

MOS devices, as well as bipolar junction transistors (BJT), find application in logic gates. In this chapter we discuss the operation of PMOS (*p*-channel), NMOS (*n*-channel), and CMOS (complementary-symmetry) gates. CMOS is rapidly becoming the most favored because of its lower power dissipation, shorter propagation delay, and shorter rise and fall times.

8.1 ANALYTIC EQUATIONS FOR MOSFETS

Within a MOSFET, by definition, the charge carriers move away from the *source* and towards the *drain*. Therefore, in an *n*-channel device, where the carriers are negative, the conventional direction of current flow within the device is from drain to source. Thus the drain is positive with respect to the source, i.e., V_{DS} is positive as is also the current I_{DS} . Typical characteristics of *n*-channel MOSFETs are shown in Fig. 8.1-1. In Fig. 8.1-1a and b the MOSFET characteristics refer to an *enhancement* device. In such a device there is no channel between source and drain at $V_{GS} = 0$ V. No drain-to-source current I_{DS} flows until the gate-to-source voltage exceeds a threshold voltage V_T . This threshold

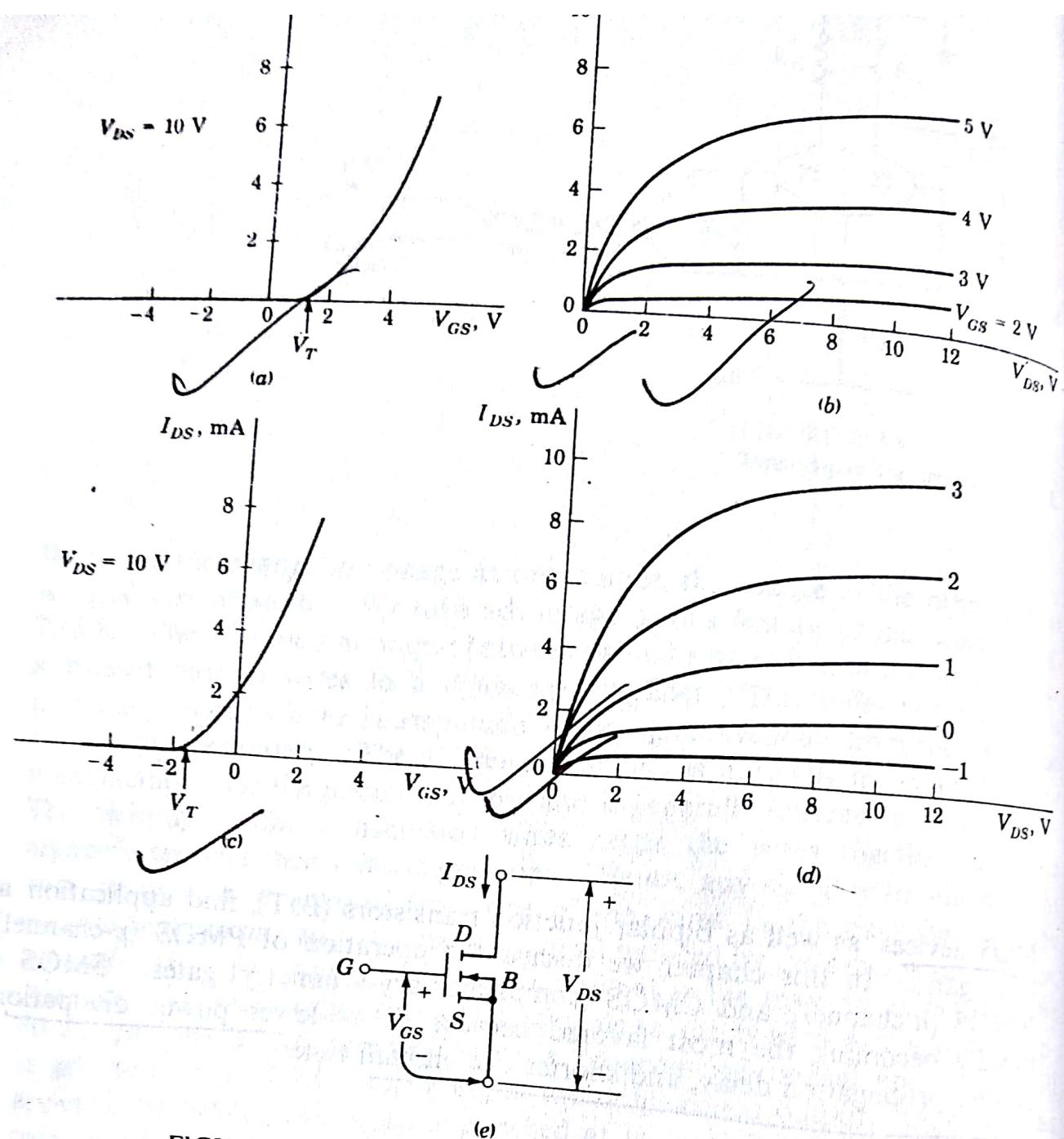


FIGURE 8.1-1
MOSFET characteristics. (a and b) Enhancement type. (c and d) Depletion type.
(e) Defining voltages and currents. The device symbol represents an *n*-channel enhancement-type transistor.

voltage is of polarity which is the same as the polarity normally applied to the drain. Thus in an *n*-channel device, where V_{DS} is positive, so also is V_T , and a channel forms when $V_{GS} > V_T$.

In a *p*-channel device, where the carriers are positive, it is V_{SD} and I_{SD} (rather than V_{DS} and I_{DS}) which are positive. Furthermore, a channel forms to allow current I_{SD} to flow, when the source-to-gate voltage V_{SG} (rather than V_{GS}) exceeds a positive threshold voltage V_T , that is, when $V_{SG} > V_T$. We are using

the symbol V_T with two meanings. In an n -channel device V_T is defined as a particular value of V_{GS} , while in a p -channel device V_T represents a particular value of V_{SG} . Where any confusion may result we shall use instead the symbols $V_T(n)$ and $V_T(p)$. The symbolism we are employing avoids inconvenient negative signs and absolute-value signs.

In Fig. 8.1-1c and d typical characteristics are shown (again for an n -channel device) for a *depletion* transistor. Here a channel exists when $V_{GS} = 0$, and the threshold voltage V_T is negative. Strictly, such an n -channel transistor operates in the depletion mode when V_{GS} is negative and in the enhancement mode when V_{GS} is positive. It is customary nonetheless to refer to such a device simply as a depletion MOSFET. Both enhancement and depletion transistors are used in logic gates.

Either transistor type (enhancement or depletion) may operate in the non-saturation region (also referred to as the *triode* region in fond memory of the days of vacuum tubes) or in the saturation region. In the triode region there is a continuous channel between source and drain, and I_{DS} varies "linearly" with V_{DS} for fixed V_{GS} . At the source, the channel depth is nominally proportional to the extent to which the gate-to-source voltage V_{GS} exceeds the threshold voltage V_T and is thus proportional to $V_{GS} - V_T$. For fixed V_{GS} the channel depth is fixed. At the drain, the channel depth is proportional to the extent to which the gate-to-drain voltage V_{GD} exceeds V_T . Hence at the drain the channel depth is proportional to $V_{GD} - V_T = V_{GS} - V_{DS} - V_T$. The channel is pinched off at the drain when $V_{GD} - V_T \leq 0$ or when

$$V_{DS} \geq V_{GS} - V_T \quad (8.1-1)$$

When $V_{DS} \geq V_{GS} - V_T$, the transistor is in *saturation*. That is, because of the channel pinch-off, the current I_{DS} remains nearly constant, increasing only very slightly with increasing V_{DS} .

Just as we found it convenient to have analytic expressions for bipolar transistors (Ebers-Moll equations), so too is it useful to have analytic expressions for the MOSFET. In the triode region it is found that for an n -channel device

$$I_{DS} = k[2(V_{GS} - V_T)V_{DS} - V_{DS}^2] \quad 0 \leq V_{DS} \leq V_{GS} - V_T \quad (8.1-2)$$

In the saturation region

$$I_{DS} = k(V_{GS} - V_T)^2 \quad 0 \leq V_{GS} - V_T \leq V_{DS} \quad (8.1-3)$$

The constant k is given by

$$k = \frac{\mu \epsilon W}{2t L} \quad (8.1-4)$$

where μ = mobility of carriers in channel (electrons in n -channel devices)

ϵ = dielectric constant of oxide insulating layer

t = thickness of oxide under gate

W = channel width

L = channel length

Typically, for *n*-channel devices $\mu e/2t \approx 12 \text{ } \mu\text{A}/\text{V}^2$ and for *p*-channel devices is smaller by about a factor of 3. The width-to-length ratio W/L may range from 0.1 for a load transistor to as high as 20 or 40 for a driver device. (See Sec. 1.13)

In a *p*-channel transistor, operating in the triode region the equations for the device current are more conveniently written in the form

$$I_{SD} = k[2(V_{SG} - V_T)V_{SD} - V_{SD}^2] \quad 0 \leq V_{SD} \leq V_{SG} - V_T \quad (8.1-5)$$

In the saturation region

$$I_{SD} = k(V_{SG} - V_T)^2 \quad 0 \leq V_{SG} - V_T \leq V_{SD} \quad (8.1-6)$$

These equations, like Eqs. (8.1-2) and (8.1-3), are approximations and do not include all effects which have an influence on device current; however, they are entirely adequate for our purposes of exploring the operation of FET logic gates.

8.2 TEMPERATURE EFFECTS

Equations (8.1-2), (8.1-3), (8.1-5), and (8.1-6) for the current I_{DS} (and I_{SD}) are affected by the temperature because both V_T , the threshold voltage, and the parameter k are temperature-sensitive. The temperature dependence of V_T is given approximately by

$$\frac{dV_T}{dT} \approx -2.5 \text{ mV/}^\circ\text{C} \quad (8.2-1)$$

The temperature sensitivity of k results almost entirely from the temperature sensitivity of μ [see Eq. (8.1-4)], the carrier mobility. The mobility decreases approximately inversely with the absolute temperature and hence so also does k . When there is a temperature increase, I_{DS} (or I_{SD}) increases because of the lowering of the magnitude of V_T and decreases because of the decreased carrier mobility. In a typical case we find that the effect of μ may be fivefold greater than the effect of V_T . The overall result is that generally the overall effect of a temperature increase is a decrease of current. In this respect the MOSFET differs from the bipolar transistor, where an increase in temperature increases the current both because the current gain h_{FE} increases and because the junction voltages decrease.

8.3 THE MOS INVERTER

As discussed in Chap. 1, the basic MOS switching-circuit configuration is an inverter which consists of a MOSFET switch driver driving a load which is itself a MOSFET device rather than a passive resistor. The driver is invariably an enhancement device since it is a great convenience that the driver be OFF when the gate voltage is at or near ground. When the driver is turned ON, it

Typically, for *n*-channel devices $\mu\varepsilon/2t \approx 12 \mu\text{A}/\text{V}^2$ and for *p*-channel devices is smaller by about a factor of 3. The width-to-length ratio W/L may range from 0.1 for a load transistor to as high as 20 or 40 for a driver device. (See Sec. 1.13)

In a *p*-channel transistor, operating in the triode region the equations for the device current are more conveniently written in the form

$$I_{SD} = k[2(V_{SG} - V_T)V_{SD} - V_{SD}^2] \quad 0 \leq V_{SD} \leq V_{SG} - V_T \quad (8.1-5)$$

In the saturation region

$$I_{SD} = k(V_{SG} - V_T)^2 \quad 0 \leq V_{SG} - V_T \leq V_{SD} \quad (8.1-6)$$

These equations, like Eqs. (8.1-2) and (8.1-3), are approximations and do not include all effects which have an influence on device current; however, they are entirely adequate for our purposes of exploring the operation of FET logic gates.

8.2 TEMPERATURE EFFECTS

Equations (8.1-2), (8.1-3), (8.1-5), and (8.1-6) for the current I_{DS} (and I_{SD}) are affected by the temperature because both V_T , the threshold voltage, and the parameter k are temperature-sensitive. The temperature dependence of V_T is given approximately by

$$\frac{dV_T}{dT} \approx -2.5 \text{ mV/}^\circ\text{C} \quad (8.2-1)$$

The temperature sensitivity of k results almost entirely from the temperature sensitivity of μ [see Eq. (8.1-4)], the carrier mobility. The mobility decreases approximately inversely with the absolute temperature and hence so also does k . When there is a temperature increase, I_{DS} (or I_{SD}) increases because of the lowering of the magnitude of V_T and decreases because of the decreased carrier mobility. In a typical case we find that the effect of μ may be fivefold greater than the effect of V_T . The overall result is that generally the overall effect of a temperature increase is a decrease of current. In this respect the MOSFET differs from the bipolar transistor, where an increase in temperature increases the current both because the current gain h_{FE} increases and because the junction voltages decrease.

8.3 THE MOS INVERTER

As discussed in Chap. 1, the basic MOS switching-circuit configuration is an inverter which consists of a MOSFET switch driver driving a load which is itself a MOSFET device rather than a passive resistor. The driver is invariably an enhancement device since it is a great convenience that the driver be OFF when the gate voltage is at or near ground. When the driver is turned ON, it

range from
(See Sec.
ations for
(8.1.5)
s and do
ever, they
ET logic

I_{SD}) are
and the
of V_T is

(8.2-1)

perature
decreases
does k .
owering
mobility.
han the
tempera-
rs from
current
voltages

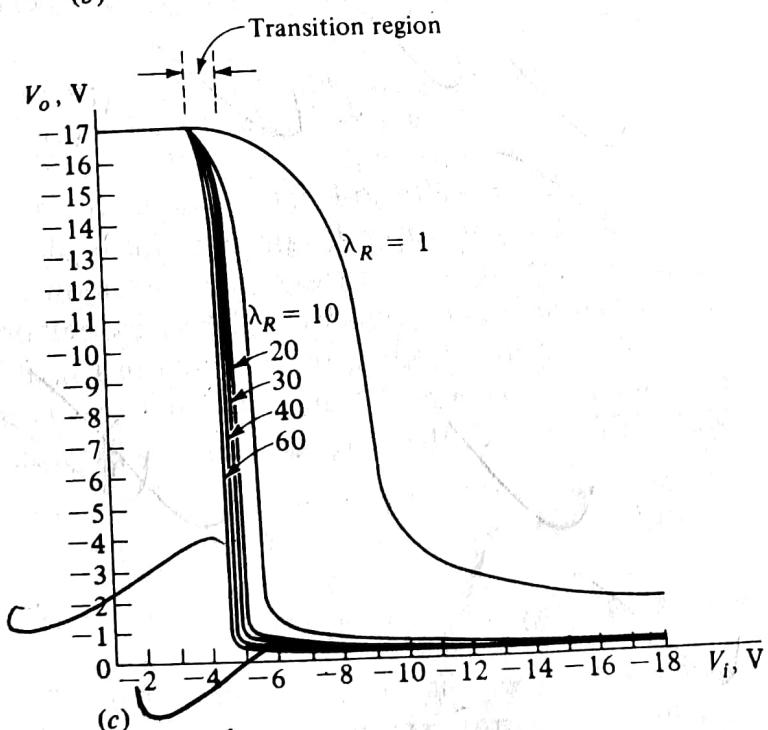
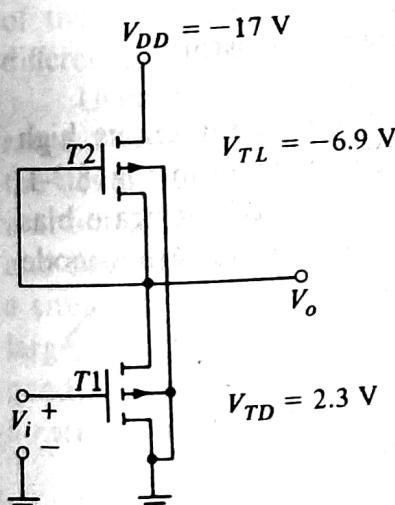
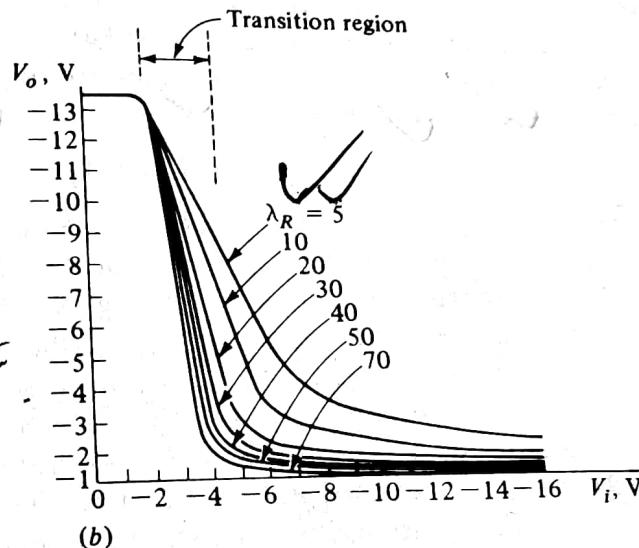
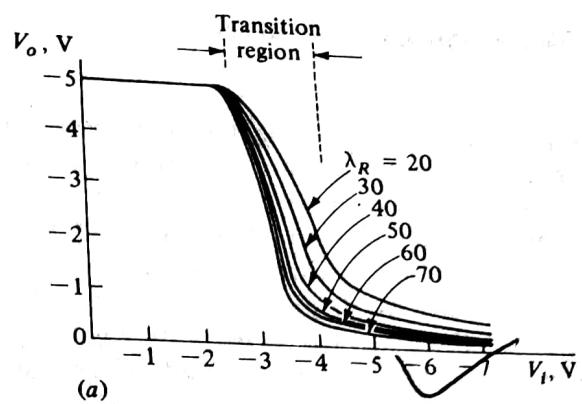
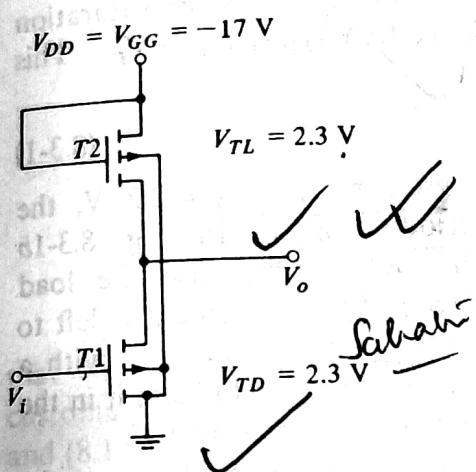
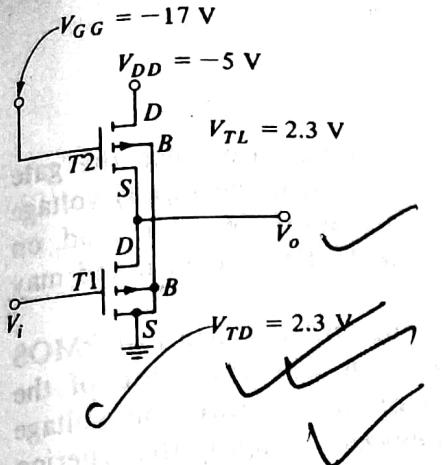


FIGURE 8.3-1
Input-output characteristics. (From "MOS/LSI Design and Applications," W. N. Carr, and J. P. Mize, McGraw-Hill, chap. 4, 1972).

invariably finds itself in the triode region. Such is the case since the gate voltage (furnished by another driving gate) will be at or near the supply voltage and the drain-to-source voltage will be at minimum magnitude. The load, on the other hand, may be an enhancement device or a depletion device and may operate in the triode or saturation region.

In Fig. 8.3-1 we display calculated input-output characteristics of PMOS inverters for three typical cases. In Fig. 8.3-1a both transistors are of the enhancement type, and both load and driver transistors have a threshold voltage $V_T = 2.3$ V. Since the transistors are *p*-channel devices, we apply the criterion given in Eq. (8.1-6) to determine whether we are in the triode or saturation region. Thus, to be in the triode region we require that $V_{SD} \leq V_{SG} - V_T$. This condition can be written as

$$V_{DG} \geq V_T \quad (8.3-1)$$

Since $V_{DG} = V_D - V_G = -5 - (-17) = 12$ V is greater than $V_T = 2.3$ V, the load transistor is biased to operate in the triode region. In Fig. 8.3-1b the driven transistor remains as in Fig. 8.3-1a, but in this case the load transistor operates in the saturation region (the proof of this statement is left to the problems). In Fig. 8.3-1c the load transistor is a depletion device with a negative threshold voltage $V_T = -6.9$ V. The biasing of the load places it in the triode region when $V_o \leq -10.1$ V and in saturation when $V_o > -10.1$ V.

As we have discussed in Sec. 1.14 (see Fig. 1.14-1), we should expect the form of the input-output characteristic to depend principally on the parameter λ_R , defined by

$$\lambda_R \equiv \frac{\lambda_D}{\lambda_L} \equiv \frac{(W/L)_D}{(W/L)_L} \quad (8.3-2)$$

where $(W/L)_D$ = width-to-length ratio of channel in driver transistor
 $(W/L)_L$ = width-to-length ratio for load

We noted that as λ_R increases, the transition of the output between its high and low levels becomes sharper. These expectations are confirmed in Fig. 8.3-1. In the calculations leading to the plots in Fig. 8.3-1 the effect of substrate bias is taken into account. Note also that the inverter using a depletion-mode MOSFET load has the steepest transition region.

8.4 THE CMOS INVERTER

The CMOS inverter is shown in Fig. 8.4-1a. The drains of a *p*-channel and an *n*-channel transistor are joined, and a supply voltage V_{ss} is applied from source to source. The output is taken at the common drain. The input V_i swings nominally through the range of V_{ss} . In the CMOS inverter shown, since we have grounded the source of the *n*-channel device, V_{ss} must be a positive voltage and V_i swings between ground and V_{ss} .

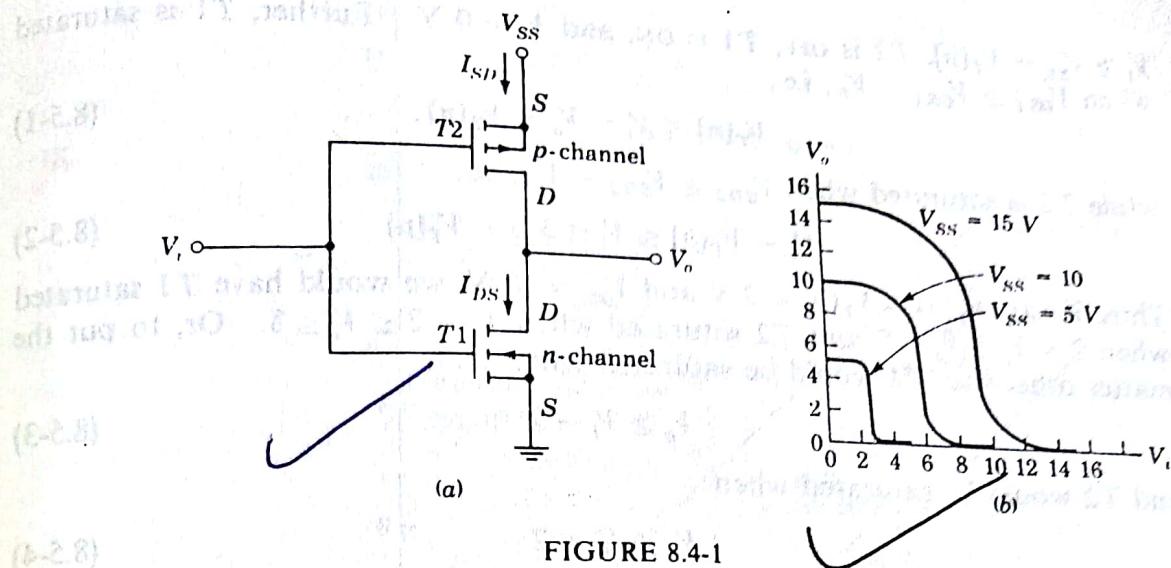


FIGURE 8.4-1

(a) A CMOS inverter and (b) its transfer characteristic.

Because of the complete symmetry of the circuit it seems intuitively clear that we shall want the two transistors to be reasonably alike. Therefore customarily it is arranged that the parameter k in Eqs. (8.1-2), (8.1-3), (8.1-5), and (8.1-6) are the same for the two transistors. The mobility of carriers in the p -channel device is smaller than the mobility in the n -channel device by a factor of 2 or 3. Hence to make the k 's equal, the ratio W/L for the p -channel must be correspondingly larger by a factor of 2 or 3 than W/L for the n -channel device [see Eq. (8.1-4)]. However, even with such an adjustment of the W/L ratio the CMOS inverter is not necessarily entirely symmetrical since the threshold voltages of the p -channel and n -channel devices generally turn out to be somewhat different.

Usually CMOS gates are designed to operate with supply voltages in the range 5 to 15 V. Typical transfer characteristics are shown in Fig. 8.4-1b. For the device to which Fig. 8.4-1b applies, the magnitude of the threshold voltage is about 2 V for each of the transistors. Observe the abruptness of the transition and that the total swing in voltage is equal to V_{ss} . In the MOS inverters such a situation prevails only when we arrange that the ratio λ_R of the λ 's be very large. In the present CMOS case, however, this situation prevails rather independently of the value of λ_R . Hence the CMOS inverter is often referred to as a ratioless inverter.

8.5 ~~CALCULATION OF CMOS-INVERTER TRANSFER CHARACTERISTIC~~

It is instructive to use the device current equations to calculate the transfer characteristic of a CMOS inverter in a typical case. Referring to Fig. 8.4-1, we have that for $V_i \leq V_T(n)$, $T1$ is OFF, $T2$ is ON, and $V_o = V_{ss}$. Similarly for

$V_i \geq V_{ss} - V_T(p)$, $T2$ is OFF, $T1$ is ON, and $V_o = 0$ V. Further, $T1$ is saturated when $V_{DS1} \geq V_{GS1} - V_T$, i.e.,

$$V_T(n) \leq V_i \leq V_o + V_T(n)$$

while $T2$ is saturated when $V_{SD2} \geq V_{SG2} - V_T$, i.e.,

$$V_o - V_T(p) \leq V_i \leq V_{ss} - V_T(p)$$

Thus, if, say, $V_T(n) = V_T(p) = 2$ V and $V_{ss} = 10$ V, we would have $T1$ saturated when $2 \leq V_i \leq V_o + 2$ and $T2$ saturated when $V_o - 2 \leq V_i \leq 8$. Or, to put the matter otherwise, $T1$ would be saturated when

$$V_o \geq V_i - 2$$

and $T2$ would be saturated when

$$V_o \leq V_i + 2$$

The currents I_{SD} and I_{DS} indicated in Fig. 8.4-1a are always equal. Accordingly, when $T1$ is in saturation and $T2$ is not, we have, using Eqs. (8.1-3),

$$k_n[V_i - V_T(n)]^2 = k_p\{2[V_{ss} - V_i - V_T(p)](V_{ss} - V_o) - (V_{ss} - V_o)^2\} \quad (8.5-5)$$

Here we have taken account of the fact that for the p -channel transistor $V_{SG} = V_{ss} - V_i$ and $V_{SD} = V_{ss} - V_o$. Similarly we find that when $T2$ is in saturation and $T1$ is not, we have

$$k_p[V_{ss} - V_i - V_T(p)]^2 = k_n\{2[V_i - V_T(n)]V_o - V_o^2\} \quad (8.5-6)$$

Finally, when both transistors are in saturation, we find that

$$k_n[V_i - V_T(n)]^2 = k_p[V_{ss} - V_i - V_T(p)]^2 \quad (8.5-7)$$

Using Eqs. (8.5-5) to (8.5-7), we have plotted in Fig. 8.5-1 the input-output characteristic of a CMOS inverter for $V_T(n) = V_T(p) = 2$ V, $V_{ss} = 10$ V, and $k_p/k_n = 1$. Above the line $V_o = V_i - 2$ [Eq. (8.5-3)] $T1$ is in saturation. Below the line $V_o = V_i + 2$ [Eq. (8.5-4)] $T2$ is in saturation. In the region between the two lines both transistors are in saturation.

Note that the simultaneous saturation of both transistors defines a unique voltage $V_i(\text{sat})$, calculated from Eq. (8.5-7) to be

$$V_i(\text{sat}) = \frac{\sqrt{k_p/k_n}[V_{ss} - V_T(p)] + V_T(n)}{1 + \sqrt{k_p/k_n}} \quad (8.5-8)$$

In Fig. 8.5-1 with $k_p/k_n = 1$, $V_T(p) = V_T(n)$, $V_i(\text{sat}) = 5$ V. This voltage, at which there occurs an abrupt transition in output voltage is midway between 0 and V_{ss} because we have selected $k_p = k_n$. If k_p were not equal to k_n , then even if $V_T(n)$ were equal to $V_T(p)$, complete symmetry would not prevail. In any event we find from Eqs. (8.5-1) and (8.5-2) that the magnitude of the abrupt transition is given by

$$\Delta V_o = V_T(n) + V_T(p) \quad (8.5-9)$$

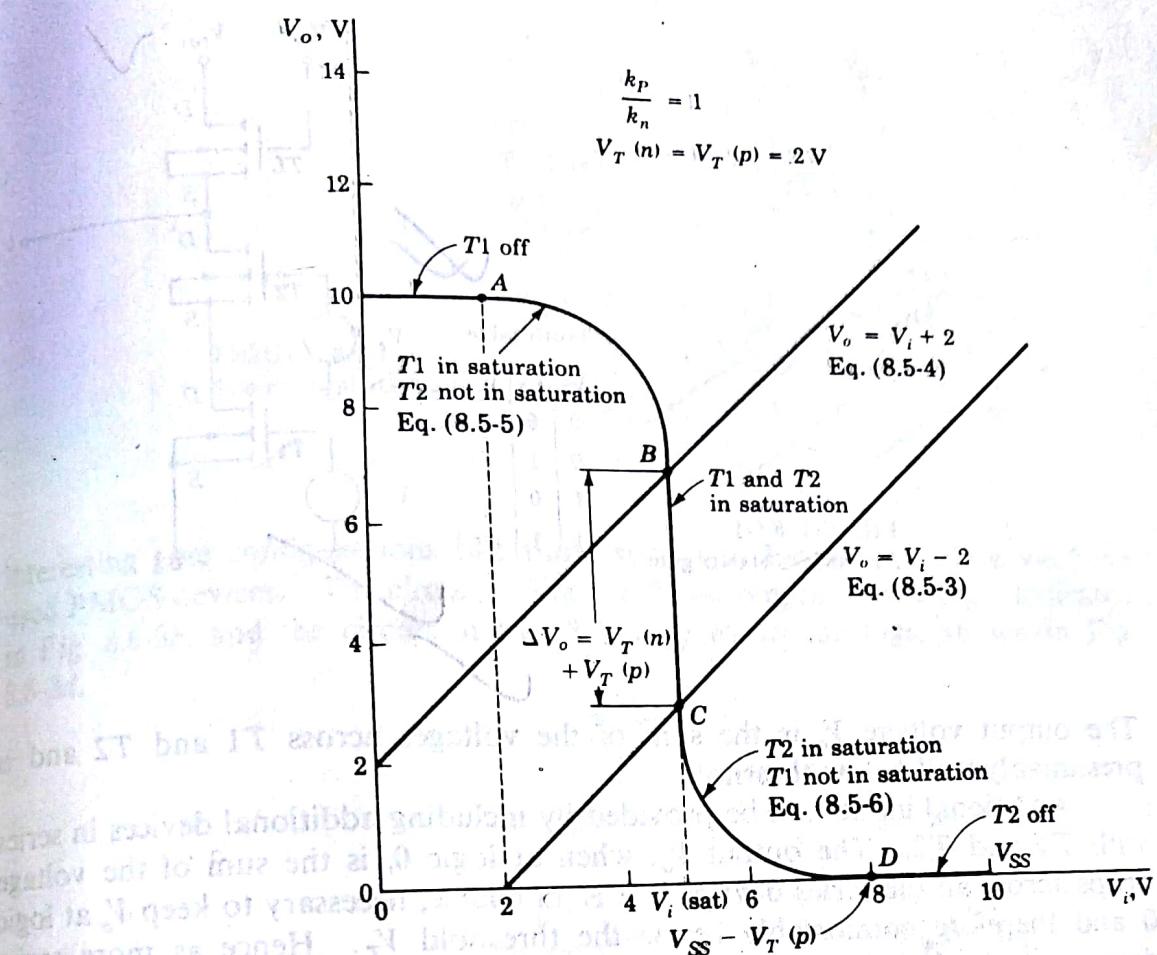


FIGURE 8.5-1 Transfer function of a CMOS inverter, $k_p/k_n = 1$.

The infinite slope displayed in Fig. 8.5-1 results from our assumption that in the saturation region the device current is absolutely independent of drain-to-source voltage, i.e., that the device is a constant current source. Such, of course, is not precisely so, and hence the transition from B to C in a physical situation would be sharp but not absolutely abrupt.

8.6 MOS GATES

Assuming that positive logic is intended, the NMOS circuit of Fig. 8.6-1 is a two-input NAND gate. The supply voltages V_{DD} and V_{GG} and the threshold voltage V_T are all positive. Logic 0 is represented by a voltage less than the threshold voltage and logic 1 by a voltage above the threshold voltage. The truth table given in Fig. 8.6-1 is readily verified. When either V_1 or V_2 or both are below threshold, only T_L conducts. In this case V_o is $V_{GG} - V_T$ or V_{DD} , whichever is lower. When both V_1 and V_2 are above threshold, both T_1 and T_2 conduct.

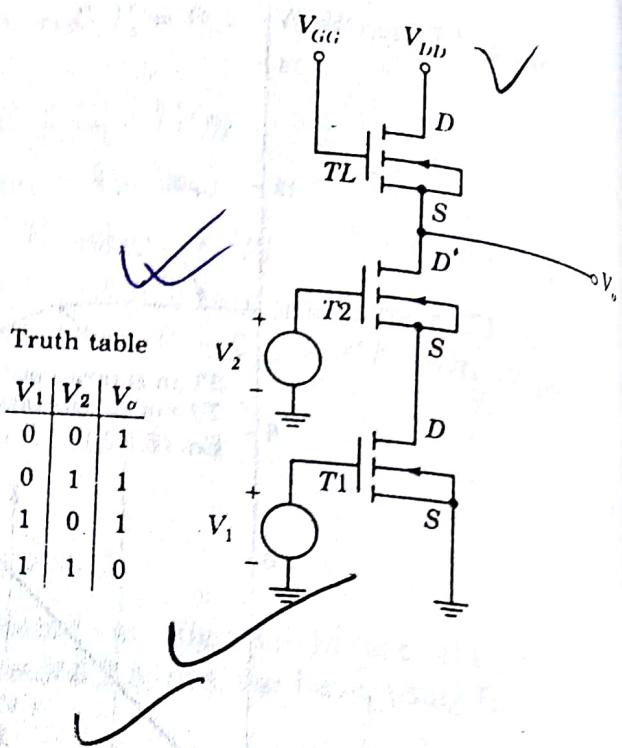


FIGURE 8.6-1
An NMOS NAND gate.

The output voltage V_o is the sum of the voltages across $T1$ and $T2$ and is presumably well below threshold.

Additional inputs can be provided by including additional devices in series with $T1$ and $T2$. The output V_o , when at logic 0, is the sum of the voltage drops across all the series devices. It is, of course, necessary to keep V_o at logic 0 and therefore comfortably below the threshold V_T . Hence as more series devices are included, the voltage drop across each one individually must be reduced. This is accomplished by increasing the width-to-length ratio W/L of the devices in order to reduce the resistance of the channel. Thus, suppose that starting with a design for a two-input gate, we wanted to modify the gate to accommodate three inputs. Then the two initial driver gates would be replaced by three driver gates, each with a W/L ratio three-halves the W/L ratio of the original devices.

Since the same input voltage (with respect to ground) is on each gate input terminal, the gate-to-source voltage on each driven FET is not the same. The gate-to-source voltage is a maximum for the driver at the bottom of the stack and decreases as we go up the stack (see Fig. 8.6-1). As noted, driver transistors operate in the triode region where the device resistance is a function of gate voltage. Thus, we can compensate for this effect of stacking by making W/L progressively larger for transistors higher up the stack.

If the NMOS driver transistors are placed in parallel, as in Fig. 8.6-2, then, as can be verified, a NOR gate results.

If PMOS devices are used, the supply voltages V_{DD} and V_{GG} must be negative with respect to ground. It is left as a student exercise to verify (again for positive logic) that with drivers in series we have a NOR gate and with drivers in parallel we have a NAND gate. The feasibility of stacking MOS in series [which is not readily permitted with bipolar devices (see Prob. 8.6-3)] allows some

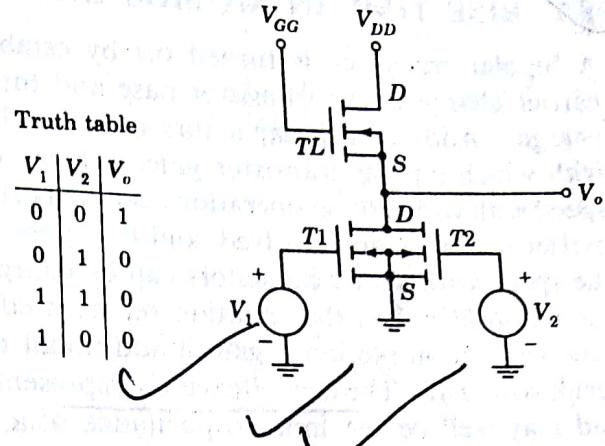
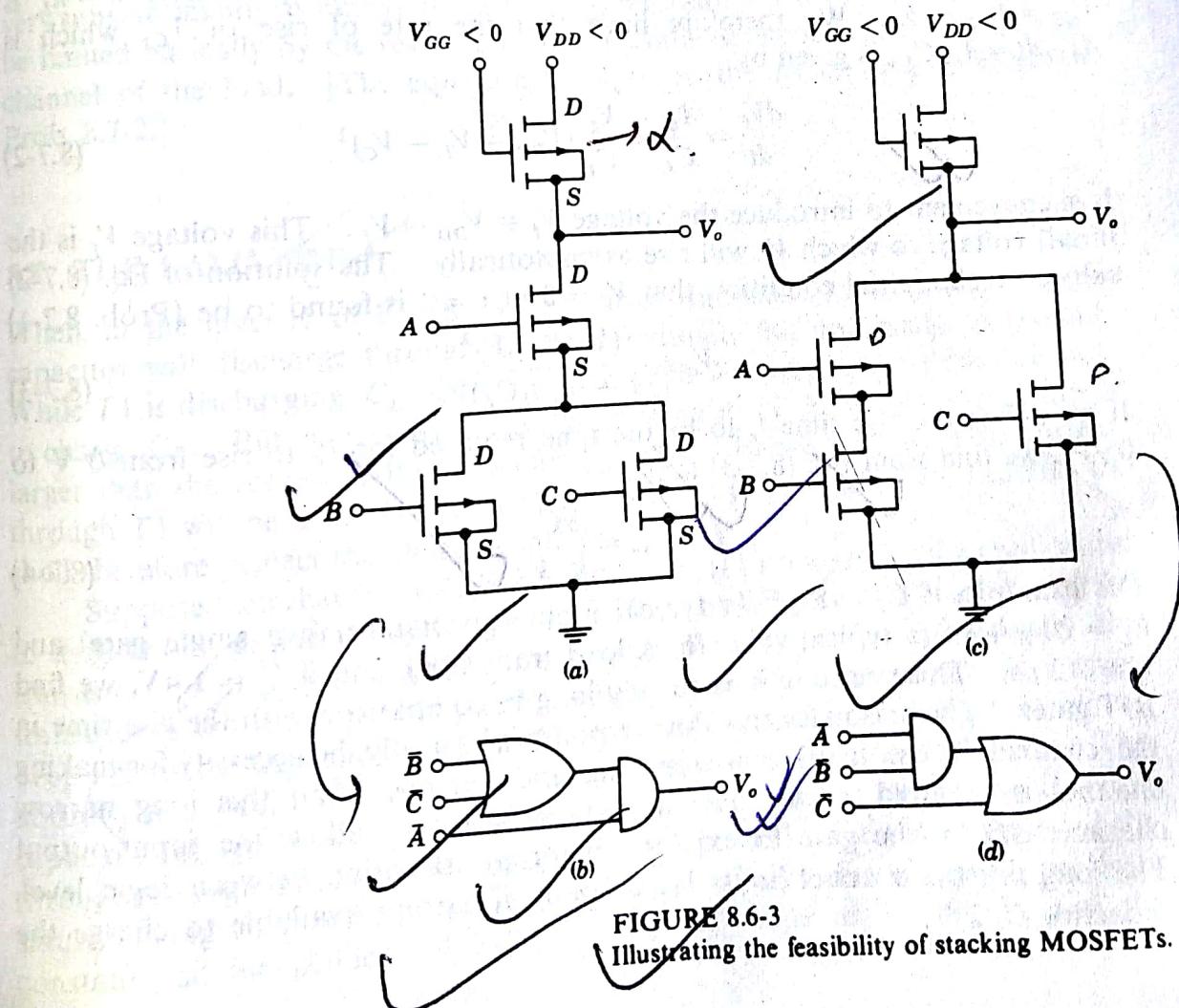


FIGURE 8.6-2
The effect of stacking.

interesting gate configurations like those shown in Fig. 8.6-3. Here we have used PMOS devices. The circuit in Fig. 8.6-3a accomplishes the logic indicated in Fig. 8.6-3b, and the circuit in Fig. 8.6-3c performs the logic shown in Fig. 8.6-3d.



8.7 RISE TIME IN AN MOS GATE

A bipolar transistor is turned ON by establishing a distribution of minority-carrier charge in the transistor base and turned OFF by removing this minority charge. Additionally, capacitors must be charged and discharged. The speed with which bipolar transistor gates can be operated is therefore limited by the speed with which these operations can be performed. In MOS devices minority-carrier charge is not involved, and the speed of operation is determined only by the speed with which capacitors can be charged.

Consider then the situation represented in Fig. 8.7-1. This basic inverter becomes a multiple input gate if additional transistors are added in parallel or series with T_1 . The capacitance C_L represents the capacitance load on the gate and may well be the input capacitance of a succeeding gate. We inquire now into the rise time of the output voltage V_o ($= V_C$, the capacitor voltage) as T_1 is turned OFF. We assume that V_C starts from 0 V. We consider first the case where the load transistor T_L operates in the saturation region, as would be the case if $V_{GG} = V_{DD}$.

In the saturation region we have, as in Eq. (8.1-3),

$$I_L = I_C = k_L(V_{GS} - V_T)^2 \quad (8.7-1)$$

The gate-to-source voltage of the load transistor with parameter $k = k_L$ is $V_{GS} = V_{DD} - V_C$. We therefore have that the rate of rise of V_C , which is $dV_C/dt = I_C/C_L$, is given by

$$\frac{dV_C}{dt} = \frac{I_C}{C_L} = \frac{k_L}{C_L}(V_{DD} - V_T - V_C)^2 \quad (8.7-2)$$

It is convenient to introduce the voltage $V_f \equiv V_{DD} - V_T$. This voltage V_f is the (final) voltage to which V_C will rise asymptotically. The solution of Eq. (8.7-2) subject to the initial condition that $V_C = 0$ at $t = 0$ is found to be (Prob. 8.7-1)

$$V_C = \frac{(k_L t/C_L)V_f^2}{1 + (k_L t/C_L)V_f} \quad (8.7-3)$$

If we define the rise time t_r to be the time required for V_C to rise from 0 V to $0.9V_f$, we find from Eq. (8.7-3) that

$$t_r = \frac{9C_L}{k_L V_f} \quad (8.7-4)$$

For example, if $C_L = 5 \text{ pF}$ (a typical input capacitance to a single gate) and $k_L = 20 \mu\text{A}/\text{V}^2$ (a typical value for a load transistor), and if $V_f = 10 \text{ V}$, we find $t_r \approx 0.2 \mu\text{s}$. This rise time is extremely long in comparison with the rise time in BJT gates. The reason for this slow response is basically the necessity for making the channel in the load transistor long and narrow. And this long narrow channel is required, as we have noted, in order to allow the input-output characteristic of the gate to exhibit an abrupt transition between logic level. The long narrow channel limits the amount of current available to charge the capacitor C_L .

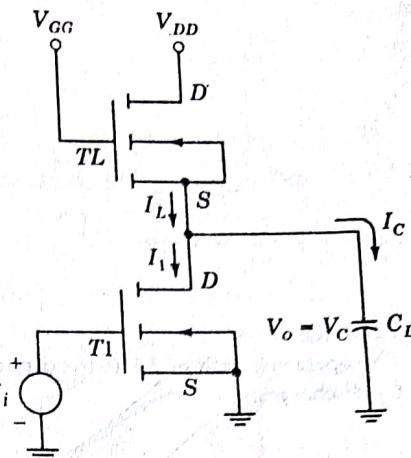


FIGURE 8.7-1
A MOS gate driving a capacitor C_L .

If, starting with a saturated load transistor, we increased the gate supply voltage V_{GG} to the point where $V_{GG} - V_T > V_{DD}$, the transistor would operate in the triode region. The increase in V_{GG} would increase the current through TL and hence the current available to charge the load capacitor. However, a modest increase in V_{GG} would result in only a small increase in current, and no very substantial improvement in rise time would result. The rise time would still be limited basically by the restricted current available through the long narrow channel of the load. [The equation for $V_C(t)$ in the triode case is given in Prob. 8.7-2.]

8.8 THE FALL TIME

When, in the inverter of Fig. 8.7-1, the driver transistor $T1$ is turned ON, the capacitor will discharge through $T1$ and eventually fall nominally to ground. While $T1$ is discharging, C_L , the current through TL , continues in the direction to charge C_L . But, as we have seen, the W/L ratio of the driver is very much larger than the corresponding ratio for the load. Hence the discharge current through $T1$ will be much larger than the charging current through TL , and we shall therefore neglect the charging current.

Suppose then that $T1$, which is initially OFF, is turned ON by the application to its gate of a gating voltage $V_{GS} > V_T$. The volt-ampere characteristic of the transistor for this gating voltage is indicated in Fig. 8.8-1. Originally the transistor is OFF and operating at P_1 , where $V_{DS} = V_{CM}$, the maximum voltage drop across the capacitor. When $T1$ is turned ON, the operating point moves abruptly to P_2 since the capacitor voltage cannot change instantaneously. The capacitor voltage $V_C = V_{DS}$ then decreases, and the transistor makes an excursion through the region of saturation as shown in Fig. 8.8-1. [We have idealized the saturation region to correspond exactly to Eq. (8.1-3), which assumes I_{DS} precisely constant and independent of V_{DS} .] At P_3 the triode region begins, and the

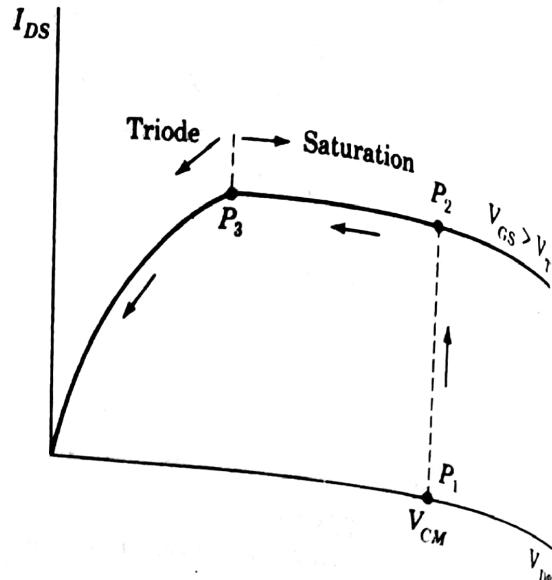


FIGURE 8.8-1
The operating path of T_1 as the capacitor C_L discharges.

transistor and capacitor follow the triode characteristic, eventually to zero voltage.

Using Eq. (8.1-3), which applies in the saturation region, we find

$$V_C(t) = V_{CM} - \frac{I_{DS}}{C_L} t = V_{CM} - \frac{k_D}{C_L} (V_{GS} - V_T)^2 t \quad (8.8-1)$$

The point P_3 is reached when $V_{DS} = V_{GS} - V_T$. The time $t \equiv t_{sat}$ at which $V_C(t) = V_{GS} - V_T$ is calculated from Eq. (8.8-1) to be

$$t_{sat} = \frac{C_L}{k_D} \left[\frac{V_{CM} - V_{GS} + V_T}{(V_{GS} - V_T)^2} \right] \quad (8.8-2)$$

In the triode region, using Eq. (8.1-2), we find

$$\frac{dV_C}{dt} = -\frac{I_{DS}}{C_L} = -\frac{k_D}{C_L} [2(V_{GS} - V_T)V_C - V_C^2] \quad (8.8-3)$$

The time in the triode region required for $V_C(t)$ to fall from V_{CM} to $0.1V_{CM}$ is

$$t_{triode} = - \int_{V_{CM}}^{0.1V_{CM}} \frac{dV_C}{(k_D/C_L)[2(V_{GS} - V_T)V_C - V_C^2]} = \frac{1.15C_L}{k_D(V_{GS} - V_T)} \quad (8.8-4)$$

Let us take $k_D = 1 \text{ mA/V}^2$. Since we have assumed $k_L = 0.5 \mu\text{A/V}^2$, we have $k_D/k_L = 50$, which is quite reasonable for an MOS gate. Assume also that $V_{CM} = 10 \text{ V}$, $V_{GS} - V_T = 5 \text{ V}$, and $C_L = 5 \text{ pF}$. Then we calculate from Eqs. (8.8-2) and (8.8-4) that $t_{sat} = 0.004 \mu\text{s}$ and $t_{triode} = 0.016 \mu\text{s}$. The total fall time is then $t_f = 20 \text{ ns}$. The fall time is thus very appreciably smaller than the rise time previously calculated to be $0.2 \mu\text{s}$. The principal reason for the large difference is the fact that $k_D \gg k_L$ because the channel of the driver is much wider and shorter than the channel of the load transistor.

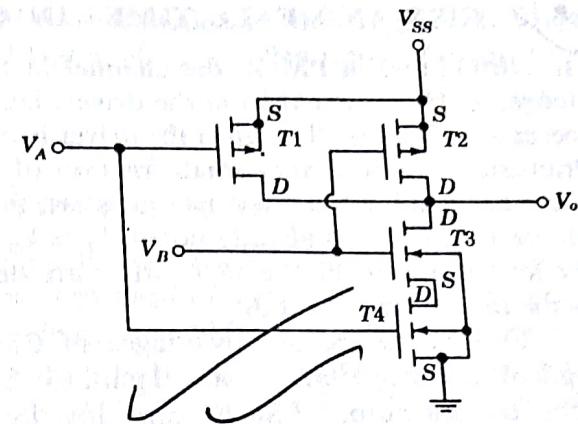


FIGURE 8.9-1
A CMOS NAND gate.

8.9 THE CMOS GATE

A two-input CMOS NAND gate is shown in Fig. 8.9-1. Note that the driver transistors are series-connected while the load transistors are paralleled. The individual input is applied simultaneously to a pair of transistors, one driver and one load. Assuming positive logic and taking logic 0 to be nominally ground voltage and logic 1 to be nominally V_{ss} , we can easily verify that the circuit is indeed a NAND gate. The output V_o will be at logic 0 (ground) only when both NMOS driver transistors are ON, in which case both PMOS load transistors will be OFF. The circuit in Fig. 8.9-2, in which the load devices are in series and the driver transistors in parallel, is readily verified to be a NOR gate.

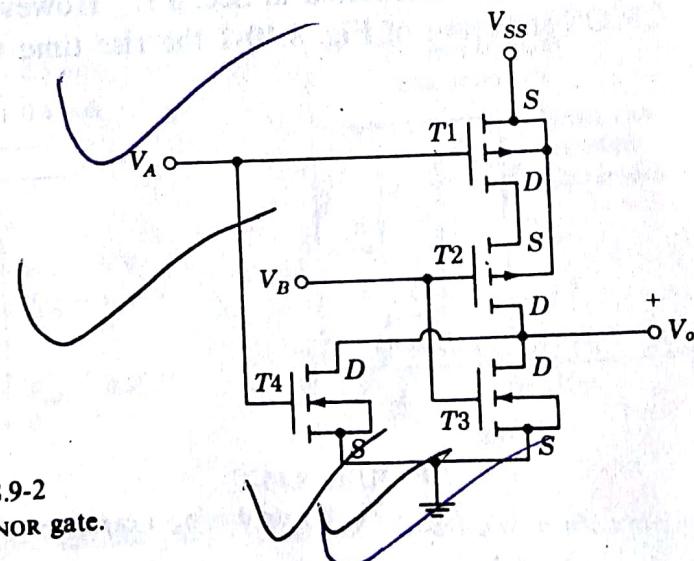


FIGURE 8.9-2
A CMOS NOR gate.

8.10 RISE AND FALL TIMES IN CMOS GATES

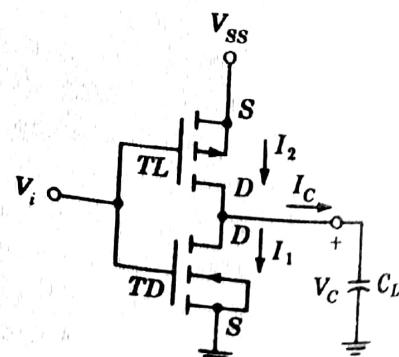
In NMOS and in PMOS the channel in the load devices must be very much longer and narrower than in the driver; i.e., we require $k_L \ll k_D$. This design is necessary to assure that when the driver is ON, the voltage drop across the driver transistor will be a very small fraction of the supply voltage. In CMOS, on the other hand, when the driver goes ON, the load is simultaneously driven OFF. Hence in CMOS, as already noted, $k_L \approx k_D$. Since the driver is *n*-channel and the load is *p*-channel, the *W/L* ratios are designed to be inversely proportional to the ratio of their mobility.

One of the major advantages of CMOS, then, is that there is always available a low-resistance channel path to charge and discharge a capacitive load across the gate output. Such a capacitive load at the output of a CMOS inverter is shown in Fig. 8.10-1. In an NMOS or in a PMOS gate the charging times of the load capacitor are widely different because of the difference in driver and load channel geometries, that is, k_D and k_L . In CMOS on the other hand, the charge and discharge times are quite comparable.

The capacitor-charging calculations given above for the NMOS inverter can be applied directly in the present case. Referring to Fig. 8.10-1, suppose that *TL* has been ON and *TD* OFF, so that $V_C = V_{SS}$. The capacitor C_L will now discharge when the input gate voltage (from the output of a preceding gate) goes to $V_{GS} = V_{SS}$, thereby turning *TL* OFF and *TD* ON. The fall time t_f is $t_f = t_{sat} + t_{triode}$, as given by Eqs. (8.8-2) and (8.8-4). In these equations $V_{C(\max)} = V_{GS} = V_{SS}$. We may note that actually these equations apply somewhat more exactly in the present case than in the case (of NMOS or PMOS) for which they were derived; for it will be recalled that in that derivation we neglected the small charging current through the load transistor while the capacitor was discharging. In the present case there is no current through the load device.

The parameter *k* for both the load and driver transistors will be comparable to the value of *k* ordinarily designed into the driver transistor of a gate without complementary symmetry. Hence in CMOS the fall time will be comparable to the fall time calculated in Sec. 8.7. However, because of the symmetry of the CMOS structure of Fig. 8.10-1 the rise time will be the same as the fall time if

FIGURE 8.10-1
A CMOS gate driving a capacitor C_L .



$k_L = k_D$. As can be verified, except for the direction in which $V_L(t)$ is changing, the equations which described the charging of C_L are identical to the equations which describe its discharge.

8.11 MANUFACTURER'S SPECIFICATIONS

The specifications provided by manufacturers for CMOS devices are similar to those provided for BJT gates. These specifications deal with input and output currents and voltages, propagation delays, rise and fall times, etc. The specifications for the Motorola 4012 low-power NAND gate are given in Fig. 8.11-1. (Specifications for the type 4001 NOR gate are identical.) These specifications apply under the circumstances that a supply voltage $V_{SS} = 5$ V is employed. The definitions of the parameters V_{iH} , V_{iL} , V_{oH} , V_{oL} are given in Sec. 4.11. The current I_{iH} stands for the minimum current which must be supplied by a driving source if the CMOS-gate input is to be held at a voltage high enough, i.e., at V_{iH} or higher, for the gate to acknowledge that its input is at logic level 1. The other current symbols have similar meanings. Observe that the input currents, being only of the order of 10 pA, may generally be ignored. Referring to the output specifications, we note that the gate output is able to sink a maximum of 0.4 mA and still stay low enough in voltage to remain in the logic 0 region, that is, $V_o \leq V_{oL}$. At the other end, the gate is able to serve as a source of at most 0.5 mA and still remain in the logic 1 region.

Figure 8.11-1b shows the worst-case transfer characteristics of the gate. The transfer characteristic of a typical gate will lie somewhere between the worst-case limiting plots shown. With a 5-V supply the manufacturer specifies that

$$\Delta 0 = V_{iL} - V_{oL} \approx 1.5 \text{ V} \quad (8.11-1a)$$

$$\Delta 1 = V_{oH} - V_{iH} \approx 1.5 \text{ V} \quad (8.11-1b)$$

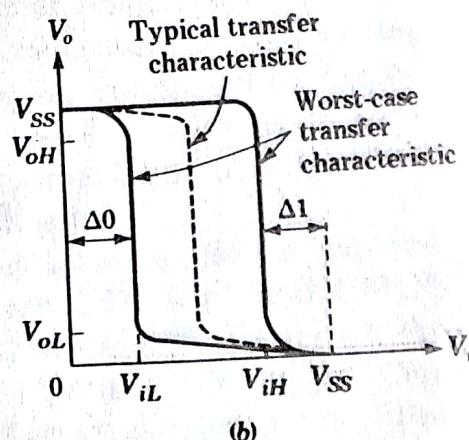
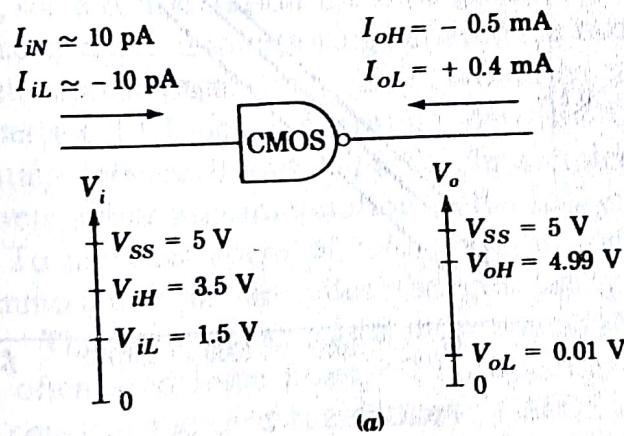


FIGURE 8.11-1
(a) Some manufacturer's specifications. (b) Typical worst-case transfer characteristic.

Table 8.11-1

	Time, ns
$t_{pd} (LH)$	30
$t_{pd} (HL)$	30
t_r	60
t_f	60

If we inquire about typical rather than worst-case noise immunities, we find that with $V_{ss} = 5$ V, these are about 2.25 V. Thus the noise immunities are larger than those encountered in BJT gates with comparable supply voltages.

The propagation delay and transition times for the Motorola gates are given in Table 8.11-1, assuming a load consisting of a capacitor $C_L = 15$ pF in parallel with a resistor $R_L = 200$ k Ω . We assume that the input impedance of a CMOS gate is a 5-pF capacitance, so that this capacitive load represents a fan-out of 3. Increasing the fan-out increases the delay and transition times linearly, since the circuit time constant is directly proportional to the total load capacitance.

The CMOS gate referred to in Table 8.11-1 is relatively slow. However, high-speed CMOS gates are available having propagation delay times of the order of 20 ns. Recent advances in fabrication techniques, in which the substrate material employed is *sapphire*, have resulted in fast, low-power CMOS gates. These gates, called SOS/CMOS (silicon-on-sapphire) have rise and fall times

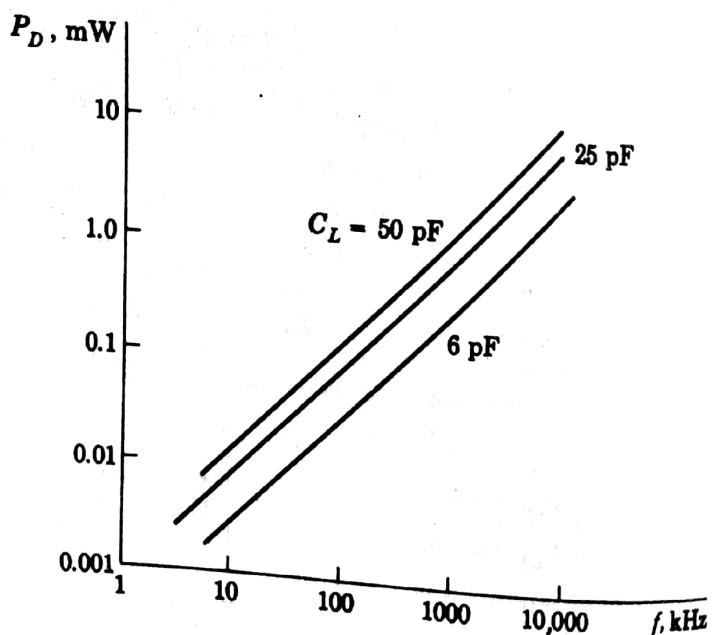


FIGURE 8.11-2
Power dissipation as a function of frequency.

which are less than 20 ns. High-speed CMOS gates are today comparable to DTL gates, and a few claim to be as fast as the slower versions of TTL.

Typical quiescent power dissipation of these low-power CMOS units is 50 nW. However, at 100-KHz operation, the power dissipated is approximately $30 \mu\text{W}$. This dissipation increases at the rate of 20 dB/decade and is also a function of the capacitive load. Plots of the dissipation of the Motorola 4012 are shown in Fig. 8.11-2. (See also Sec. 1.15.)

Buffers The CMOS gate is a low-current gate designed to drive other CMOS gates. When a CMOS gate is to be used to drive a TTL or DTL gate, a CMOS buffer is often employed. An RCA 4009A buffer is capable of sinking a load current of 4 mA (I_{OL}) when it is in the low state and can source 1.75 mA when the output voltage is 2.5 V. This 2.5-V value is specified since it represents V_{IH} for many TTL gates.

The circuit configuration of the buffer is the same as the standard gate configuration. However, to obtain extra current capability, the dimensions of the CMOS devices used are increased.

8.12 INTERFACING BJT AND CMOS GATES

MOS gates and (at the present writing) CMOS gates are slower than BJT gates. On the other hand, the MOS gates can be fabricated with a component density on a silicon die which exceeds that possible with BJT devices. There is a merit to conserving "real estate" on the silicon die. For as the area involved on the die increases, so does the likelihood that a crystal imperfection will render the device defective, i.e., a reject. As a result there is an advantage in using both BJT and MOS devices in combination. The BJT devices are used where speed is required, and the MOS devices are used where slower operation is allowed. Since BJT and MOS devices generally operate at different voltage and current levels, some consideration must be given to proper interfacing. In many cases, interfacing requires interposing between the two types of devices circuits involving discrete bipolar transistors. Such would be the case, for example, if we needed to interface ECL logic operating between 0 and -5.2 V with MOS logic operating between 0 and $+14$ V. In simpler cases, interfacing may require a relatively minor accommodation in the BJT or MOS gate.

To illustrate some of the measures which may serve to effect a required accommodation we consider the interfacing between TTL logic and CMOS logic. This is the most common interfacing employed since TTL gates are the most often used logic types. TTL operates from a positive 5-V supply. We shall consider the case in which the CMOS as well operates from +5 V. For this purpose we have listed in Fig. 8.12-1 some of the relevant specifications for the TTL gate.

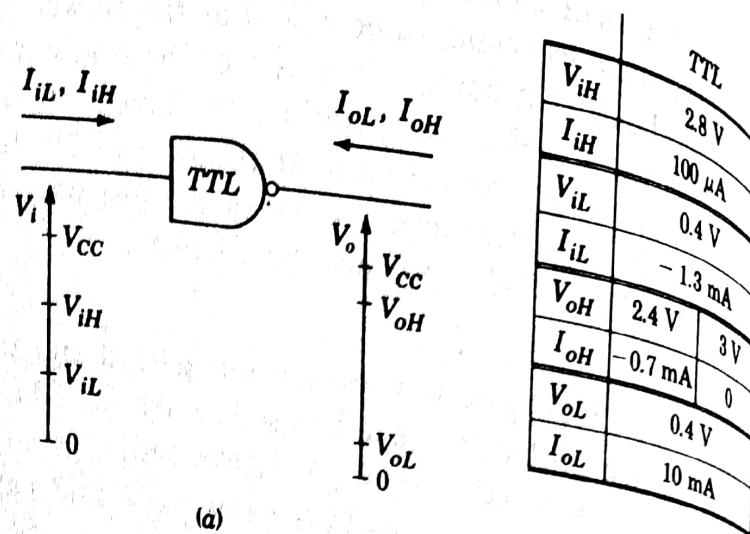


FIGURE 8.12-1
(a) Input and output characteristics. (b) Table of typical values.

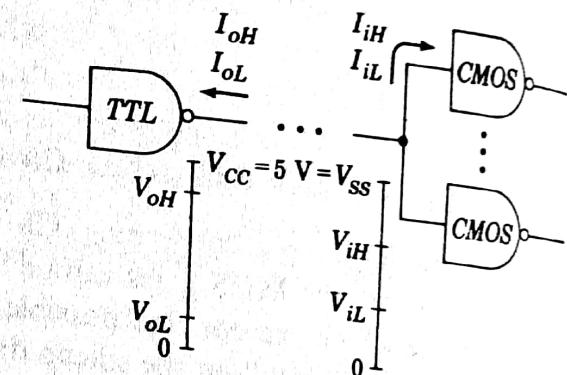


FIGURE 8.12-2
A TTL gate driving N CMOS gates.

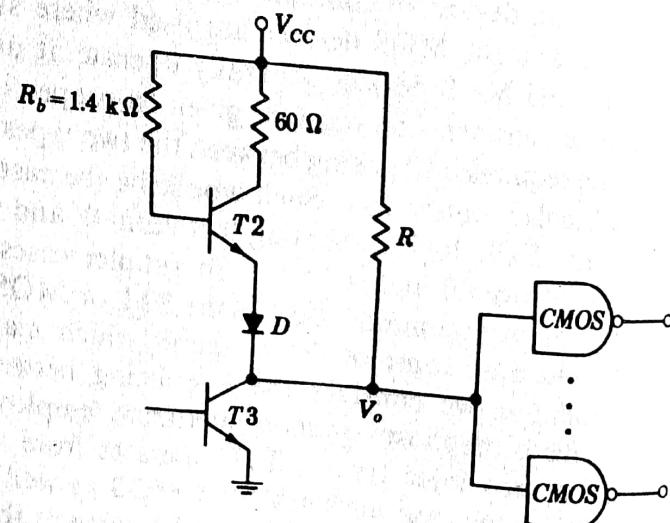


FIGURE 8.12-3
TTL with a passive pull-up.

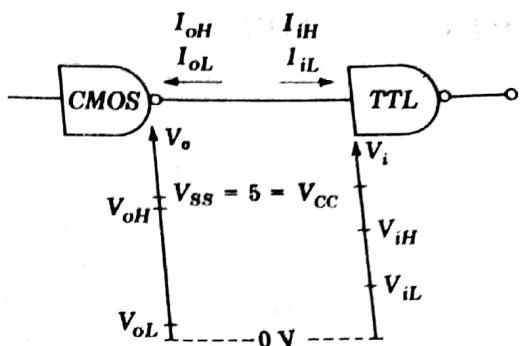


FIGURE 8.12-4
CMOS driving TTL.

We consider first the case of a TTL gate driving N CMOS gates, as in Fig. 8.12-2. Thus, the fan-out is N . For such an arrangement to operate successfully it is required that

$$-I_{oH}(\text{TTL}) \geq NI_{iH}(\text{CMOS}) \quad (8.12-1a)$$

$$I_{oL}(\text{TTL}) \geq -NI_{iL}(\text{CMOS}) \quad (8.12-1b)$$

$$V_{oL}(\text{TTL}) \leq V_{iL}(\text{CMOS}) \quad (8.12-1c)$$

$$V_{oH}(\text{TTL}) \geq V_{iH}(\text{CMOS}) \quad (8.12-1d)$$

As is readily verified from the data in Fig. 8.11-1 and in Fig. 8.12-1, Eqs. (8.12-1a) and (8.12-1b) are satisfied for any reasonable fan-out N . In addition, Eq. (8.12-1c) is also satisfied. However we find that Eq. (8.12-1d) is not satisfied since, even at "no load" $V_{oH}(\text{TTL}) = 3\text{ V}$ while $V_{iH}(\text{CMOS}) = 3.5\text{ V}$. A frequently employed circuit modification used to raise $V_{oH}(\text{TTL})$ above 3.5 V is shown in Fig. 8.12-3 where an external resistor R has been bridged between V_{cc} and the output. Typically R is in the range 2 to 6 k Ω .

When we consider a CMOS gate driving a TTL gate (see Fig. 8.12-4), we find that the condition $I_{oL}(\text{CMOS}) \geq -NI_{iL}(\text{TTL})$ is not satisfied even for $N = 1$ since $I_{oL}(\text{CMOS}) = 0.4\text{ mA}$ while $-I_{iL}(\text{TTL}) = 1.34\text{ mA}$. There are available, however, a number of CMOS buffers having adequate available output current (up to 6 mA)

REFERENCES

- 1 Carr, W. N., and J. P. Mize: "MOS/LSI Design and Applications," McGraw-Hill, chap. 4, 1972.
- 2 RCA Solid State Databook: SSD-203B, COS/MOS Digital Integrated Circuits.