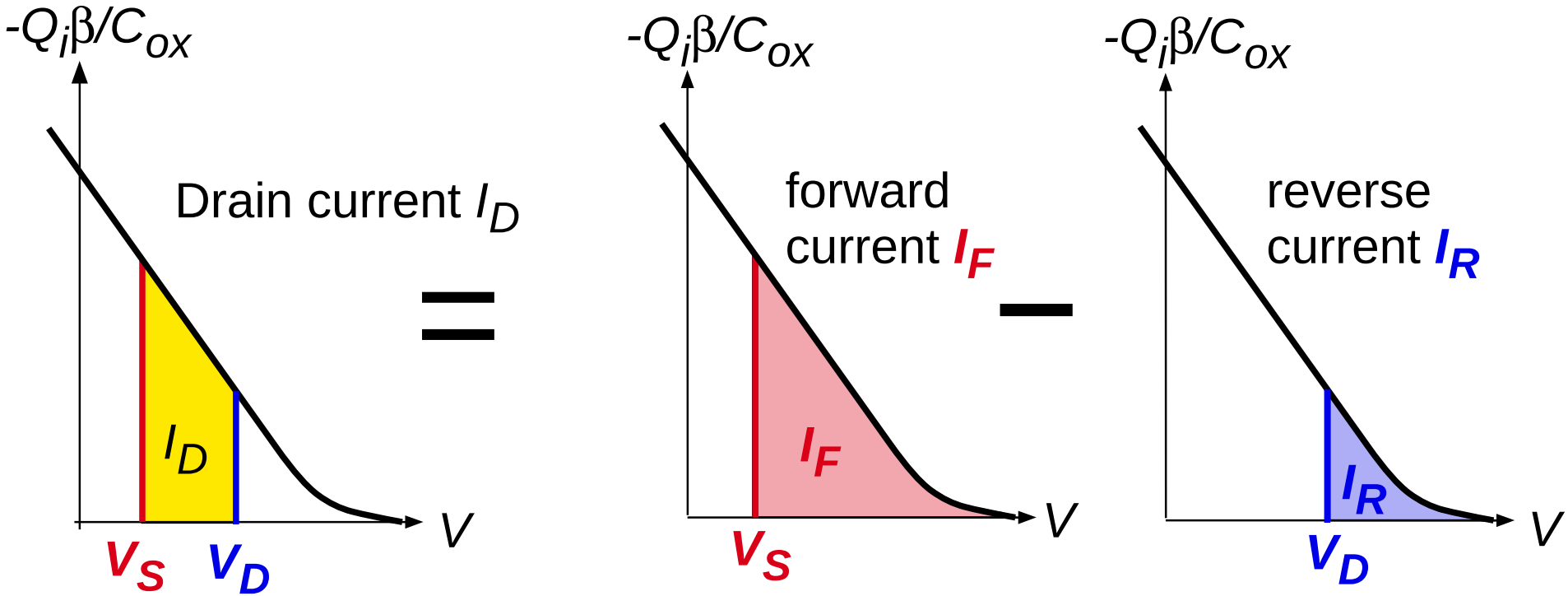
- Weak inversion already possible for V_S=0 if V_G<V_{T0} ("subthreshold")

DRAIN CURRENT IN WEAK INVERSION

(vertical axis magnified)



FORWARD AND REVERSE CURRENTS



$$I_D(V_G,V_S,V_D) = F(V_G,V_S) - F(V_G,V_D) = I_F - I_R$$

- Drain current is the **superposition** of **independent** and **symmetrical** effects of source and drain voltages.
  - **basic property** of long-channel transistors, independent of current **[6]**.
- Transistor saturated if $I_R \ll I_F$, then $I_D = I_F$.

DRAIN CURRENTEXPRESSION IN WEAK INVERSION

- $-Q_i/C_{ox} = 2nU_T\, e^{\frac{V_P - V}{U_T}} \ll 2nU_T$ thus: $I_{F,R} = I_S\, e^{\frac{V_P - V_{S,D}}{U_T}}$

$$\downarrow$$

- Definition: **specific current** of the transistor: $\boxed{I_S = 2n\beta U_T^2}$

(10 to 300 nA for W=L)

- Introducing $V_P \cong (V_G - V_{T0})/n$ and $I_D = I_F - I_R$, this yields:

$$\boxed{I_D = I_S\, e^{\frac{V_G - V_{T0}}{nU_T}} \left(e^{-\frac{V_S}{U_T}} - e^{-\frac{V_D}{U_T}} \right)}$$

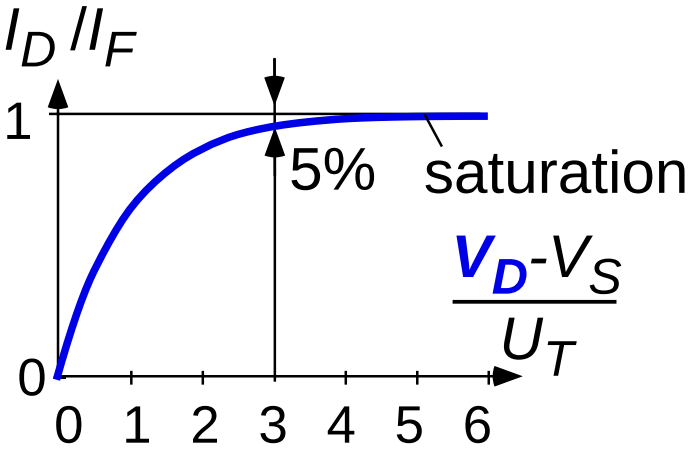
I_F I_R

for I_F and $I_R \ll I_S$

FORWARD CHARACTERISTICS IN WEAK INVERSION

$$I_D = I_{D0}\, e^{\frac{V_G}{nU_T}} \left(e^{-\frac{V_S}{U_T}} - e^{-\frac{V_D}{U_T}} \right) \qquad \text{where } I_{D0} = I_S\, e^{-\frac{V_{T0}}{nU_T}}$$

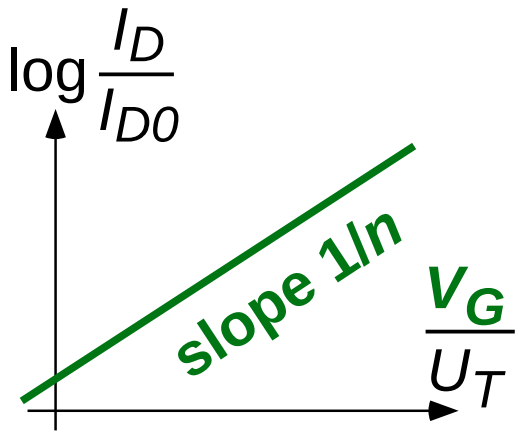
- output
  V_G, V_S = const.



$$I_D \sim 1 - e^{-\frac{V_D - V_S}{U_T}}$$

minimum V_{DSsat}

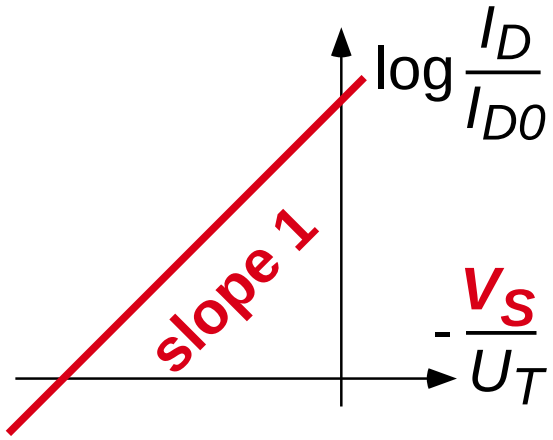
- transfer from gate
  V_S, V_D const.



$$I_D \sim e^{\frac{V_G}{nU_T}}$$

exponential, slope 1/*n*

- transfer from source
  V_G, V_D const.



$$I_D \sim e^{-\frac{V_S}{U_T}}$$

exponential, slope 1

CONTINUOUS MODELS WEAK-STRONG INVERSION

a. From charge analysis [7,8]:

$$\frac{V_P - V_{S,D}}{U_T} = \sqrt{1 + 4\frac{I_{F,R}}{I_S}} - 1 + \ln\frac{\sqrt{1 + 4\frac{I_{F,R}}{I_S}} - 1}{2}$$

cannot be inverted to express $I_{F,R}(V_P, V_{S,D})$

b. Interpolation formula: $\frac{I_{F,R}}{I_S} = \ln^2\left(1 + e^{\frac{V_P - V_{S,D}}{2U_T}}\right)$ [9]

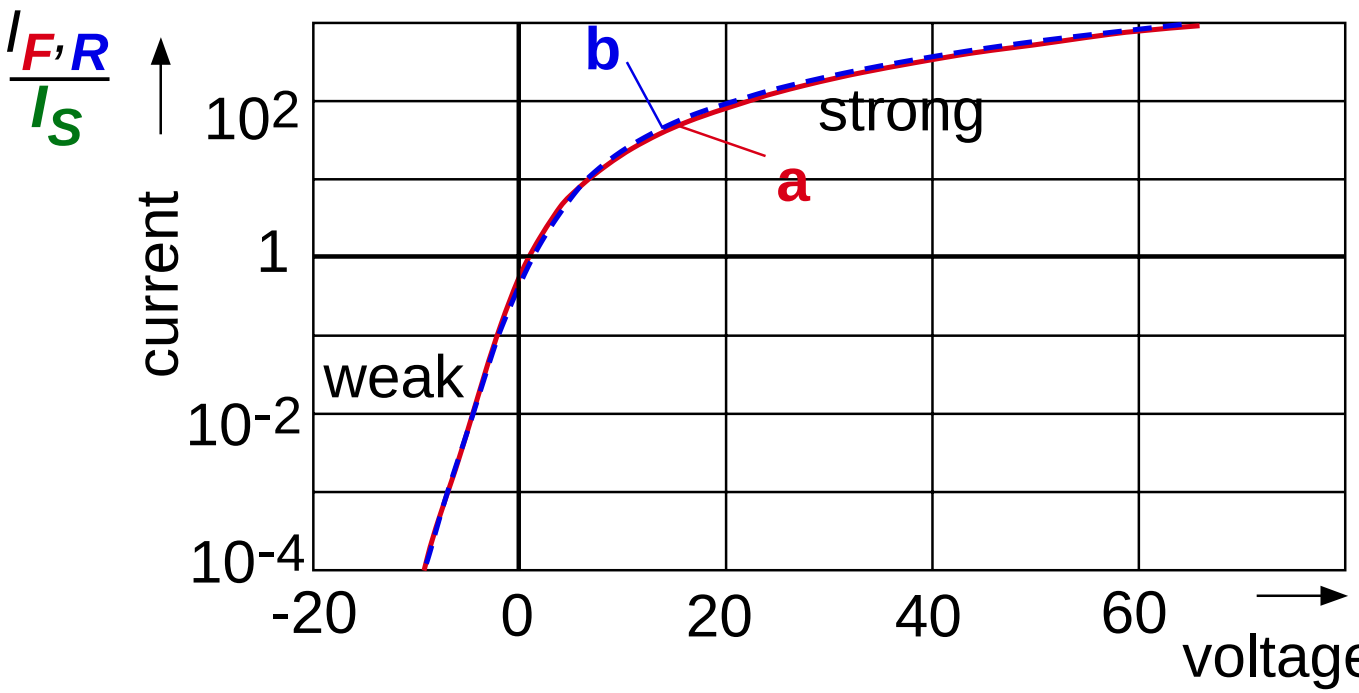
Both converge asymptotically towards:

$$\frac{I_{F,R}}{I_S} = e^{\frac{V_P - V_{S,D}}{U_T}} \quad \text{for } V_P - V_{S,D} \ll U_T \text{ (weak inversion)}$$

$$\frac{I_{F,R}}{I_S} = \left(\frac{V_P - V_{S,D}}{2U_T}\right)^2 \quad \text{for } V_P - V_{S,D} \gg U_T \text{ (strong inversion)}$$

- Only 3 parameters: V_{T0}, n (inside V_P) and I_S (or β) to model the current from weak to strong inversion.

CONTINUOUS MODELS WEAK-STRONG INVERSION



with:

$V_P = (V_G - V_{T0})/n$

$I_D = I_F - I_R$

- Definition: **Inversion coefficient**: ***IC*** = the larger of I_F/I_S and I_R/I_S

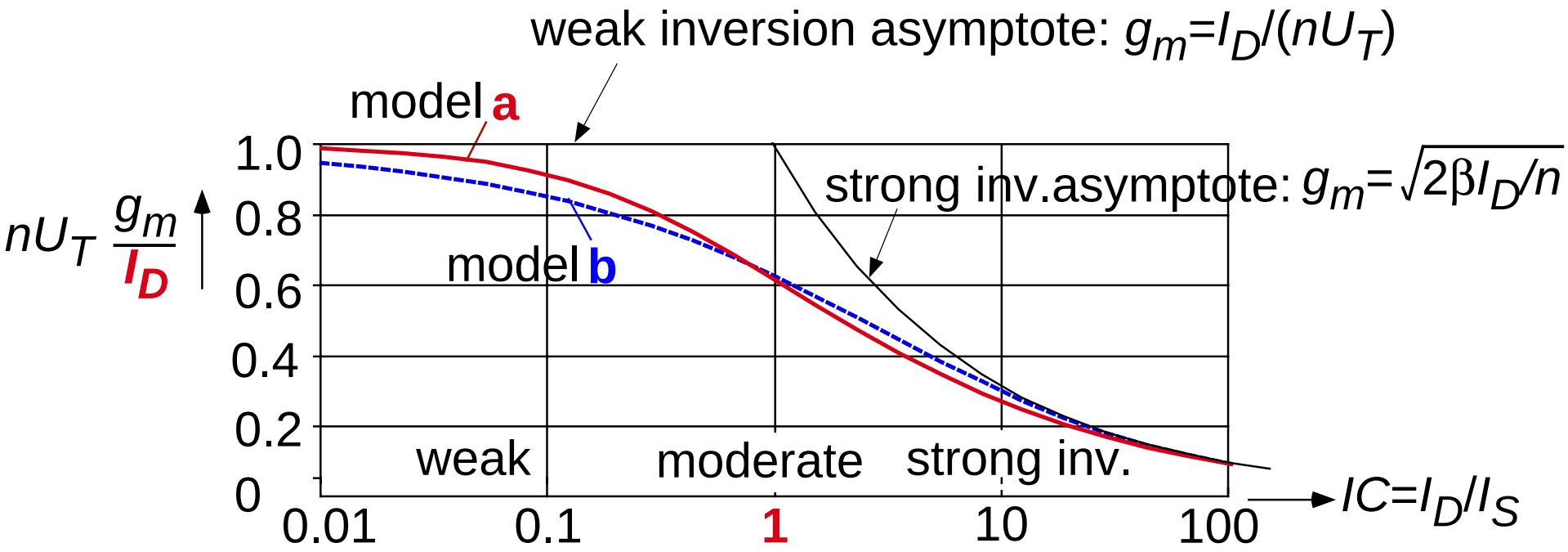
weak inversion: ***IC*** « 1

moderate inversion: ***IC*** ≅ 1

strong inversion: ***IC*** = $\left(\frac{V_{DSsat}}{2U_T}\right)^2$ » 1

TRANCONDUCTANCE FROM WEAK TO STRONG INVERSION

- Transonductance g_m from gate in **saturation**



- g_m/I_D decreases with increasing inversion coefficient *IC.*
- g_m/I_D is **max**imum in **weak** inversion.

SUMMARY OF FEATURES OF WEAK INVERSION

- Large-signal DC model:

$$I_D = I_S\, e^{\frac{V_G - V_{T0}}{n U_T}} \left(e^{-\frac{V_S}{U_T}} - e^{-\frac{V_D}{U_T}} \right)$$

+ exponential : $\Longrightarrow$
  - + translinear circuits and log domain filters
  - + max. I_{on}/I_{off} for given voltage swing
  - – intermodulation in RF front ends

+ min. V_{DSsat}
+ min. gate voltage
+ min. gate capacitance
+ max. g_m/I_D : $\Longrightarrow$
  - + max. intrinsic voltage gain
  - + min. input noise density for given I_D
  - + max. bandwidth for given kT/C and I_D
  - + min. input offset voltage
  - – max. output noise current for given I_D
  - – max. **current mismatch** : dominated by V_T -mismatch: $\frac{\Delta I_D}{I_D} = \frac{\Delta V_{T0}}{n U_T}$
+ $g_m(I_D)$ linear
± g_m independent of β
– Low speed: $f_T \cong \frac{\mu U_T}{2\pi L^2}$

EVOLUTION OF *IC* WITH SCALED-DOWN PROCESSES

- Scaling-down of process:
  - dimension scaling by factor k
  - all voltages decreased by k, **except U_T**:

    - analog circuits: V_{DSsat} must be decreased by k, thus

$$IC = \left(\frac{V_{DSsat}}{2U_T}\right)^2 \quad \textbf{decreased by } k^2$$

    - digital circuits: V_B decreased by k, thus

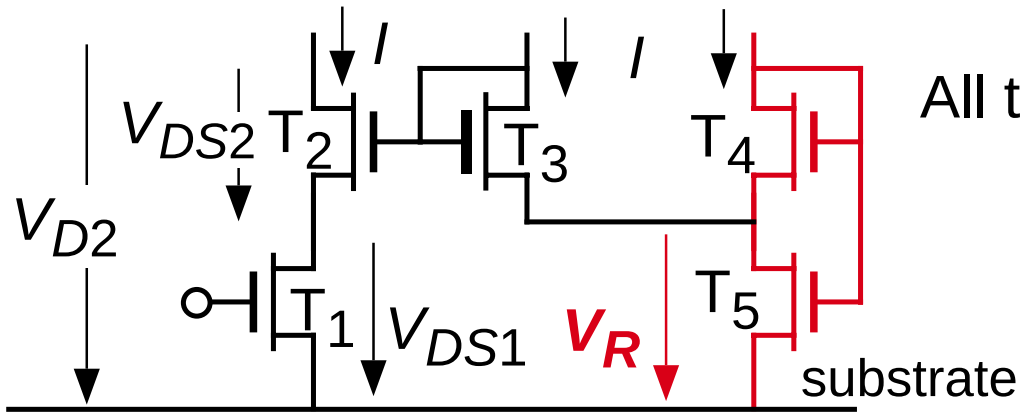
$$IC_{on} = \left(\frac{V_B - V_{T0}}{2nU_T}\right)^2 \quad \textbf{decreased by } k^2$$

- **Weak inversion approached** for constant temperature T.
- Transition frequency: $f_T = \frac{\mu V_{DSsat}}{2\pi L^2}$ **increased by k**

  - weak inversion with L=100nm : f_T >4 GHz

LOW-VOLTAGE CASCODE IN WEAK INVERSION [2]

V_{DSsat} = 4 to 6U_T per transistor



All transistors in weak inversion with:

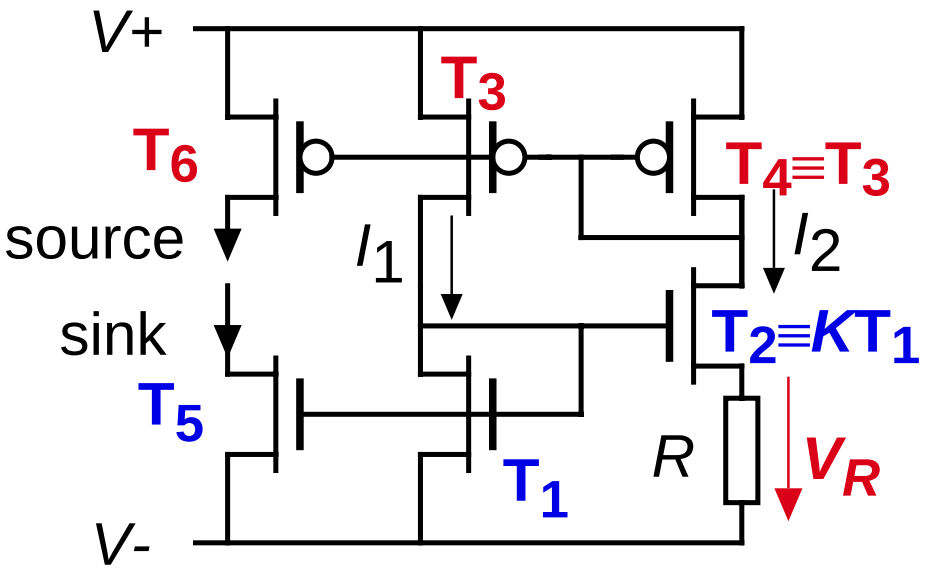
$$\frac{\beta_2}{\beta_3} = P \quad \text{and} \quad \frac{\beta_4}{\beta_5} = M$$

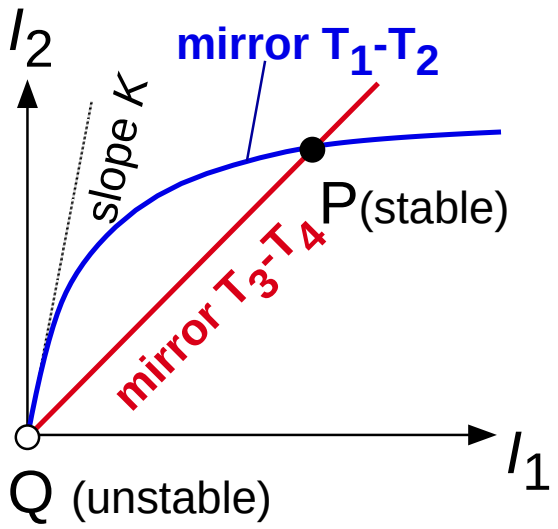
- Model in weak inversion yields: $V_{DS1} = U_T \ln [\, P\,(1+ 2M)]$

for $P = M = 8$: $V_{DS1}=5U_T$,

thus V_{D2} = **$10U_T$ sufficient** to saturate T_1 and T_2

EXTRACTION OF U_T AND CURRENT REFERENCE [1]



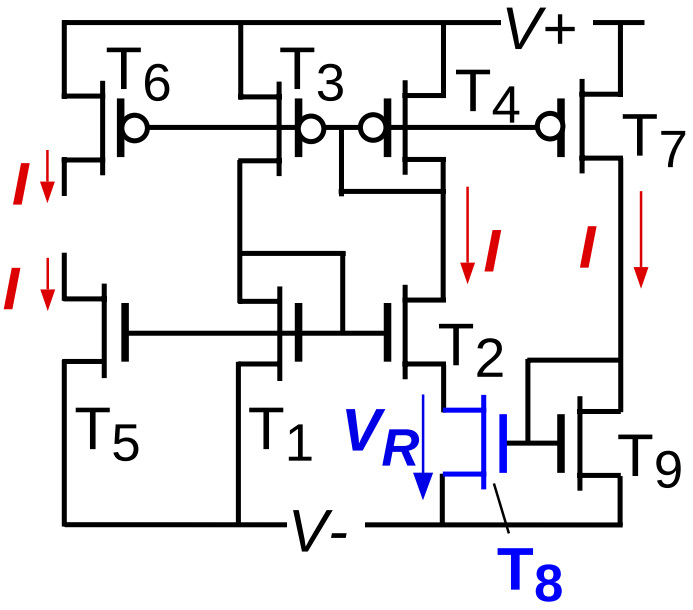


- For T_1 and T_2 in weak inversion: $V_R = RI_2 = U_T \ln K$

- Self-starting if leakage of T_2 larger than that of T_1.

CURRENT GENERATION WITHOUT RESISTOR

- Resistor replaced by transistor T_8 in conduction [10]:
  - $T_6 = T_3 = T_4 = T_7$ and $T_5 = T_1$
  - T_8 and T_9 in **strong** inversion with $\beta_8 = A\beta_9$ (A»1 to have T_8 in conduction)
  - T_2 and T_1 in **weak** inversion with $\beta_2 = K\beta_1$



yields: $$I = 2n\beta_8 U_T^2 . A \ln^2 K = I_{S8} . A \ln^2 K$$

- Reference current I proportional to specific current I_{S8}
- Useful to bias transistors at inversion coef.IC independently of process.
- If mobility ~ T^{-2}, then compensation by U_T^2 : $I \sim I_S$ independent of T

MOS TRANSISTOR OPERATED AS A PSEUDO-RESISTOR

[11,12,13,6]

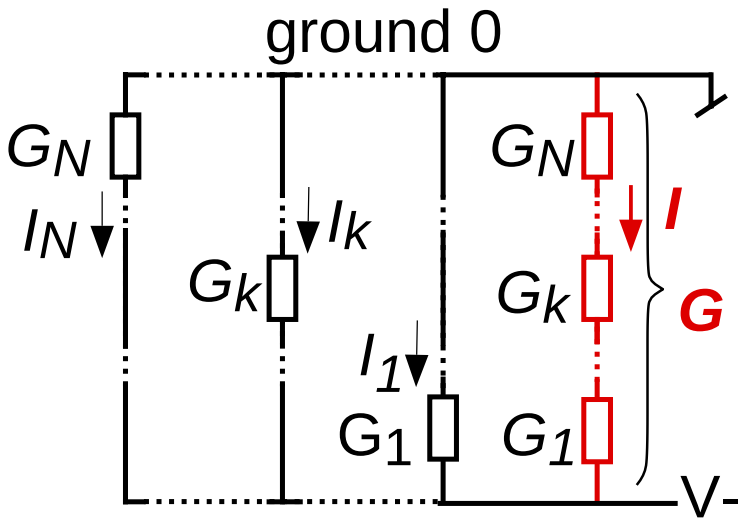
Consequence of basic property $I_D = F(V_S) - F(V_D)$:

- Networks of transistors with **same gate voltage** are
  - **linear** with respect to **currents**
  - thus equiv. for currents to a **resistive prototype**, with $G_i = 1/R_i \sim I_{Si}$
  - **ground** in res. prototype correspond to **saturated transistors.**
  - example of application: current-mode linear attenuator (e.g. *R*-2*R*).
- In weak inversion:
  - **linearity** of currents even for **different gate voltages**

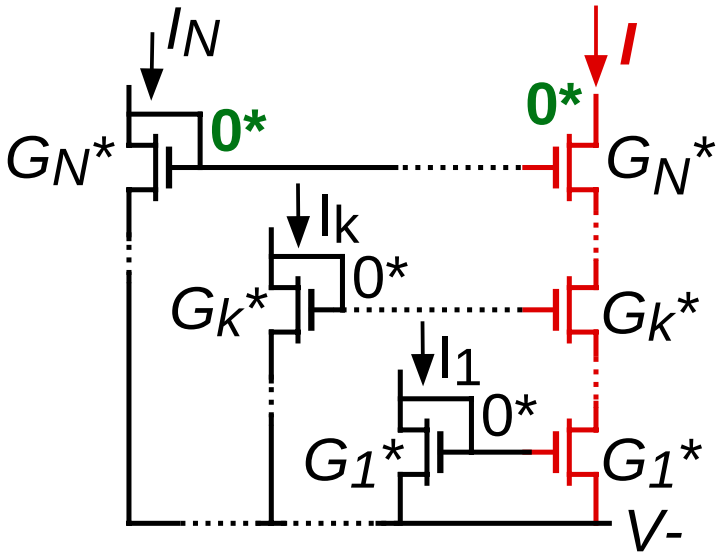
with $$G_i = 1/R_i \sim I_{Si} \exp \frac{V_{Gi}}{nU_T}$$

simple example of pseudo-R network in weak inversion:

CALCULATION OF HARMONIC MEAN [14,13]



resistive prototype



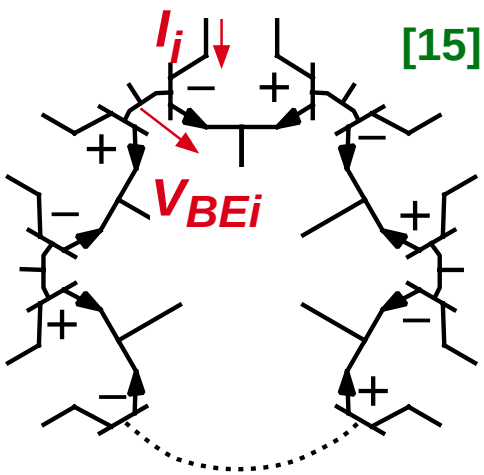
pseudo-resistive version

(**0***=pseudo-ground)

- Series combination of G_i : ***G*** = $\frac{1}{\sum 1/G_i}$ **harmonic mean**
- Same voltage across ***G*** and G_i, thus ***I*** = $\frac{1}{\sum 1/I_i} = \frac{I_{hm}}{N}$
- Can be used as a fuzzy AND gate.

TRANSLINEAR CIRCUITS

With bipolar transistors:



[15]

+ and - directions of BE_i (any sequence)

$$\sum_{+} V_{BEi} = \sum_{-} V_{BEi}$$

with: $V_{BEi} = U_T \ln \frac{I_i}{I_{si}}$

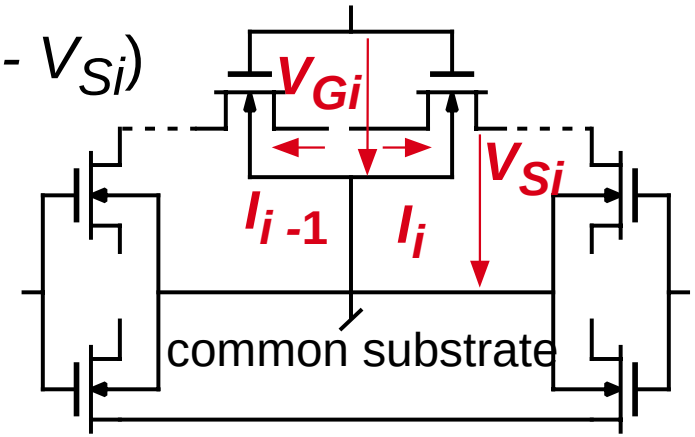
$$\frac{\prod_{+} I_i}{\prod_{-} I_i} = \frac{\prod_{+} I_{si}}{\prod_{-} I_{si}} = \lambda$$

With MOS transistors in weak inversion: [16,17]

$$\sum_{+}(V_{Gi} - V_{Si}) = \sum_{-}(V_{Gi} - V_{Si})$$

with:

$$\frac{V_{Gi}}{n} - V_{Si} = U_T \ln \frac{I_i}{I_{D0i}}$$



- If **+ and -** are **alternated** then: pairs of equal V_{Gi} both sides of equation:

  $V_{Gi} \Rightarrow V_{Gi}/n$ for each pair,

  and then

- Otherwise: **separate wells** connected to sources to impose $V_{Si} = 0$
- Precision degraded by V_{T0} mismatch

BASIC CONSIDERATIONS FOR WEAK INVERSION LOGIC [5]

- Dynamic power consumption: $P_{dyn} = f\, C\, \Delta V\, V_B$ (supply voltage; logic swing)
- Weak inversion model can be rewritten as

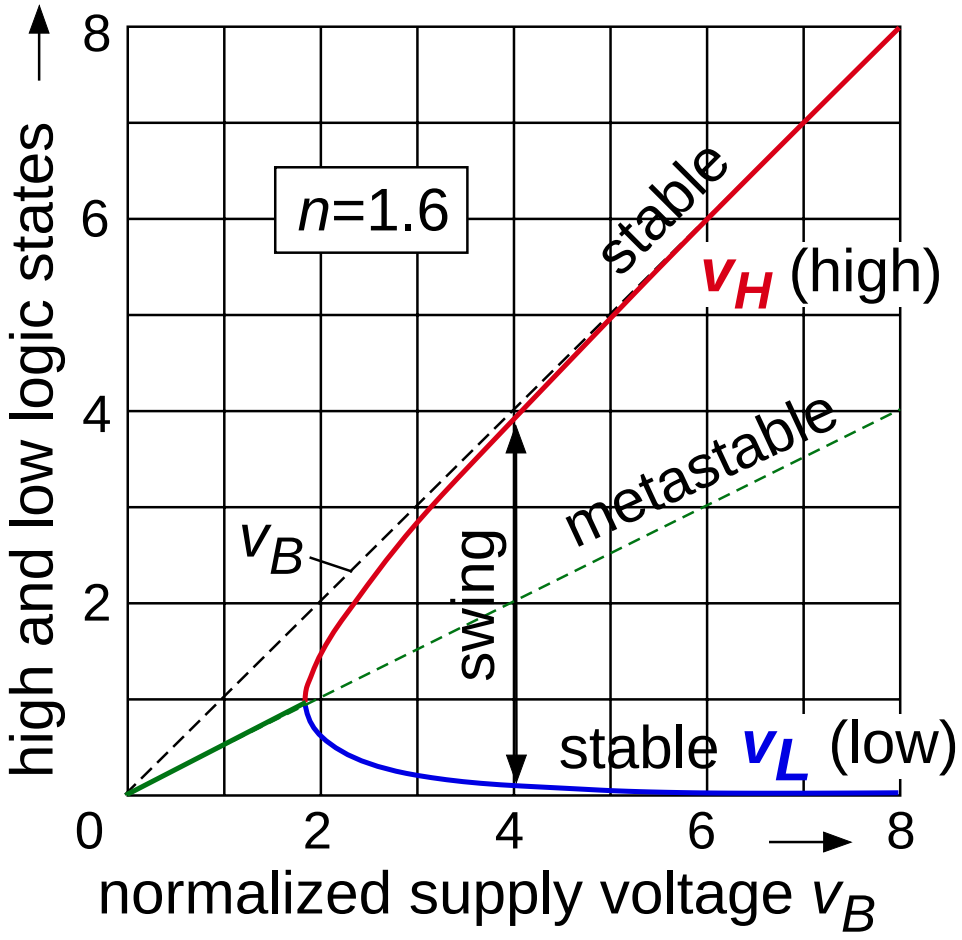
$$I_D = I_0\, e^{\frac{V_{GS}}{nU_T}} \left(1 - e^{-\frac{V_{DS}}{nU_T}}\right)$$

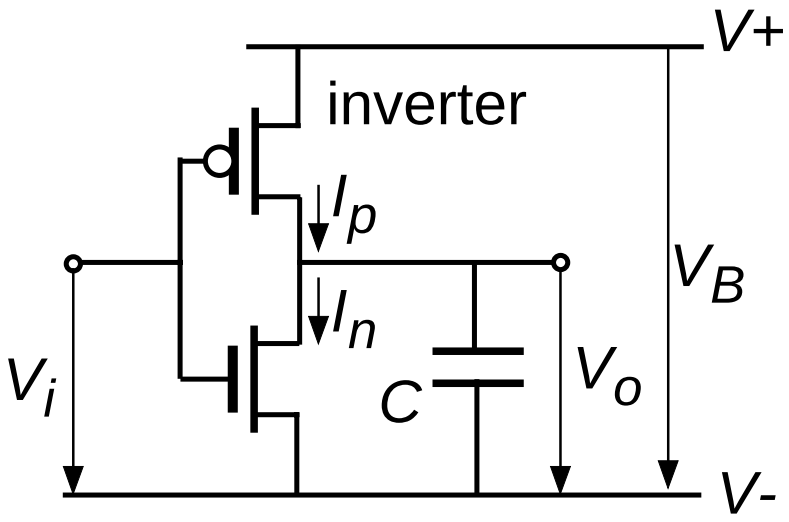
  - exponential in V_{GS}, with maximum g_m/I_D, thus:
    - **minimum swing ΔV for given I_{on}/I_{off}**, hence
      - **minimum P_{dyn} for given I_{off}**
    - with: $I_0 = I_S\, e^{-\frac{V_{T0}+(n-1)V_S}{nU_T}}$ **adjustable by V_S.**

- Assumptions on process:
  1. Threshold V_{T0} close to 0 (V_S cannot be too negative).
  2. Triple well (true twin well): separate local p and n substrates
     - adjustment of I_0 by V_S for n- and p-channel.

STABLE STATES OF CMOS FLIP-FLOP

- Simplifying assumptions: $n_n=n_p=n$, $I_{0n}=I_{0p}=I_0$
- Normalized voltages $v_k=V_k/U_T$

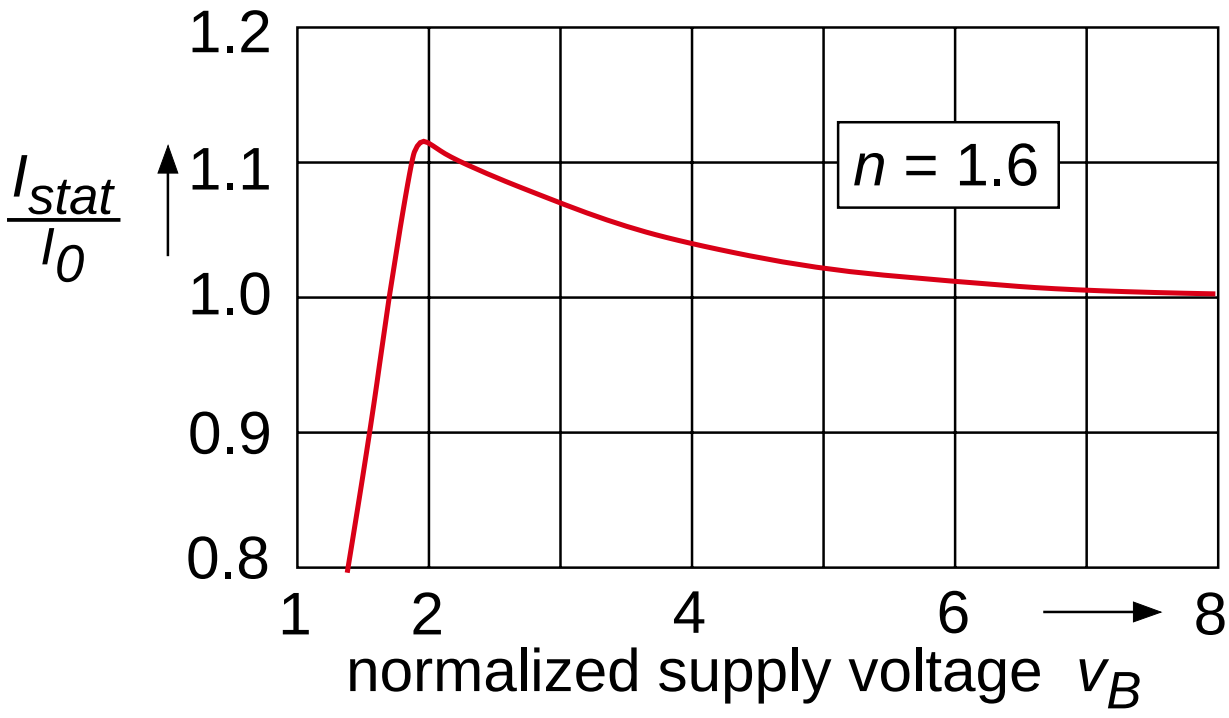




- bistable for $V_B > 1.91U_T$
- 95% swing for $V_B = 4U_T$

STATIC CURRENT AT LOGIC STATES

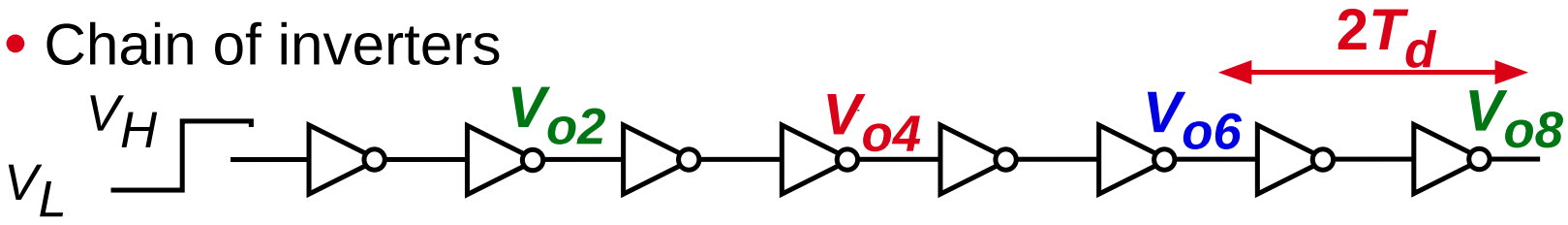
- Since $V_L = V_B - V_H > 0$, static current I_{stat} at each state is larger than I_0

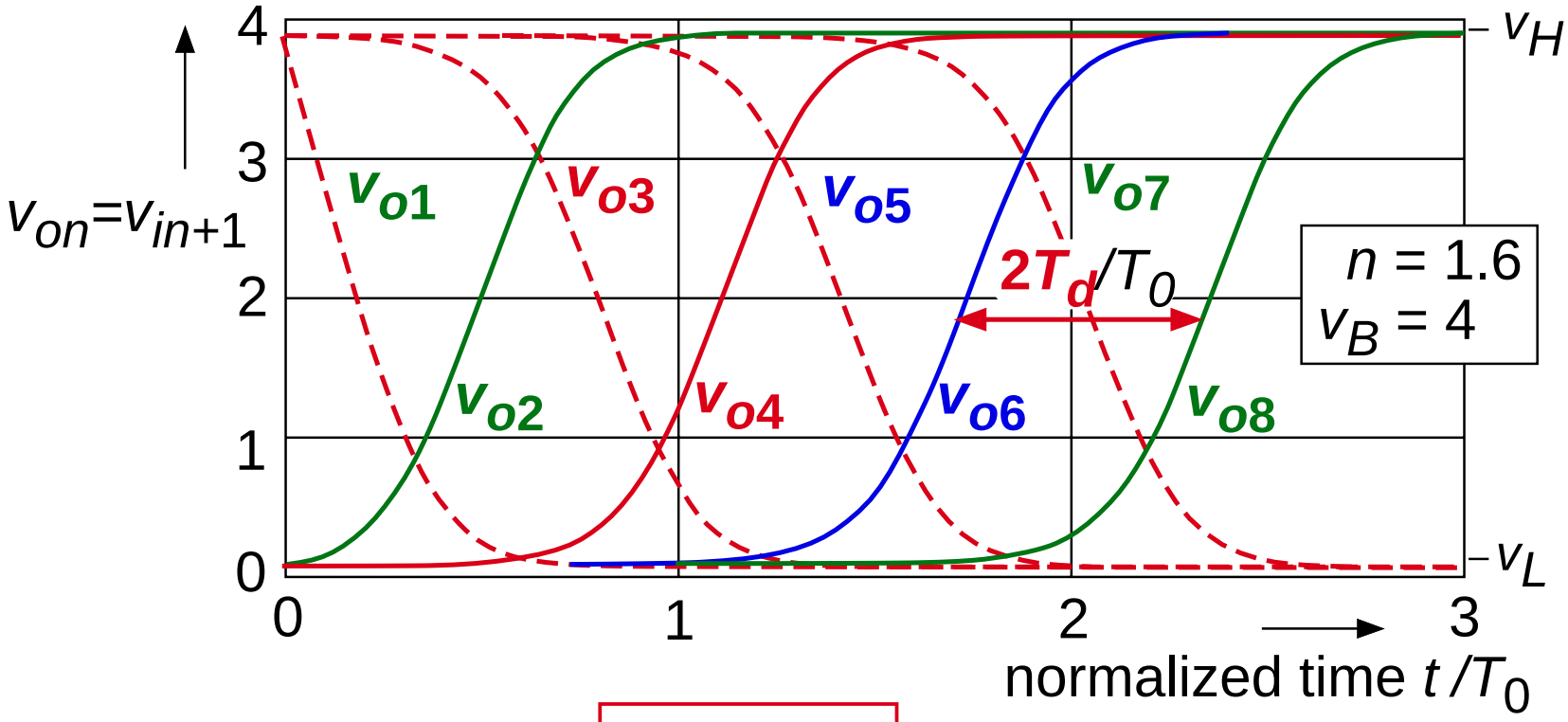


- I_{stat} <4% above I_0 for $v_B \geq 4$: the difference can be **neglected** thus:
- Static power : $P_{stat} \cong I_0 V_B$

STANDARD TRANSITIONS IN HOMOGENEOUS SYSTEM

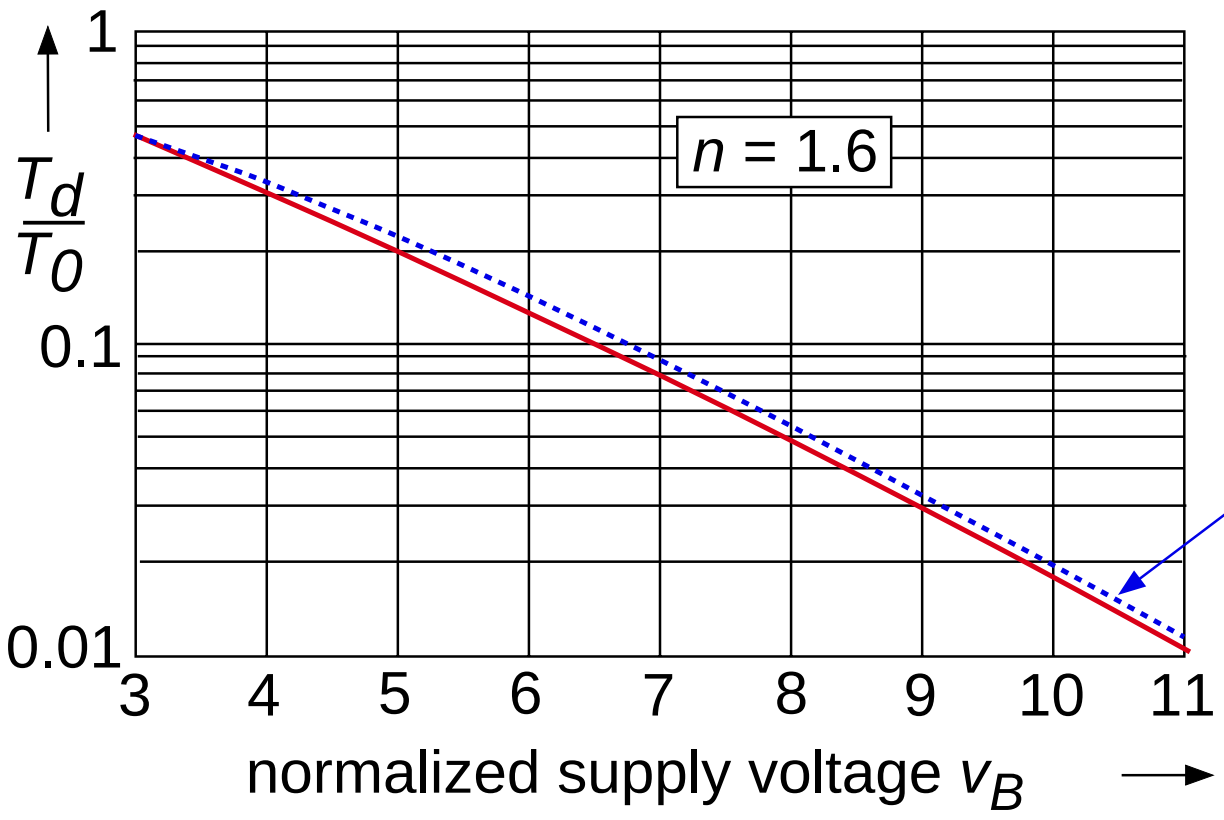
- Chain of inverters





- Characteristic time : $T_0=CU_T/I_0$
- Transitions become standard after a few stages
- Normalized delay time T_d/T_0 only depends on V_B and n.

DELAY TIME FOR STANDARD TRANSITIONS



- Approximation:

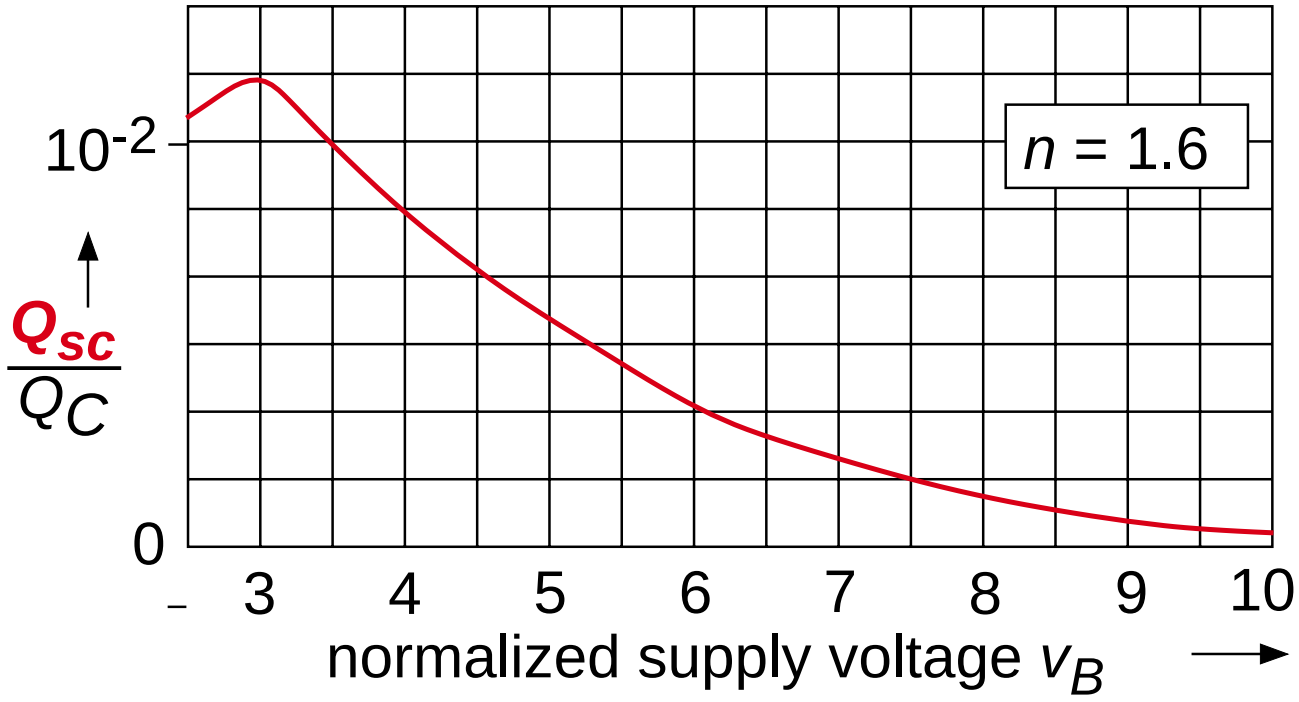
$$T_d \cong \frac{CV_B}{I_{on}} \cong \frac{CV_B}{I_0 e^{V_B/nU_T}}$$

or $I_0 \cong \frac{CV_B}{T_d} e^{-V_B/nU_T}$

(for calcul. of P_{stat})

- T_d decreases approximately **exponentially** with increasing V_B.

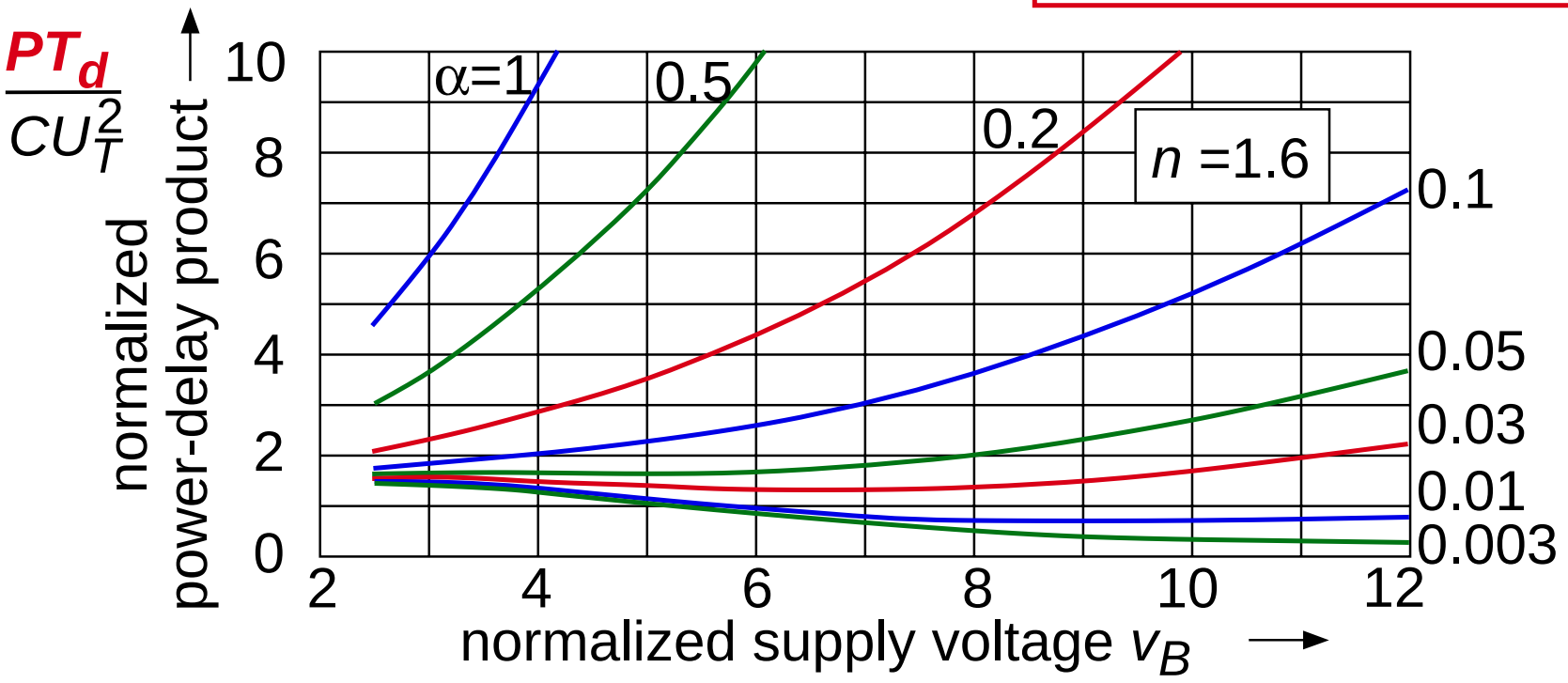
PROPORTION OF SHORT-CIRCUIT CHARGE FOR STANDARD TRANSITIONS



- Short-circuit charge Q_{sc} < 1.4% capacitor charge Q_C : **negligible**, thus:
  - dynamic power $P_{dyn} \cong fQ_CV_B \cong fCV_B^2$
  - with static power $P_{stat} = I_{stat}V_B \cong I_0V_B$

POWER-DELAY PRODUCT

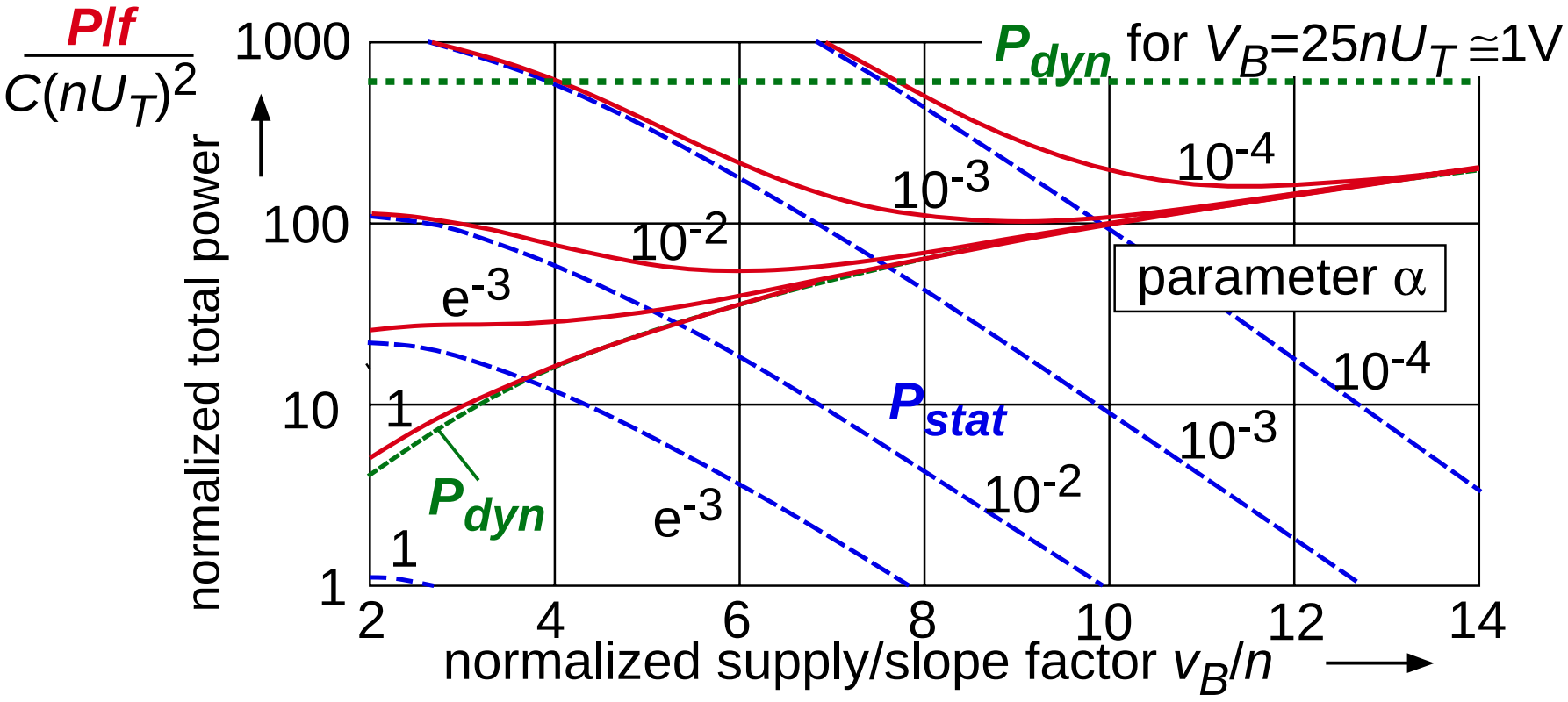
- Definition: **duty factor** $\alpha = 2f\,T_d \leq 1$
  - proportion of time during which the gate is in transition.
- Then, total power $P = P_{dyn} + P_{stat} \Rightarrow$ $P = \frac{CU_T^2}{T_d} v_B^2(\alpha/2 + e^{-v_B/n})$



- P_{dyn} dominates for large $\alpha \Rightarrow$ min. V_B for min. PT_d
- P_{stat} dominates for small $\alpha \Rightarrow$ increase V_B to increase I_{on}/I_{off}

POWER/FREQUENCY RATIO

- By re-using $\alpha = 2f\,T_d$: $P/f = C(nU_T)^2\,(v_B/n)^2\left(1+\frac{\alpha}{2}\,e^{-v_B/n}\right)$



- V_{Bopt} and P_{min} increase for decreasing α
- At P_{min} : $P_{dyn} \gg P_{stat}$
  - Increasing I_0 does not allow to reduce V_B significantly for T_d const.
- For α > 5%, power reduction by >**20** compared to P_{dyn} at 1V.

MAXIMUM SPEED

- Since $T_d \cong \frac{CV_B}{I_{on}}$ and $I_{onmax} \cong IC_{on}\, I_S$ (inv. coeff* spec. current), thus:

$$T_{dmin} \cong \frac{V_B}{IC_{on}} \frac{C}{I_S}$$

process

- Limit of weak inversion: $IC_{on} \cong 1$, thus

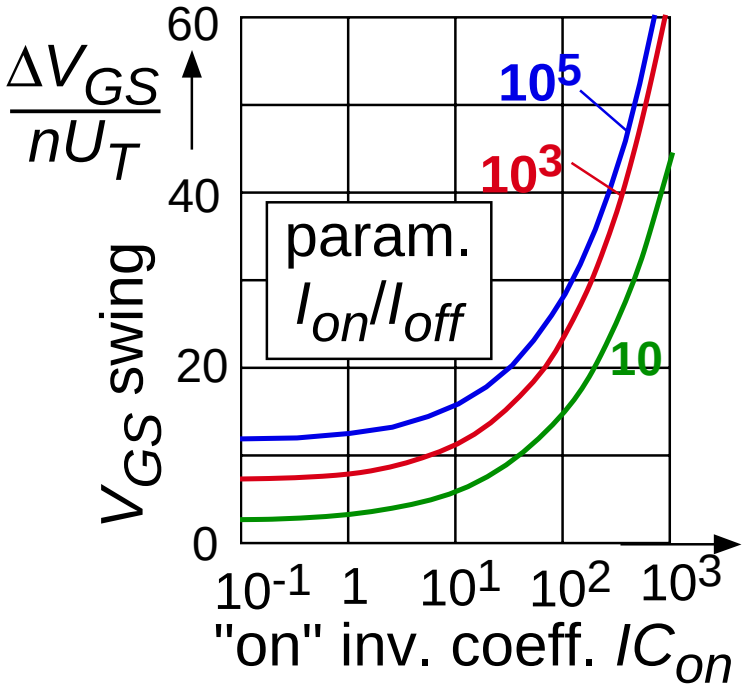
$$T_{dmin(weak)} \cong V_B \frac{C}{I_S}$$

- Higher speed can only be obtained by entering moderate or strong inv.

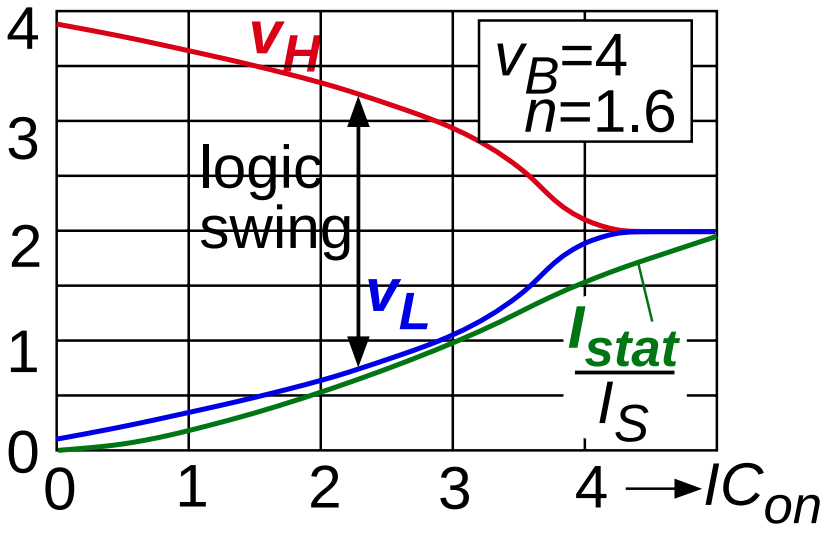
EFFECT OF ENTERING MODERATE AND STRONG INVERSION

(using continuous model of I_D)

- More **voltage swing** needed
  to obtain I_{on}/I_{off}
  - from continuous current model:



- **Degeneration** of logic states:
  - reduction of logic swing
  - large increase of static current I_{stat}
  - loss of bistability
  - more supply voltage needed.

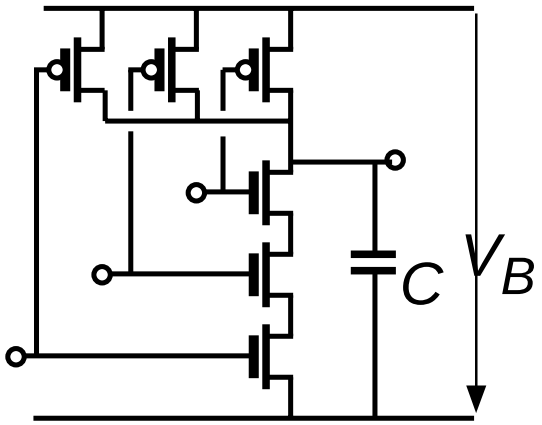


NUMERICAL RESULTS

- Simple inverter replaced by 3-input NAND-gate:
  - approx. equivalent to inverter with

    L = 3-time that of n-ch transistor

    C = 6-time that of min. inverter (includes $C_{interconnect}=C/2$).



parameter		process A	process B	unit
min. channel length	L_{min}	500	180	nm
equiv. spec. current	I_S	200	400	nA
equiv. load capac.	C	20	4	fF
specific energy	$C(nU_T)^2$	28	4.2	aJ
P/f for α=1 $V_B=4U_T$		228	44	aJ
$(P/f)_{min}$ for α=0.01 and $V_{Bopt}=6nU_T$		1.46	0.22	fJ
P_{dyn}/f at V_B=1V		20	4	fJ
f_{max1} for α=1 and $V_B=4U_T$		50	500	MHz
f_{max2} for α=0.01 and $V_B=V_{Bopt}$		0.22	2.56	MHz
P_{min} at f_{max2}		32.5	56.3	nW

PRACTICAL CONSIDERATIONS AND LIMITATIONS

- Low-voltage power source
  - should be proportional to U_T (PTAT)
  - need for power-efficient adapter from higher supply voltage.
- Asymmetry
  - p/n asymmetry may result in speed reduction.
- Mismatch
  - dominated by threshold mismatch δV_T
  - may result in speed reduction proportional to $\delta V_T / V_B$.
- Short channel effects: should not drastically degrade the results.
- Gate leakage current : should be alleviated by very low V_B.
- Adjustment of I_0 or T_d to required value
  - control by V_S with charge pump in loop **[18]**; n>1 needed (no SOI!)
  - corresponds to threshold adjustment unavoidable at very low V_B.
- System architectures and applications.

SYSTEM ARCHITECTURE AND APPLICATIONS

- Duty factor α **must be maximized** to reach minimum P/f,
  (where f is the average transition frequency), thus
  - avoid idling gates (contrary to traditional CMOS culture)
  - new architectures needed:
    - **maximally active** gates of **minimum speed** (max. delay time T_d)
    - particular problem with RAMs (short T_d but sparse activity)
    - how? new constraints should result in novel solutions.
  - partition the system in blocks of comparable α and T_d
    - **optimum V_B and I_0** for each block (separate I_0 control).
- Maximum frequency much lower than for strong inversion:
  - best applicable when no high local speed is required
  - m-parallelize: mT_d but same power if same α (m units with P/m)
    - digital image processing ?

CONCLUSION

- Weak inversion permits very low supply voltage V_B
  - approached with scaled-down V_B: $IC \sim V_B^2$
  - limit for scaled-down V_B.
- Analog:
  - $V_B > 10U_T = 250$ mV
  - provides maximum g_m/I_D
  - bipolar-like behaviour can be exploited in new schemes.
- Digital:
  - $V_B > 4U_T = 100$mV
  - transistor not a switch but a current modulator (I_{on}/I_{off})
  - new architectural approaches for max. duty factor α.
  - ultimum (asymptotic) limit for low power$*$delay.
- Low speed, but keeps increasing with $1/L^2$ in scaled down processes.

REFERENCES

[1] E.Vittoz and J.Fellrath, "CMOS analog integrated circuits based on weak inversion operation", IEEE J.Solid-State Circuits, vol.SC-12, pp.224-231, June 1977.

[2] E.Vittoz, "Micropower techniques", in *Design of VLSI Circuits for Telecommunications and Signal Processing*, J.E.Franca and Y.P.Tsividis Editors, Prentice Hall, 1991

[3]. C.Enz, F.Krummenacher and E.Vittoz, "An analytical MOS transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications", Analog Integrated Circuits and Signal Processing, Vol.8, pp.83-114, 1995.

[4] R.M Swanson and J.D.Meindl,"Ion-implanted complementary MOS transistors in low-voltage circuits", IEEE J.Solid-State Circuits, vol.SC-7, pp.146-153, April 1972.

[5] E. Vittoz, "Weak inversion for ultimate low-power logic", to be published in *Low-Power Electronic Design*, ed. C. Piguet, CRC Press LLC (2003?), Chapter 16.

[6] E. Vittoz, C. Enz and F. Krummenacher, "A basic property pf MOS transistors and its circuit implications", Workshop on Compact Modeling, WCM MSM.2003, Febr. 23-27, San F.rancisco, pp. 246-249. Slide of presentation can be downloaded at www.nanotech2003.com/WCM2003.html#Slides.

[7] M.A. Maher and C. Mead, "A physical charge-controlled model for the MOS transistors", Advanced research in VLSI, Proc. of the 1987 Stanford Conference, MIT Press, Cambridge MA, 1987.

[8] A. Cunha *et al*., "An MOS transistor model for analog circuit design", IEEE J.Solid-State Circuits, vol.33, pp.1510-1519, Oct. 1998.

[9] H. Oguey and S. Cserveny, "MOS modelling at low current density", Summer Course on "Process and Device Modelling", ESAT Leuven-Heverlee, Belgium, June 1983.

[10] H.J.Oguey and D.Aebischer. "CMOS current without resistance."IEEE Journal of Solid-State Circuits, vol 32, pp.1132-1135 July 1997.

[11] K.Bult and G.Geelen, "A inherently linear and compact MOST-only current division technique", Dig. ISSCC Tech. Papers, February 1992, pp.198-199.

[12] E.Vittoz and X.Arreguit,"Linear networks based on transistors", Electronics Letters, vol.29, pp.297-299, 4th Febr. 1993.

[13] E.Vittoz, ÒPseudo-resistive networks and their applications to analog collective computationÓ, Proc. MicroNeuroÕ97, Dresden, pp.163-173.

[14] T. DelbrŸck, "Bump circuit for computing similarity and dissimilarity of analog voltages", Proc. of International Joint Cœnf. on Neural Networks, vol.1, pp. I - 475-479. 1991.

[15] B. Gilbert, "Translinear circuits: a proposed classification", Electron. Letters, vol.11, p.14, 1975.

[16] A. Andreou and K. Boahen, "Neural information processing II" in *Analog VLSI Signal and Information Processing*, M. Ismail and T. Fiez, editors, pp.358-409, McGraw-Hill, 1994.

[17] E.Vittoz, "Analog VLSI implementation of neural networks", published in the *Handbook of Neural Computation*, Institute of Physics Publishing and Oxford University Press, USA, 1996.

[18] V. von Kaenel *et al*. "Automatic adjustment of threshold and supply voltage for minimum power consumption in CMOS digital circuits", Proc. IEEE Symposium on Low Power Electronics, San Diego, 1994, pp.78-79.