Chapter 2 MOS Transistor Theory

- 2.1 Introduction
  - Examine the characteristics of MOS transistors to lay the foundation for predicting the performance of the MOS circuits.
  - Circuit symbol for MOS transistors

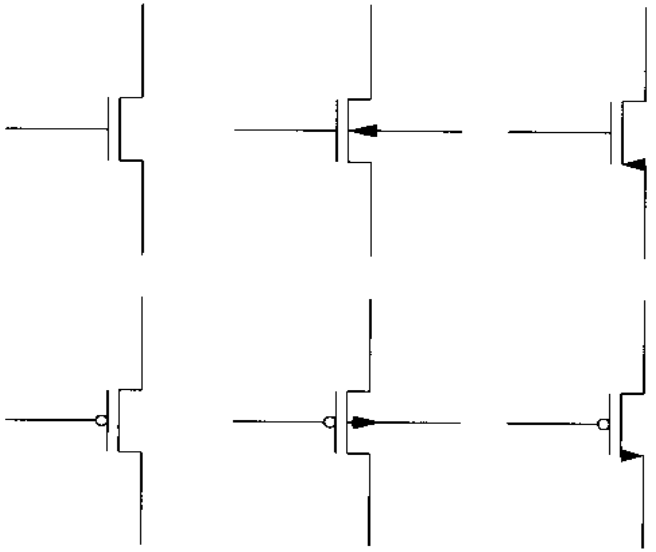


FIGURE 2.1 MOS transistor symbols

  - A MOS transistor is called a majority carrier device, in which the current in a conducting channel (the region immediately under the gate) between the source and the drain is modulated by a voltage applied to the gate.

– The majority carriers of an nMOS transistor: Electrons.

– The majority carriers of a pMOS transistor: Holes.

– Symbol Definitions

  - V_t: the threshold voltage of an nMOS or a pMOS transistor.
  - V_{tn}: the threshold voltage of an nMOS transistor.
  - V_{tp}: the threshold voltage of a pMOS transistor.
  - V_{ds}: the voltage difference between the drain and the source for an nMOS or a pMOS transistor.
  - V_{dsn}: the voltage difference between the drain and the source for an nMOS transistor.
  - V_{dsp}: the voltage difference between the drain and the source for a pMOS transistor.
  - V_{gs}: the voltage difference between the gate and the source for an nMOS or a pMOS transistor.
  - V_{gsn}: the voltage difference between the gate and the source for an nMOS transistor.
  - V_{gsp}: the voltage difference between the gate and the source for a pMOS transistor.
  - I_{ds}: the current between the drain and the source for an nMOS or a pMOS transistor.
  - I_{dsn}: the current between the drain and the source for an nMOS transistor.
  - I_{dsp}: the current between the drain and the source for a pMOS transistor.
  - V_{in}: the input voltage.
  - V_{inp}: the input voltage for a pMOS transistor.
  - V_{inn}: the input voltage for an nMOS transistor.
  - V_{out}: the output voltage.
  - V_{dd}: power supply.
  - V_{ss}: ground.

– **Four modes of transistors**

– Enhancement mode nMOS transistor:

  • $V_{tn} > 0$
  • If $V_{gs} > V_{tn}$, the transistor starts to conduct. The number of electrons in the channel increases so that I_{dsn} increases accordingly. If $V_{gs} < V_{tn}$, the transistor is cut off and I_{ds} is almost zero.

– Depletion mode nMOS transistor:

  • $V_{tn} < 0$ (in the textbook it is referred as $-V_{tn}$ and $V_{tn} > 0$)
  • Even if $V_{gs} = 0 > V_{tn}$, the transistor is “on”.
  • If $V_{gs} < V_{tn} < 0$, the transistor is cut off.

– Enhancement mode pMOS transistor:

  • $V_{tp} < 0$ (in the textbook it is referred as $-V_{tp}$ and $V_{tp} > 0$)
  • If $V_{gs} < V_{tp} < 0$, the transistor starts to conduct. The number of holes in the channel increases so that I_{dsp} increases accordingly.
  • If $V_{gs} > V_{tp}$, the transistor is cut off.

– Depletion mode pMOS transistor:

  • $V_{tp} > 0$
  • Even if $V_{gs} = 0 < V_{tp}$, the transistor is “on”.
  • If $V_{gs} > V_{tp} > 0$, the transistor is cut off.

– Conduction characteristics of MOS transistors

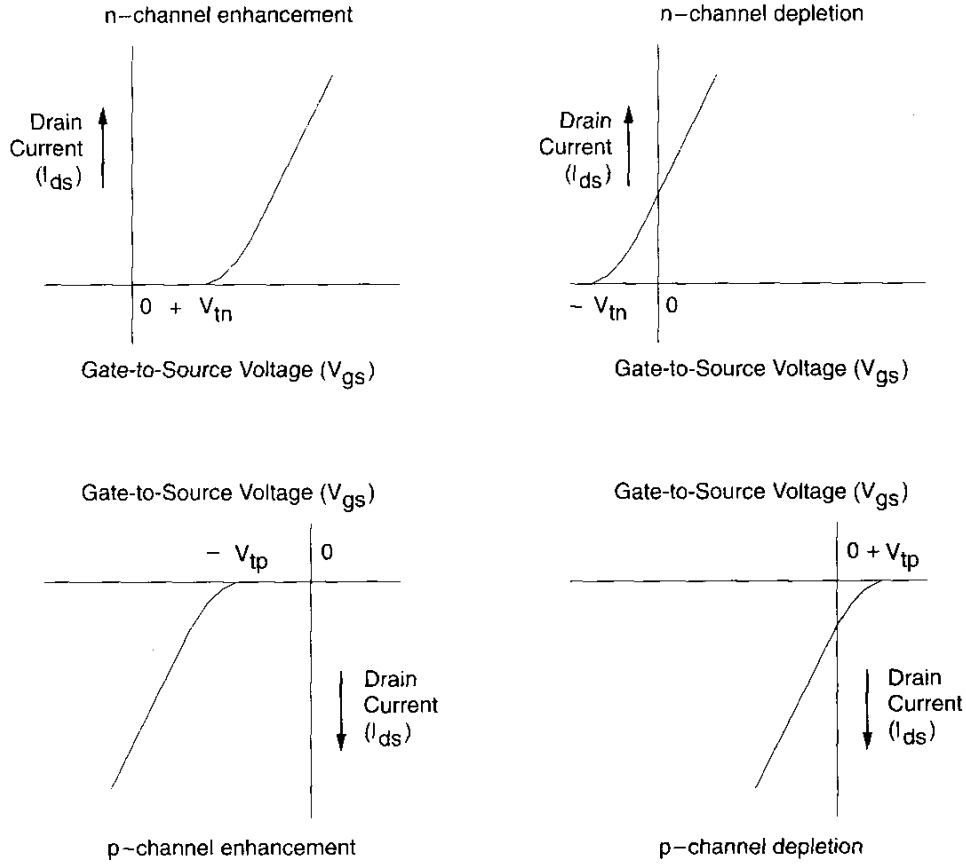


FIGURE 2.2 Conduction characteristics for enhancement and depletion mode MOS transistors (assuming fixed V_{ds})

– *Note that the minus sign attached to V_{tp} and V_{tn} in Figure 2.2 should be deleted.*

– The n-channel transistors and p-channel transistors are the duals of each other; that is , the voltage polarities required for correct operation are the opposite.

– Most CMOS integrated circuits at present use enhancement mode transistors.

- 2.1.1 nMOS Enhancement Transistor
  - Structure of an n-channel enhancement-type transistor. (Figure 2.3)
    - Moderately doped p-type silicon substrate.
    - Heavily doped n^+ regions (the source and the drain).
    - A channel sandwiched between a thin insulating layer of silicon dioxide (SiO_2) and p-substrate.
    - A polysilicon gate over the gate oxide (the SiO_2 over the channel).
    - No DC current flows from gate to channel due to insulating of gate oxide.

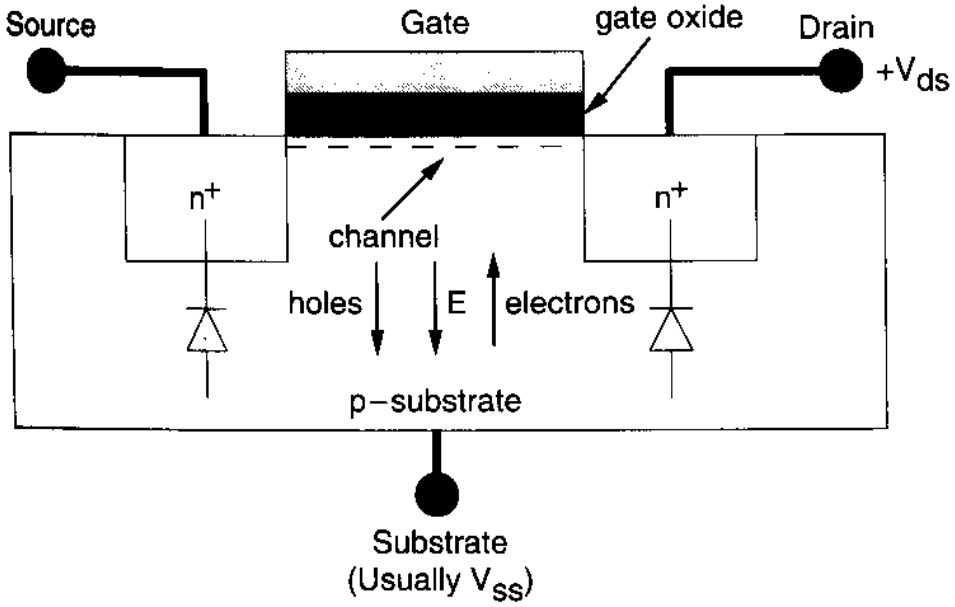


FIGURE 2.3 Physical structure of an nMOS transistor

– Operation of nMOS transistor

- With zero gate bias, i.e. $V_{gs} = 0$, $I_{ds} = 0$ because the source and the drain are effectively insulated from each other by the two reversed-bias pn junctions (indicated as the diode symbol in Figure 2.3).
- Accumulation mode: With positive gate bias with respect to the source and substrate (generally denoted by $V_{gs} > 0$), an electric field E across the substrate is established such that electrons are attracted to the gate and holes are repelled from the gate.(See Figure 2.4 (a))
- Depletion mode: If $V_g \cong V_{tn}$, a depletion channel under the gate free of charges is established.(See Figure 2.4 (b))
- Inversion mode: If $V_{gs} > V_{tn}$, an inversion channel (region) consisting of electrons is established just under the gate oxide and a depletion channel (region) is also established just under the inversion region.(See Figure 2.4 (c))
- Hence the term “n-channel” is applied to the nMOS structure.

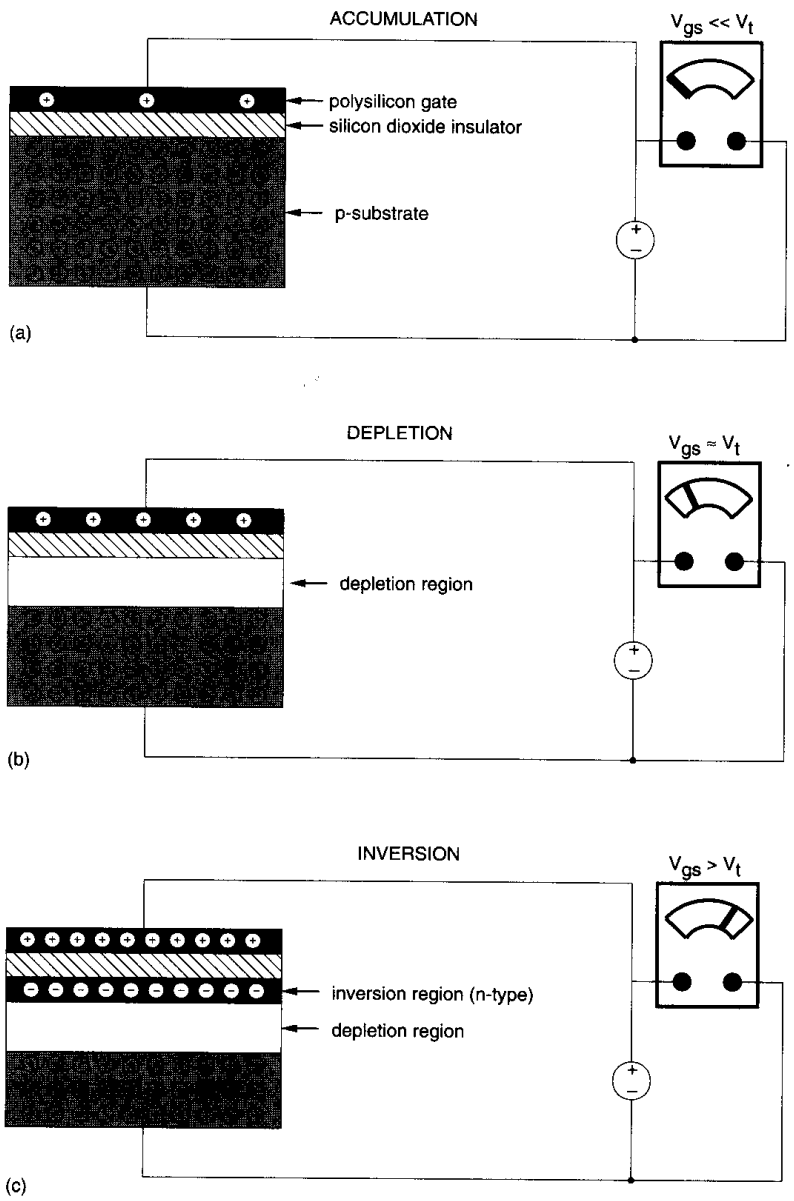
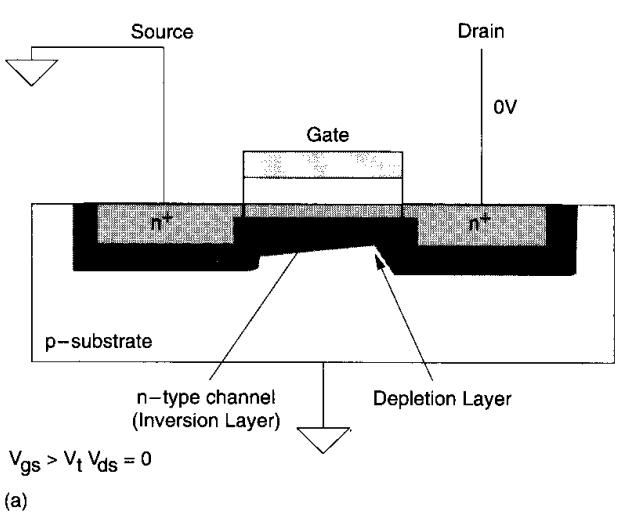


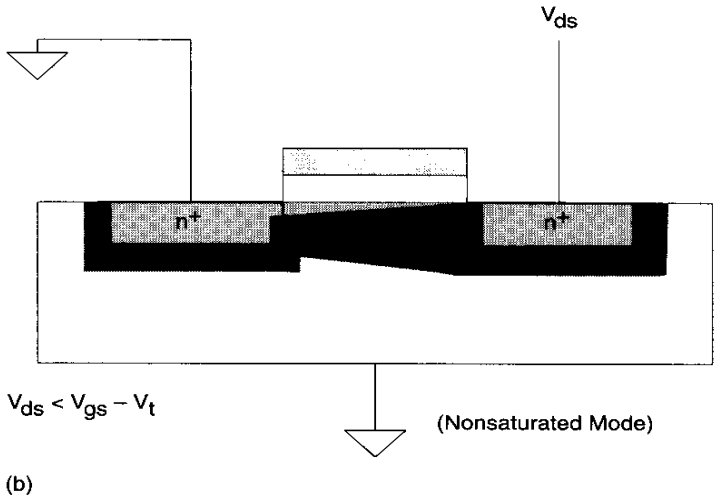
FIGURE 2.4 Accumulation, Depletion and Inversion modes in an MOS structure.

– Electrically, a MOS device acts as a voltage-controlled switch.

- It conducts initially when $V_{gs} = V_t$.
- V_{gs} establishes a conducting channel, while V_{ds} is responsible for sweeping the electrons from the source to the drain. Thus, establish a current flow between the drain and the source.
- The electric field established by V_{gs} is orthogonal to the electric field established by V_{ds}.
- When $V_{gs} \geq V_t$ and $V_{ds} = 0$, the width of the n-type channel at the source end is equal to that at the drain end. This is due to $V_{gs} = V_{gd}$ (See Figure 2.5 (a)).
- Nonsaturated (resistive or linear) mode: when $V_{gs} - V_t > V_{ds} > 0$, the width of the n-type channel at the source end is larger than that at the drain end. This is due to $V_{gs} \geq V_{gd} > V_t$. (See Figure 2.5 (b))
- Saturated mode: When $V_{ds} > V_{gs} - V_t > 0$, the n-type channel no longer reaches the drain. That is, the channel is pinched off. This is due to $V_{gs} > V_t$ and $V_{gd} < V_t$. (See Figure 2.5 (c))

- In non-saturated mode, I_{ds} is a function of gate and drain voltage, while in saturated mode, I_{ds} is a function of gate voltage.
- In saturated mode, the movement of electrons in the channel is brought about under the influence of positive drain voltage. After the electrons leave the channel and inject into the drain depletion region, they are accelerated toward the drain. Because the voltage across the pinched-off channel tends to remain fixed at $V_{gs} - V_t$, the drifting speed of electrons in the channel is controlled by $V_{gs} - V_t$, but almost independent of V_{ds}. This is what “saturation” means.





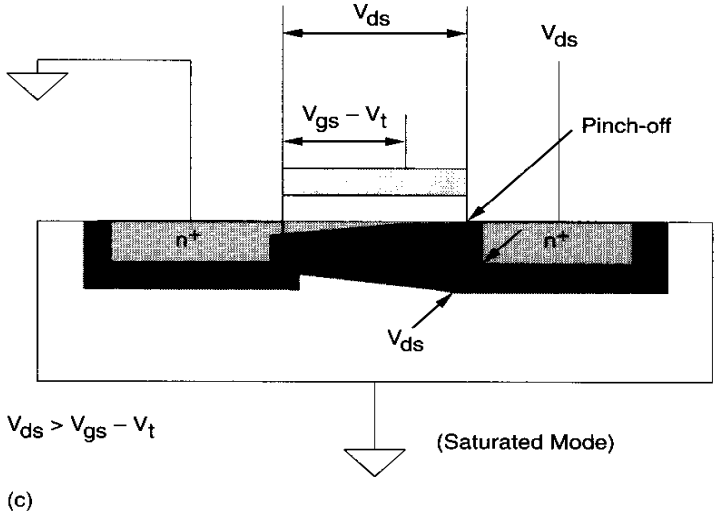


FIGURE 2.5 nMOS device behavior under the influence of different terminal voltages

– Normal conduction characteristics of a MOS transistor:

  • "Cut-off" region: $I_{ds} \cong 0$
  • "Non-saturated" region: The channel is weakly inverted. I_{ds} is dependent on the gate and drain voltage with respect to the substrate.
  • "Saturated" region: The channel is strongly inverted. I_{ds} is ideally independent of V_{ds}.

– For a fixed V_{ds} and V_{gs}, the factors that influence I_{ds}:

  • The distance between source and drain.
  • The channel width
  • V_t
  • The thickness of gate oxide.
  • The dielectric constant of the gate oxide.
  • The carrier mobility £ g

- 2.1.2 pMOS Enhancement Transistor
  - A reversal of n-type and p-type regions yields a p-channel channel transistor. (See Figure 2.6)

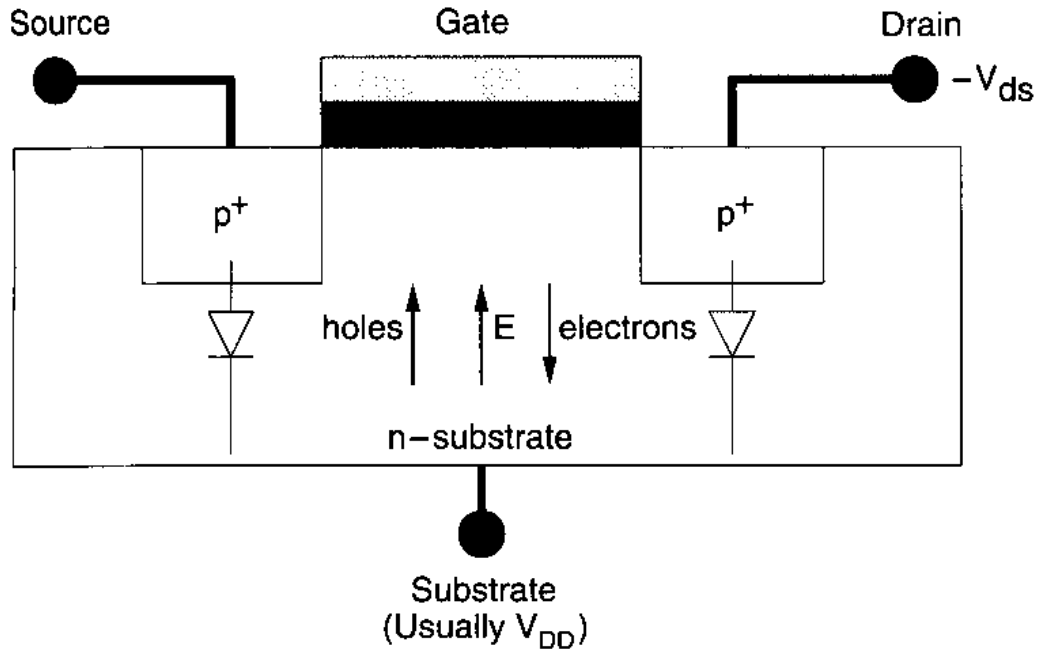


FIGURE 2.6 Physical structure of a pMOS transistor

- 2.1.3 Threshold Voltage
  - V_t: the threshold voltage for a MOS transistor can be defined as the voltage applied between the gate and the source of a MOS device below which the drain-to-source current I_{ds} "effectively" drops to zero.
  - In general, V_t is a function of the following parameters.
    - Gate conduction material
    - Gate insulation material
    - Gate insulator thickness
    - Channel doping
    - Impurities at the silicon-insulator interface
    - Voltage between the source and the substrate, V_{sb}.
  - Threshold voltage equations

$$V_t = V_{t\text{-}mos} + V_{fb} \quad (2.1)$$

$$V_{t\text{-}mos} = 2\phi_b + \frac{Q_b}{C_{ox}} \quad (2.2)$$

$$\phi_b = \frac{kT}{q}\ln\left(\frac{N_A}{N_i}\right)$$

$$Q_b = \sqrt{2\varepsilon_{Si} q N_A 2\phi_b}$$

$$V_{fb} = \phi_{ms} - \frac{Q_{fc}}{C_{ox}} \quad (2.3)$$

- $V_{t\text{-}mos}$ is the ideal threshold voltage for an ideal MOS capacitor
- V_{fb} is the flat-band voltage
- k: Boltzmann's constant = 1.38 * 10^{-23} J/^{o}K.
- q: electronic charge = 1.602 * 10^{-19} Coulomb.
- T: Temperature (^{o}K).
- N_A: the density of carriers in the doped substrate.
- N_i: the density of carriers in the undoped substrate.
- £ `si: the permittivity of silicon = 1.06 * 10^{-12} (F/cm)
- C_{ox}: the gate-oxide capacitance, which is inversely proportional to the gate oxide thickness (t_{ox}).
- Q_{fc}: the fixed charge due to surface states that arise due to imperfections in the silicon-oxide interface and doping.
- ϕ_{ms}: the work function difference between the gate material and the silicon substrate.

– More details can be found in the text book by Weste.

– Two common techniques for the adjustment of V_t.

- Affecting Q_{fc} by varying the doping concentration at the silicon-insulator interface through ion implantation.
- Affecting C_{ox} by using different insulating material for the gate. A layer of silicon nitride (Si_3N_4) combined with a layer of silicon oxide can effectively increase the relative permittivity of gate insulator from 3.9 to 6.

- 2.1.4 Body Effect
  - When connecting several devices in series as shown in Figure 2.7, the source-to-substrate of each individual devices may be different. For example, $V_{sb2} > V_{sb1} = 0$.
  - As V_{sb} (V_{source} - $V_{substrate}$) is increased, the density of the trapped carriers in the depletion layer also increases. The overall effect is an increase in the threshold voltage, V_t ($V_{t2} > V_{t1}$).

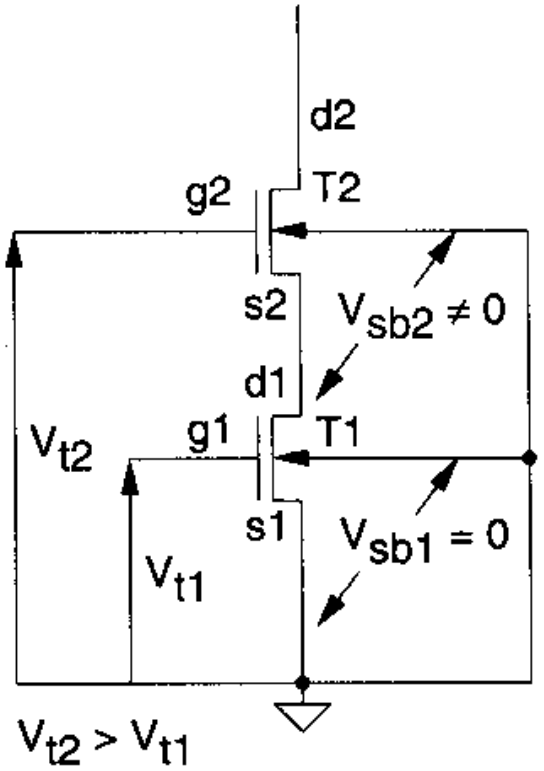


FIGURE 2.7
The effect of substrate bias on series-connected n-transistors

- 2.2 MOS Device Design Equations
- 2.2.1 Basic DC Equations
  - Cutoff region: $I_{ds} = 0$, $V_{gs} \le V_t$ (2.5a)
  - Nonsaturation, linear or resistive region:
    - $I_{ds} = \beta\left[(V_{gs} - V_t)V_{ds} - \frac{V_{ds}^2}{2}\right]$ $0 < V_{ds} < V_{gs} - V_t$ (2.5b)
  - Saturation region:
    - $I_{ds} = \beta\frac{(V_{gs} - V_t)}{2}, 0 < V_{gs} - V_t < V_{ds}$ (2.5c)
  - $\beta = \frac{\mu\varepsilon}{t_{ox}}(\frac{W}{L})$ is the MOS transistor gain factor, where
    - μ the effective surface mobility of the carriers in the channel.
    - ε : the permittivity of the gate insulator.
    - t_{ox}: the thickness of the gate insulator.
    - L: the length of the channel
    - W: the width of the channel
    - $\frac{\mu\varepsilon}{t_{ox}} = \mu C_{ox}$ is a process dependent term.
    - W/L is a geometry dependent term. (See Figure 2.8)

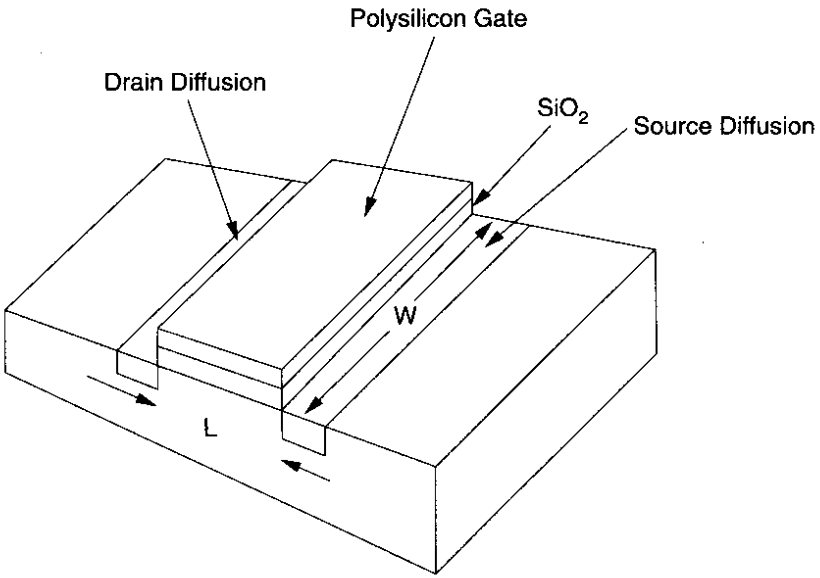


FIGURE 2.8 Geometric terms in the MOS device equation

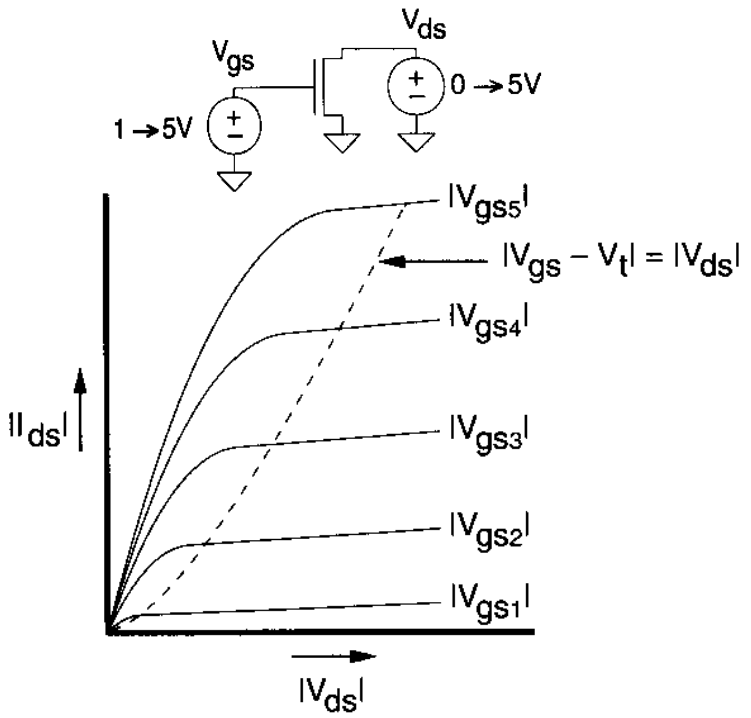


FIGURE 2.9 *VI* characteristics for n- and p-transistors

- Figure 2.9 shows the voltage-current characteristics of pMOS and nMOS transistor.
- The boundary between the linear and saturation regions is defined by the condition $|V_{ds}| = |V_{gs} - V_t|$.

Example

Typical values (for an n-device) for current (~1μ) processes are as follows:

$\mu_n = 500\ cm^2 / V\text{-}sec$

$\varepsilon = 3.9\varepsilon_0 = 3.9 \times 8.85 \times 10^{-14}\ F/cm$ (permittivity of silicon dioxide, S_iO_2)

$t_{ox} = 200$ Å

Hence a typical n-device β would be

$$\frac{500 \times 3.9 \times 8.85 \times 10^{-14}}{.2 \times 10^{-5}} \frac{W}{L} = 88.5 \frac{W}{L} \mu A / V^2$$

On the other hand, p-devices have hole mobilities (μ_p) of about 180 $cm^2/V\text{-}sec$, yielding a β of

$$= 31.9 \frac{W}{L} \mu A / V^2$$

Thus the ratio of n-to-p gain factors in this example is about 2.8. This ratio varies from about 2 to 3 depending on the process.

- 2.2.2 Second Order Effects
  - The circuit simulation program such as SPICE generally use a parameter called LEVEL to specify which model equations are used for circuit simulation.
    - LEVEL 1: Use the EQ. (2.5) and include some important second order effects.
    - LEVEL 2: Calculate the currents based on device physics.
    - LEVEL 3: Is a semiempirical approach that relies on parameters selected on the basis of matching the equations to real circuits.
  - Process gain factor:
    - $KP = \frac{me}{t_{ox}} = mC_{ox}$: the process gain factor may vary from 10 ~ 100 £gA/V^2, depending on the characteristics of a process and the type of transistor.
  - Threshold voltage change due to body effect: the increase in threshold voltage leads to lower device currents, which in turn leads to slower circuits.
    - $V_t = V_{t0} + \tilde{a}\left[\sqrt{(2f_b + |V_{sb}|} - \sqrt{2f_b}\right]$.
    - V_{t0}: the threshold voltage for V_{sb} = 0 (SMP ¡ ¥VTO)
    - SMP: SPICE Model Parameter.
    - V_{sb}: the substrate bias, i.e., the voltage difference between source and substrate.
    - ϕ_b : defined in Eq. 2.2. (ϕ_s=2ϕ_b, SMP of ϕ_s is PHI)
    - £^the constant describing the substrate bias effect: (SMP ¡ ¢GAMMA)
    - $\tilde{a} = \frac{t_{ox}}{å_{ox}}\sqrt{2qå_{si}N_A} = \frac{1}{C_{ox}}\sqrt{2qå_{si}N_A}$.

- £ `$_{ox}$: the dielectric constant of SiO_2.
- £ `$_{si}$: the dielectric constant of the silicon substrate.
- N_A: doping density of the substrate. (SMP ¡ ₦SUB)
- **Example**

For with $N_A = 3 \times 10^{16}$ cm^{-3}, $t_{ox} = 200$Å, $\varepsilon_{ox} = 3.9 \times 8.85 \times 10^{-14}$ *F/cm,* $\varepsilon_{Si} = 11.7 \times 8.85 \times 10^{-14}$ *F/cm,* and $q = 1.6 \times 10^{-19}$ Coulomb

$$\gamma = \frac{0.2 \times 10^{-5}}{3.9 \times 8.85 \times 10^{-14}} \sqrt{2 \times 1.6 \times 10^{-19} \times 11.7 \times 8.85 \times 10^{-14} \times 3 \times 10^{16}}$$

$$= .57$$

$$\phi_b = .02586 \; ln \left(\frac{3 \times 10^{16}}{1.5 \times 10^{10}} \right)$$

$$= .375$$

At a V_{sb} of 2.5 volts, and with

$$V_{t2.5} = V_{t0} + .57 \left[\sqrt{.75 + 2.5} - \sqrt{.75} \right]$$

$$= V_{t0} + .53$$

– Subthreshold Region (Cutoff region): Although I_{ds} is very small ($I_{ds} \cong 0$) in this region, it increase exponentially with V_{ds} and V_{gs}. The finite value of I_{ds} may be used to construct very low power circuits. LEVEL 1 SPICE models set the subthreshold current to 0.

– Channel-length Modulation: The variation of channel length is due to the changes in drain-to-source voltage, V_{ds}.

  • When an MOS device is in saturation, the effective channel length (channel pinched-off) is actually decreased such that $L_{eff} = L - L_{short}$, where $L_{short} = \sqrt{2\frac{\varepsilon_{si}}{qN_A}(V_{ds} - (V_{gs} - V_t))}$

  • Reduction in channel length increases the (W/L) ratio, there by increasing I_{ds} as the drain voltage increases.

  • Take this behavior into account, $I_{ds} = \frac{k}{2}\frac{W}{L}(V_{gs} - V_t)^2(1 + \lambda V_{ds})$ (2,10)

    – $k = \frac{\mu\varepsilon}{t_{ox}}$: process gain factor.

    – λ: empirical channel-modulation factor. (SMP 參數 LAMBDA)

– Mobility variation: Mobility decreases with increasing doping-concentration and increasing temperature.

  • Mobility (SMP→ 參數 UO) $= \frac{\text{average carrier drift velocity (V)}}{\text{Electric Field (E)}}$

– Fower-Nordheim Tunneling: When the gate oxide is very thin, a current can flow from gate to source or drain by electron tunneling through gate oxide.
  - $I_{FN} = C_1 W L E_{ox}^2 e^{\frac{-E_0}{E_{ox}}}$
  - $E_{ox} \cong \frac{V_{gs}}{t_{ox}}$ is the electric field across the gate oxide.
  - E_0 and C_1 are constants
  - This effect limits the gate oxide thickness. However, it is of great use in electrically alterable programmable logic devices.

– Drain Punchthrough: When the drain is at a high enough voltage with respect to the source, the depletion region around the drain may extend to the source, thus causing current to flow irrespective of the gate voltage.
  - Can be used for I/O protection
  - Limit the power supply voltage when process is scaling.

– Impact Ionization - Hot Electrons: As the gate length of an MOS transistor is reduced, the electric field at the drain of a transistor in saturation increases. The field can become so high that electrons have enough energy to become what is termed “hot”. These hot electrons impact the drain, dislodging holes that are then swept toward the negatively charged substrate and appear as a substrate current. Moreover the elections can penetrate the gate oxide, causing a gate current. This cause reliability problem.

As an illustration of the relative magnitude of the substrate current, the following equation is representative[18] (for an $L = 0.8\ \mu$, $t_{ox} = 160$Å CMOS process):

$$I_{substrate} = I_{ds} C1\ \ (V_{ds} - V_{dsat})^{C2} \tag{2.13}$$

where

$$C1 = 2.24 \times 10^{-5} - .1 \times 10^{-5} V_{ds}$$

$$C2 = 6.4$$

$$V_{dsat} = \frac{V_{tm} L_{eff} E_{sat}}{V_{tm} + L_{eff} E_{sat}}$$

with

$$V_{tm} = V_{gs} - V_{tn} - 0.13 V_{bs} - 0.25 V_{gs}$$

$$E_{sat} = 1.10 \times 10^{7} + 0.25 \times 10^{7} V_{gs}$$

L_{eff} is the effective channel length in meters.

- 2.2.3 MOS Models
  - Most CMOS digital foundry operations have been standardized on the LEVEL 3 models in SPICE as the level of circuit modeling that is required for CMOS digital system design.
  - Table 2.1 is a summary of the main SPICE DC parameters that are used in LEVELS 1, 2, and 3 with representative values for a 1μ n-well CMOS process.

Table 2.1 SPICE DC Parameters

Parameter	nMOS	pMOS	Units	Description
VTO	0.7	0.7	volt	Threshold voltage
KP	8×10^{-5}	2.5×10^{-5}	A/V^2	Transconductance coefficient
GAMMA	.4	.5	$V^{0.5}$	Bulk threshold parameter
PHI	.37	.36	volt	Surface potential at strong inversion
LAMBDA	.01	.01	$volt^{-1}$	Channel length modulation parameter
LD	0.1×10^{-6}	0.1×10^{-6}	meter	Lateral diffusion
TOX	2×10^{-8}	2×10^{-8}	meter	Oxide thickness
NSUB	2×10^{16}	4×10^{16}	$1/cm^3$	Substrate doping density

- 2.3 The Complementary CMOS Inverter-DC Characteristics
  - A CMOS inverter is realized by the series connection of a p-device and an n-device as show in Figure 2.11.
  - DC-transfer characteristics: the output voltage V_{out} is drawn as a function of the input voltage V_{in}.

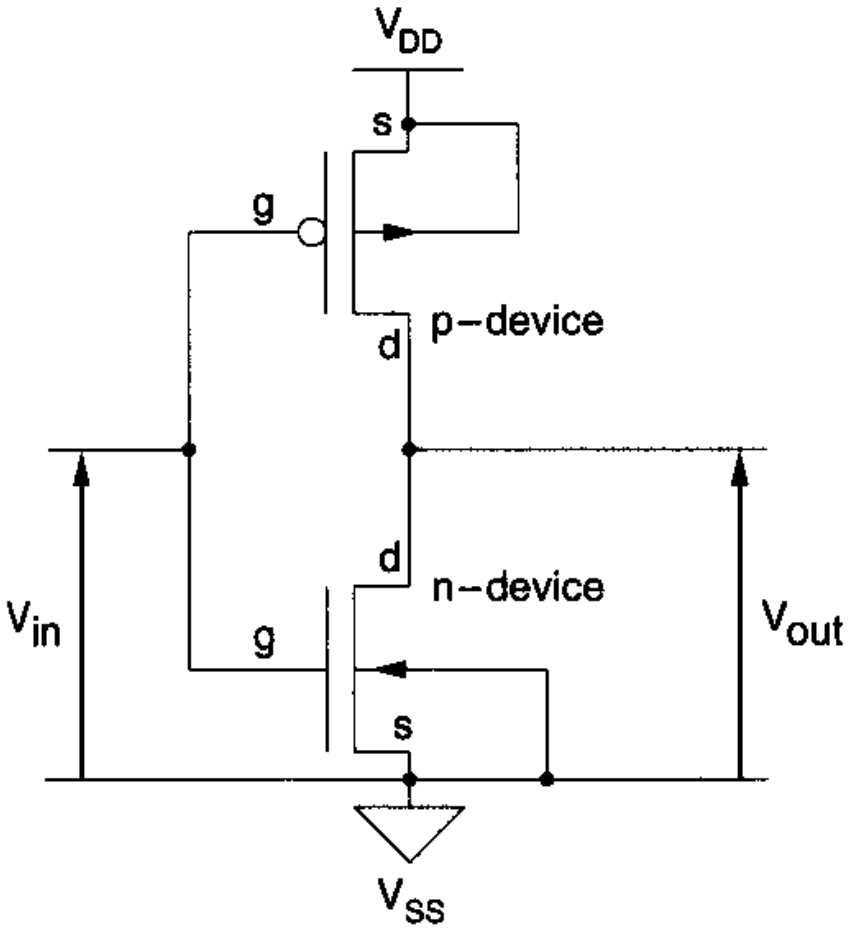


FIGURE 2.11 A CMOS inverter (with substrate connections)

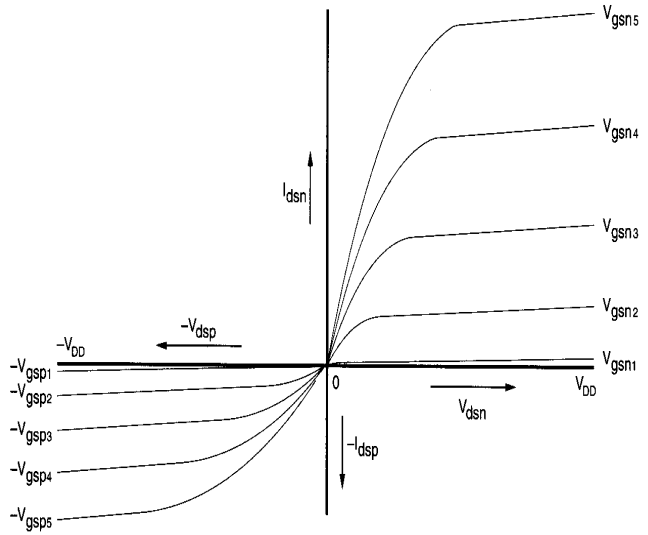
– Table 2.2 outlines the various regions of operation for the n- and p-transistors of an inverter.

TABLE 2.2 Relations Between Voltages for the Three Regions of Operation of a CMOS Inverter

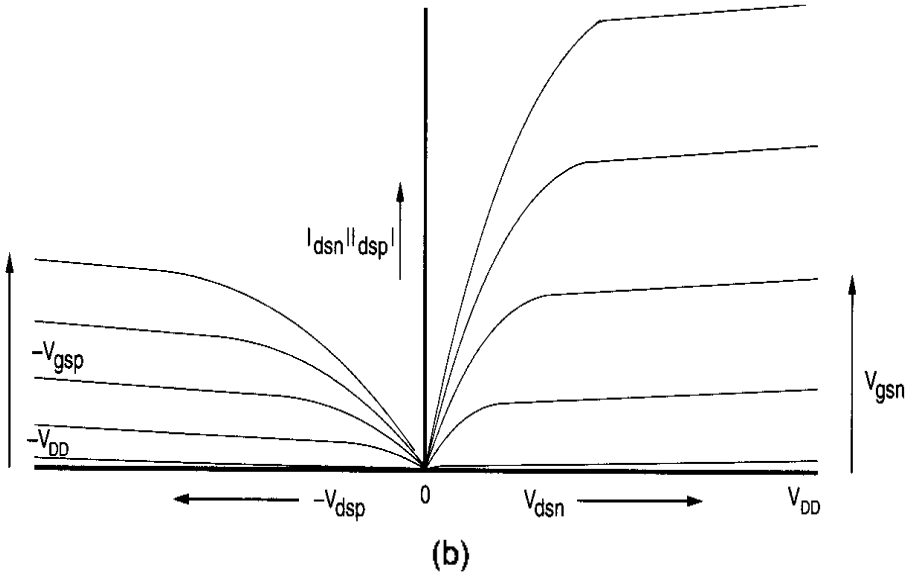
	CUTOFF	NONSATURATED	SATURATED
p-device	$V_{gsp} > V_{tp}$ $V_{in} > V_{tp} + V_{DD}$	$V_{gsp} < V_{tp}$ $V_{in} < V_{tp} + V_{DD}$ $V_{dsp} > V_{gsp} - V_{tp}$ $V_{out} > V_{in} - V_{tp}$	$V_{gsp} < V_{tp}$ $V_{in} < V_{tp} + V_{DD}$ $V_{dsp} < V_{gsp} - V_{tp}$ $V_{out} < V_{in} - V_{tp}$
n-device	$V_{gsn} < V_{tn}$ $V_{in} < V_{tn}$	$V_{gsn} > V_{tn}$ $V_{in} > V_{tn}$ $V_{dsn} < V_{gs} - V_{tn}$ $V_{out} < V_{in} - V_{tn}$	$V_{gsn} > V_{tn}$ $V_{in} > V_{tn}$ $V_{dsn} > V_{gs} - V_{tn}$ $V_{out} > V_{in} - V_{tn}$

– Find the DC-transfer characteristics of an inverter.

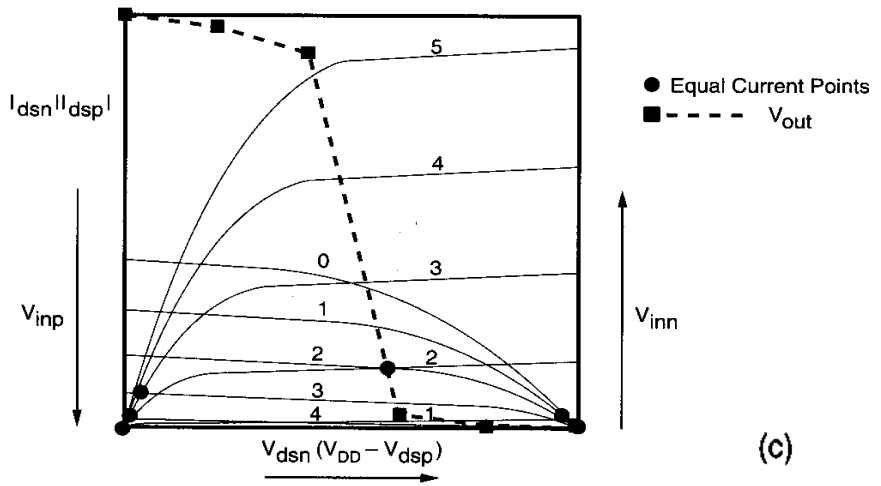
- Step1: obtain *VI* characteristics for p- and n-transistor respectively based on the equation (2.5). (see Figure 2.12(a))
- Step2: reflect the *VI* characteristics for p-transistor about the x-axis (see Figure 2.12(b))
- Step3: the input/output transfer curve may now be determined by the points of common V_{gs} intersection in Figure 2.12(c).



(a)



(b)



(c)

FIGURE 2.12 Graphical derivation of CMOS inverter characteristic

– Solving for $V_{inn}=V_{inp}$ and $I_{dsn}=I_{dsp}$ gives the desired transfer characteristics of the inverter. (see figure 2.13)

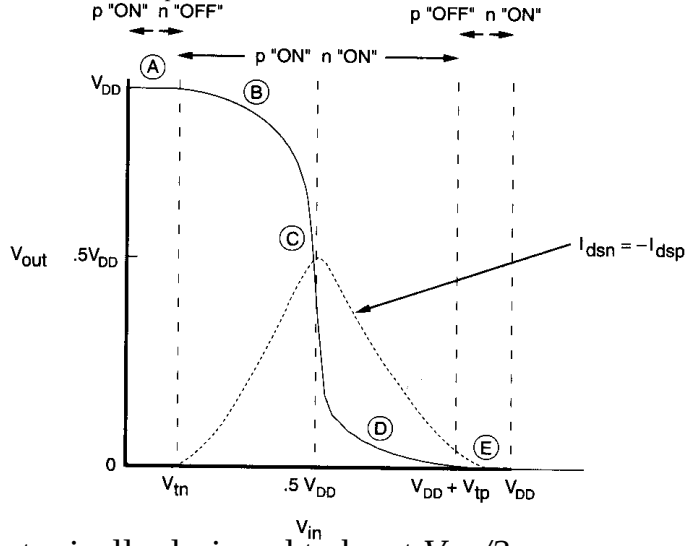


FIGURE 2.13 CMOS inverter DC transfer characteristic and operating regions

– The switching point is typically designed to be at $V_{DD}/2$.

– Operation regions of an inverter (see Figure 2.13 and Table 2.3)

TABLE 2.3 Summary of CMOS Inverter Operation

REGION	CONDITION	p-device	n-device	OUTPUT
A	$0 \le V_{in} < V_{tn}$	nonsaturated	cutoff	$V_{out} = V_{DD}$
B	$V_{tn} \le V_{in} < \frac{V_{DD}}{2}$	nonsaturated	saturated	Eq. (2.23)
C	$V_{in} = \frac{V_{DD}}{2}$	saturated	saturated	$V_{out} \neq f(V_{in})$
D	$\frac{V_{DD}}{2} < V_{in} \le V_{DD} - \lvert V_{tp} \rvert$	saturated	nonsaturated	Eq. (2.27)
E	$V_{in} > V_{DD} - \lvert V_{tp} \rvert$	cutoff	nonsaturated	$V_{out} = V_{SS}$

– **Region A:** defined by $0 \le V_{in} \le V_{tn}$, where the n-device is cutoff and the p-device is in the linear region.
  - $V_{out}=V_{DD}$ because $I_{dsn}=-I_{dsp}=0 \Rightarrow V_{dsp}=V_{out}-V_{DD}=0$ Þ $V_{out}=V_{DD}$
– **Region B:** defined by $V_{tn} < V_{in} < \frac{V_{DD}}{2}$, where the p-device is in the nonsaturated region while the n-device is in saturation.
  - The equivalent circuit in this region can be represented by a resistor for the p-transistor and a current source for the n-transistor as shown in Figure 2.14(a).
  - $I_{dsn} = \beta_n \frac{(V_{in}-V_{tn})^2}{2}$ By setting $V_{gs}=V_{in}$, where $\beta_n = \frac{\mu_n \varepsilon}{t_{ox}}(\frac{W_n}{L_n})$
  - $I_{dsp} = -\beta_p[(V_{in}-V_{DD}-V_{tp})(V_{out}-V_{DD}) - (\frac{(V_{out}-V_{DD})^2}{2})]$ by setting that $V_{gs}=V_{in}-V_{DD}$ and $V_{ds}=V_{out}-V_{DD}$, where $\beta_p = \frac{\mu_p \varepsilon}{t_{ox}}(\frac{W_p}{L_p})$

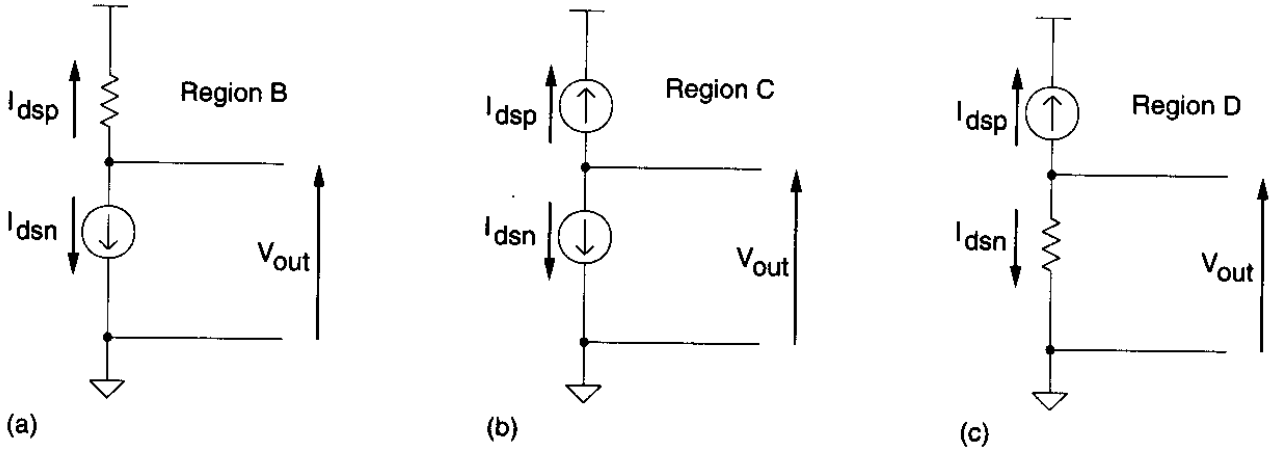


FIGURE 2.14 Equivalent circuits for operating regions of a CMOS inverter

- Let I_{dsp}=-I_{dsn}, the output voltage V_{out} can be expressed as

$$V_{out} = (V_{in} - V_{tp}) + \sqrt{(V_{in} - V_{tp})^2 - 2(V_{in} - \frac{V_{DD}}{2} - V_{tp})V_{DD} - \frac{\beta_n}{\beta_p}(V_{in} - V_{tn})^2}$$

– **Region C:** defined by $V_{in} \approx \frac{V_{DD}}{2}$, where the p- and n-devices are in saturation.

- Its equivalent circuit is shown in Figure 2.14(b).
- $I_{dsp} = -\frac{\beta_p}{2}(V_{in} - V_{DD} - V_{tp})^2$
- $I_{dsn} = \frac{\beta_n}{2}(V_{in} - V_{tn})^2$
- $V_{in} = \frac{V_{DD} + V_{tp} + V_{tn}\sqrt{\frac{\beta_n}{\beta_p}}}{1 + \sqrt{\frac{\beta_n}{\beta_p}}}$ with I_{dsp}=-I_{dsn}. (2.24)
- $V_{in} = \frac{V_{DD}}{2}$ by setting $\beta_n = \beta_p$ and V_{tn}=-V_{tp}.
- For n-transistor:
  – V_{out}=V_{dsn}>V_{in}-V_{tn} $\Rightarrow$ V_{in}-V_{out}<V_{tn} $\Rightarrow$ V_{out}>V_{in}-V_{tn}

  For p-transistor:
  – V_{dsp}<V_{gs}-V_{tp} $\Rightarrow$ V_{out}-V_{DD}<V_{in}-V_{DD}-V_{tp} $\Rightarrow$ V_{out}<V_{in}-V_{tp}

  Combining the two inequalities results in $\Rightarrow$ V_{in}-V_{tn}<V_{out}<V_{in}-V_{tp}

- In this region, we have two current sources in series, which is an “unstable” condition. Thus a small input voltage has a large effect at the output.
- The input equation shown in EQ. (2.24) can be used for defining the gate threshold V_{inv} , which corresponds to the state where $V_{out}=V_{in}$.

– **Region D:** defined by $\frac{V_{DD}}{2} < V_{in} < V_{DD} + V_{tp}$, where the p-device is in saturation while the n-device is in the nonstaturated region. (see Figure 2.14(c) for its equivalent circuit)

- $I_{dsp} = -\frac{\beta_p}{2}(V_{in} - V_{DD} - V_{tp})^2$
- $I_{dsn} = \beta_n[(V_{in} - V_{tn})V_{out} - \frac{V_{out}^2}{2}]$
- $V_{out} = (V_{in} - V_{tn}) - \sqrt{(V_{in} - V_{tn})^2 - \frac{\beta_p}{\beta_n}(V_{in} - V_{DD} - V_{tp})^2}$ with I_{dsp}=-I_{dsn} .

– **Region E:** defined by $V_{in}>V_{DD}+V_{tp}$ where the p-device is cutoff ($I_{dsp}=0$) and the n-device is in the linear region.

– Because $V_{gsp} = V_{in} - V_{DD}$ and $V_{in} \geq V_{DD} + V_{tp} \Rightarrow V_{gsp} + V_{DD} \geq V_{DD} + V_{tp} \Rightarrow V_{gsp} \geq V_{tp}$.

$\Rightarrow$ the p-device is in cutoff region $\Rightarrow$ $V_{out}=0$

- 2.3.1 β_n/β_p Ratio
  - $\hat{a} = \frac{me}{t_{ox}}(\frac{W}{L})$
  - The transfer curves of an inverter plotted as a function of β_n/β_p are shown in Figure 2.15(a).
  - V_{inv}(gate threshold voltage) where $V_{in} = V_{out}$ is dependent on β_n/β_p.
  - As the ration β_n/β_p is decreased, as shown in Figure 2.15(a) the transition region shifts from left to right.
  - For the CMOS invert a ratio of $\beta_n/\beta_p = 1$ may be desirable since it provides equal current-source and -sink capability.
  - Change channel dimension W and L of the p and n devices would change the value β and thus would change the ratio β_n/β_p for a given process.
  - The inverter transfer curve is also plotted for W_n/ W_p as shown in Figure 2.15(b).
  - $\beta_n \propto T^{-1.5} \Rightarrow I_{ds} \propto T^{-1.5}$. The effective carrier mobility μ decreases when temperature increases, but β_n/β_p ratio is relatively independent of temperature to a good approximation.

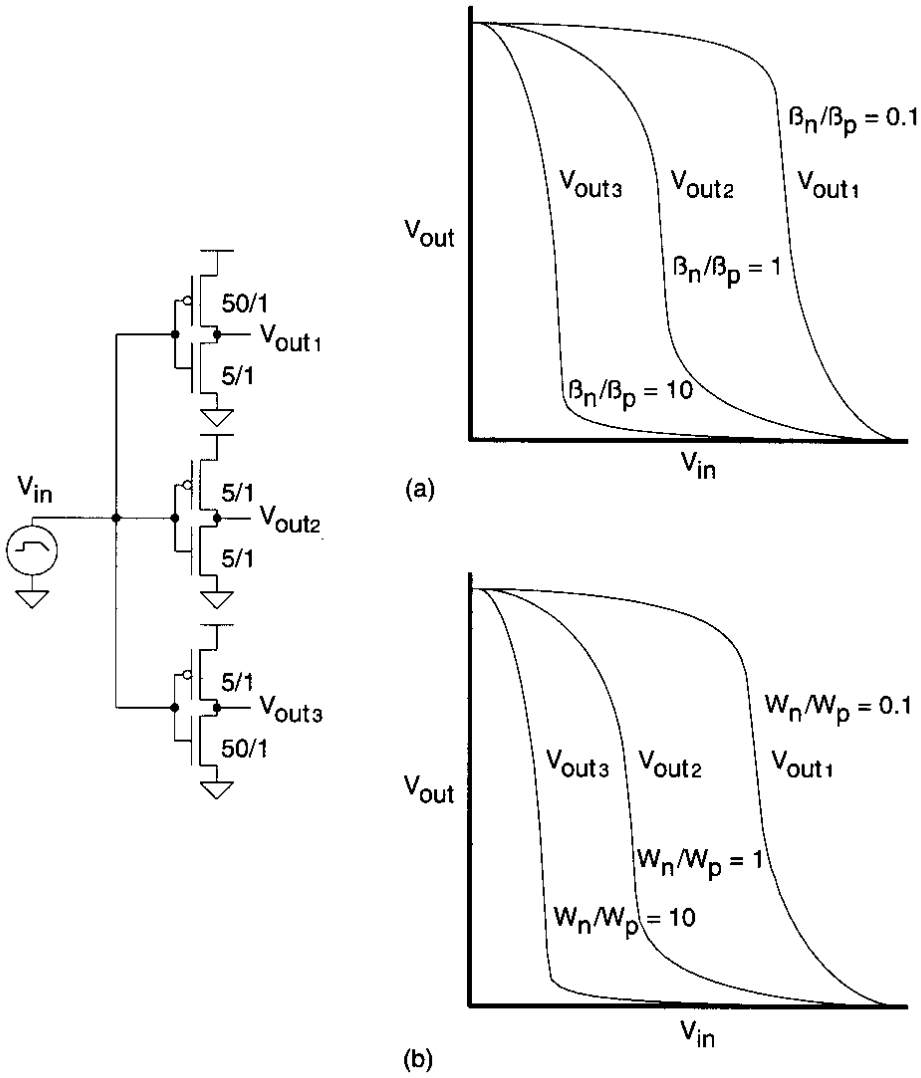


FIGURE 2.15 Influence of $\frac{\beta_n}{\beta_p}$ on inverter DC transfer characteristic

- 2.3.2 Noise Margin
  - Noise margin allows us to determine the allowable noise voltage on the input of a gate so that the output will not be affected. It is closely related to the input-output voltage characteristics.
  - The LOW noise margin $NM_L = |V_{IL\,\max} - V_{OL\,\max}|$:defined as the difference in magnitude between the maximum LOW output voltage(V_{OLmax}) of the driving gate and the maximum input LOW voltage(V_{ILmax}) recogniaed by the driven gate (Figure 2.16).
  - The HIGH noise margin $NM_H = |V_{OH\,\min} - V_{IH\,\min}|$: defined as the difference in magnitude between the minimum HIGH output voltage (V_{OHmin}) of the driving gate and minimum HIGH voltage (V_{Hmin}) recognized by the driven gate.(see Figure 2.16)

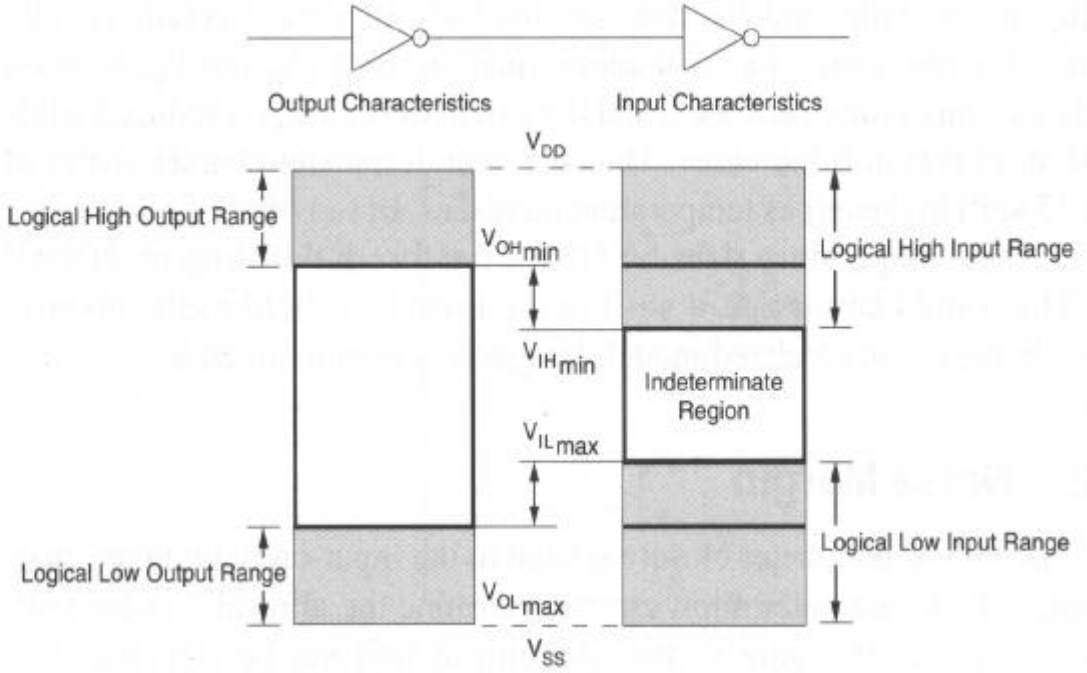


FIGURE 2.16 Noise margin definitions

– $V_{IH} = V_{IL}$ is desirable. This implies the transfer characteristic should switch abruptly. That is, there should be a high gain in the transition region.

– For the purpose of calculating noise margins, the transfer characteristics of a typical inverter and the definition of voltage levels V_{IL}, V_{OL}, V_{IN}, V_{OH} are sown in Figure 2.17.

– The noise margin defined for the inverter shown in Figure 2.17 is $NM_L = 2.3V$ and $NM_H = 1.7V$. Note that $V_{IH\,\min} = 3.3V$, $V_{OH\,\min} = 3.3V$, $V_{L\,\max} = 0$, and $V_{IL\,\max} = 2.3V$.

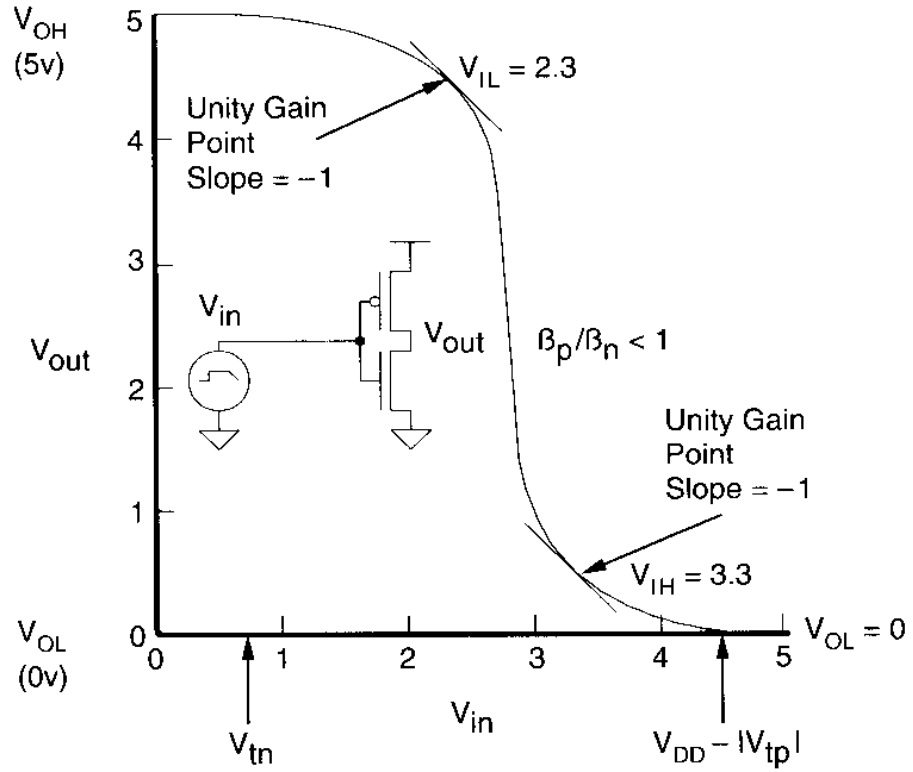


FIGURE 2.17 CMOS inverter noise margins

- 2.4 Static Load MOS Inverters
  - A generic nMOS inverter that uses either a resistive load or a constant current source is shown in Figure 2.19(a).
  - For the resistor case, VI characteristics of n-device and the resistor is shown in 2.19(b). The DC transfer characteristics is shown in 2.19(c) .

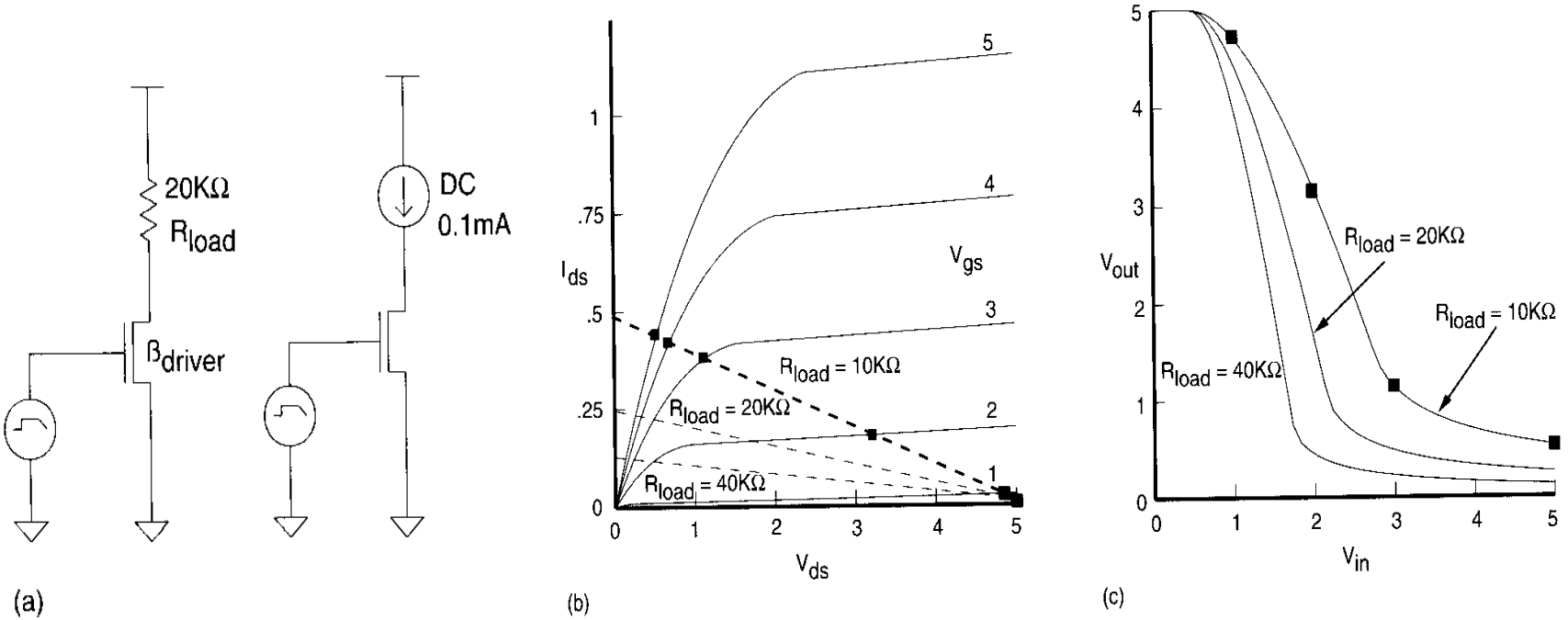


Figure 2.19 A generic static load inverter

- If R_{load} increases, the V_{OL} decreases (NM_L increases) and the "ON" current decreases; if R_{load} decreases, the V_{OL} rises (NM_L decreases) and the "ON" current rises.
- Selection of the resistor value would consider a compromise between V_{OL}, the current drawn and the pull-up speed.
- Resistors can be implemented using highly resitive undoped polysilicon.
- When transistors are used as a current-source load, the inverter is called a saturated load inverter if the load transistor is operated in saturation; if the load transistor is biased for use as a resistor, it is called an unsaturated inverter.
- The reason to use static load inverter is to reduce the number of transistors used for a gate to improve density and/or to lower dynamic power consumption

- 2.4.1 The Pseudo-nMOS Inverter
  - An inverter that uses a p-device pull-up or load that has its gate permanently grounded is shown in Figure 2.20 (a). An n-device pull-down or driver is driven with the input signal. This circuit is called ”pseudo-nMOS” inverter.
  - Figure 2.20 (b) and (c) shows transfer characteristics in terms of β_n/β_p .
  - While Figure 2.20(d) shows the I_{ds} in terms of V_{in} (note that there is a DC current flow when the output is in LOW state).
  - Although pseudo-nMOS gates are not used for low power applications, they do find wide application in high-speed circuits and circuits that require large fan-in NOR gates.

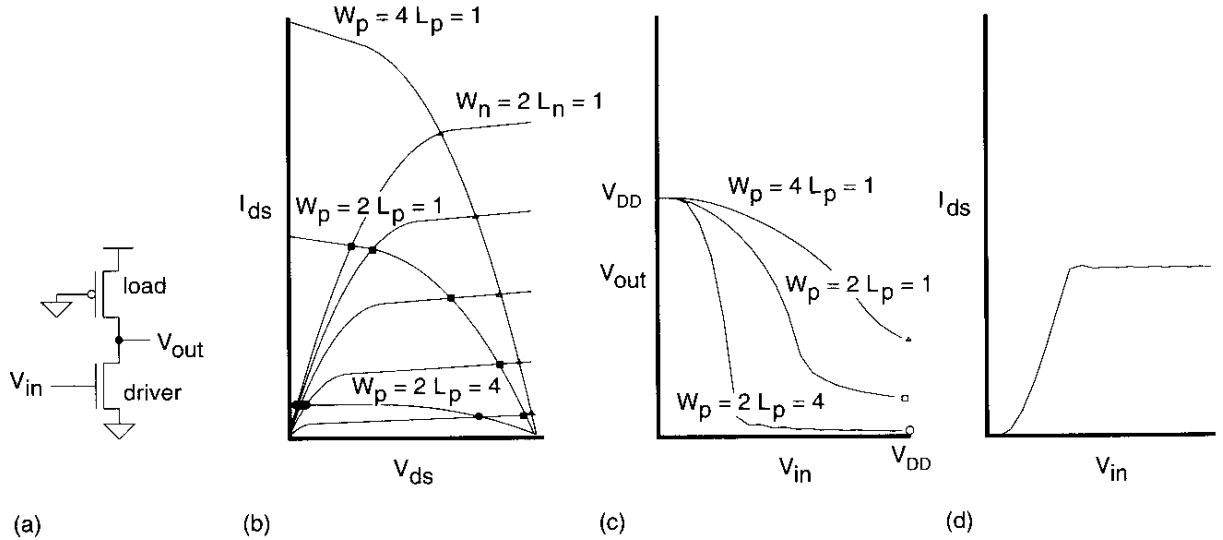


FIGURE 2.20 The pseudo-nMOS inverter and DC transfer characteristics

– For the circuit shown in Figure 2.20 with V_{in} increasing from V_{SS} to V_{DD},

- For the saturated n-device: $I_{dsn} = \frac{\beta_n}{2}(V_{in} - V_{tn})^2$ for $(V_{out} = V_{dsn} > V_{in} - V_{tn})$
- For the non-saturated p-device ($V_{dsp} = -V_{DD}$) : $I_{dsp} = -\beta_p\left[(-V_{DD} - V_{tp})(V_{out} - V_{DD}) - \frac{(V_{out} - V_{DD})^2}{2}\right]$
- Let $I_{dsp} = -I_{dsn}$, we obtain $V_{out} = -V_{tp} + \sqrt{(V_{DD} + V_{tp})^2 - C}$
- where $C = \frac{\beta_n}{\beta_p}(V_{in} - V_{tn})^2$ and $\frac{\beta_n}{\beta_p} = \frac{(V_{DD} + V_{tp})^2 - (V_{out} + V_{tp})^2}{(V_{in} - V_{tn})^2}$

– The noise margin of this kind of inverters is dependent on the ratio $\beta n/\beta p$ and generally is poor than that of the CMOS inverters.

β_n/β_p	V_{IL}	V_{IH}	V_{OL}	V_{OH}	NM_L	NM_H
2	3.4	4.5	1.4	5	2.0	0.5
4	1.8	3.3	0.6	5	1.2	1.7
6	1.4	2.8	0.35	5	1.05	2.2
8	1.1	2.4	0.24	5	0.86	2.6
100	0.5	1.1	0.00	5	0.5	3.9

– Figure 2.21(a) shows an example of noise margin and 2.21(b) shows the transfer curves in terms of b*n*/ b*p* .

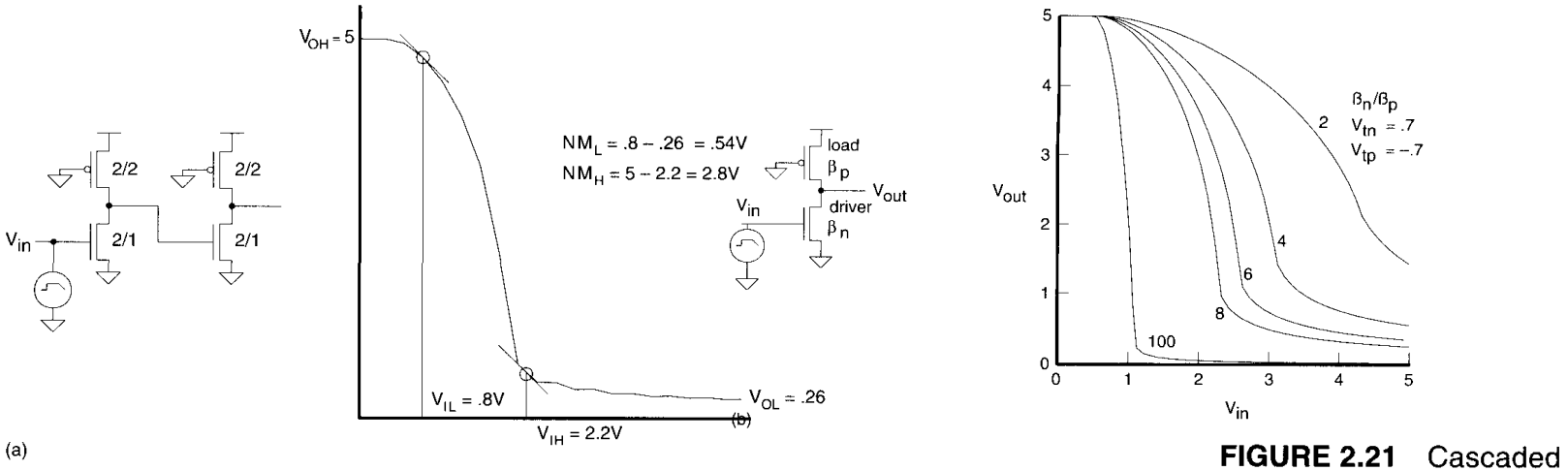


FIGURE 2.21 Cascaded pseudo-nMOS inverters

– The inverter shown in Figure 2.21(a) is typically used in static ROMs and PLAs.

– Figure 2.22(a) shows an inverter with a p-transistor biased to be a constant current source.

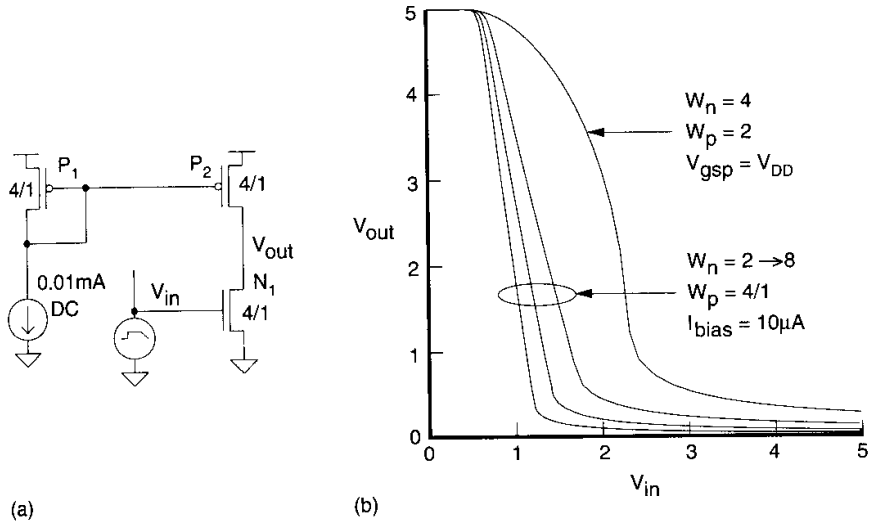


FIGURE 2.22 Constant current source load pseudo nMOS inverter

- 2.4.3 More Saturated Load Inverters
  - Figure 2.24(a) shows another "pseudo-nMOS" inverter that has a p-device load with its gate connected to the output.

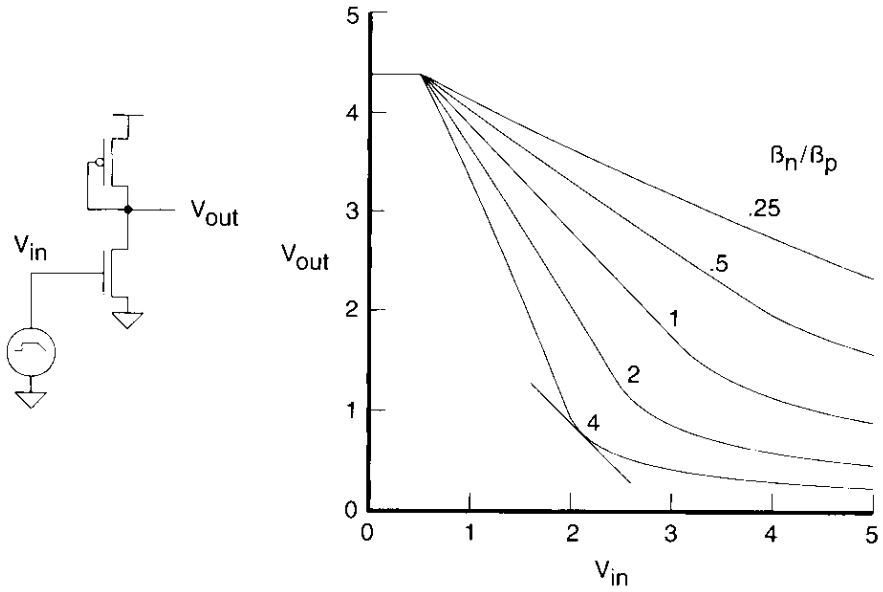


FIGURE 2.24 Saturated load inverter

  - While $V_{out} > V_{in} - V_{tn}$, the driver transistor is in saturation, $I_{dsdriver} = \frac{\hat{a}_{driver}}{2}(V_{in} - V_{tn})^2$, and the load transistor is permanently in the cutoff ($I_{ds} = 0$) or saturated, $I_{dsload} = \frac{-\hat{a}_{load}}{2}(V_{out} - V_{DD} - V_{tp})^2$. Let $I_{dsdriver} = -I_{dsload}$, we obtain $V_{out} = V_{DD} + V_{tp} - \sqrt{k}(V_{in} - V_{tn})$, where $k = \frac{\beta_{driver}}{\beta_{load}}$. If $V_{in} = V_{tn}$, $V_{out} = V_{DD} + V_{tp} < V_{DD}$; that is the maximum output is reached when $V_{in} = V_{tn}$.

– Figure 2.25 shows an nMOS depletion load inverter. The output of this inverter can rise to a full V_{DD} level, because the depletion mode transistor is always "ON".

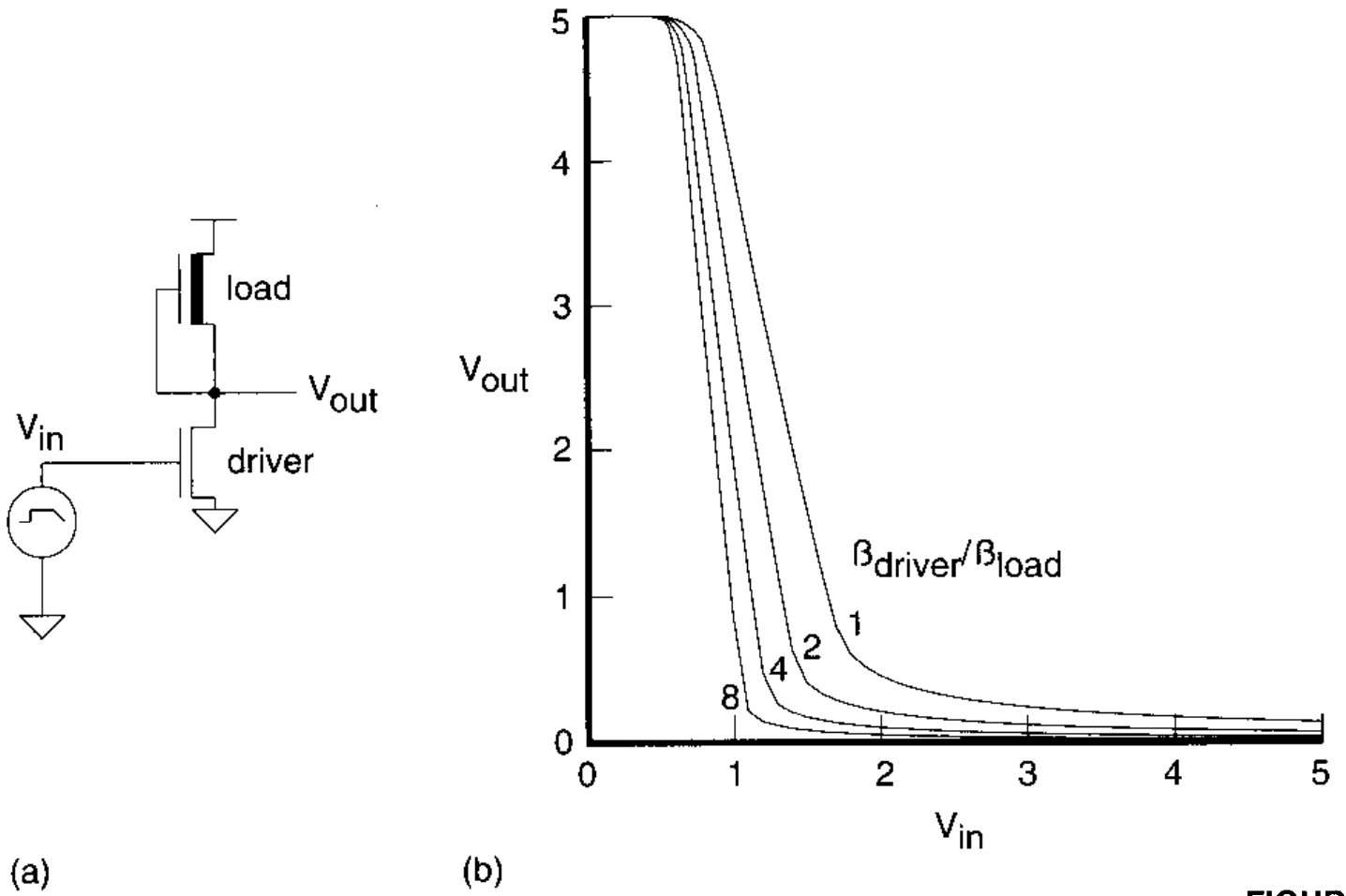


FIGURE 2.25 Depletion load inverter

- 2.6 The Transmission gate
  - Transmission gates find used as a multiplexing element, a logic structure, a latch element and an analog switch.
  - The transistor schematic of a transmission gate is reviewed in Figure 2.32. The operation of the transmission gate can be best explained by considering the n-device and p-device as pass transistors individually.

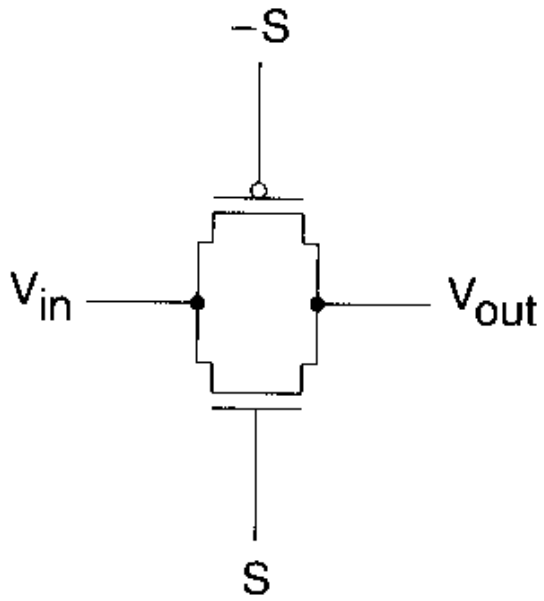


FIGURE 2.32
Transistor connection for CMOS transmission gate

– nMos pass transistor

- Figure 2.33(a) shows an nMOS pass transistor. Initially, the load capacitor C_{load} is discharged. Thus, $V_{out} = V_{ss}$
- With $S = 0 \Rightarrow V_{gs} = 0 \Rightarrow I_{ds} = 0 \Rightarrow V_{out} = V_{ss} = 0$ independent of V_{in}.
- When $S = 1$ and $V_{in} = 1$, the pass transistor starts to conduct. When $V_{out} = V_{DD} - V_{tn(Vdd)}$, (ie., $V_{gs} < V_{tn(Vdd)}$) the n-device begins to turn off where $V_{tn(Vdd)}$ is the body effected threshold with the source at $V_{DD} - V_{tn(Vdd)}$. (i.e., the n-device passes a poor ”1”).
- With $V_{in} = 0$, $S = 1$ and $V_{out} = V_{DD} - V_{tn(Vdd)}$, the n-device begins to discharge the C_{load}. Finally $V_{out} = V_{in} = 0$. (i.e., n-device passes a good “0”)

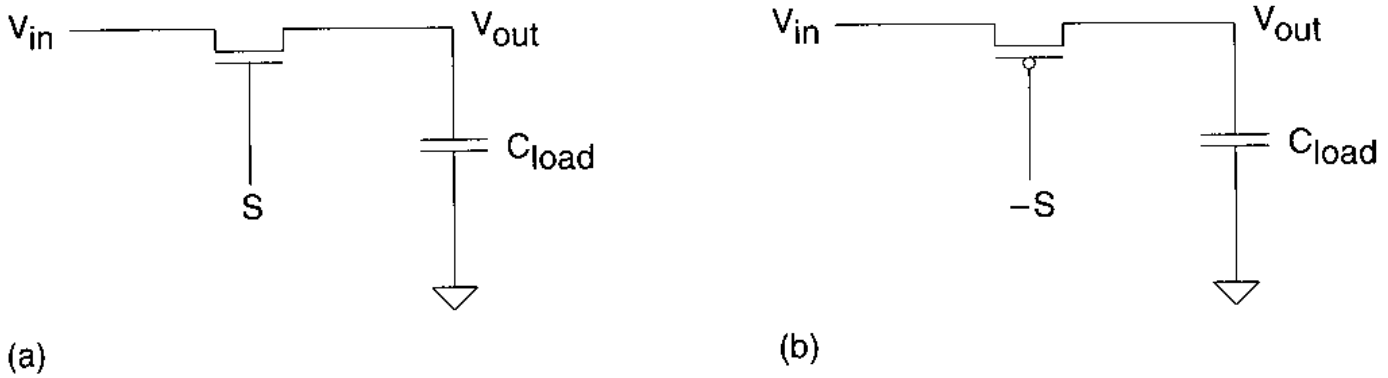


FIGURE 2.33 nMOS and pMOS transistor operation transmission gate

– pMOS transistor

- Figure 2.33(b) shows a pMOS transistor. Initially C_{load} is discharged and $V_{out} = V_{ss}$.
- With S = 0 (-S = 1), $V_{in} = V_{DD}$, and $V_{out} = V_{ss}$, C_{load} remains uncharged.
- With S = 1 (-S = 0), the p-device starts to charge the C_{load} toward V_{DD}.(i.e., p-device passes a good "1")
- When $V_{in} = V_{ss}$, $V_{out} = V_{DD}$, and S = 1 (-S = 0), the C_{load} discharges until $V_{out} = |V_{tp(Vss)}|$ (i.e., p-device passes a poor "0").
- The resultant behavior of the n-device and p-device are shown in Table 2.4

TABLE 2.4 Transmission Gate Characteristics

DEVICE	TRANSMISSION OF '1'	TRANSMISSION OF '0'
n	poor	good
p	good	poor

– The overall behavior a transmission gate can be expressed as:

$$S = 0\,(-S = 1);\begin{cases} \text{n-device} = \text{off} \\ \text{p-device} = \text{off} \\ V_{in} = V_{SS},\ V_{out} = Z \\ V_{in} = V_{DD},\ V_{out} = Z \end{cases}$$

$$S = 1\,(-S = 0);\begin{cases} \text{n-device} = \text{on} \\ \text{p-device} = \text{on} \\ V_{in} = V_{SS},\ V_{out} = V_{SS}. \\ V_{in} = V_{DD},\ V_{out} = V_{DD} \end{cases}$$

– Figure 2.34 (a) a typical circuit configuration for a transmission gate. Two cases of the operation have to be considered :

① The transmission gate acts as a resistor :

when the control input changes rapidly, the inverter input has been low (V_{ss}), the inverter output has been high (V_{DD}), and the capacitor on the transmission gate output is initially discharged (V_{ss}).

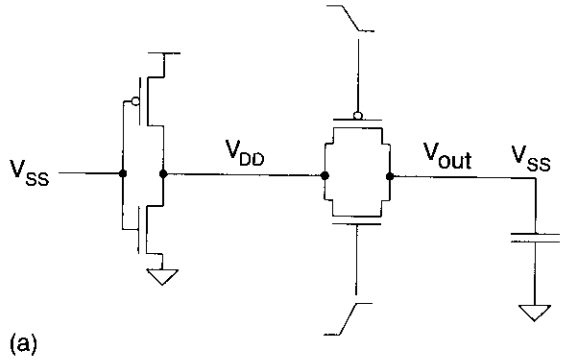
– Three regions of operation :

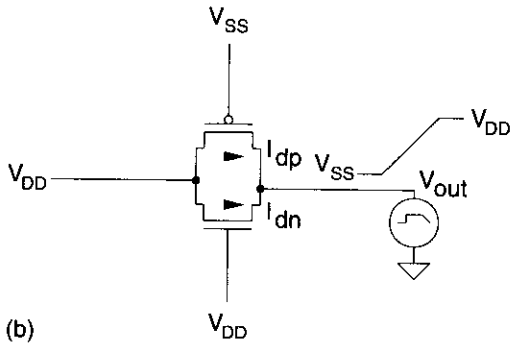
» Region A : n saturated , p saturated
($V_{out} < |V_{tp}|$)

» Region B : n saturated, p nonsaturated
($|V_{tp}| < V_{out} < V_{DD} - V_{tn}$)

» Region C : n off, p nonsaturated
($V_{DD} - V_{tn} < V_{out}$)

– The currents of the n- device , p-device , and the n- and p-device are respectively shown in Figure 2.34 (c), where it can be see that the combined current is linear with respect to the V_{out}.
($R = \frac{V_{DD} - V_{out}}{I}$)





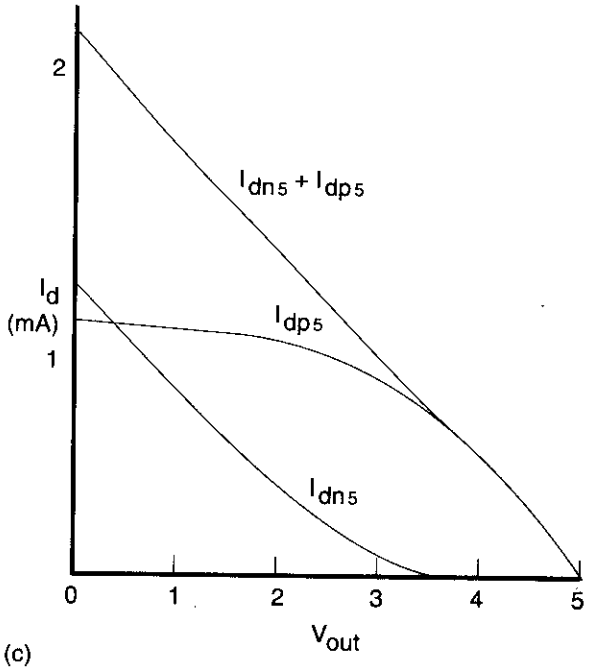


FIGURE 2.34 Transmission gate output characteristic for control input changing

② Another operation mode that the transmission gate encounters in lightly loaded circuits is where the output closely follows the input, such as shown in Figure 2.35 (c). That is the control input remains at the logice value that makes p- and n-devices “ON” and the input changes from 0 to 1 or 1 to 0.

– Three regions of operation:
  » Region A : n nonsaturated, p off
    ($V_{in} < |V_{tp}|$)
  » Region B : n nonsaturated, p nonsaturated
    ($|V_{tp}| < V_{in} < V_{DD} - V_{in}$)
  » Region C : n off, p nonsaturated
    ($V_{in} > V_{DD} - V_{tn}$)
– Figure 2.35 (a) shows the n- and p- pass transistor currents for $V_{out} - V_{in} = -0.1V$
– Figure 2.36 shows a plot of the effective “ON” resistance of the transimission gate for the test circuit shown in Figure 2.35 (c).

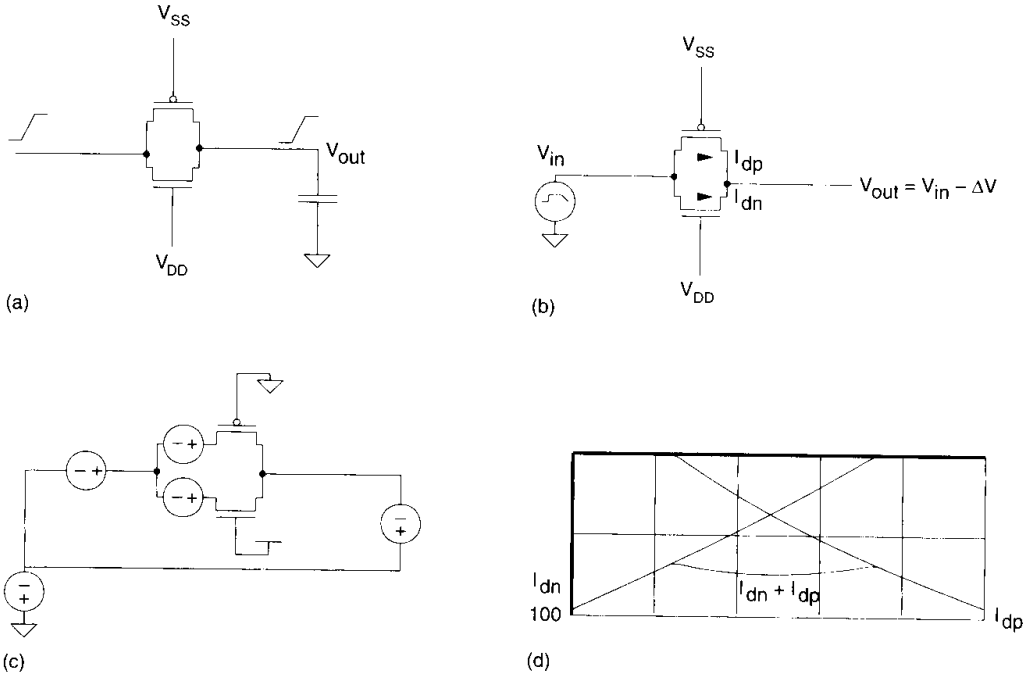


FIGURE 2.35 Transmission gate output characteristic for switched input changing

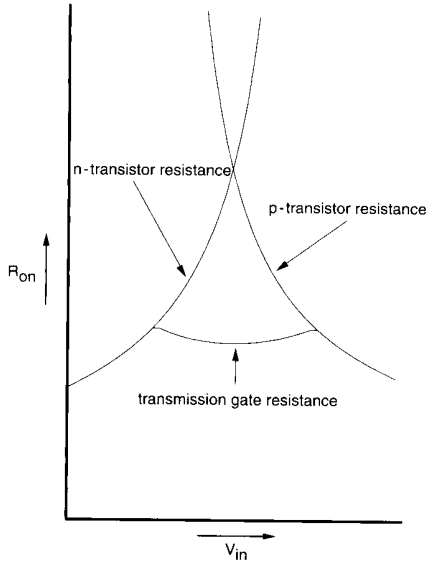


FIGURE 2.36 Resistance of a transmission gate for conditions in Figure 2.35

- 2.7 The Tristate Inverter
  - A tristate inverter can be formed by cascading a transmission gate with an inverter as shown in Figure 2.37(a).
  - When C = 0 and -C = 1, the output Z is in high impedance.

    When C =1 and -C = 0, the output Z is equal to the complement of the input A.
  - The connection between the n- and p- driver transistor may be omitted as shown in Figure 2.37 (b) and the logic operation remains the same.

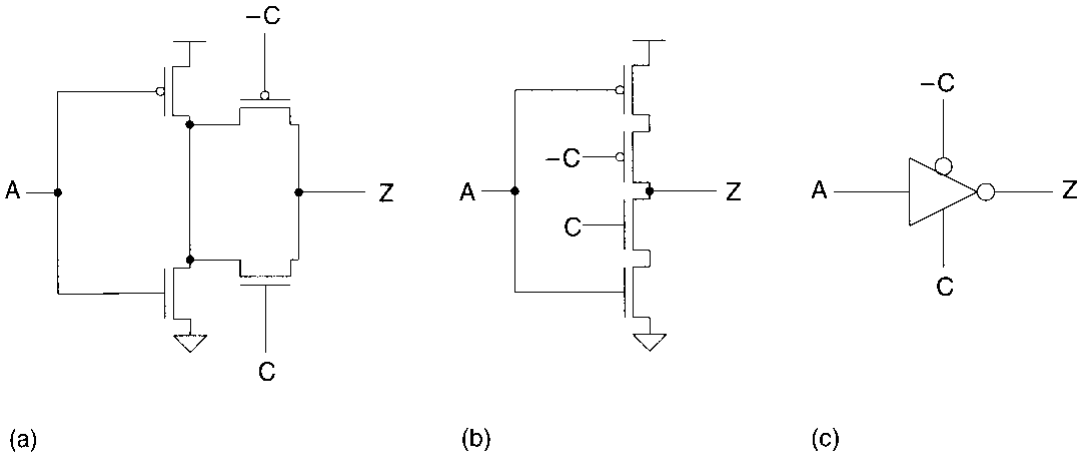


FIGURE 2.37 Tristate inverter

- 2.9 Summary
  - Examine the DC and AC characteristics of MOS transistors and COMS inverters.
  - Study the operation of the CMOS transmission gate.