Thus from the third column of Fig. 3.1-19*b*,

$$C_{GD} = 1.58\,\text{fF}$$

$$C_{GS} = 24.9\,\text{fF}$$

$$C_{BG} = 0$$

$$C_{BD} = 19.8\,\text{fF}$$

$$C_{BS} = 44.0\,\text{fF}$$

Note the small size of these parasitic capacitors relative to those associated with discrete electronic circuitry. Although they are small, it will be seen later that the effects of these parasitics are significant at high frequencies.

3.1.4 Measurement of MOSFET Model Parameters

Typical values of process parameters are generally available at the start of a design. Knowledge of the process parameters themselves, as well as of statistical and/or worst case information, is necessary during the design stage so that tradeoffs between performance, complexity, and yield can be considered by the designer. Experimentally measured process parameters also serve as an interface between the fabrication and design groups. A check to verify that the process parameters fall within previously agreed-upon bounds is typically made after fabrication and prior to extensive circuit testing by the designers. These measurements are made over a large enough number of dies to establish confidence in the measurements.

Automated device characterization equipment is readily available for measuring many of the parameters that are used to characterize devices. In the absence of this equipment, some of the more important parameters can be measured with reasonable accuracy using standard laboratory instruments. Techniques for measuring the parameters K', V_{T0}, γ, and λ that appear in the model of the MOSFET previously introduced are discussed in this section. Since models are not perfect, it is advisable to measure model parameters in the same regions where the device will be operated and near the intended operating point if possible. The following discussion assumes that the reader is familiar with the basic operation of an operational amplifier. For those lacking background in this area, most electronics texts have a brief, self-contained section discussing operational amplifiers (e.g., References 4 and 5).

Measurement of λ. Consider the expression for the drain current of the MOSFET in the saturation region as given in Table 3.1-1:

$$I_D = \frac{K'W}{2L}(V_{GS} - V_T)^2(1 + \lambda V_{DS}) \qquad \text{for} \quad V_{DS} > V_{GS} - V_T, V_{GS} > V_T \tag{3.1-18}$$

This equation is plotted in Fig. 3.1-21 for different values of $(V_{GS} - V_T)$ with constant V_{BS}. Projection of these curves into quadrant 2 shows that all curves

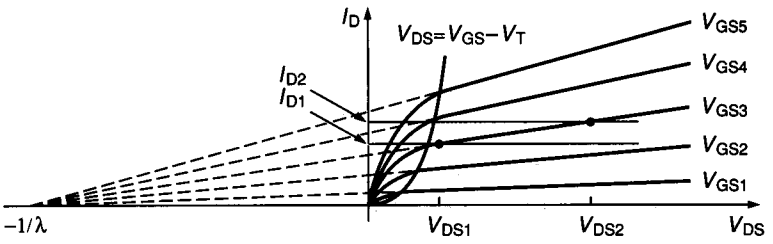
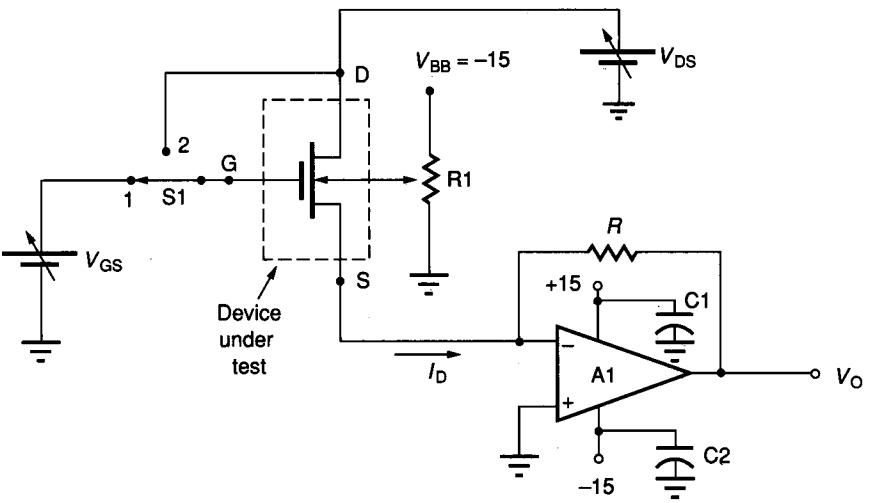


FIGURE 3.1-21
Measurement of λ.

intersect the V_{DS} axis at $V_{DS} = -1/\lambda$. The parameter λ can thus be determined by fixing V_{GS} and V_{BS}, thus fixing $(V_{GS} - V_T)$, and determining the $-V_{DS}$ axis intercept of I_D versus V_{DS}. If I_{D1} and I_{D2} are the currents corresponding to two distinct values of V_{DS}, V_{DS1}, and V_{DS2} (see Fig. 3.1-21), then it readily follows that

$$\lambda = \frac{I_{D2} - I_{D1}}{I_{D1}V_{DS2} - I_{D2}V_{DS1}} \tag{3.1-19}$$

The circuit of Fig. 3.1-22 with S1 in position 1 is useful for measuring λ. V_{GS} and V_{BS} are fixed so that $V_{GS} > V_T$. The operational amplifier (op amp) is used to establish the relationship $V_O = -I_D R$ while maintaining a source voltage of nearly 0 V. If V_{DS1} and V_{DS2} are the two distinct values of V_{DS} required in



A 356 op amp is a reasonable choice for A1. C1 and C2, which may be needed for stability, should be as close to the supply terminals of the op amp as possible to prevent oscillation. A value of 1μF is reasonable for these capacitors.

FIGURE 3.1-22
Circuit for measuring MOSFET parameters. V_{GS}, V_{DS}, and V_{BB} should be of opposite polarity when measuring p-channel parameters.

(3.1-19) and V_{O1} and V_{O2} respectively are the output voltages of the op amp for these inputs, it follows that

$$\lambda = \frac{V_{O2} - V_{O1}}{[V_{O1}V_{DS2} - V_{O2}V_{DS1}]} \tag{3.1-20}$$

Measurement of V_T and V_{T0}. With V_{BS} and V_{DS} fixed, the values of I_D (I_{D1} and I_{D2}) corresponding to two distinct values of V_{GS} (V_{GS1} and V_{GS2}) can be measured. If the device is operating in saturation for both values of V_{GS}, a routine calculation shows

$$V_T = \frac{V_{GS2}(I_{D1}/I_{D2})^{1/2} - V_{GS1}}{(I_{D1}/I_{D2})^{1/2} - 1} \tag{3.1-21}$$

If V_{BS} is set to 0 V, it follows by definition that the expression in (3.1-21) yields V_{T0}. The circuit of Fig. 3.1-22 with S1 in position 1 can again be used to measure V_T.

Measurement of K'. Once λ and V_T have been measured, it follows from (3.1-18) that K' can be determined from a measurement of I_D for fixed values of V_{GS} and V_{DS} from the expression

$$K' = \frac{2LI_D}{W(V_{GS} - V_T)^2(1 + \lambda V_{DS})} \tag{3.1-22}$$

The circuit of Fig. 3.1-22 is again useful for this measurement.

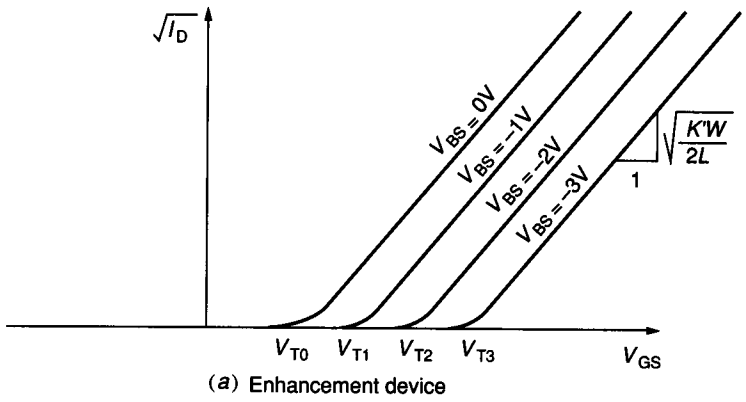
Another scheme is often used for measuring K' and V_T. Since λV_{DS} is typically much less than 1, it follows from (3.1-18) that

$$I_D \simeq \frac{K'W}{2L}(V_{GS} - V_T)^2 \qquad \text{for} \quad V_{GS} > V_T, V_{DS} > V_{GS} - V_T$$

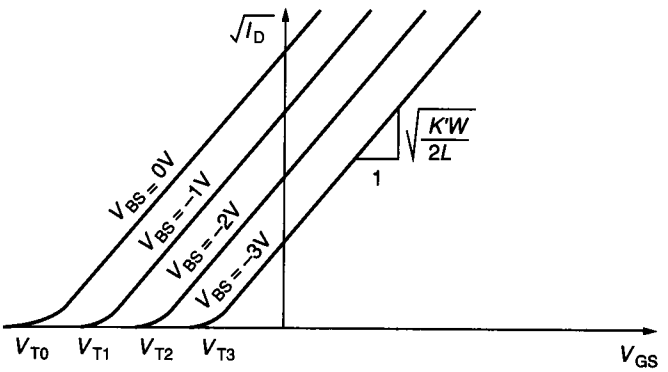
To guarantee operation in the saturation region for enhancement devices, it is convenient to tie the gate to the drain by setting switch S1 to position 2 and for depletion devices to establish a large dc bias on V_{DS} with S1 in position 1, yielding

$$I_D = \frac{K'W}{2L}(V_{GS} - V_T)^2 \qquad V_{GS} > V_T \tag{3.1-23}$$

A typical plot of $\sqrt{I_D}$ versus V_{GS} is shown in Fig. 3.1-23 for both enhancement and depletion MOSFETs for several different values of V_{BS}. In all cases the slopes of the straight lines are $[(K'W)/(2L)]^{1/2}$, enabling a ready determination of K'. Note the bending of the curves for low values of V_{GS}. This bending is due to subthreshold currents and is not included in the models discussed up to this point. A discussion of subthreshold operation appears in Section 3.1.6. The drain current must be large enough so that the slope is not affected by the subthreshold effects when measuring V_T. The horizontal axis intercept of the projection of the straight lines is the threshold voltage. The intercept corresponding to $V_{BS} = 0$ is V_{T0}. The circuit of Fig. 3.1-22 is useful for measuring V_T and K' following this approach.



(*a*) Enhancement device



(*b*) Depletion device

FIGURE 3.1-23
Alternate method of measuring K' and V_T: (*a*) Enhancement device, (*b*) Depletion device.

Measurement of γ. From Table 3.1-1 recall that

$$V_T = V_{T0} + \gamma\left(\sqrt{\phi - V_{BS}} - \sqrt{\phi}\right) \tag{3.1-24}$$

Since methods of measuring both V_T and V_{T0} have been presented, both γ and ϕ can be determined by obtaining values of V_T from two distinct nonzero values of V_{BS}. Details about obtaining these parameters from either of these two distinct values or by applying some minimum mean squared error approximation from several measured values is left to the reader.

For large values of V_{BS}, the threshold voltage is only weakly dependent upon ϕ. In that case, using a typical value of ϕ, ϕ_{typ}, a good approximation to γ can be obtained with a single large value of V_{BS} from the expression

$$\gamma \simeq \frac{V_T - V_{T0}}{\sqrt{\phi_{typ} - V_{BS}} - \sqrt{\phi_{typ}}} \tag{3.1-25}$$

Information about the parasitic capacitance parameters is often required for accurate time domain analysis of MOS circuits. The parasitic capacitors discussed in the previous section are quite small for minimum-sized MOSFETs (well under 1 pF). Some were also seen to be voltage dependent. Due primarily to their small size, it is nearly impossible to measure directly these capacitances with any acceptable degree of accuracy using standard laboratory equipment. Special test circuits are often included on an IC for the purpose of obtaining information about the parasitic capacitances. These test circuits may contain either geometrically large devices so that the total parasitic capacitance is large enough to measure easily or they may contain more complicated circuit structures (such as ring oscillators) from which these parasitics can be calculated indirectly from measurements of other characteristics of the circuits.

Other measurement considerations. The methods presented above provide simple measurement of some of the key parameters of a MOSFET. Questions about such problems as accuracy and Q-point were avoided. Regardless of what measurement technique is used, it is advisable to measure these parameters as close to the intended device operating point (Q-point) as possible because model inaccuracies can cause significant global variations in these parameters whereas local variations are usually quite small. Also, measurements at several data points in the neighborhood of the Q-point, followed by data smoothing or curve fitting, usually give better results than single-point or two-point measurements. This is easily achieved if computer-controlled measurement equipment is available. The extension of these measurement methods to smoothing or curve fitting is straightforward and is left to the reader.

Finally, the question of how accurately one has measured these parameters and how much accuracy is needed naturally arises. Unfortunately, the functional form of the model presented here is not perfect, and thus the parameters, along with the functional form of the model, only approximate the device performance. Because of this imperfect fit, it is inherent that different measurement algorithms will give slightly different values for these parameters. This problem is analogous to the problem of trying to fit a second-order polynomial to data points obtained from a third-order polynomial. Nevertheless, the model for the MOSFET is quite good, and the measured parameters are useful for predicting device performance. In applications where matching or ratio matching of parameters is particularly important, the same parameter measurement algorithms should be employed for each to avoid introducing systematic errors.

3.1.5 Short Channel Devices

Downward trends in device scaling have caused minimum MOSFET device lengths to decrease from the 10 μ range of the mid-1970s to the 1 μ range of the mid-1980s. Device lengths in the 0.5 μ range or below are projected by the early 1990s. The models that we discussed earlier in this chapter are quite good for channel lengths longer than 5 μ. The models gradually deteriorate for shorter channel lengths. Transistors with channel lengths less than 3 to 5 μ are termed

short channel devices. With short channel devices the ratio between lateral and vertical dimensions is reduced. Geometrically, the channel region changes from a rather uniform, thin right-rectangular region to a much more irregular structure. The effects of the transition region from the drain and source diffusions to the channel become significant, causing need for increased complexity in the device model.

Short channel transistors offer some significant advantages over larger transistors but also have serious limitations. The major advantages are the reduced area requirements and improvements in speed that are attainable with circuits employing short channel transistors. The improvements in speed are mostly attributable to the reduced input capacitance associated with the smaller devices. The major limitation is a deterioration in the output impedance characteristics. Good matching of short channel transistor characteristics is also more difficult to achieve.

The typical output characteristics of a short channel transistor are compared with that of a longer channel device in Fig. 3.1-24. A simple empirical model of the MOSFET which is essentially an extension of the long channel model summarized in Table 3.1-1 is given by the following equations:

$$I_G = 0 \tag{3.1-26}$$

$$I_D = \begin{cases} 0 & V_{GS} < V_T \text{ (cutoff)} \\ \dfrac{K'W_{eff}}{L_{eff}}\left(V_{GS} - V_T - \dfrac{V_{DS}}{2}\right)V_{DS}\left[1 + \lambda\left(\dfrac{1 + \theta L_{eff}}{\theta L_{eff}}\right)V_{DS}\right] & \\ \qquad V_{GS} > V_T,\ 0 < V_{DS} < V_{GS} - V_T \text{ (ohmic)} & (3.1\text{-}27) \\ \dfrac{K'W_{eff}}{2L_{eff}}(V_{GS} - V_T)^2\left[1 + \lambda\left(\dfrac{1 + \theta L_{eff}}{\theta L_{eff}}\right)V_{DS}\right] & \\ \qquad V_{GS} > V_T,\ V_{DS} > V_{GS} - V_T \text{ (saturation)} & (3.1\text{-}28) \end{cases}$$

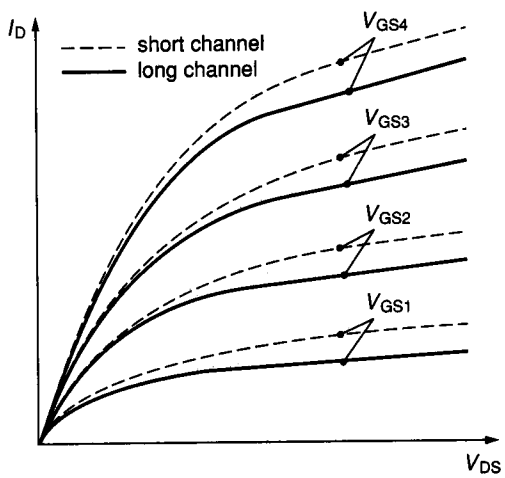


FIGURE 3.1-24
Short channel effects on MOSFET performance. Solid curves for 5 μ device, dashed curves for 1 μ device.

The parameter θ has units (length)$^{-1}$. A typical value of θ may be 0.2 μ^{-1}. The parameters W_{eff} and L_{eff} denote the effective channel width and length respectively. These are given by

$$W_{eff} = W - W_R \tag{3.1-29}$$

$$L_{eff} = L - 2L_D - L_R \tag{3.1-30}$$

where W and L are the drawn width and length, W_R and L_R are constants representing width and length reduction due to processing, and L_D is the lateral diffusion of the source or drain under the gate. W_{eff} and L_{eff} could also be used to model W and L in the model of the long channel MOSFET presented earlier and are, in fact, included in most good models used for simulation. The use of W_{eff} and L_{eff} for hand analysis of large structures is not so important since the difference between these values and the drawn width and length is relatively small. For short channel devices, however, the relative effects of the width and length reduction and the lateral diffusions become more pronounced. Short channel effects should be considered for channel lengths less than 5 μ. This model is still usable in hand calculations. It should be noted that a size adjust of the appropriate mask is often done so that the effective length and/or width ideally agrees with the drawn length and/or width. The designer should be aware of what size adjusts, if any, will be used at mask generation.

More accurate models of the short channel effects are available.[6–7,19] These models are much more involved and not amenable to hand calculations but do give reasonable results when used for computer simulations. The model in Reference 7 is incorporated into the popular circuit simulator, SPICE. For the reasons discussed below, a simple model such as that of (3.1-27) and (3.1-28) is often adequate for hand analysis of circuits utilizing short channel transistors.

Were it not for the additional notational complexity, the model presented in Table 3.1-1 could be replaced with the short channel model introduced in this section. Note, in particular, that for large L the short channel model is essentially equivalent to that for the longer transistors. It can be concluded that the short channel model can be appropriately used for short as well as long device sizes if desired. The empirical short channel model parameter θ was included in Table 3.1-2.

We close this section by comments about where short channel devices are used. Short channel devices are widely used in digital logic, where a net improvement in speed and reduction in circuit area are readily achievable in spite of the performance limitations of the short channel MOSFET. In analog applications, where good matching characteristics and high output impedance are often critical, short channel transistors find fewer applications in spite of the improved frequency response. When analog circuits are built in processes where short channel lengths are available, most devices will be somewhat larger than minimum size; some performance benefits in terms of device matching will, however, be derived in these processes because of the more stringent lithography requirements that must be maintained in the small feature size processes.

3.1.6 Subthreshold Operation

In the dc model of the MOSFET introduced in Sec. 3.1.1, the drain current for positive V_{DS} was assumed to be zero for $V_{GS} < V_T$ and nonzero for $V_{GS} > V_T$ as indicated by Eq. 3.1-1. In physical devices such an abrupt transition is not anticipated and does not occur experimentally. The drain current is, however, much smaller for $V_{GS} < V_T$ than for $V_{GS} > V_T$ and hence in most applications the assumption that $I_D = 0$ for $V_{GS} < V_T$ is justifiable.

Applications do exist, however, where it is crucial that current levels be extremely small. These include, but are not limited to, biomedical applications such as pacemakers and other implantable devices that must operate for several years with small nonrechargeable batteries. Useful circuits with extremely low supply currents in which $V_{GS} < V_T$ find applications in such situations. If $V_{GS} > V_T$, the devices are said to be operating in strong inversion. If $V_{GS} < V_T$, the devices are said to be operating in weak inversion, or equivalently, in the subthreshold region. At room temperature, the transition between strong inversion and weak inversion actually occurs around $V_{GS} \simeq V_T + 100$ mV. The expression

$$V_{GS} = V_T + 2n V_t \tag{3.1-31}$$

where n is a constant between 1 and 2, can be used to predict the transition at other temperatures.[8] The term V_t is equal to kT/q where k is Boltzmann's constant ($k = 1.387 \times 10^{-23}$ V · C/°K), T is the device temperature in degrees Kelvin, and q is the charge of an electron ($q = 1.6 \times 10^{-19}$ C). At room temperature, $V_t = 26$ mV.

A plot of I_D versus V_{GS} for the MOSFET with $V_{DS} = V_{GS}$ is shown in Fig. 3.1-25 for a wide range of V_{GS} values. As can be seen from this figure, the current in weak inversion is much smaller than the current in strong inversion, as previously stated. The behavior is, however, well characterized far into weak inversion and varies exponentially rather than quadratically with V_{GS}. The potential for extremely low power dissipation should be apparent.

Example 3.1-6.

(*a*) Compare the power dissipation required for a single device operating with $V_{GS} = 3V_T$ driven from a 5 V source to the power dissipation required for the same device operating from a 1 V supply biased at $V_{GS} = 0.85V_T$. Assume the device has characteristics as shown in Fig. 3.1-25.

(*b*) If ideal batteries with a total stored energy of 1 watt-hr are used to power the two devices, what is the battery life of each?

Solution.

(*a*) From Fig. 3.1-25*a*, it follows that with $V_{GS} = 3V_T$, $I_{D1} \simeq 150\ \mu$A, and from Fig. 3.1-25*b*, with $V_{GS} = 0.85V_T$, $I_{D2} \simeq 1$ nA. The power dissipation ratio is thus $(5I_{D1})/(I_{D2}) = 7.5 \times 10^5$.

(*b*) The battery life with the 5 V system would be

$$T = 1 \text{ watt-hr}/(150\ \mu\text{A} \cdot 5\text{ V}) = 1333 \text{ hr} \simeq 8 \text{ weeks}$$

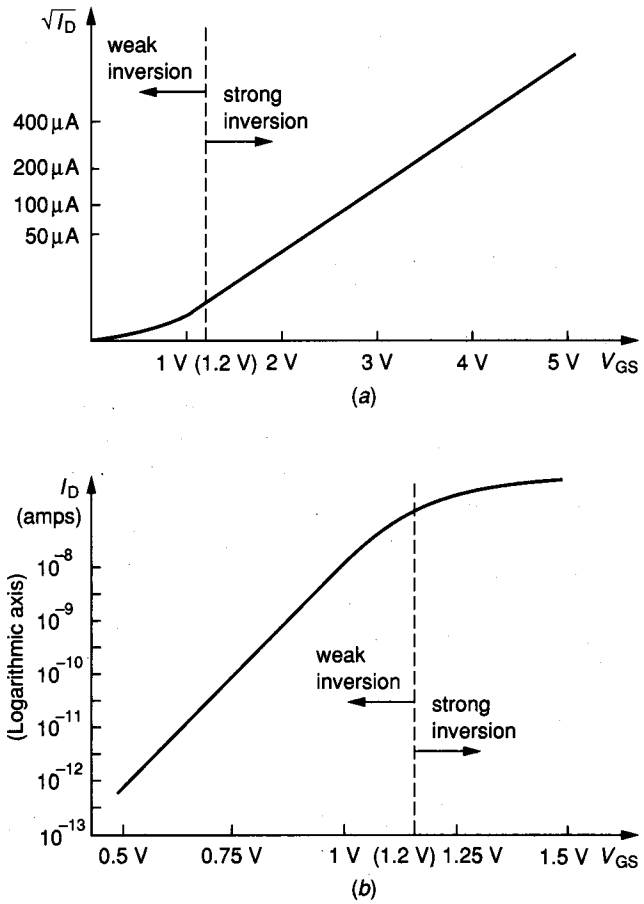


FIGURE 3.1-25
Typical I_D–V_{DS} characteristics for MOSFET operating in weak inversion: (*a*) Quadratic vertical axis, (*b*) Logarithmic vertical axis ($V_T = 1\ V$, $V_{DS} = V_{GS}$).

and the battery life with the 1 V system would be

$$T = 1 \text{ watt-hr}/[(1 \text{ nA}) \cdot 1 \text{ V}] \simeq 10^9 \text{ hr} \simeq 114,000 \text{ years}$$

It should be apparent that the subthreshold system may be practical whereas the strong inversion system would be impractical in many applications.

A dc model for the MOSFET operating in weak inversion is needed for both design and simulation. The following model is useful in weak inversion.[9]

$$\left.\begin{aligned} I_G &= 0 \\ I_D &= \frac{W}{L} I_{DO} e^{-V_{BS}[(1/(nV_t))-(1/V_t)]}(1 - e^{-V_{DS}/V_t}) e^{(V_{GS}-V_T)/(nV_t)} \end{aligned}\right\} \tag{3.1-32}$$

where V_T is the threshold voltage and n and V_t are as defined in Eq. 3.1-31. The constants I_{DO} and n are process parameters. Typical values for these parameters are $I_{DO} \simeq 20$ nA and $n = 2$.

A typical plot of I_D versus V_{DS} is shown in Fig. 3.1-26. Note that for $V_{DS} > 3V_t$, the term e^{-V_{DS}/V_t} in Eq. 3.1-32 becomes negligible, and one obtains the equivalent of the strong inversion saturation region of operation. Often $V_{BS} = 0$. Under the assumption that $V_{BS} = 0$ and $V_{DS} > 3V_t$, Eq. 3.1-32 simplifies considerably, to

$$\left.\begin{aligned} I_G &= 0 \\ I_D &\simeq \frac{W}{L} I_{DO} e^{(V_{GS}-V_T)/(nV_t)} \end{aligned}\right\} \tag{3.1-33}$$

It is often the case, however, that power supply restrictions will require operation of subthreshold devices with smaller values of V_{DS}.

The parameter I_{DO} is related to the transconductance parameter and is approximately given by the expression

$$I_{DO} \simeq \frac{K'2(nV_t)^2}{e^2}$$

This approximation for I_{DO} provides for continuity in the device model of (3.1-33) and (3.1-4) and in its derivative, $(\partial I_D/\partial V_{GS})$, at the strong inversion–weak inversion interface defined in (3.1-31)—see Problem 3.30.

There are several practical limitations of devices operating in weak inversion that deserve noting. First, the frequency response of devices operating far into weak inversion is poor. This can be qualitatively seen by observing that the parasitic device capacitances are geometrically determined and hence nearly equal to those discussed in the strong inversion model. The maximum current available to charge and discharge these capacitors is significantly less, causing a significant deterioration of frequency response. Second, the drain and source substrate currents associated with the reverse-biased moat–substrate junction are not necessarily negligible compared to subthreshold drain currents. Third, the linearity

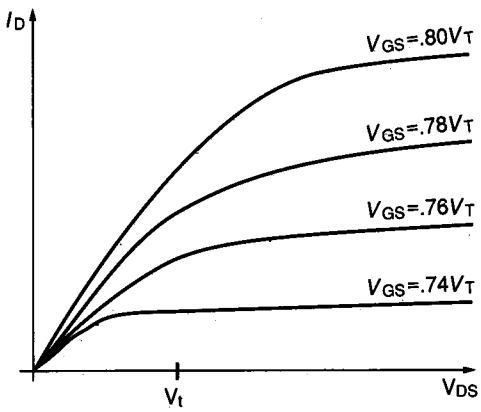


FIGURE 3.1-26
Typical I_D–V_{DS} characteristics for MOSFET operating in subthreshold.

is quite poor for $V_{DS} < 3V_t$, making linear designs more challenging. Fourth, the deterioration of matching characteristics of MOS transistors with decreasing drain currents further complicates linear design.

Because of these practical limitations, it is often desirable to operate near the transition region between strong inversion and weak inversion, where some of the benefits of reduced power associated with weak inversion can be derived but where these other limitations are not too problematic. Whereas diffusion current dominates weak inversion operation and drift current dominates strong inversion operation, both mechanisms interact in the transition region and thus complicate the modeling problem. A discussion of the modeling problem in this region is beyond the scope of this text and continues to receive attention in the technical literature. Even weak inversion models used in simulation programs such as SPICE have major limitations. Nevertheless, some designers do design useful circuits that operate in this transition region.

3.1.7 Operation in the Third Quadrant of the I_D–V_{DS} Plane

Although MOS devices are typically operated with $I_D > 0$ and $V_{DS} > 0$, (Quadrant 1 in Fig. 3.1-7), there are applications where the device is biased to operate in Quadrant 3 of the I_D–V_{DS} plane for some inputs. For notational convenience, a MOS transistor without drain and source designations is shown in Fig. 3.1-27, along with plots of I_1 versus V_{12} for different values of V_{32} and fixed V_B. The normal convention for nodal designations if the device were operating in Quadrant 1 would be ① → drain, ② → source, and ③ → gate. For operation in Quadrant 3, the drain and source designations would be reversed, based upon the device models discussed earlier in this chapter.

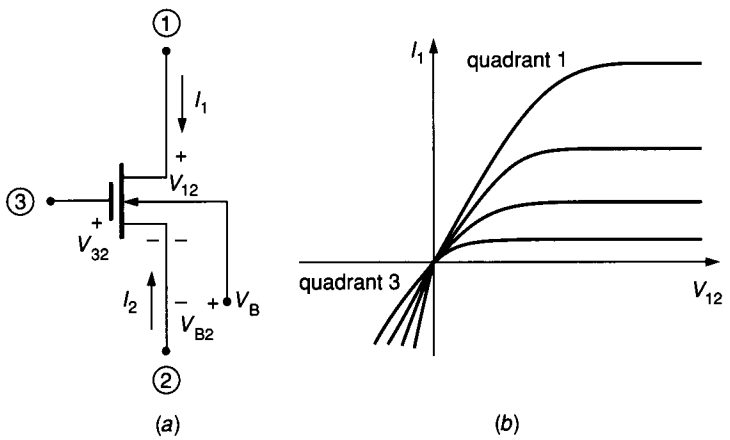


FIGURE 3.1-27
Operation of the MOSFET in the third quadrant: (*a*) Terminal convention, (*b*) I–V characteristics.

It should be apparent that the I–V characteristics do not display symmetry from Quadrant 1 to Quadrant 3. The asymmetry becomes more pronounced with decreasing reverse bulk bias. It should also be apparent that, based upon the device model introduced earlier in this chapter, the drain–source designations are not arbitrary.

To obtain a qualitative appreciation for how the device behaves in Quadrant 3, assume that $0 < V_{32} < V_T$ and that subthreshold currents are negligible. By the previous model, we are tempted to conclude that the device is cutoff since "$V_{GS} < V_T$" and hence conclude that $I_1 = 0$. If, however, the voltage at terminal ① in Fig. 3.1-27 were made negative relative to terminal ③ and if V_{13} is less than $-V_T$, it can be argued that the gate–channel voltage near terminal ① is greater than V_T, causing the formation of an inversion layer which would, in turn, cause a positive current I_2 to flow. This is what happens in an actual MOSFET. This biasing strategy is consistent with that discussed earlier in this chapter, provided the nodal designations ① → source, ② → drain, and ③ → gate are made.

For consistency with the model summarized in Table 3.1-1, the MOS device will be modeled in Quadrant 3 by using the model of Table 3.1-1 with the agreement that the designations *source* and *drain* will be made so that $V_{DS} > 0$. This means that in some applications, the drain and source designations change during normal operation of the device.

As an alternative to allowing the drain and source designations to change so that $V_{DS} \geq 0$ at all times, the designations can be arbitrarily fixed and the device model in Quadrant 3 can be defined to be equal to that which would be attained were the designations changed. Following this approach, the equivalent fixed drain–source designation model (for the n-channel transistor) summarized in Table 3.1-4 is obtained. The parameter λ has been included in the ohmic region expressions in this model to maintain continuity in the device characteristics. The regions of operation in the V_{DS}–V_{GS} plane are depicted in Fig. 3.1-28 for a fixed bulk bias.

Example 3.1-7. The device of Fig. 3.1-29 is to be used as an active grounded resistor. Verify that the I–V characteristics are continuous and differentiable at the origin ($V = 0$).

Solution. Since for small V, this depletion device is operating in either the forward ohmic or reverse ohmic regions, it follows from Table 3.1-4 that

$$I = K' \frac{W}{L} \left(-V_{Ti} - \frac{V}{2} \right) V (1 + \alpha_i \lambda V) \qquad i = 1, 2$$

where $i = 1$ for $V > 0$ and $i = 2$ for $V < 0$, where $\alpha_1 = +1$ and $\alpha_2 = -1$, and where $V_{DS} = V$.

To verify continuity, we must show that

$$\lim_{V \to 0^+} I = \lim_{V \to 0^-} I$$

From the expression for I,

$$\lim_{V \to 0^+} I = 0 = \lim_{V \to 0^-} I .$$

TABLE 3.1-4
Low-frequency fixed-terminal MOSFET model in four quadrants.

$$I_D = \begin{cases} 0 & V_{GS} < V_{T1} \\ & V_{DS} > V_{GS} - V_{T2} \end{cases} \quad \text{cutoff}$$

$$\frac{K'W}{2L}(V_{GS} - V_{T1})^2(1 + \lambda V_{DS}) \qquad \begin{matrix} V_{GS} > V_{T1} \\ V_{DS} > V_{GS} - V_{T1} \end{matrix} \quad \text{forward saturation}$$

$$\frac{K'W}{L}\left(V_{GS} - V_{T1} - \frac{V_{DS}}{2}\right)V_{DS}(1 + \lambda V_{DS}) \qquad \begin{matrix} 0 < V_{DS} < V_{GS} - V_{T1} \\ V_{GS} > V_{T1} \end{matrix} \quad \text{forward ohmic}$$

$$\frac{K'W}{L}\left(V_{GS} - V_{T2} - \frac{V_{DS}}{2}\right)V_{DS}(1 - \lambda V_{DS}) \qquad \begin{matrix} V_{GS} > V_{T2} \\ V_{DS} < 0 \end{matrix} \quad \text{reverse ohmic}$$

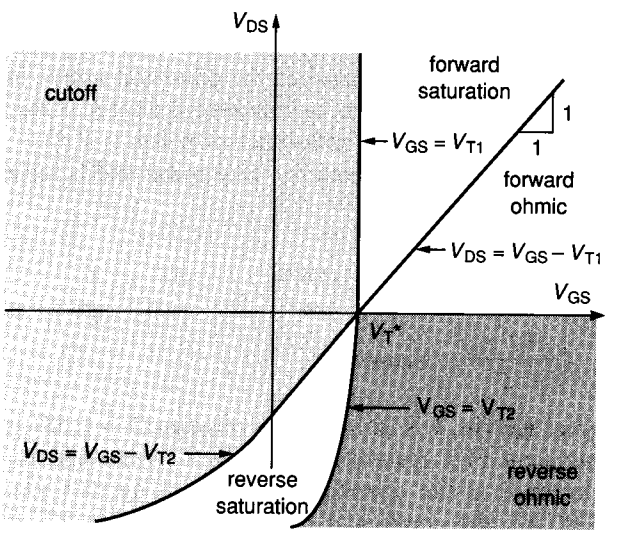
$$-\frac{K'W}{2L}(V_{GS} - V_{DS} - V_{T2})^2(1 - \lambda V_{DS}) \qquad \begin{matrix} V_{GS} < V_{T2} \\ V_{DS} < V_{GS} - V_{T2} \end{matrix} \quad \text{reverse saturation}$$

$$V_{T^*}{}^a = V_{T0} + \gamma\left(\sqrt{\phi - V_{BB}} - \sqrt{\phi}\right)$$

$$V_{T1} = V_{T0} + \gamma\left(\sqrt{\phi - V_{BS}} - \sqrt{\phi}\right)$$

$$V_{T2} = V_{T0} + \gamma\left(\sqrt{\phi - V_{BS} + V_{DS}} - \sqrt{\phi}\right)$$

[a] V_{T^*} characterizes the V_{DS} transition in Fig. 3.1-28.



$V_{T^*} = V_{T0} + \gamma\left(\sqrt{\phi - V_{BB}} - \sqrt{\phi}\right)$
$V_{T1} = V_{T0} + \gamma\left(\sqrt{\phi - V_{BS}} - \sqrt{\phi}\right)$
$V_{T2} = V_{T0} + \gamma\left(\sqrt{\phi - V_{BS} + V_{DS}} - \sqrt{\phi}\right)$

FIGURE 3.1-28
Model extension regions of operation of MOSFET (fixed V_{BS}).

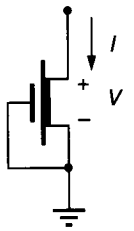


FIGURE 3.1-29
Grounded simulated resistor of Example 3.1-7.

To verify differentiability, we must show

$$\lim_{V\to 0^+} \frac{\partial I}{\partial V} = \lim_{V\to 0^-} \frac{\partial I}{\partial V}$$

From the expression for I, it follows that

$$\frac{\partial I}{\partial V} = -\frac{K'W}{2L}\left[V^2\left(3 + 2\frac{\partial V_{Ti}}{\partial V}\right)(\alpha_i\lambda) + V\left(2 + 4\alpha_i\lambda V_{Ti} + \frac{2\partial V_{Ti}}{\partial V}\right) + 2V_{Ti}\right]\Bigg|_{V_{DS}=0V}$$

and hence that

$$\lim_{V\to 0}\frac{\partial I}{\partial V} = -\frac{K'W}{L}V_{Ti}\Bigg|_{V_{DS}=0V} \qquad V = \frac{K'W}{L}\left(V_{T0} + \gamma\left[\sqrt{\phi - V_{BB}} - \sqrt{\phi}\right]\right)$$

for $i = 1,2$. The model is therefore differentiable at the origin.

3.1.8 Modeling Noise Sources in MOSFETs

Two mechanisms are the primary contributors to the presence of noise in MOSFETs. One is thermal noise associated with the carriers in the channel. The second is flicker noise associated with the trapping and releasing of electrons in the Si–SiO_2 interface region. These noise sources contribute to the total drain current and can thus be modeled as a current source between the drain and source in either the large signal or small signal device model, as indicated in Fig. 3.1-30*a*, where this current source combines the effects of both types of noise.

The thermal noise current is white noise, which has zero mean and is most easily characterized by its spectral density:

$$S_{IW} = \begin{cases} \dfrac{4kT}{R_{FET}} & \text{ohmic region} \\ \dfrac{8kTg_m}{3} & \text{saturation} \end{cases} \tag{3.1-34}$$

where T is temperature in degrees Kelvin, k is Boltzmann's constant ($k = 1.381 \times 10^{-23}$ J/°K), R_{FET} is the equivalent FET resistance given by (3.1-8), and g_m is the small signal transconductance at the operating point given by (3.1-11). At room temperature ($T = 300$°K), the coefficient $8kT/3$ equals 1.1×10^{-20} V · A · sec.

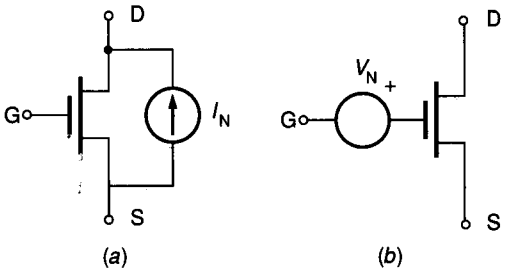


FIGURE 3.1-30
Model of noise sources in the MOSFET: (*a*) Current source in output, (*b*) Input-referred voltage source for small signal operation in the saturation region.

It can be seen that the units of S_{IW} are $A^2 \cdot$ sec. The RMS noise current will be obtained from the spectral density later in this section. Both current and voltage spectral densities are used to characterize noise, and these are often termed the *power spectral densities*. Although it may appear from the saturation region expression in (3.1-34) that the small signal MOSFET model plays a role in characterizing the noise of the MOSFET in the saturation region due to the presence of g_m, this may be somewhat misleading. It can be shown that the expression for the thermal noise contains a $\sqrt{V_{GS} - V_T}$ term, which also appears in the small signal g_m expression of (3.1-11), thus motivating the use of the g_m term in the noise current Eq. 3.1-34. Actually, the thermal noise current is just that associated with the channel source resistance of the MOSFET when operating in the saturation region (see Problems 3.22 and 3.23). Thus, in the ohmic regions, the thermal noise current is identical to that of a resistor of value R_{FET}, and in the saturation region it is equal to that in a resistor of value $3/(2g_m)$.

The flicker noise current in both the saturation and ohmic regions is characterized by the spectral density

$$S_{If} = \frac{2K_f K' I_{DQ}}{C_{ox} L^2 f} \tag{3.1-35}$$

where K_f is the flicker noise coefficient, I_{DQ} is the quiescent current, f is frequency, and K', C_{ox}, and L are the MOSFET model parameters introduced in Sec. 3.1.1. K_f is typically[10] about 3×10^{-24} $V^2 \cdot F$. The flicker noise is often termed $1/f$ noise because of the $1/f$ dependence in (3.1-35).

The white noise and $1/f$ noise sources are statistically uncorrelated, and hence the spectral current density of the noise current source, I_N in Fig. 3.1-30*a*, can be obtained by adding S_{IW} and S_{If} to obtain

$$S_N = S_{IW} + S_{If} \tag{3.1-36}$$

It is well known that the RMS noise current source in the frequency band $[f_1, f_2]$ can be obtained from the spectral density and is given by

$$I_{NB} = \sqrt{\int_{f_1}^{f_2} S_N \, df} \tag{3.1-37}$$

It should be pointed out that if we define the RMS white noise and flicker noise currents by the square root of the integrals of the spectral densities

$$I_{WB} = \sqrt{\int_{f_1}^{f_2} S_{IW}\, df} \tag{3.1-38}$$

and

$$I_{FB} = \sqrt{\int_{f_1}^{f_2} S_{If}\, df} \tag{3.1-39}$$

then these currents add in the RMS sense by the equation

$$I_{NB} = \sqrt{I_{WB}^2 + I_{FB}^2} \tag{3.1-40}$$

The current I_{NB} is the contribution to I_N in Fig. 3.1-30 in the frequency band $[f_1, f_2]$.

The relationship between small signal voltages and/or currents in a circuit and the total output noise voltage or current is often of interest because it characterizes the impact noise has on the performance of a circuit. Two commonly used indicators of this relationship are the signal-to-noise ratio (SNR) and the dynamic range. Since the noise signal amplitudes are usually sufficiently small to be considered small signals, the small signal model of the MOSFET is often used to analyze the noise performance of circuits. This analysis is usually much simpler than a noise analysis based on the large signal models of the MOSFETs.

It can be readily shown that if the MOSFET in Fig. 3.1-30*a* is replaced by its small signal model, then an *equivalent* small signal model that characterizes the noise performance can be obtained by replacing the noise current source from drain to source with a noise voltage source in series with the gate, as denoted by V_N in Fig. 3.1-30*b*. This is termed an *input-referred* noise source. The spectral density of the input-referred noise voltage source relates to the noise spectral density of the noise current source by the expression

$$S_{VN} = \frac{S_{IN}}{g_m^2} \tag{3.1-41}$$

Correspondingly, the input-referred noise voltage source relates to the noise current source by the expression

$$V_N = \frac{I_N}{g_m} \tag{3.1-42}$$

where g_m is the small signal transconductance gain of the MOSFET. In the common case where the MOSFET is operating in the saturation region, the parameter g_m is given by (3.1-11). The input-referred model is useful for simplifying the analysis of some circuits.

Two points about the input-referred noise model of Fig. 3.1-30*b* deserve mention. First, the reader must keep in mind that this is *merely* a small signal

equivalent circuit characterization. The mechanisms that contribute to noise are inherently associated with the drain current. Second, whereas the noise model of Fig. 3.1-30*a* was valid for either a large signal or small signal model of the MOSFET, the input-referred model of Fig. 3.1-30*b* was developed under the assumption of small signal operation of the MOSFET. The equivalence of the circuit in Fig. 3.1-30*b* as characterized by (3.1-42) is not valid for large signal MOSFET operation.

Example 3.1-8. A simple transconductance amplifier is shown in Fig. 3.1-31.

(*a*) Calculate the output noise current spectral density, the input-referred voltage spectral density and the RMS output thermal noise current, the RMS output flicker noise current, the output RMS noise current, and the input-referred RMS noise voltage in the flat frequency band from 100 Hz to 1 MHz if the small signal input, v_i, is zero. Assume $W = 5\ \mu$, $L = 5\ \mu$, $V_{gg} = 2$ V, $K_f = 3 \times 10^{-24}$, $V^2 \cdot F$. Use the typical CMOS parameters of Table 2B.4 in Appendix 2B.

(*b*) Repeat for $W = 200\ \mu$, $L = 5\ \mu$, and $V_{gg} = 1.25$ V.

Solution.

(*a*) Observe that the MOSFET is operating in the saturation region since $V_{DS} = V_{DD} > V_{GS} - V_T$. From (3.1-34) and (3.1-35), both g_m and I_{DQ} are needed.

$$I_{DQ} = \frac{K'W}{2L}(V_{GSQ} - V_T)^2 = 18.75\ \mu A$$

$$g_m = \sqrt{\frac{2K'W}{L}}\sqrt{I_{DQ}} = 30.0\ \mu A/V$$

Thus, from (3.1-34), (3.1-35), and (3.1-36),

$$S_{IW} = 3.3 \times 10^{-25}\ A^2 \cdot sec$$

$$S_{If} = \frac{1.54 \times 10^{-19}}{f}\ A^2 \cdot sec$$

$$S_N = 3.3 \times 10^{-25} + \frac{1.54 \times 10^{-19}}{f}\ A^2 \cdot sec$$

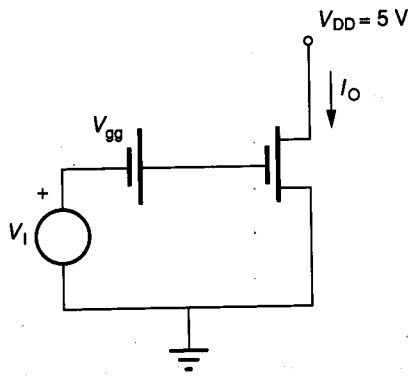


FIGURE 3.1-31
Example 3.1-8.

From (3.1-38), (3.1-39), and (3.1-40),

$$I_{WB} = \sqrt{\int_{10^2}^{10^6} S_{IW}\, df} = 0.489 \text{ nA (RMS)}$$

$$I_{FB} = \sqrt{\int_{10^2}^{10^6} S_{If}\, df} = 1.190 \text{ nA (RMS)}$$

$$I_{NB} = \sqrt{I_{WB}^2 + I_{FB}^2} = 1.29 \text{ nA (RMS)}$$

and from (3.1-42)

$$V_N = \frac{I_N}{g_m} = 43\ \mu\text{V (RMS)}$$

(*b*) Repeating with the new device sizes,

$$I_{DQ} = 120\ \mu\text{A}$$
$$g_m = 48 \text{ mA/V}$$
$$S_{IW} = 5.28 \times 10^{-24} \text{ A}^2 \cdot \text{sec}$$
$$S_{If} = (9.87 \times 10^{-19}/f) \text{ A}^2 \cdot \text{sec}$$
$$I_{WB} = 2.29 \text{ nA (RMS)}$$
$$I_{FB} = 3.01 \text{ nA (RMS)}$$
$$I_{NB} = 3.78 \text{ nA (RMS)}$$
$$V_N = 7.88\ \mu\text{V (RMS)}$$

Note that in the previous example the $1/f$ noise is strongly dominant in the first part whereas the relative contributions are comparable in the second part. Device sizes and bias current affect both the noise current and the relative significance of the two types of noise. At higher frequencies the thermal noise effects dominate whereas the flicker noise, because of its $1/f$ type spectral density, dominates at lower frequencies.

Example 3.1-9. Obtain an expression in terms of the quiescent excess gate bias, $V_{GS} - V_T$, that shows where the spectral density of the $1/f$ noise crosses over that of the thermal noise. Assume operation in the saturation region.

Solution. Equating spectral densities from Eqs. 3.1-34 and 3.1-35, we obtain

$$\frac{8kTg_m}{3} = \frac{K_f 2K' I_{DQ}}{fC_{ox}L^2}$$

From (3.1-4) and (3.1-11), this simplifies to

$$f = \frac{3}{8}\frac{K'K_f}{kTC_{ox}}\left(\frac{V_{GS} - V_T}{L^2}\right)$$

Using typical values of these parameters from Table 2B.4, we obtain

$$f \simeq \left(9.5 \frac{\mu^2}{\text{V}}\right)\left(\frac{V_{GS} - V_T}{L^2}\right) \text{ MHz}$$

From Example 3.1-9 it can be observed that the relative impact of the flicker noise and thermal noise are strongly dependent on the quiescent operating point and the device size.

3.1.9 Simple MOSFET Models for Digital Applications

Many digital applications involve large numbers of gates and transistors, which make hand analysis using the device models, presented in Sec. 3.1.1 and summarized in Table 3.1-1, totally impractical. A very simple model that makes hand analysis tractable is needed for these applications. Such a model will, of necessity, not provide highly accurate results. We are willing to accept the tradeoffs between model accuracy and simplicity in these applications. Models that are widely used for this purpose are shown in Fig. 3.1-32.

In the circuit of Fig. 3.1-32*a*, the switch S1 is open if $V_{GS} < V_H/2$ and closed if $V_{GS} > V_H/2$ where V_H is the logical *high* voltage in the digital logic family and where the logical *low* voltage is assumed to be near zero. The values of C_μ and R are given by

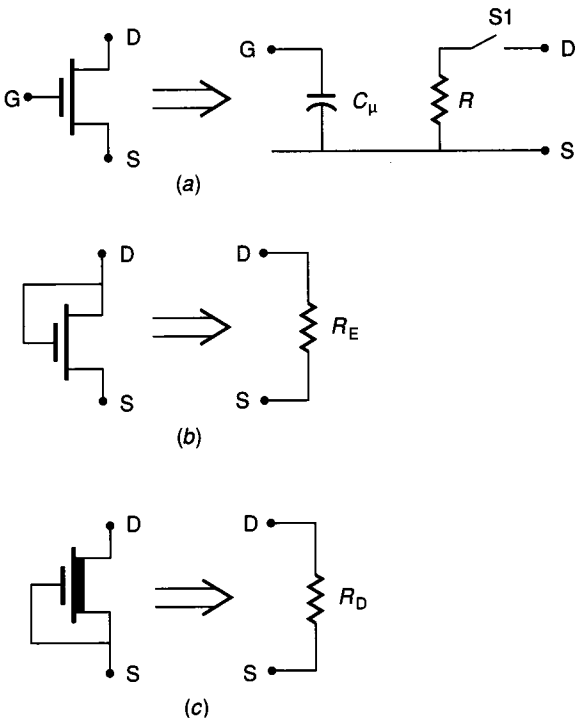


FIGURE 3.1-32
Models for the MOSFET used for simple analysis of digital circuits: (*a*) Three-terminal MOSFET, (*b*) Two-terminal enhancement, (*c*) Two-terminal depletion.

$$C_\mu = C_{ox} WL \tag{3.1-43}$$

$$R = \frac{L}{K'W\,(V_H - V_T)} \tag{3.1-44}$$

where V_T is the threshold voltage of the MOSFET. The two-terminal enhancement and depletion configurations of Figs. 3.1-32*b* and *c* are characterized by

$$R_E = \frac{L}{K'W\,(V_H - V_{TE})} \tag{3.1-45}$$

$$R_D = \frac{L}{K'W\,|\,V_{TD}\,|} \tag{3.1-46}$$

where V_{TE} and V_{TD} denote the threshold voltages of the enhancement and depletion transistors.

The model of Fig. 3.1-32*a* is motivated by the observation that in most digital circuits, the inputs to the logic gates are typically inputs to the gates of transistors, which swing between the high and low logic levels. From the results presented in Section 3.1.3, it follows that the gate input port is capacitive with a value near $C_{ox}W\,L$ throughout most of the input signal swing. When the input is low, the output port looks nearly like an open circuit; and when the gate of the MOSFET is driven to V_H, which is typically much larger than V_T, the output voltage is sufficiently low through most of the high to low transition to force operation in the ohmic region. The ohmic region impedance which is modeled by R in Fig. 3.1-32 is characterized by (3.1-8), which is repeated in (3.1-44). In applications where the gate is permanently connected to the drain or source as shown in Fig. 3.1-32*b* and *c*, the approximation of (3.1-8) is less justifiable but still widely used (see Problem 3.17). These models are used for simple dc and timing analysis in Chapter 7 with modest modifications of the resistances given in (3.1-45) and (3.1-46) to obtain closer agreement between theoretical and experimental results.

Example 3.1-10. Using the simple models of Fig. 3.1-32, obtain the dc transfer characteristics, V_O versus V_I, for the circuit of Fig. 3.1-33. Use the typical process parameters of Table 3.1-2.

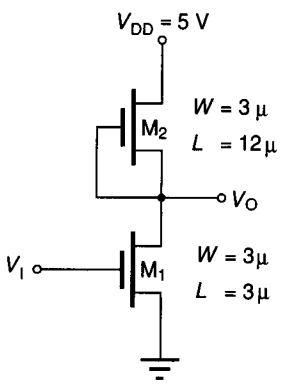


FIGURE 3.1-33
Simple digital inverter.

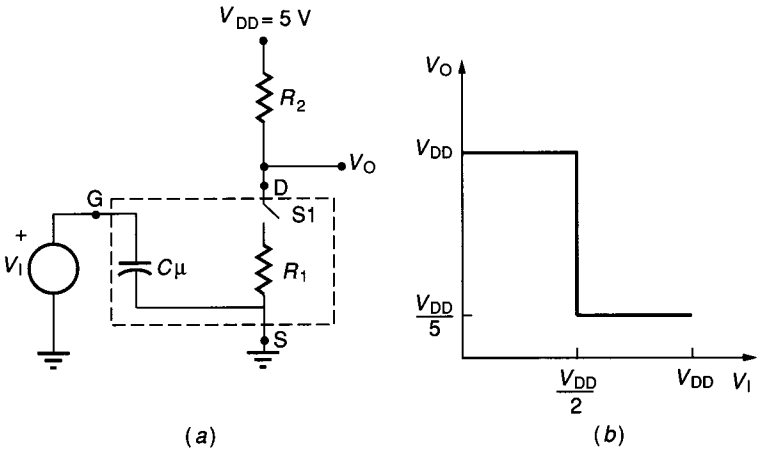


FIGURE 3.1-34
Solution of Example 3.1-10: (*a*) Equivalent circuit, (*b*) Plotting of results.

Solution. It follows from the models of Fig. 3.1-32 that the equivalent circuit can be drawn as shown in Fig. 3.1-34*a*. Assuming the logic high, V_H, is 5 V, it follows from Table 3.1-2 and Eqs. 3.1-44 and 3.1-45 that $R_2 = 39.2$ kΩ and $R_1 = 9.8$ kΩ. For $V_I < V_{DD}/2$, switch S1 is open, so $V_O = V_{DD} = 5$ V. For $V_I > V_{DD}/2$, the switch S1 is closed and the resistive voltage divider yields an output of $V_O = V_{DD}/5 = 1$ V. These idealized results are plotted in Fig. 3.1-34*b*.

3.2 DIODE MODELS

The pn junction diode with the convention used here for the electrical variables is shown in Fig. 3.2-1. The junction diode is characterized by a region of n-type material adjacent to a region of p-type material. A depletion region is formed at the interface. With a sufficiently large forward bias, considerable current will flow, but very little current flows under reverse bias. Such a device, which passes current in one direction and blocks it in the opposite direction, has many practical applications.

3.2.1 dc Diode Model

The diode depicted in Fig. 3.2-1 is characterized at low frequencies by the equation

$$I = I_S\left(e^{\frac{V}{nV_t}} - 1\right) \tag{3.2-1}$$

where the parameters I_S, n, and V_t characterize the device. The parameters in the equation are designated by

I_S = saturation current

n = emission coefficient

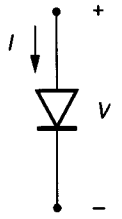


FIGURE 3.2-1
Variable convention for a diode.

and the parameter V_t is given by the equation

$$V_t = \frac{kT}{q} \tag{3.2-2}$$

where k is Boltzmann's constant ($k = 1.381 \times 10^{-23}$ V·C/°K), T is absolute temperature in degrees Kelvin, and q is the charge of an electron ($q = 1.6 \times 10^{-19}$ C). At room temperature (300 °K), $V_t = 26$ mV. Equation 3.2-1 is often called the "diode equation."

The emission coefficient takes on a value of about 1 for silicon. The saturation current can be expressed as

$$I_S = J_S A \tag{3.2-3}$$

where J_S is a process parameter equal to the saturation current density and A is the cross-sectional area of the junction.

The I–V characteristics of a typical diode are shown in Fig. 3.2-2. The current scale for $I < 0$ is considerably smaller than that for $I > 0$ to allow for plotting of the entire diode characteristics, in a meaningful manner, on one set of axes.

The breakdown that occurs for a large reverse bias is not predicted by the diode equation. In discrete diodes this may be quite large (often −100 V or higher). Since the knee is typically very sharp, the breakdown can be gainfully used in voltage-reference applications. Diodes that are intentionally designed to operate in breakdown are called zener diodes. The breakdown voltage for zener diodes is intentionally reduced. Discrete zener diodes with breakdown voltage specifications from the low volt range to over 20 V are readily available.

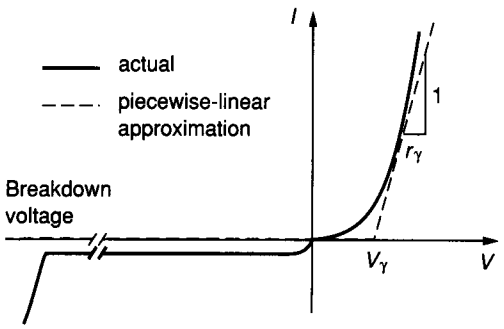


FIGURE 3.2-2
I–V characteristics of a diode.

In IC applications, diodes are often parasitic devices formed in the process of making other devices, as can be seen from the descriptions of the typical CMOS and bipolar processes in Chapter 2. As such, the process is tailored to optimize performance of the MOSFET (in a MOS process) or the BJT (in a bipolar process) rather than the diode. The reverse breakdown voltage of the base-emitter junction in a bipolar process is typically in the 7 V range, making this diode a useful parasitic zener diode. The reverse breakdowns of the parasitic diodes in the MOS processes and for the base-collector junction in a bipolar process are typically much larger. The latter junctions are seldom operated intentionally in the breakdown region.

The diode equation simplifies considerably for $V >> nV_t$. Under forward biases greater than 0.25 V, essentially no accuracy is lost in modeling the diode by

$$I = I_S e^{\frac{V}{nV_t}} \tag{3.2-4}$$

Likewise, for $V < -0.25\,V$, the diode is accurately modeled by

$$I = -I_S \tag{3.2-5}$$

The piecewise-linear model indicated by the dashed curves in Fig. 3.2-2 is often used for biasing and/or large signal applications. The piecewise-linear model is mathematically characterized by the equations

$$\left.\begin{aligned} V &= I r_\gamma + V_\gamma \quad \text{for } I > 0 \\ I &= 0 \quad\quad\quad\quad\;\; \text{for } V < V_\gamma \end{aligned}\right\} \tag{3.2-6}$$

where r_γ is the *on* resistance of the diode and V_γ is the diode cut-in voltage.

The equivalent circuit corresponding to this piecewise-linear model is shown in Fig. 3.2-3. The value of r_γ is obtained by measuring the slope of the I–V characteristics near the desired point of operation. V_γ is the V-axis intercept of the corresponding tangent line; r_γ is typically in the tens of ohms range while $V_\gamma \simeq 0.7$ V. An ideal diode is one in which $r_\gamma = 0$ and $V_\gamma = 0$. In many applications, little accuracy is lost by assuming $r_\gamma = 0$.

Example 3.2-1. Assume $I_S = 10^{-14}$A, T = 300°K, and $n = 1$.
(*a*) Show that $V \simeq 0.7$ V for diode currents in the 1 mA to 100 mA range.
(*b*) Determine V_γ and r_γ at $I = 1$ mA and $I = 100$ mA.

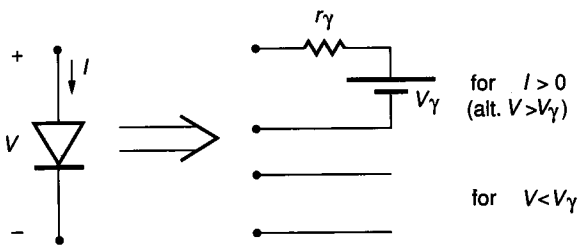


FIGURE 3.2-3
Piecewise-linear diode model.

Solution.

(*a*) Referring to the variable convention of Fig. 3.2-1, assume that $V > 0.25$ V (this must be later verified). It thus follows from Eq. 3.2-4 that

$$V = nV_t \ln\left(\frac{I}{I_S}\right)$$

At about room temperature, we thus have for 1 mA $\leq I \leq$ 100 mA

$$0.658 \text{ V} \leq V \leq 0.778 \text{ V}$$

The initial assumption that $V > 0.25$ V is verified and the conclusion that $V_\gamma \simeq 0.7$ V follows.

(*b*) From part (*a*), it suffices to consider the model of (3.2-4) at $I = 1$ mA and $I = 100$ mA. Upon differentiation of (3.2-4) we obtain

$$r_\gamma = \frac{nV_t}{I} \tag{3.2-7}$$

which at 1 mA and 100 mA becomes 26 Ω and 0.26 Ω respectively. The corresponding V_γs are given by

$$V_\gamma = V - I r_\gamma$$

where V and I are the coordinates of the operating point. It follows from part (*a*) and the values obtained for r_γ that V_γ is 0.632 V for $I = 1$ mA and V_γ is 0.752 V for $I = 100$ mA. In practical applications, V_γ is often approximated by 0.7 V, independent of I.

3.2.2 Small Signal Diode Model

The diode is seldom used in low-frequency small signal applications and, as such, a small signal model need not be developed. Since the small signal model of the diode is, however, very simple, it will be mentioned in passing. For reverse bias, the low-frequency small signal model is an open circuit. Under forward bias, it is modeled by a single resistor of value determined by (3.2-7), evaluated at the dc operating point. The diode is occasionally used in high-frequency small signal applications as a capacitor. When used as a capacitor, the diode is often termed a vari-cap or a varactor diode. These diode capacitances are strongly dependent upon the quiescent voltage. A discussion of the high-frequency diode model follows.

3.2.3 High-Frequency Diode Model

At high frequencies, the capacitance formed by the depletion region in the pn junction becomes significant. This capacitance is modeled by a capacitor connected between the terminals of the diode of value

$$C = \begin{cases} \dfrac{C_{j0}A}{\left(1 - \dfrac{V}{\phi_B}\right)^n} & V < \dfrac{\phi_B}{2} \qquad (3.2\text{-}8) \\ 2^n C_{j0} A\left[\dfrac{2nV}{\phi_B} + (1-n)\right] & V > \dfrac{\phi_B}{2} \qquad (3.2\text{-}9) \end{cases}$$

where the parameters C_{j0}, A, ϕ_B, and n are as defined for the parasitic junction capacitances of the MOSFET in Section 3.1.3. Note that for $V < \phi_B/2$, this parasitic capacitance agrees with that of the MOSFET given by (3.1-14). It can be readily shown (see Problem 3.31) that the value for $V > \phi_B/2$ is the continuous and differentiable linear extension of the reverse-bias value across the boundary defined by $V = \phi_B/2$.

3.3 BIPOLAR MODELS

The npn and pnp bipolar transistors and the convention for the electrical variables established in Chapter 2 are shown in Fig. 3.3-1, along with a top view of a typical geometrical layout of a bipolar transistor. The following model development will

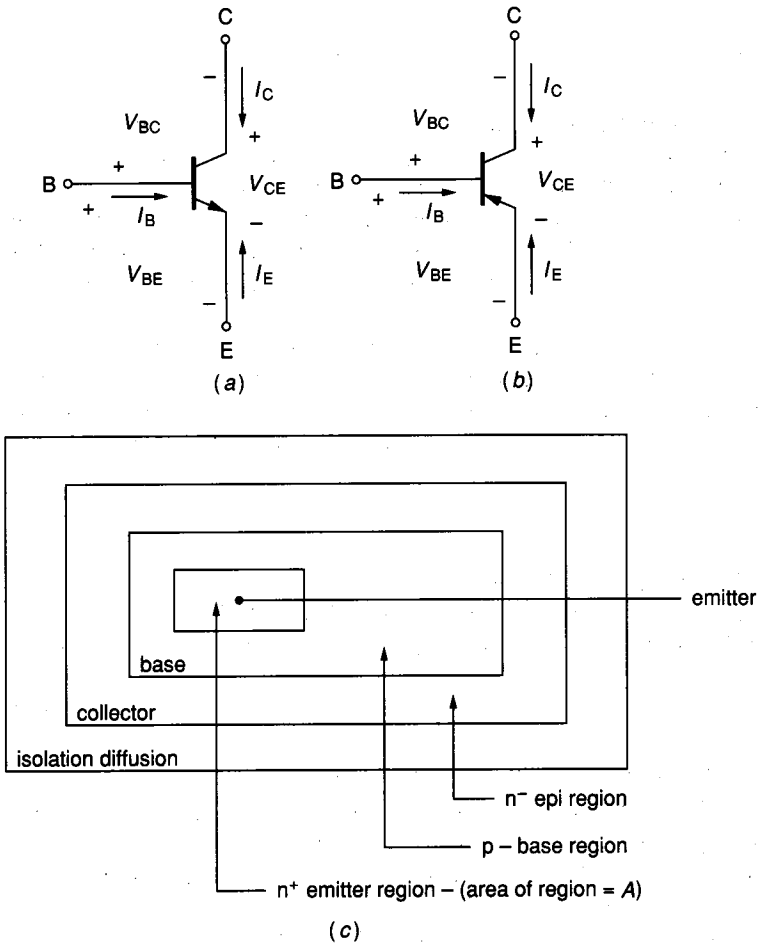


FIGURE 3.3-1
Convention for electrical variables of BJT transistors: (*a*) npn, (*b*) pnp, (*c*) Top view of vertical npn transistor geometry.

be restricted to the npn transistor. The pnp development is identical with the exception of sign changes in some of the equations. Results only for the pnp devices will be presented following the npn discussion.

3.3.1 dc BJT Model

The starting point will be the dc model introduced by Ebers and Moll in 1954[11]. In this model the relationship between the terminal variables is given by

$$I_C = I_S(e^{(V_{BE}/V_t)} - 1) - \frac{I_S}{\alpha_R}(e^{(V_{BC}/V_t)} - 1) \tag{3.3-1}$$

$$I_E = \frac{-I_S}{\alpha_F}(e^{(V_{BE}/V_t)} - 1) + I_S(e^{(V_{BC}/V_t)} - 1) \tag{3.3-2}$$

where the four parameters I_S, α_R, α_F, and V_t characterize the device. The parameters in these equations are designated by

I_S = transport saturation current

α_R = large signal reverse current gain of common base configuration

α_F = large signal forward current gain of common base configuration

and

$$V_t = \frac{kT}{q} \tag{3.3-3}$$

where k is Boltzmann's constant (k $= 1.381 \times 10^{-23}$ V· C/°K), T is absolute temperature in degrees Kelvin, and q is the charge of an electron ($q = 1.602 \times 10^{-19}$ C). At room temperature (300 °K), $V_t \simeq 26$ mV. This is the same V_t that appeared in the diode model and the subthreshold MOSFET model.

The two Ebers-Moll equations along with KVL and KCL applied to the transistor itself

$$I_B = -I_C - I_E \tag{3.3-4}$$

$$V_{CE} = V_{BE} - V_{BC} \tag{3.3-5}$$

provide four independent equations relating the six terminal variables, $I_C, I_B, I_E, V_{CE}, V_{BC}$, and V_{BE}. The BJT is characterized by these four equations.

It should be noted that there is complete functional symmetry in the device model with respect to the collector and emitter terminals. This is comforting since the npn transistor can be thought of as a sandwich of a p-type region between two n-channel layers. Since the shape and impurity concentrations of the collector and emitter differ, however, one would not expect these terminals to be interchangeable. The lack of geometrical symmetry causes large differences in α_F and α_R.

The parameters α_F and α_R are determined by impurity concentrations and junction depths, and as such are process parameters. The parameter V_t is a function of two physical constants and temperature and thus cannot be considered

a design parameter. The parameter V_t should not be confused with the threshold voltage for a MOSFET, which uses the symbol V_T. Because of the temperature dependence, V_t can be considered an environmental parameter. The transport saturation current can be expressed as

$$I_S = J_S A \tag{3.3-6}$$

where A is the area of the emitter and J_S, the transport saturation current density, is a process parameter. The parameter J_S can be further expressed as

$$J_S = q\bar{D}_n n_i^2 / Q_B \tag{3.3-7}$$

where $\bar{D}_n$ is the average effective electron diffusion constant, n_i is the intrinsic carrier concentration in silicon ($n_i = 1.45 \times 10^{10}$ atoms/cm^3 at 300 °K), and Q_B is the number of doping atoms in the base per unit area.

From the model presented, it can be seen that the emitter area, A, is the only design parameter available to the design engineer for the vertical bipolar transistor. The shape of the emitter, the size and shape of the base, and the size and shape of the collector have no effect on the dc model of the BJT just introduced.

Typical values of $\bar{D}_n$ and Q_B are 13 cm^2/sec and 7.25×10^{12} atoms/cm^2, respectively. α_R is typically between 0.3 and 0.8. α_F is generally close to unity, falling in the range of 0.9 to 0.999. Typical values for the bipolar parameters that will be used in examples in this text are given in Table 3.3-1.

Labeling of the emitter and collector terminals is arbitrary at this point. Convention has been established, however, for making this distinction. The typical values for α_R and α_F given in the previous paragraph follow this convention. Specifically, the emitter and collector designations are established so that $\alpha_F > \alpha_R$.

An equivalent circuit for the bipolar transistor based upon the Ebers-Moll model is sometimes useful. The circuit of Fig. 3.3-2 serves as such an equivalent circuit where the currents in the two diodes satisfy the standard diode equations

$$I_F = \frac{I_S}{\alpha_F}(e^{(V_{BE}/V_t)} - 1) \tag{3.3-8}$$

$$I_R = \frac{I_S}{\alpha_R}(e^{(V_{BC}/V_t)} - 1) \tag{3.3-9}$$

TABLE 3.3-1
Typical process parameters for vertical npn transistor in bipolar technology

$J_S = 6 \times 10^{-10}$	μA/mil^2
$\alpha_F = 0.99$	($\beta_F = 100$)
$\alpha_R = 0.3$	($\beta_R = 0.43$)
$V_{AF} = 200$ V	
$V_{AR} = 200$ V	

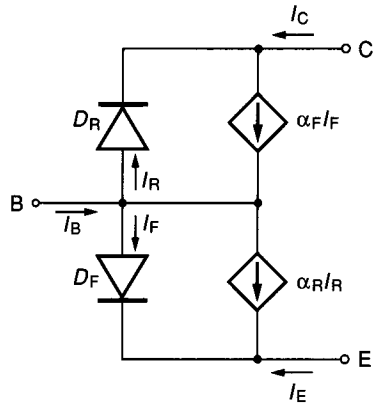


FIGURE 3.3-2
Equivalent circuit of BJT based upon Ebers-Moll model.

Several different regions of operation for the BJT are readily identifiable. Those that are of the most interest are depicted in the V_{BC}–V_{BE} plane shown in Fig. 3.3-3. The forward active, reverse active, and saturation regions are defined by the dashed curves. The cutoff region is defined to be the third quadrant of the V_{BC}–V_{BE} plane. No particular designation is made of the lightly shaded region. Most applications involve operation in one of the four well-defined regions, with possible excursions through the lightly shaded region in digital circuits during switching transitions. For power dissipation reasons, operation in the reverse and forward active regions is generally restricted further to the hatched regions and in the saturation region to the heavily shaded region (see Problem 3.40). It will be shown that major simplifications in the Ebers-Moll equations are possible in all except the lightly shaded region of Fig. 3.3-3. These simplifications are crucial for hand analysis because of the unwieldiness of the Ebers-Moll Eqs. 3.3-1 and 3.3-2 in even relatively simple problems (see Problem 3.38). Simplified models for the forward and reverse active, saturation, and cutoff regions will now be discussed.

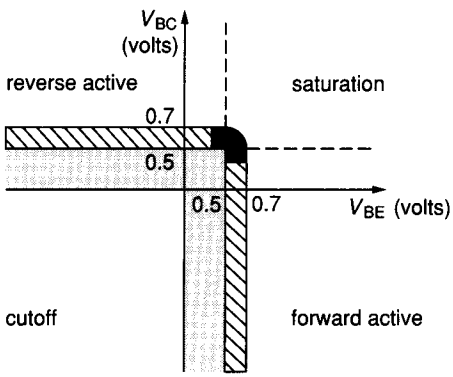


FIGURE 3.3-3
Regions of operation of npn silicon bipolar transistor.

FORWARD AND REVERSE ACTIVE REGIONS OF OPERATION. The BJT is typically operated in the forward active region in most linear applications. If $V_{BE} > 0.5$ V and $V_{BC} < 0.3$ V, then it can be shown that (3.3-1) and (3.3-2) simplify to

$$I_C = I_S e^{(V_{BE}/V_t)} \tag{3.3-10}$$

$$I_E = -\frac{I_S}{\alpha_F} e^{(V_{BE}/V_t)} \tag{3.3-11}$$

Note that "=" was used rather than "≃" in these equations. A simple calculation (see Problem 3.39) will show that the terms dropped in obtaining (3.3-10) and (3.3-11) are several orders of magnitude smaller than the terms that were retained!

It is often more convenient to use I_B and I_C rather than I_C and I_E as the dependent variables when modeling the BJT. Defining β_F by

$$\beta_F = \frac{\alpha_F}{1 - \alpha_F} \tag{3.3-12}$$

it follows from (3.3-10) and (3.3-11) that I_C and I_B can be written as

$$I_C = I_S e^{(V_{BE}/V_t)} \tag{3.3-13}$$

$$I_B = \frac{I_S}{\beta_F} e^{(V_{BE}/V_t)} \tag{3.3-14}$$

Note that in the forward active region the BJT acts as a good current amplifier from the base to the collector with the gain relationship

$$I_C = \beta_F I_B \tag{3.3-15}$$

This should be contrasted to the MOSFET, which is inherently a good transconductance amplifier.

Similarly, if β_R is defined as

$$\beta_R = \frac{\alpha_R}{1 - \alpha_R} \tag{3.3-16}$$

then the Ebers-Moll equations in the reverse active region ($V_{BC} > 0.5$ V, $V_{BE} < 0.3$ V) simplify to

$$I_E = I_S e^{(V_{BC}/V_t)} \tag{3.3-17}$$

$$I_B = \frac{I_S}{\beta_R} e^{(V_{BC}/V_t)} \tag{3.3-18}$$

As before, it follows that the base-to-emitter current gain in the reverse active region is given by

$$I_E = \beta_R I_B \tag{3.3-19}$$

Actually, it can be shown both theoretically and experimentally that the collector current of the BJT in the forward active region increases slightly in

approximately a linear manner with V_{CE}. Defining V_{AF} to be the coefficient that represents the linear dependence on V_{CE}, the following improvement on (3.3-13) can be made:

$$I_C \simeq I_S e^{(V_{BE}/V_t)}\left(1 + \frac{V_{CE}}{V_{AF}}\right) \tag{3.3-20}$$

Correspondingly, this type of linear dependence in the reverse active region can be characterized by the parameter V_{AR}, resulting in the following improvement on (3.3-17):

$$I_E \simeq I_S e^{(V_{BC}/V_t)}\left(1 - \frac{V_{CE}}{V_{AR}}\right) \tag{3.3-21}$$

The parameters V_{AF} and V_{AR} are process parameters and are referred to as the forward and reverse Early voltages, respectively. V_{AF} and V_{AR} for the BJT serve the same role as the parameter λ (actually $1/\lambda$) in the MOSFET, as can be seen by comparing Eqs. 3.3-20 and 3.1-5.

Many IC applications require base currents in the 0.1 μA to 1 mA range. It is shown in Example 3.3-1 that in the forward active region the base–emitter voltage is typically in the 0.5 V to 0.75 V range. For hand manipulations it is often sufficient to obtain the base–emitter voltage to within 0.1 V ($V_{BE} = 0.6$ V is often used) in this region.

Example 3.3-1. Show that at around room temperature the base–emitter voltage is between 0.5 V and 0.75 V for base currents in the 0.1 μA to 1 mA range for transistors with 1 mil^2 emitter areas and reverse-biased base–collector junctions. Use the process parameters of Table 3.3-1.

Solution. From Eqs. 3.3-6 and 3.3-14 it follows that

$$I_B \simeq \frac{AJ_S}{\beta_F} e^{(V_{BE}/V_t)}$$

With $A = 1$ mil^2 and $J_S = 6 \times 10^{-10}$ μA/mil^2, we can write

$$0.1\ \mu\text{A} \leq 6 \times 10^{-10} e^{(V_{BE}/.026\ \text{V})} \leq 1\ \text{mA}$$

It thus follows that

$$0.5\ \text{V} < V_{BE} < 0.75\ \text{V} \tag{3.3-22}$$

It can be shown that even with substantial changes in temperature, J_S, or A, the changes in V_{BE} for the current range considered in this example are modest.

An even simpler model is often used for characterizing the BJT in the forward and reverse active regions. As indicated in Example 3.3-1, the base–emitter voltage is about 0.6 V in the forward active region for a wide range of base currents. The model

$$V_{BE} = 0.6\ \text{V} \tag{3.3-23}$$

$$I_C = \beta_F I_B \tag{3.3-24}$$