7.1 Basic Principles

7.1.1 The Basis for Amplifier Operation

The basis for the application of the transistor (a MOSFET or a BJT) in amplifier design is that when the device is operated in the active region, a voltage-controlled current source is realized. Specifically, when a MOSFET is operated in the saturation or pinch-off region, also referred to in this chapter as the active region, the voltage between gate and source, v_{GS}, controls the drain current i_D according to the square-law relationship which, for an NMOS transistor, is expressed as

$$i_D = \frac{1}{2} k_n (v_{GS} - V_{tn})^2 \tag{7.1}$$

We note that in this first-order model of MOSFET operation, the drain current i_D does not depend on the drain voltage v_{DS} because the channel is pinched off at the drain end,[1] thus "isolating" the drain.

Similarly, when a BJT is operated in the active region, the base-emitter voltage v_{BE} controls the collector current i_C according to the exponential relationship which, for an *npn* transistor, is expressed as

$$i_C = I_S e^{v_{BE}/V_T} \tag{7.2}$$

Here, this first-order model of BJT operation indicates that the collector current i_C does not depend on the collector voltage v_{CE} because the collector–base junction is reverse biased, thus "isolating" the collector.

Figure 7.1 shows an NMOS transistor and an *npn* transistor operating in the active mode. Observe that for the NMOS transistor, the pinch-off condition is ensured by keeping $v_{DS} \geq v_{OV}$. Since the overdrive voltage $v_{OV} = v_{GS} - V_{tn}$, this condition implies that $v_{GD} \leq V_{tn}$, which indeed ensures channel pinch-off at the drain end.

Similarly, for the *npn* transistor in Fig. 7.1(b), the CBJ reverse-bias condition is ensured by keeping $v_{CE} \geq 0.3$ V. Since v_{BE} is usually in the vicinity of 0.7 V, v_{BC} is thus kept

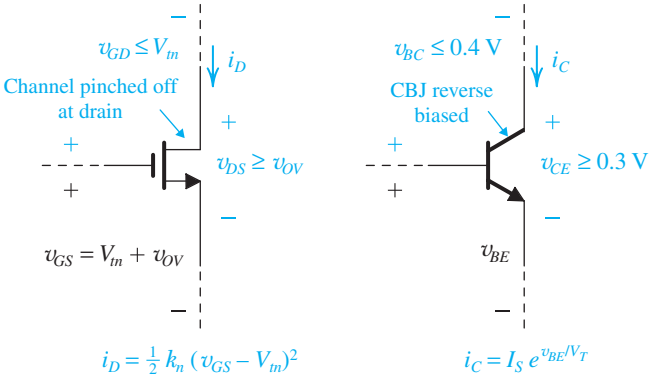


Figure 7.1 Operating (**a**) an NMOS transistor and (**b**) an *npn* transistor in the active mode. Note that $v_{GS} = V_{tn} + v_{OV}$ and $v_{DS} \geq v_{OV}$; thus $v_{GD} \leq V_{tn}$, which ensures channel pinch-off at the drain end. Similarly, $v_{BE} \simeq 0.7$ V, and $v_{CE} \geq 0.3$ V results in $v_{BC} \leq 0.4$ V, which is sufficient to keep the CBJ from conducting.

[1] To focus on essentials, we shall neglect the Early effect until a later point.

smaller than 0.4 V, which is sufficient to prevent this relatively large-area junction from conducting.

Although we used NMOS and *npn* transistors to illustrate the conditions for active-mode operation, similar conditions apply for PMOS and *pnp* transistors, as studied in Chapters 5 and 6, respectively.

Finally, we note that the control relationships in Eqs. (7.1) and (7.2) are nonlinear. Nevertheless, we shall shortly devise a technique for obtaining almost-linear amplification from these fundamentally nonlinear devices.

7.1.2 Obtaining a Voltage Amplifier

From the above we see that the transistor is basically a transconductance amplifier: that is, an amplifier whose input signal is a voltage and whose output signal is a current. More commonly, however, one is interested in voltage amplifiers. A simple way to convert a transconductance amplifier to a voltage amplifier is to pass the output current through a resistor and take the voltage across the resistor as the output. Doing this for a MOSFET results in the simple amplifier circuit shown in Fig. 7.2(a). Here v_{GS} is the input voltage, R_D (known as a **load resistance**) converts the drain current i_D to a voltage ($i_D R_D$), and V_{DD} is the supply voltage that powers up the amplifier and, together with R_D, establishes operation in the active region, as will be shown shortly.

In the amplifier circuit of Fig. 7.2(a) the output voltage is taken between the drain and ground, rather than simply across R_D. This is done because of the need to maintain a common ground reference between the input and the output. The output voltage v_{DS} is given by

$$v_{DS} = V_{DD} - i_D R_D \tag{7.3}$$

Thus it is an inverted version (note the minus sign) of $i_D R_D$ that is shifted by the constant value of the supply voltage V_{DD}.

An exactly similar arrangement applies for the BJT amplifier, as illustrated in Fig. 7.2(c). Here the output voltage v_{CE} is given by

$$v_{CE} = V_{CC} - i_C R_C \tag{7.4}$$

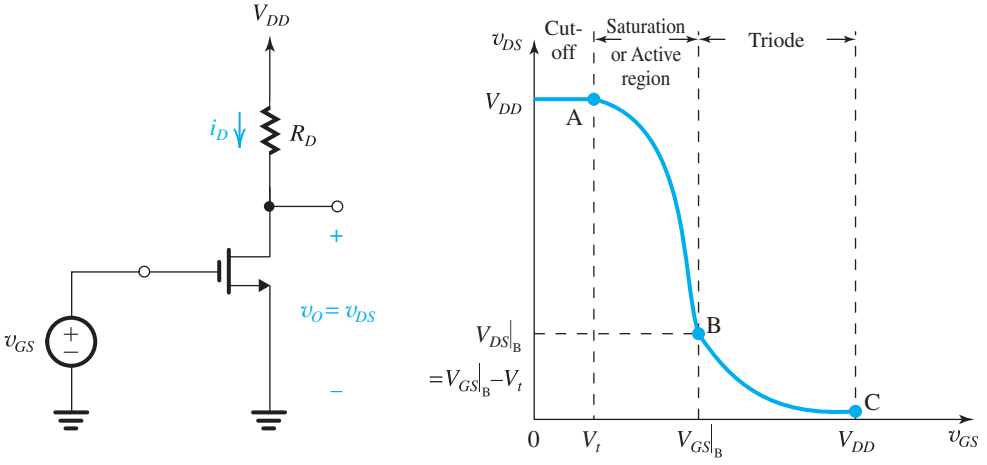


Figure 7.2 (**a**) An NMOS amplifier and (**b**) its VTC; and (**c**) an *npn* amplifier and (**d**) its VTC.

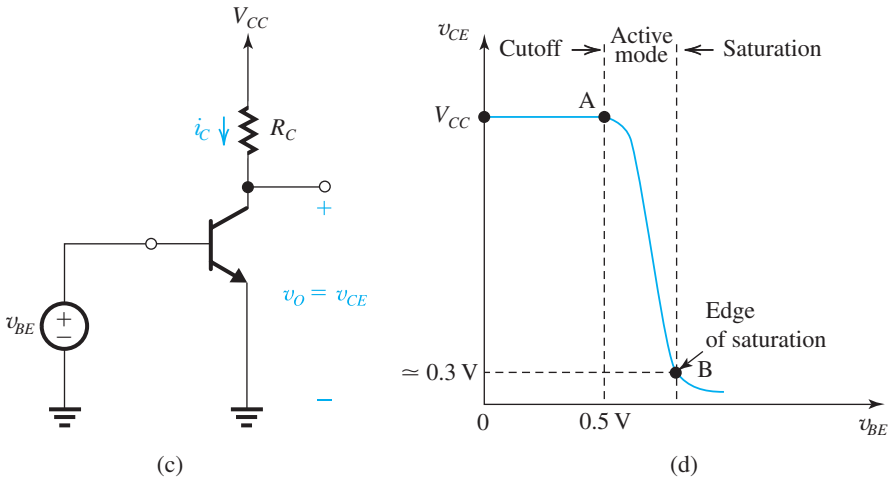


Figure 7.2 *continued*

7.1.3 The Voltage-Transfer Characteristic (VTC)

A useful tool that provides insight into the operation of an amplifier circuit is its voltage-transfer characteristic (VTC). This is simply a plot (or a clearly labeled sketch) of the output voltage versus the input voltage. For the MOS amplifier in Fig. 7.2(a), this is the plot of v_{DS} versus v_{GS} shown in Fig. 7.2(b).

Observe that for $v_{GS} < V_t$, the transistor is cut off, $i_D = 0$ and, from Eq. (7.3), $v_{DS} = V_{DD}$. As v_{GS} exceeds V_t, the transistor turns on and v_{DS} decreases. However, since initially v_{DS} is still high, the MOSFET will be operating in saturation or the active region. This continues as v_{GS} is increased until the value of v_{GS} is reached that results in v_{DS} becoming lower than v_{GS} by V_t volts [point B on the VTC in Fig. 7.2(b)]. For v_{GS} greater than that at point B, the transistor operates in the triode region and v_{DS} decreases more slowly.

The VTC in Fig. 7.2(b) indicates that the segment of greatest slope (hence potentially the largest amplifier gain) is that labeled AB, which corresponds to operation in the active region. When a MOSFET is operated as an amplifier, its operating point is confined to the segment AB at all times. An expression for the segment AB can be obtained by substituting for i_D in Eq. (7.3) by its active-region value from Eq. (7.1), thus

$$v_{DS} = V_{DD} - \frac{1}{2} k_n R_D (v_{GS} - V_t)^2 \tag{7.5}$$

This is obviously a nonlinear relationship. Nevertheless, linear (or almost-linear) amplification can be obtained by using the technique of biasing the MOSFET. Before considering biasing, however, it is useful to determine the coordinates of point B, which is at the boundary between the saturation and the triode regions of operation. These can be obtained by substituting in Eq. (7.5), $v_{GS} = V_{GS}\big|_B$ and $v_{DS} = V_{DS}\big|_B = V_{GS}\big|_B - V_t$. The result is

$$V_{GS}\big|_B = V_t + \frac{\sqrt{2k_n R_D V_{DD} + 1} - 1}{k_n R_D} \tag{7.6}$$

Point B can be alternatively characterized by the overdrive voltage

$$V_{OV}\big|_{\mathrm{B}} \equiv V_{GS}\big|_{\mathrm{B}} - V_t = \frac{\sqrt{2k_nR_DV_{DD}+1}-1}{k_nR_D} \tag{7.7}$$

and

$$V_{DS}\big|_{\mathrm{B}} = V_{OV}\big|_{\mathrm{B}} \tag{7.8}$$

EXERCISE

7.1 Consider the amplifier of Fig. 7.2(a) with $V_{DD} = 1.8$ V, $R_D = 17.5$ kΩ, and with a MOSFET specified to have $V_t = 0.4$ V, $k_n = 4$ mA/V^2, and $\lambda = 0$. Determine the coordinates of the end points of the active-region segment of the VTC. Also, determine $V_{DS}\big|_C$ assuming $V_{GS}\big|_C = V_{DD}$.
Ans. A: 0.4 V, 1.8 V; B: 0.613 V, 0.213 V; $V_{DS}\big|_C = 18$ mV

An exactly similar development applies to the BJT case. This is illustrated in Fig. 7.2(c) and (d). In this case, over the active-region or amplifier segment AB, the output voltage v_{CE} is related to the input voltage v_{BE} by

$$v_{CE} = V_{CC} - R_C I_S e^{v_{BE}/V_T} \tag{7.9}$$

Here also, the input–output relationship is nonlinear. Nevertheless, linear (or almost-linear) amplification can be obtained by using the biasing technique discussed next.

7.1.4 Obtaining Linear Amplification by Biasing the Transistor

Biasing enables us to obtain almost-linear amplification from the MOSFET and the BJT. The technique is illustrated for the MOSFET case in Fig. 7.3(a). A dc voltage V_{GS} is selected to obtain operation at a point Q on the segment AB of the VTC. How to select an appropriate location for the bias point Q will be discussed shortly. For the time being, observe that the coordinates of Q are the dc voltages V_{GS} and V_{DS}, which are related by

$$V_{DS} = V_{DD} - \frac{1}{2}k_nR_D(V_{GS} - V_t)^2 \tag{7.10}$$

Point Q is known as the **bias point** or the **dc operating point**. Also, since at Q no signal component is present, it is also known as the **quiescent point** (which is the origin of the symbol Q).

Next, the signal to be amplified, v_{gs}, a function of time t, is superimposed on the bias voltage V_{GS}, as shown in Fig. 7.4(a). Thus the total instantaneous value of v_{GS} becomes

$$v_{GS}(t) = V_{GS} + v_{gs}(t)$$

The resulting $v_{DS}(t)$ can be obtained by substituting for $v_{GS}(t)$ into Eq. (7.5). Graphically, we can use the VTC to obtain $v_{DS}(t)$ point by point, as illustrated in Fig. 7.4(b). Here we show

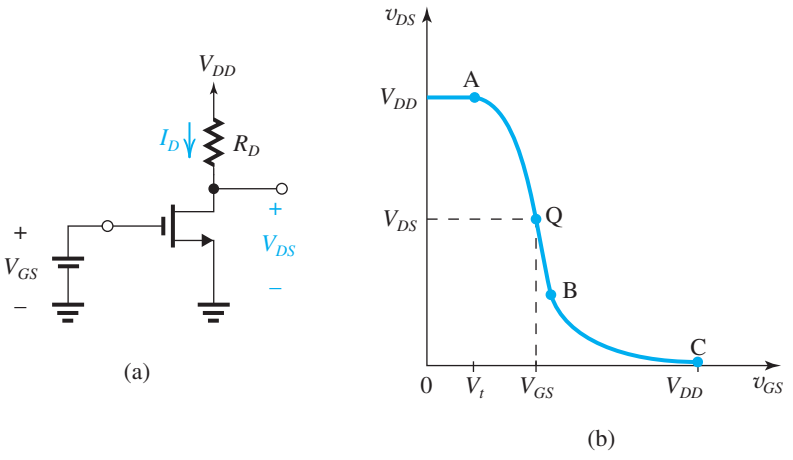


Figure 7.3 Biasing the MOSFET amplifier at a point Q located on the segment AB of the VTC.

the case of v_{gs} being a triangular wave of "small" amplitude. Specifically, the amplitude of v_{gs} is small enough to restrict the excursion of the instantaneous operating point to a short, almost-linear segment of the VTC around the bias point Q. The shorter the segment, the greater the linearity achieved, and the closer to an ideal triangular wave the signal component at the output, v_{ds}, will be. This is the essence of obtaining linear amplification from the nonlinear MOSFET.

Before leaving Fig. 7.4(b) we wish to draw the reader's attention to the consequence of increasing the amplitude of the signal v_{gs}. As the instantaneous operating point will no longer be confined to the almost-linear segment of the VTC, the output signal v_{ds} will deviate from its ideal triangular shape; that is, it will exhibit nonlinear distortion. Worse yet, if the input signal amplitude becomes sufficiently large, the instantaneous operating point may leave the segment AB altogether. If this happens at the negative peaks of v_{gs}, the transistor will cut off for a portion of the cycle and the positive peaks of v_{ds} will be "clipped off." If it occurs at the positive peaks of v_{gs}, the transistor will enter the triode region for a portion of the cycle, and the negative peaks of v_{ds} will become flattened. It follows that the selection of the location of the bias point Q can have a profound effect on the maximum allowable amplitude of v_{ds}, referred to as the *allowable signal swing at the output*. We will have more to say later on this important point.

An exactly parallel development can be applied to the BJT amplifier. In fact, all we need to do is replace the NMOS transistor in Figs. 7.3 and 7.4 with an *npn* transistor and change the voltage and current symbols to their BJT counterparts. The resulting bias point Q will be characterized by dc voltages V_{BE} and V_{CE}, which are related by

$$V_{CE} = V_{CC} - R_C I_S e^{V_{BE}/V_T} \tag{7.11}$$

and a dc current I_C,

$$I_C = I_S e^{V_{BE}/V_T} \tag{7.12}$$

Also, superimposing a small-signal v_{be} on the dc bias voltage V_{BE} results in

$$v_{BE}(t) = V_{BE} + v_{be}(t)$$

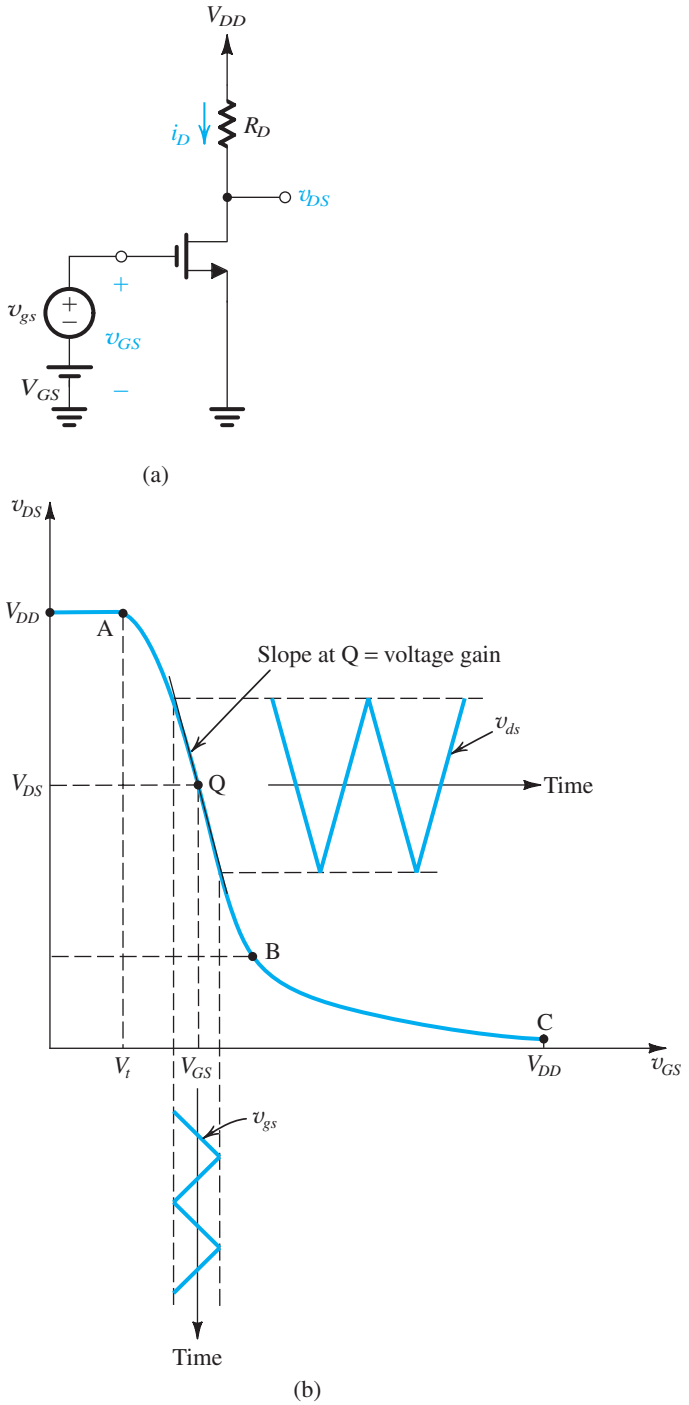


Figure 7.4 The MOSFET amplifier with a small time-varying signal $v_{gs}(t)$ superimposed on the dc bias voltage V_{GS}. The MOSFET operates on a short almost-linear segment of the VTC around the bias point Q and provides an output voltage $v_{ds} = A_v v_{gs}$.

which can be substituted into Eq. (7.9) to obtain the total instantaneous value of the output voltage $v_{CE}(t)$. Here again, almost-linear operation is obtained by keeping v_{be} small enough to restrict the excursion of the instantaneous operating point to a short, almost-linear segment of the VTC around the bias point Q. Similar comments also apply to the maximum allowable signal swing at the output.

7.1.5 The Small-Signal Voltage Gain

The MOSFET Case Consider the MOSFET amplifier in Fig. 7.4(a). If the input signal v_{gs} is kept small, the corresponding signal at the output v_{ds} will be nearly proportional to v_{gs} with the constant of proportionality being the slope of the almost-linear segment of the VTC around Q. This is the voltage gain of the amplifier, and its value can be determined by evaluating the slope of the tangent to the VTC at the bias point Q,

$$A_v = \left.\frac{dv_{DS}}{dv_{GS}}\right|_{v_{GS}=V_{GS}} \tag{7.13}$$

Utilizing Eq. (7.5) we obtain

$$A_v = -k_n(V_{GS} - V_t)R_D \tag{7.14}$$

which can be expressed in terms of the overdrive voltage at the bias point, V_{OV}, as

$$A_v = -k_n V_{OV} R_D \tag{7.15}$$

We make the following observations on this expression for the voltage gain.

1. The gain is negative, which signifies that the amplifier is inverting; that is, there is a 180° phase shift between the input and the output. This inversion is obvious in Fig. 7.4(b) and should have been anticipated from Eq. (7.5).
2. The gain is proportional to the load resistance R_D, to the transistor transconductance parameter k_n, and to the overdrive voltage V_{OV}. This all makes intuitive sense.

Another simple and insightful expression for the voltage gain A_v can be derived by recalling that the dc current in the drain at the bias point is related to V_{OV} by

$$I_D = \frac{1}{2}k_n V_{OV}^2$$

This equation can be combined with Eq. (7.15) to obtain

$$A_v = -\frac{I_D R_D}{V_{OV}/2} \tag{7.16}$$

That is, the gain is simply the ratio of the dc voltage drop across the load resistance R_D to $V_{OV}/2$. It can be expressed in the alternative form

$$A_v = -\frac{V_{DD} - V_{DS}}{V_{OV}/2} \tag{7.17}$$

Since the maximum slope of the VTC in Fig. 7.4(b) occurs at point B, the maximum gain magnitude $|A_{v\max}|$ is obtained by biasing the transistor at point B,

$$|A_{v\max}| = \frac{V_{DD} - V_{DS}|_{\text{B}}}{V_{OV}|_{\text{B}}/2}$$

and since $V_{DS}\big|_B = V_{OV}\big|_B$,

$$|A_{v\max}| = \frac{V_{DD} - V_{OV}\big|_B}{V_{OV}\big|_B/2} \tag{7.18}$$

where $V_{OV}\big|_B$ is given by Eq. (7.7). Of course, this result is only of theoretical importance since biasing at B would leave no room for negative signal swing at the output. Nevertheless, the result in Eq. (7.18) is valuable as it provides an upper bound on the magnitude of voltage gain achievable from this basic amplifier circuit. As an example, for a discrete-circuit amplifier operated with $V_{DD} = 5$ V and $V_{OV}\big|_B = 0.5$ V, the maximum achievable gain is 18 V/V. An integrated-circuit amplifier utilizing a modern submicron MOSFET operated with $V_{DD} = 1.3$ V and with $V_{OV}\big|_B = 0.2$ V realizes a maximum gain of 11 V/V.

Finally, note that to maximize the gain, the bias point Q should be as close to point B as possible, consistent with the required signal swing at the output. This point will be explored further in the end-of-chapter problems.

Example 7.1

Consider the amplifier circuit shown in Fig. 7.4(a). The transistor is specified to have $V_t = 0.4$ V, $k_n' = 0.4$ mA/V^2, $W/L = 10$, and $\lambda = 0$. Also, let $V_{DD} = 1.8$ V, $R_D = 17.5$ kΩ, and $V_{GS} = 0.6$ V.

(a) For $v_{gs} = 0$ (and hence $v_{ds} = 0$), find V_{OV}, I_D, V_{DS}, and A_v.
(b) What is the maximum symmetrical signal swing allowed at the drain? Hence, find the maximum allowable amplitude of a sinusoidal v_{gs}.

Solution

(a) With $V_{GS} = 0.6$ V, $V_{OV} = 0.6 - 0.4 = 0.2$ V. Thus,

$$\begin{aligned} I_D &= \frac{1}{2}k_n'\left(\frac{W}{L}\right)V_{OV}^2 \\ &= \frac{1}{2} \times 0.4 \times 10 \times 0.2^2 = 0.08 \text{ mA} \\ V_{DS} &= V_{DD} - R_D I_D \\ &= 1.8 - 17.5 \times 0.08 = 0.4 \text{ V} \end{aligned}$$

Since V_{DS} is greater than V_{OV}, the transistor is indeed operating in saturation. The voltage gain can be found from Eq. (7.15),

$$\begin{aligned} A_v &= -k_n V_{OV} R_D \\ &= -0.4 \times 10 \times 0.2 \times 17.5 \\ &= -14 \text{ V/V} \end{aligned}$$

An identical result can be found using Eq. (7.17).

(b) Since $V_{OV} = 0.2$ V and $V_{DS} = 0.4$ V, we see that the maximum allowable negative signal swing at the drain is 0.2 V. In the positive direction, a swing of +0.2 V would not cause the transistor to

Example 7.1 *continued*

cut off (since the resulting v_{DS} would be still lower than V_{DD}) and thus is allowed. Thus the maximum symmetrical signal swing allowable at the drain is ±0.2 V. The corresponding amplitude of v_{gs} can be found from

$$\hat{v}_{gs} = \frac{\hat{v}_{ds}}{|A_v|} = \frac{0.2\text{ V}}{14} = 14.2\text{ mV}$$

Since $\hat{v}_{gs} \ll V_{OV}$, the operation will be reasonably linear (more on this in later sections).

Greater insight into the issue of allowable signal swing can be obtained by examining the signal waveforms shown in Fig. 7.5. Note that for the MOSFET to remain in saturation at the negative peak of v_{ds}, we must ensure that

$$v_{DS\min} \geq v_{GS\max} - V_t$$

that is,

$$0.4 - |A_v|\hat{v}_{gs} \geq 0.6 + \hat{v}_{gs} - 0.4$$

which results in

$$\hat{v}_{gs} \leq \frac{0.2}{|A_v| + 1} = 13.3\text{ mV}$$

This result differs slightly from the one obtained earlier.

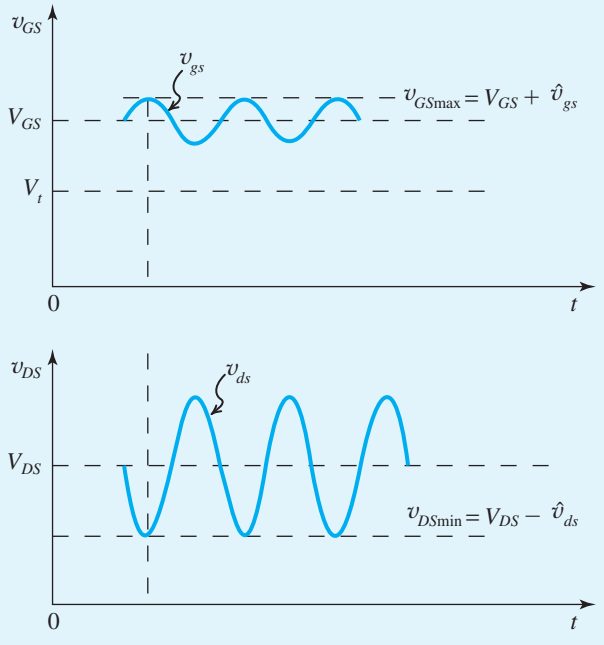


Figure 7.5 Signal waveforms at gate and drain for the amplifier in Example 7.1. Note that to ensure operation in the saturation region at all times, $v_{DS\min} \geq v_{GS\max} - V_t$.

EXERCISE

D7.2 For the amplifier circuit studied in Example 7.1, create two alternative designs, each providing a voltage gain of −10 by (a) changing R_D while keeping V_{OV} constant and (b) changing V_{OV} while keeping R_D constant. For each design, specify V_{GS}, I_D, R_D, and V_{DS}.
Ans. (a) 0.6 V, 0.08 mA, 12.5 kΩ, 0.8 V; (b) 0.54 V, 0.04 mA, 17.5 kΩ, 1.1 V

The BJT Case A similar development can be used to obtain the small-signal voltage gain of the BJT amplifier shown in Fig. 7.6,

$$A_v = \left.\frac{dv_{CE}}{dv_{BE}}\right|_{v_{BE}=V_{BE}} \tag{7.19}$$

Utilizing Eq. (7.9) together with Eq. (7.12), we obtain

$$A_v = -\left(\frac{I_C}{V_T}\right)R_C \tag{7.20}$$

We make the following observations on this expression for the voltage gain:

1. The gain is negative, which signifies that the amplifier is inverting; that is, there is a 180° phase shift between the input and the output. This inversion should have been anticipated from Eq. (7.9).
2. The gain is proportional to the collector bias current I_C and to the load resistance R_C.

Additional insight into the voltage gain A_v can be obtained by expressing Eq. (7.20) as

$$A_v = -\frac{I_C R_C}{V_T} \tag{7.21}$$

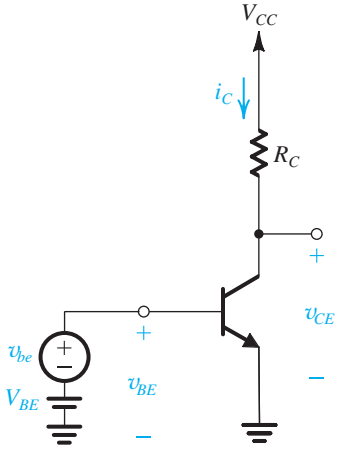


Figure 7.6 BJT amplifier biased at a point Q, with a small voltage signal v_{be} superimposed on the dc bias voltage V_{BE}. The resulting output signal v_{ce} appears superimposed on the dc collector voltage V_{CE}. The amplitude of v_{ce} is larger than that of v_{be} by the voltage gain A_v.

That is, the gain is the ratio of the dc voltage drop across the load resistance R_C to the physical constant V_T (recall that the thermal voltage $V_T \simeq 25$ mV at room temperature). This relationship is similar in form to that for the MOSFET (Eq. 7.16) except that here the denominator is a physical constant (V_T) rather than a design parameter ($V_{OV}/2$). Usually, $V_{OV}/2$ is larger than (V_T), thus we can obtain higher voltage gain from the BJT amplifier than from the MOSFET amplifier. This should not be surprising, as the exponential i_C–v_{BE} relationship is much steeper than the square-law relationship i_D–v_{GS}.

The gain A_v in Eq. (7.21) can be expressed alternately as

$$A_v = -\frac{V_{CC} - V_{CE}}{V_T} \tag{7.22}$$

from which we see that maximum gain is achieved when V_{CE} is at its minimum value of about 0.3 V,

$$|A_{v\max}| = \frac{V_{CC} - 0.3}{V_T} \tag{7.23}$$

Here again, this is only a theoretical maximum, since biasing the BJT at the edge of saturation leaves no room for negative signal swing at the output. Equation (7.23) nevertheless provides an upper bound on the voltage gain achievable from the basic BJT amplifier. As an example, for $V_{CC} = 5$ V, the maximum gain is 188 V/V, considerably larger than in the MOSFET case. For modern low-voltage technologies, a V_{CC} of 1.3 V provides a gain of 40 V/V, again much larger than the MOSFET case. The reader should not, however, jump to the conclusion that the BJT is preferred to the MOSFET in the design of modern integrated-circuit amplifiers; in fact, the opposite is true, as we shall see in Chapter 8 and beyond.

Finally, we conclude from Eq. (7.22) that to maximize $|A_v|$ the transistor should be biased at the lowest possible V_{CE} consistent with the desired value of negative signal swing at the output.

Example 7.2

Consider an amplifier circuit using a BJT having $I_S = 10^{-15}$ A, a collector resistance $R_C = 6.8$ kΩ, and a power supply $V_{CC} = 10$ V.

(a) Determine the value of the bias voltage V_{BE} required to operate the transistor at $V_{CE} = 3.2$ V. What is the corresponding value of I_C?

(b) Find the voltage gain A_v at this bias point. If an input sine-wave signal of 5-mV peak amplitude is superimposed on V_{BE}, find the amplitude of the output sine-wave signal (assume linear operation).

(c) Find the positive increment in v_{BE} (above V_{BE}) that drives the transistor to the edge of saturation, where $v_{CE} = 0.3$ V.

(d) Find the negative increment in v_{BE} that drives the transistor to within 1% of cutoff (i.e., to $v_{CE} = 0.99V_{CC}$).

Solution

(a)

$$I_C = \frac{V_{CC} - V_{CE}}{R_C}$$
$$= \frac{10 - 3.2}{6.8} = 1 \text{ mA}$$

The value of V_{BE} can be determined from

$$1 \times 10^{-3} = 10^{-15} e^{V_{BE}/V_T}$$

which results in

$$V_{BE} = 690.8 \text{ mV}$$

(b)

$$A_v = -\frac{V_{CC} - V_{CE}}{V_T}$$
$$= \frac{10 - 3.2}{0.025} = -272 \text{ V/V}$$
$$\hat{v}_{ce} = 272 \times 0.005 = 1.36 \text{ V}$$

(c) For $v_{CE} = 0.3$ V,

$$i_C = \frac{10 - 0.3}{6.8} = 1.617 \text{ mA}$$

To increase i_C from 1 mA to 1.617 mA, v_{BE} must be increased by

$$\triangle v_{BE} = V_T \ln\left(\frac{1.617}{1}\right)$$
$$= 12 \text{ mV}$$

(d) For $v_{CE} = 0.99V_{CC} = 9.9$ V,

$$i_C = \frac{10 - 9.9}{6.8} = 0.0147 \text{ mA}$$

To decrease i_C from 1 mA to 0.0147 mA, v_{BE} must change by

$$\triangle v_{BE} = V_T \ln\left(\frac{0.0147}{1}\right)$$
$$= -105.5 \text{ mV}$$

EXERCISE

7.3 For the situation described in Example 7.2, while keeping I_C unchanged at 1 mA, find the value of R_C that will result in a voltage gain of −320 V/V. What is the largest negative signal swing allowed at the output (assume that v_{CE} is not to decrease below 0.3 V)? What (approximately) is the corresponding input signal amplitude? (Assume linear operation.)

Ans. 8 kΩ; 1.7 V; 5.3 mV

7.1.6 Determining the VTC by Graphical Analysis

Figure 7.7 shows a graphical method for determining the VTC of the amplifier of Fig. 7.4(a). Although graphical analysis of transistor circuits is rarely employed in practice, it is useful to us at this stage for gaining greater insight into circuit operation, especially in answering the question of where to locate the bias point Q.

The graphical analysis is based on the observation that for each value of v_{GS}, the circuit will be operating at the point of intersection of the i_D–v_{DS} graph corresponding to the particular

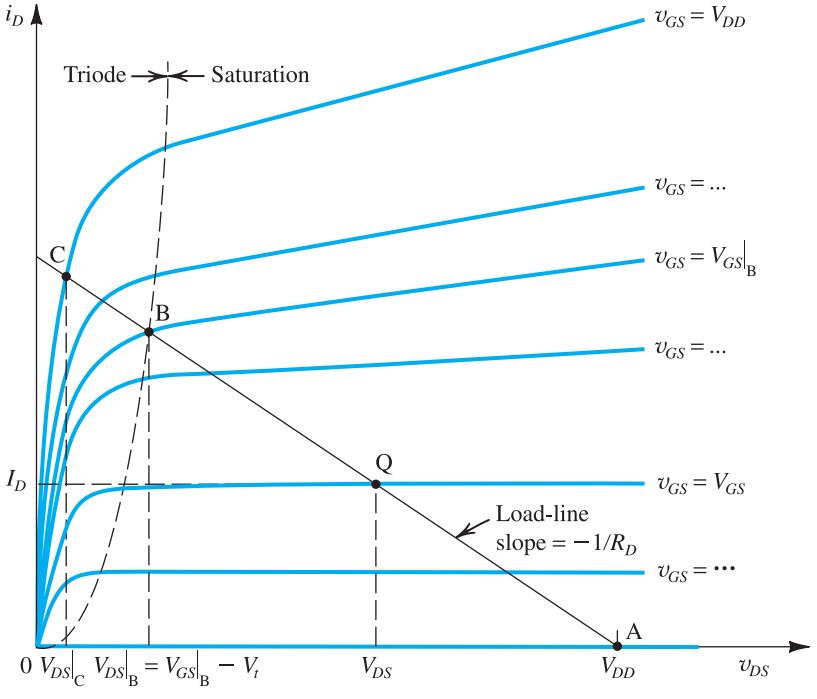


Figure 7.7 Graphical construction to determine the voltage-transfer characteristic of the amplifier in Fig. 7.4(a).

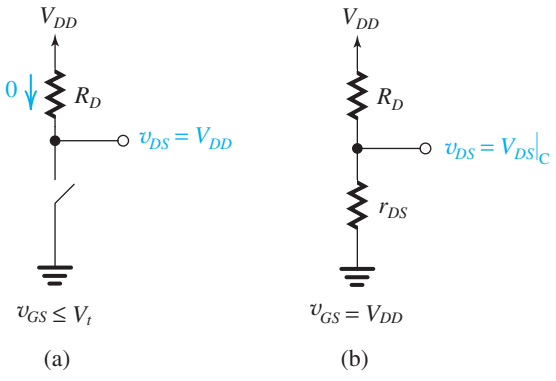


Figure 7.8 Operation of the MOSFET in Fig. 7.4(a) as a switch: **(a)** open, corresponding to point A in Fig. 7.7; **(b)** closed, corresponding to point C in Fig. 7.7. The closure resistance is approximately equal to r_{DS} because V_{DS} is usually very small.

value of v_{GS} and the straight line representing Eq. (7.3), which can be rewritten in the form

$$i_D = \frac{V_{DD}}{R_D} - \frac{1}{R_D} v_{DS} \tag{7.24}$$

The straight line representing this relationship is superimposed on the $i_D - v_{DS}$ characteristics in Fig. 7.7. It intersects the horizontal axis at $v_{DS} = V_{DD}$ and has a slope of $-1/R_D$. Since this straight line represents in effect the load resistance R_D, it is called the **load line**. The VTC is then determined point by point. Note that we have labeled four important points: point A at which $v_{GS} = V_t$, point Q at which the MOSFET can be biased for amplifier operation ($v_{GS} = V_{GS}$ and $v_{DS} = V_{DS}$), point B at which the MOSFET leaves saturation and enters the triode region, and point C, which is deep into the triode region and for which $v_{GS} = V_{DD}$. If the MOSFET is to be used as a switch, then operating points A and C are applicable: At A the transistor is off (open switch), and at C the transistor operates as a low-valued resistance r_{DS} and has a small voltage drop (closed switch). The incremental resistance at point C is also known as the **closure resistance**. The operation of the MOSFET as a switch is illustrated in Fig. 7.8. A detailed study of the application of the MOSFET as a switch is undertaken in Chapter 14, dealing with CMOS digital logic circuits.

The graphical analysis method above can be applied to determine the VTC of the BJT amplifier in Fig. 7.2(c). Here point A, Fig. 7.2(d), corresponds to the BJT just turning on ($v_{BE} \simeq 0.5$ V) and point B corresponds to the BJT leaving the active region and entering the saturation region. If the BJT is to be operated as a switch, the two modes of operation are cutoff (open switch) and saturation (closed switch). As discussed in Section 6.2, in saturation, the BJT has a small closure resistance $R_{CE\text{sat}}$ as well as an offset voltage. More seriously, switching the BJT out of its saturation region can require a relatively long delay time to ensure the removal of the charge stored in the BJT base region. This phenomenon has made the BJT much less attractive in digital logic applications relative to the MOSFET.[2]

7.1.7 Deciding on a Location for the Bias Point Q

For the MOSFET amplifier, the bias point Q is determined by the value of V_{GS} and that of the load resistance R_D. Two important considerations in deciding on the location of Q

[2]The only exception is a nonsaturating form of BJT logic circuits known as emitter-coupled logic (ECL).

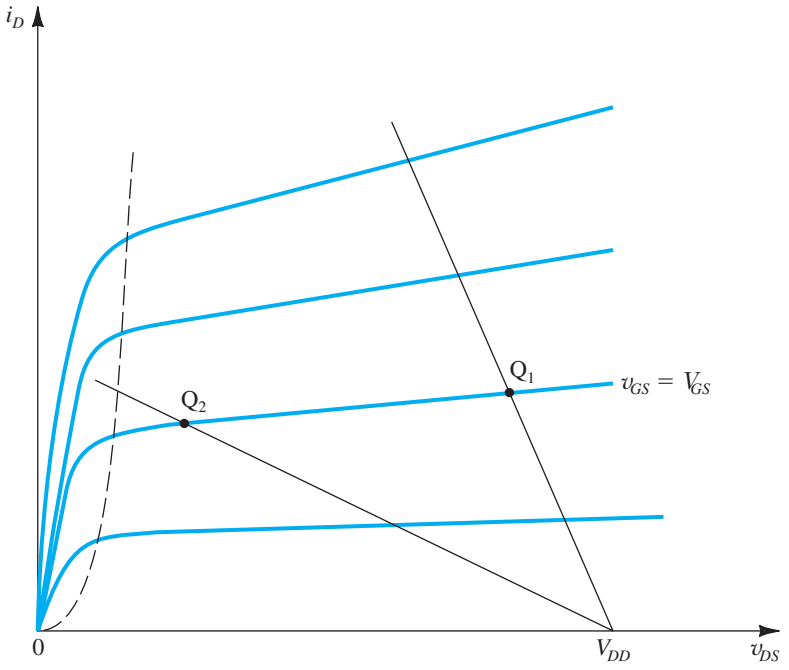


Figure 7.9 Two load lines and corresponding bias points. Bias point Q_1 does not leave sufficient room for positive signal swing at the drain (too close to V_{DD}). Bias point Q_2 is too close to the boundary of the triode region and might not allow for sufficient negative signal swing.

are the required gain and the desired signal swing at the output. To illustrate, consider the VTC shown in Fig. 7.4(b). Here the value of R_D is fixed and the only variable remaining is the value of V_{GS}. Since the slope increases as we move closer to point B, we obtain higher gain by locating Q as close to B as possible. However, the closer Q is to the boundary point B, the smaller the allowable magnitude of negative signal swing. Thus, as often happens in engineering design, we encounter a situation requiring a trade-off. The answer here is relatively simple: For a given R_D, locate Q as close to the triode region (point B) as possible to obtain high gain but sufficiently distant to allow for the required negative signal swing.

In deciding on a value for R_D, it is useful to refer to the $i_D - v_{DS}$ plane. Figure 7.9 shows two load lines resulting in two extreme bias points: Point Q_1 is too close to V_{DD}, resulting in a severe constraint on the positive signal swing of v_{ds}. Exceeding the allowable positive maximum results in the positive peaks of the signal being clipped off, since the MOSFET will turn off for the part of each cycle near the positive peak. We speak of this situation by saying that the circuit does not have sufficient "headroom." Similarly, point Q_2 is too close to the boundary of the triode region, thus severely limiting the allowable negative signal swing of v_{ds}. Exceeding this limit would result in the transistor entering the triode region for part of each cycle near the negative peaks, resulting in a distorted output signal. In this situation we say that the circuit does not have sufficient "legroom." We will have more to say on bias design in Section 7.4.

Finally, we note that exactly similar considerations apply to the case of the BJT amplifier.

7.2 Small-Signal Operation and Models

In our study of the operation of the MOSFET and BJT amplifiers in Section 7.1 we learned that linear amplification can be obtained by biasing the transistor to operate in the active region and by keeping the input signal small. In this section, we explore the small-signal operation in greater detail.

7.2.1 The MOSFET Case

Consider the conceptual amplifier circuit shown in Fig. 7.10. Here the MOS transistor is biased by applying a dc voltage[3] V_{GS}, and the input signal to be amplified, v_{gs}, is superimposed on the dc bias voltage V_{GS}. The output voltage is taken at the drain.

The DC Bias Point The dc bias current I_D can be found by setting the signal v_{gs} to zero; thus,

$$I_D = \frac{1}{2}k_n(V_{GS} - V_t)^2 = \frac{1}{2}k_n V_{OV}^2 \tag{7.25}$$

where we have neglected channel-length modulation (i.e., we have assumed $\lambda = 0$). Here $V_{OV} = V_{GS} - V_t$ is the overdrive voltage at which the MOSFET is biased to operate. The dc voltage at the drain, V_{DS}, will be

$$V_{DS} = V_{DD} - R_D I_D \tag{7.26}$$

To ensure saturation-region operation, we must have

$$V_{DS} > V_{OV}$$

Furthermore, since the total voltage at the drain will have a signal component superimposed on V_{DS}, V_{DS} has to be sufficiently greater than V_{OV} to allow for the required negative signal swing.

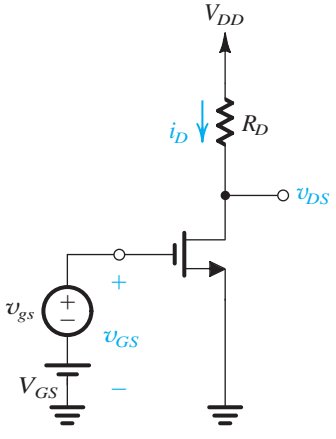


Figure 7.10 Conceptual circuit utilized to study the operation of the MOSFET as a small-signal amplifier.

[3] Practical biasing arrangements will be studied in Section 7.4.

The Signal Current in the Drain Terminal Next, consider the situation with the input signal v_{gs} applied. The total instantaneous gate-to-source voltage will be

$$v_{GS} = V_{GS} + v_{gs} \tag{7.27}$$

resulting in a total instantaneous drain current i_D,

$$\begin{aligned} i_D &= \frac{1}{2}k_n\left(V_{GS} + v_{gs} - V_t\right)^2 \\ &= \frac{1}{2}k_n(V_{GS} - V_t)^2 + k_n(V_{GS} - V_t)v_{gs} + \frac{1}{2}k_n v_{gs}^2 \end{aligned} \tag{7.28}$$

The first term on the right-hand side of Eq. (7.28) can be recognized as the dc bias current I_D (Eq. 7.25). The second term represents a current component that is directly proportional to the input signal v_{gs}. The third term is a current component that is proportional to the square of the input signal. This last component is undesirable because it represents *nonlinear distortion*. To reduce the nonlinear distortion introduced by the MOSFET, the input signal should be kept small so that

$$\frac{1}{2}k_n v_{gs}^2 \ll k_n(V_{GS} - V_t)v_{gs}$$

resulting in

$$v_{gs} \ll 2(V_{GS} - V_t) \tag{7.29}$$

or, equivalently,

$$v_{gs} \ll 2V_{OV} \tag{7.30}$$

If this **small-signal condition** is satisfied, we may neglect the last term in Eq. (7.28) and express i_D as

$$i_D \simeq I_D + i_d \tag{7.31}$$

where

$$i_d = k_n(V_{GS} - V_t)v_{gs}$$

The parameter that relates i_d and v_{gs} is the MOSFET **transconductance** g_m,

$$g_m \equiv \frac{i_d}{v_{gs}} = k_n(V_{GS} - V_t) \tag{7.32}$$

or in terms of the overdrive voltage V_{OV},

$$g_m = k_n V_{OV} \tag{7.33}$$

Figure 7.11 presents a graphical interpretation of the small-signal operation of the MOSFET amplifier. Note that g_m is equal to the slope of the i_D–v_{GS} characteristic at the bias point,

$$g_m \equiv \left.\frac{\partial i_D}{\partial v_{GS}}\right|_{v_{GS}=V_{GS}} \tag{7.34}$$

This is the formal definition of g_m, which can be shown to yield the expressions given in Eqs. (7.32) and (7.33).

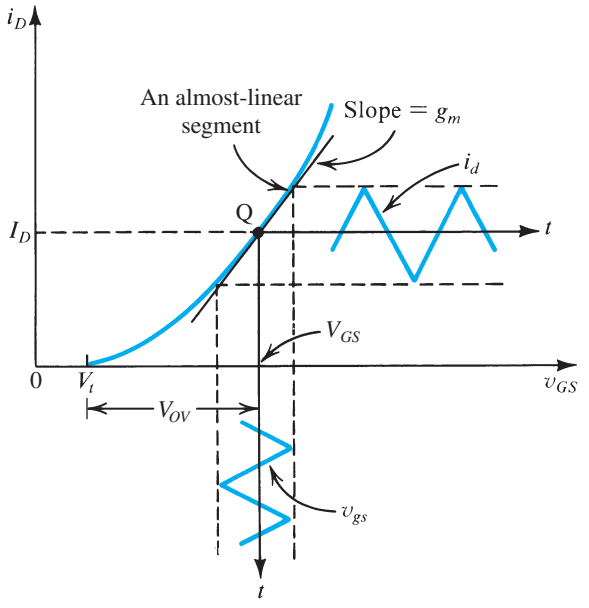


Figure 7.11 Small-signal operation of the MOSFET amplifier.

The Voltage Gain Returning to the circuit of Fig. 7.10, we can express the total instantaneous drain voltage v_{DS} as follows:

$$v_{DS} = V_{DD} - R_D i_D$$

Under the small-signal condition, we have

$$v_{DS} = V_{DD} - R_D(I_D + i_d)$$

which can be rewritten as

$$v_{DS} = V_{DS} - R_D i_d$$

Thus the signal component of the drain voltage is

$$v_{ds} = -i_d R_D = -g_m v_{gs} R_D \tag{7.35}$$

which indicates that the voltage gain is given by

$$A_v \equiv \frac{v_{ds}}{v_{gs}} = -g_m R_D \tag{7.36}$$

The minus sign in Eq. (7.36) indicates that the output signal v_{ds} is 180° out of phase with respect to the input signal v_{gs}. This is illustrated in Fig. 7.12, which shows v_{GS} and v_{DS}. The input signal is assumed to have a triangular waveform with an amplitude much smaller than $2(V_{GS} - V_t)$, the small-signal condition in Eq. (7.29), to ensure linear operation. For operation in the saturation (active) region at all times, the minimum value of v_{DS} should not fall below the corresponding value of v_{GS} by more than V_t. Also, the maximum value of v_{DS} should be

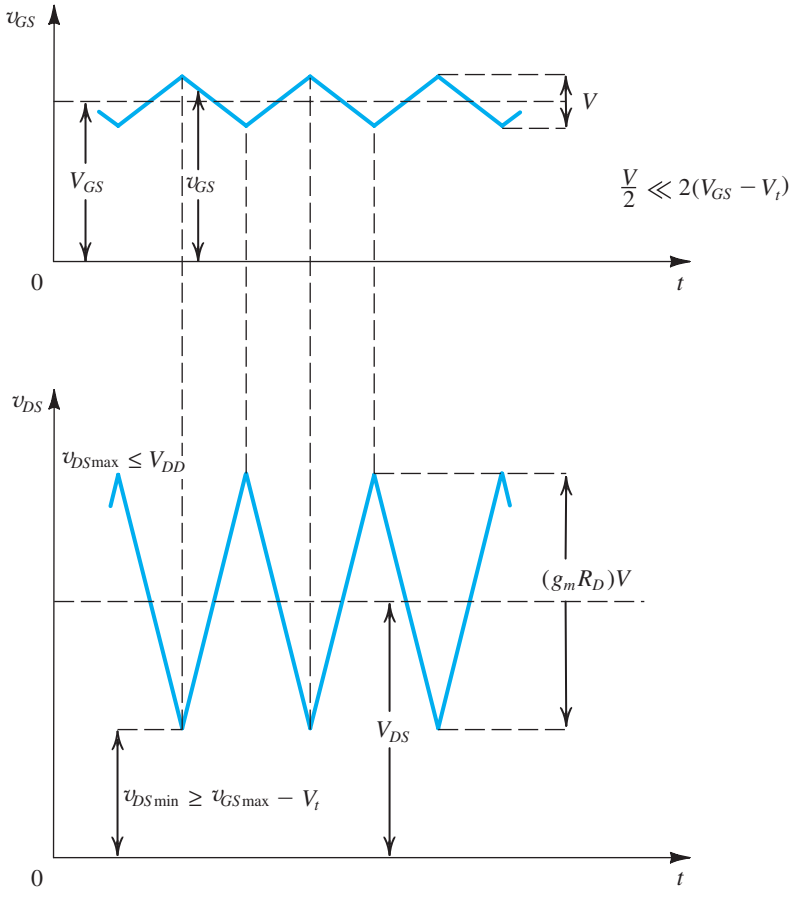


Figure 7.12 Total instantaneous voltages v_{GS} and v_{DS} for the circuit in Fig. 7.10.

smaller than V_{DD}; otherwise the FET will enter the cutoff region and the peaks of the output signal waveform will be clipped off.

Finally, we note that by substituting for g_m from Eq. (7.33) the voltage-gain expression in Eq. (7.36) becomes identical to that derived in Section 7.1—namely, Eq. (7.15).

Separating the DC Analysis and the Signal Analysis From the preceding analysis, we see that under the small-signal approximation, signal quantities are superimposed on dc quantities. For instance, the total drain current i_D equals the dc current I_D plus the signal current i_d, the total drain voltage $v_{DS} = V_{DS} + v_{ds}$, and so on. It follows that the analysis and design can be greatly simplified by separating dc or bias calculations from small-signal calculations. That is, once a stable dc operating point has been established and all dc quantities calculated, we may then perform signal analysis ignoring dc quantities.

Small-Signal Equivalent-Circuit Models From a signal point of view, the FET behaves as a voltage-controlled current source. It accepts a signal v_{gs} between gate and source and provides a current $g_m v_{gs}$ at the drain terminal. The input resistance of this controlled source is very high—ideally, infinite. The output resistance—that is, the resistance looking into the

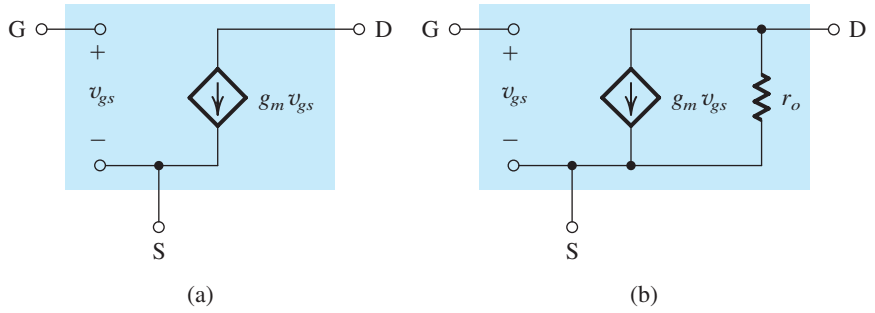


Figure 7.13 Small-signal models for the MOSFET: **(a)** neglecting the dependence of i_D on v_{DS} in the active region (the channel-length modulation effect) and **(b)** including the effect of channel-length modulation, modeled by output resistance $r_o = |V_A|/I_D$. These models apply equally well for both NMOS and PMOS transistors.

drain—also is high, and we have assumed it to be infinite thus far. Putting all of this together, we arrive at the circuit in Fig. 7.13(a), which represents the small-signal operation of the MOSFET and is thus a **small-signal model** or a **small-signal equivalent circuit**.

In the analysis of a MOSFET amplifier circuit, the transistor can be replaced by the equivalent-circuit model shown in Fig. 7.13(a). The rest of the circuit remains unchanged except that *ideal constant dc voltage sources are replaced by short circuits*. This is a result of the fact that the voltage across an ideal constant dc voltage source does not change, and thus there will always be a zero voltage signal across a constant dc voltage source. A dual statement applies for constant dc current sources; namely, the signal current of an ideal constant dc current source will always be zero, and thus *an ideal constant dc current source can be replaced by an open circuit* in the small-signal equivalent circuit of the amplifier. The circuit resulting can then be used to perform any required signal analysis, such as calculating voltage gain.

The most serious shortcoming of the small-signal model of Fig. 7.13(a) is that it assumes the drain current in saturation to be independent of the drain voltage. From our study of the MOSFET characteristics in saturation, we know that the drain current does in fact depend on v_{DS} in a linear manner. Such dependence was modeled by a finite resistance r_o between drain and source, whose value was given by Eq. (5.27) in Section 5.2.4, which we repeat here (with the prime on I_D dropped) as

$$r_o = \frac{|V_A|}{I_D} \tag{7.37}$$

where $V_A = 1/\lambda$ is a MOSFET parameter that either is specified or can be measured. It should be recalled that for a given process technology, V_A is proportional to the MOSFET channel length. The current I_D is the value of the dc drain current without the channel-length modulation taken into account; that is,

$$I_D = \frac{1}{2} k_n V_{OV}^2 \tag{7.38}$$

Typically, r_o is in the range of 10 kΩ to 1000 kΩ. It follows that the accuracy of the small-signal model can be improved by including r_o in parallel with the controlled source, as shown in Fig. 7.13(b).

It is important to note that the small-signal model parameters g_m and r_o depend on the dc bias point of the MOSFET.

Returning to the amplifier of Fig. 7.10, we find that replacing the MOSFET with the small-signal model of Fig. 7.13(b) results in the voltage-gain expression

$$A_v = \frac{v_{ds}}{v_{gs}} = -g_m(R_D \| r_o) \tag{7.39}$$

Thus, the finite output resistance r_o results in a reduction in the magnitude of the voltage gain.

Although the analysis above is performed on an NMOS transistor, the results, and the equivalent-circuit models of Fig. 7.13, apply equally well to PMOS devices, except for using $|V_{GS}|, |V_t|, |V_{OV}|$, and $|V_A|$ and replacing k_n with k_p.

The Transconductance g_m We shall now take a closer look at the MOSFET transconductance given by Eq. (7.32), which we rewrite with $k_n = k_n'(W/L)$ as follows:

$$g_m = k_n'(W/L)(V_{GS} - V_t) = k_n'(W/L)V_{OV} \tag{7.40}$$

This relationship indicates that g_m is proportional to the process transconductance parameter $k_n' = \mu_n C_{ox}$ and to the W/L ratio of the MOS transistor; hence to obtain relatively large transconductance the device must be short and wide. We also observe that for a given device the transconductance is proportional to the overdrive voltage, $V_{OV} = V_{GS} - V_t$, the amount by which the bias voltage V_{GS} exceeds the threshold voltage V_t. Note, however, that increasing g_m by biasing the device at a larger V_{GS} has the disadvantage of reducing the allowable voltage signal swing at the drain.

Another useful expression for g_m can be obtained by substituting for V_{OV} in Eq. (7.40) by $\sqrt{2I_D/(k_n'(W/L))}$ [from Eq. (7.25)]:

$$g_m = \sqrt{2k_n'}\sqrt{W/L}\sqrt{I_D} \tag{7.41}$$

This expression shows two things:

1. For a given MOSFET, g_m is proportional to the square root of the dc bias current.
2. At a given bias current, g_m is proportional to $\sqrt{W/L}$.

In contrast, as we shall see shortly, the transconductance of the bipolar junction transistor (BJT) is proportional to the bias current and is independent of the physical size and geometry of the device.

To gain some insight into the values of g_m obtained in MOSFETs consider an integrated-circuit device operating at $I_D = 0.5$ mA and having $k_n' = 120\ \mu\text{A/V}^2$. Equation (7.41) shows that for $W/L = 1, g_m = 0.35$ mA/V, whereas a device for which $W/L = 100$ has $g_m = 3.5$ mA/V. In contrast, a BJT operating at a collector current of 0.5 mA has $g_m = 20$ mA/V.

Yet another useful expression for g_m of the MOSFET can be obtained by substituting for $k_n'(W/L)$ in Eq. (7.40) by $2I_D/(V_{GS} - V_t)^2$:

$$g_m = \frac{2I_D}{V_{GS} - V_t} = \frac{2I_D}{V_{OV}} \tag{7.42}$$

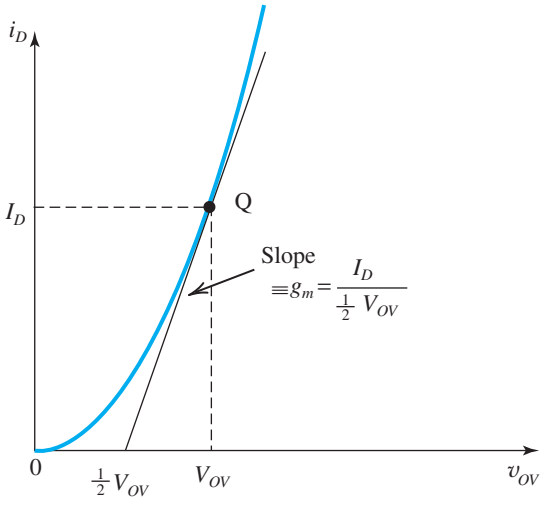


Figure 7.14 The slope of the tangent at the bias point Q intersects the v_{OV} axis at $\frac{1}{2}V_{OV}$. Thus, $g_m = I_D/(\frac{1}{2}V_{OV})$.

A convenient graphical construction that clearly illustrates this relationship is shown in Fig. 7.14.

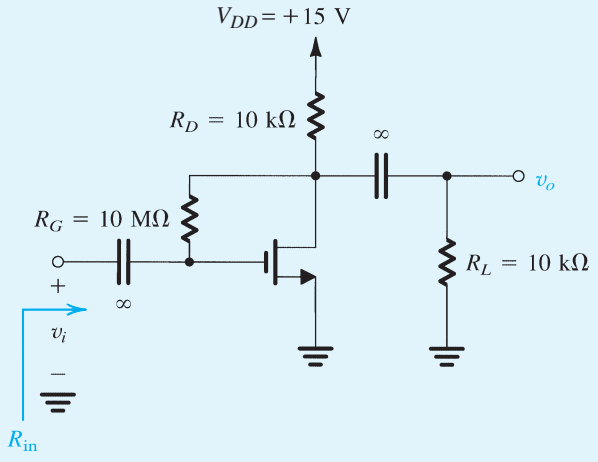
In summary, there are three different relationships for determining g_m—Eqs. (7.40), (7.41), and (7.42)—and there are three design parameters—(W/L), V_{OV}, and I_D, any two of which can be chosen independently. That is, the designer may choose to operate the MOSFET with a certain overdrive voltage V_{OV} and at a particular current I_D; the required W/L ratio can then be found and the resulting g_m determined.[4]

Example 7.3

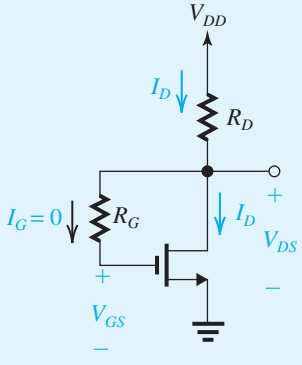
Figure 7.15(a) shows a discrete MOSFET amplifier utilizing a drain-to-gate resistance R_G for biasing purposes. Such a biasing arrangement will be studied in Section 7.4. The input signal v_i is coupled to the gate via a large capacitor, and the output signal at the drain is coupled to the load resistance R_L via another large capacitor. We wish to analyze this amplifier circuit to determine its small-signal voltage gain, its input resistance, and the largest allowable input signal. The transistor has $V_t = 1.5$ V, $k_n'(W/L) = 0.25$ mA/V^2, and $V_A = 50$ V. Assume the coupling capacitors to be sufficiently large so as to act as short circuits at the signal frequencies of interest.

[4]This assumes that the circuit designer is also designing the device, as is typically the case in IC design. On the other hand, a circuit designer working with a discrete-circuit MOSFET obviously does not have the freedom to change its W/L ratio. Thus, in this case there are only two design parameters—V_{OV} and I_D, and only one can be specified by the designer.

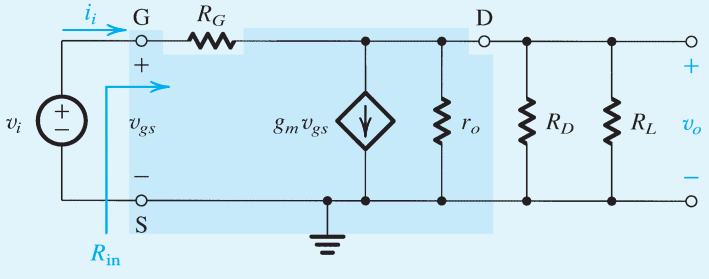
Example 7.3 *continued*



(a)

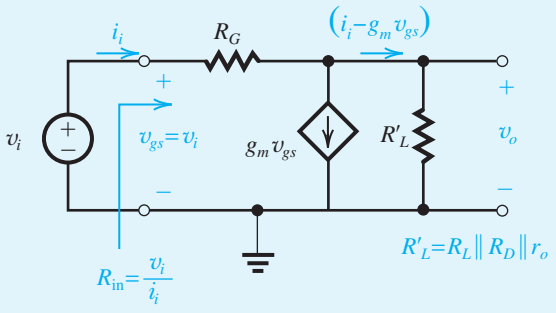


(b)



(c)

Figure 7.15 Example 7.3: **(a)** amplifier circuit; **(b)** circuit for determining the dc operating point; **(c)** the amplifier small-signal equivalent circuit; **(d)** a simplified version of the circuit in **(c)**.



(d)

Figure 7.15 *continued*

Solution

We first determine the dc operating point. For this purpose, we eliminate the input signal v_i, and open-circuit the two coupling capacitors (since they block dc currents). The result is the circuit shown in Fig. 7.14(b). We note that since $I_G = 0$, the dc voltage drop across R_G will be zero, and

$$V_{GS} = V_{DS} = V_{DD} - R_D I_D \tag{7.43}$$

With $V_{DS} = V_{GS}$, the NMOS transistor will be operating in saturation. Thus,

$$I_D = \frac{1}{2} k_n (V_{GS} - V_t)^2 \tag{7.44}$$

where, for simplicity, we have neglected the effect of channel-length modulation on the dc operating point. Substituting $V_{DD} = 15$ V, $R_D = 10$ kΩ, $k_n = 0.25$ mA/V^2, and $V_t = 1.5$ V in Eqs. (7.43) and (7.44), and substituting for V_{GS} from Eq. (7.43) into Eq. (7.44) results in a quadratic equation in I_D. Solving the latter and discarding the root that is not physically meaningful yields the solution

$$I_D = 1.06 \text{ mA}$$

which corresponds to

$$V_{GS} = V_{DS} = 4.4 \text{ V}$$

and

$$V_{OV} = 4.4 - 1.5 = 2.9 \text{ V}$$

Next we proceed with the small-signal analysis of the amplifier. Toward that end we replace the MOSFET with its small-signal model to obtain the small-signal equivalent circuit of the amplifier, shown in Fig. 7.15(c). Observe that we have replaced the coupling capacitors with short circuits. The dc voltage supply V_{DD} has also been replaced with a short circuit to ground.

Example 7.3 *continued*

The values of the transistor small-signal parameters g_m and r_o can be determined by using the dc bias quantities found above, as follows:

$$\begin{aligned} g_m &= k_n V_{OV} \\ &= 0.25 \times 2.9 = 0.725 \text{ mA/V} \\ r_o &= \frac{V_A}{I_D} = \frac{50}{1.06} = 47 \text{ k}\Omega \end{aligned}$$

Next we use the equivalent circuit of Fig. 7.15(c) to determine the input resistance $R_{\text{in}} \equiv v_i/i_i$ and the voltage gain $A_v = v_o/v_i$. Toward that end we simplify the circuit by combining the three parallel resistances r_o, R_D, and R_L in a single resistance R_L',

$$\begin{aligned} R_L' &= R_L||R_D||r_o \\ &= 10||10||47 = 4.52 \text{ k}\Omega \end{aligned}$$

as shown in Fig. 7.15(d). For the latter circuit we can write the two equations

$$v_o = (i_i - g_m v_{gs})R_L' \tag{7.45}$$

and

$$i_i = \frac{v_{gs} - v_o}{R_G} \tag{7.46}$$

Substituting for i_i from Eq. (7.46) into Eq. (7.45) results in the following expression for the voltage gain $A_v \equiv v_o/v_i = v_o/v_{gs}$:

$$A_v = -g_m R_L' \frac{1 - (1/g_m R_G)}{1 + (R_L'/R_G)}$$

Since R_G is very large, $g_m R_G \gg 1$ and $R_L'/R_G \ll 1$ (the reader can easily verify this), and the gain expression can be approximated as

$$A_v \simeq -g_m R_L' \tag{7.47}$$

Substituting $g_m = 0.725$ mA/V and $R_L' = 4.52$ kΩ yields

$$A_v = -3.3 \text{ V/V}$$

To obtain the input resistance, we substitute in Eq. (7.46) for $v_o = A_v v_{gs} = -g_m R_L' v_{gs}$, then use $R_{\text{in}} \equiv v_i/i_i = v_{gs}/i_i$ to obtain

$$R_{\text{in}} = \frac{R_G}{1 + g_m R_L'} \tag{7.48}$$

This is an interesting relationship: The input resistance decreases as the gain $(g_m R_L')$ is increased. The value of R_{in} can now be determined; it is

$$R_{in} = \frac{10\ \text{M}\Omega}{1+3.3} = 2.33\ \text{M}\Omega$$

which is still very large.

The largest allowable input signal $\hat{v}_i$ is constrained by the need to keep the transistor in saturation at all times; that is,

$$v_{DS} \geq v_{GS} - V_t$$

Enforcing this condition with equality at the point v_{GS} is maximum and v_{DS} is minimum, we write

$$v_{DS\,\min} = v_{GS\,\max} - V_t$$

$$V_{DS} - |A_v|\hat{v}_i = V_{GS} + \hat{v}_i - V_t$$

Since $V_{DS} = V_{GS}$, we obtain

$$\hat{v}_i = \frac{V_t}{|A_v| + 1}$$

This is a general relationship that applies to this circuit irrespective of the component values. Observe that it simply states that the maximum signal swing is determined by the fact that the bias arrangement makes $V_D = V_G$ and thus, to keep the MOSFET out of the triode region, the signal between D and G is constrained to be equal to V_t. For our particular design,

$$\hat{v}_i = \frac{1.5}{3.3+1} = 0.35\ \text{V}$$

Since $V_{OV} = 2.9$ V, a v_i of 0.35 is indeed much smaller than $2V_{OV} = 5.8$ V; thus the assumption of linear operation is justified.

A modification of this circuit that increases the allowable signal swing is investigated in Problem 7.103.

EXERCISE

D7.4 Consider the amplifier circuit of Fig. 7.15(a) without the load resistance R_L and with channel-length modulation neglected. Let $V_{DD} = 5$ V, $V_t = 0.7$ V, and $k_n = 1$ mA/V^2. Find V_{OV}, I_D, R_D, and R_G to obtain a voltage gain of -25 V/V and an input resistance of 0.5 M Ω. What is the maximum allowable input signal, $\hat{v}_i$?

Ans. 0.319 V; 50.9 μA; 78.5 kΩ; 13 MΩ; 27 mV

The T Equivalent-Circuit Model Through a simple circuit transformation it is possible to develop an alternative equivalent-circuit model for the MOSFET. The development of such a model, known as the T model, is illustrated in Fig. 7.16. Figure 7.16(a) shows the equivalent circuit studied above without r_o. In Fig. 7.16(b) we have added a second $g_m v_{gs}$ current source in series with the original controlled source. This addition obviously does not change the terminal currents and is thus allowed. The newly created circuit node, labeled X, is joined to the gate terminal G in Fig. 7.16(c). Observe that the gate current does not change—that is, it remains equal to zero—and thus this connection does not alter the terminal characteristics. We now note that we have a controlled current source $g_m v_{gs}$ connected across its control voltage v_{gs}. We can replace this controlled source by a resistance as long as this resistance draws an equal current as the source. (See the source-absorption theorem in Appendix D.) Thus the value of the resistance is $v_{gs}/g_m v_{gs} = 1/g_m$. This replacement is shown in Fig. 7.16(d), which depicts

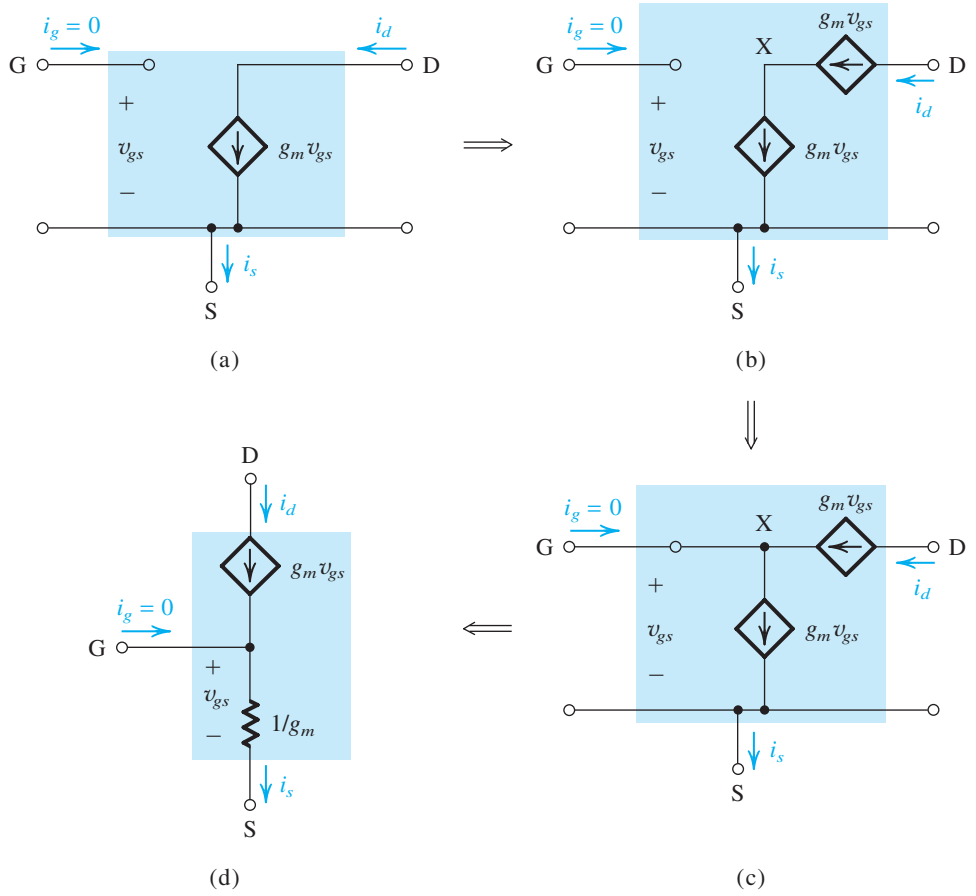


Figure 7.16 Development of the T equivalent-circuit model for the MOSFET. For simplicity, r_o has been omitted; however, it may be added between D and S in the T model of **(d)**.

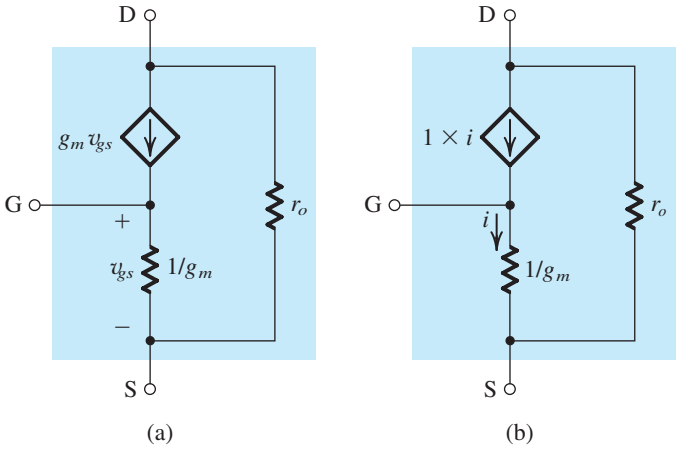


Figure 7.17 **(a)** The T model of the MOSFET augmented with the drain-to-source resistance r_o. **(b)** An alternative representation of the T model.

the alternative model. Observe that i_g is still zero, $i_d = g_m v_{gs}$, and $i_s = v_{gs}/(1/g_m) = g_m v_{gs}$, all the same as in the original model in Fig. 7.16(a).

The model of Fig. 7.16(d) shows that the resistance between gate and source looking into the source is $1/g_m$. This observation and the T model prove useful in many applications. Note that the resistance between gate and source, looking into the gate, is infinite.

In developing the T model we did not include r_o. If desired, this can be done by incorporating in the circuit of Fig. 7.16(d) a resistance r_o between drain and source, as shown in Fig. 7.17(a). An alternative representation of the T model, in which the voltage-controlled current source is replaced with a current-controlled current source, is shown in Fig. 7.17(b).

Finally, we should note that in order to distinguish the model of Fig. 7.13(b) from the equivalent T model, the former is sometimes referred to as the **hybrid-π model**, a carryover from the bipolar transistor literature. The origin of this name will be explained shortly.

Example 7.4

Figure 7.18(a) shows a MOSFET amplifier biased by a constant-current source I. Assume that the values of I and R_D are such that the MOSFET operates in the saturation region. The input signal v_i is coupled to the source terminal by utilizing a large capacitor C_{C1}. Similarly, the output signal at the drain is taken through a large coupling capacitor C_{C2}. Find the input resistance R_{in} and the voltage gain v_o/v_i. Neglect channel-length modulation.

Example 7.4 *continued*

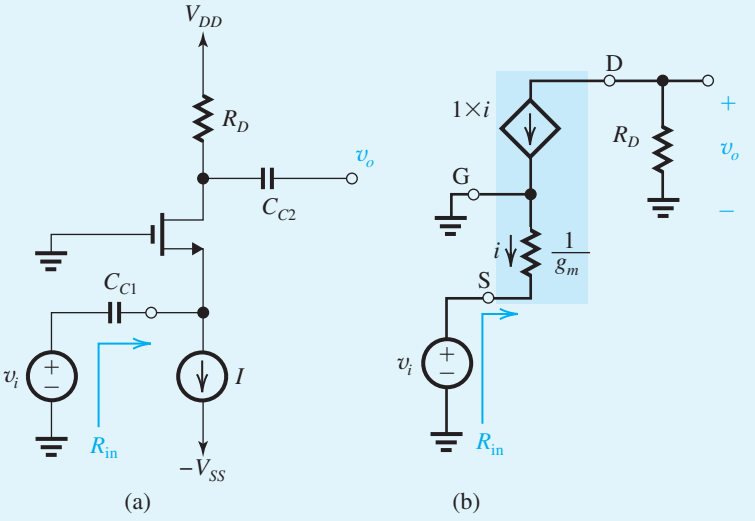


Figure 7.18 **(a)** Amplifier circuit for Example 7.4. **(b)** Small-signal equivalent circuit of the amplifier in **(a)**.

Solution

Replacing the MOSFET with its T equivalent-circuit model results in the amplifier equivalent circuit shown in Fig. 7.18(b). Observe that the dc current source I is replaced with an open circuit and the dc voltage source V_{DD} is replaced by a short circuit. The large coupling capacitors have been replaced by short circuits. From the equivalent-circuit model we determine

$$R_{\text{in}} = \frac{v_i}{-i} = 1/g_m$$

and

$$v_o = -iR_D = \left(\frac{v_i}{1/g_m}\right)R_D = g_m R_D v_i$$

Thus,

$$A_v \equiv \frac{v_o}{v_i} = g_m R_D$$

We note that this amplifier, known as the common-gate amplifier because the gate at ground potential is common to both the input and output ports, has a low input resistance $\left(1/g_m\right)$ and a noninverting gain. We shall study this amplifier type in Section 7.3.5.

EXERCISE

7.5 Use the T model of Fig. 7.17(b) to show that a MOSFET whose drain is connected to its gate exhibits an incremental resistance equal to $[(1/g_m) \| r_o]$.
Ans. See Fig. E7.5.

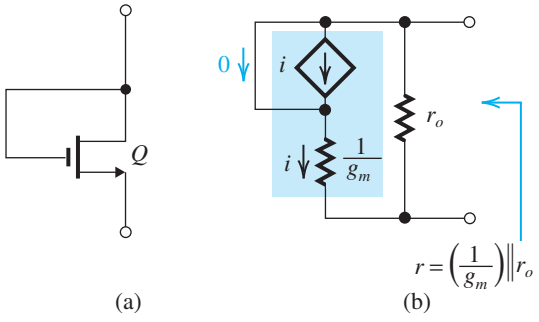


Figure E7.5 Circuits for Exercise 7.5. Note that the bias arrangement of Q is not shown.

Modeling the Body Effect As mentioned earlier (see Section 5.4), the body effect occurs in a MOSFET when the source is not tied to the substrate (which is always connected to the most negative power supply in the integrated circuit for n-channel devices and to the most positive for p-channel devices). Thus the substrate (body) will be at signal ground, but since the source is not, a signal voltage v_{bs} develops between the body (B) and the source (S). The substrate then acts as a "second gate" or a **backgate** for the MOSFET. Thus the signal v_{bs} gives rise to a drain-current component, which we shall write as $g_{mb}v_{bs}$, where g_{mb} is the **body transconductance**, defined as

$$g_{mb} \equiv \left.\frac{\partial i_D}{\partial v_{BS}}\right|_{\substack{v_{GS}=\text{constant} \\ v_{DS}=\text{constant}}} \tag{7.49}$$

Recalling that i_D depends on v_{BS} through the dependence of V_t on V_{BS}, we can show that

$$g_{mb} = \chi g_m \tag{7.50}$$

where

$$\chi \equiv \frac{\partial V_t}{\partial V_{SB}} = \frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} \tag{7.51}$$

Typically the value of χ lies in the range 0.1 to 0.3.

Figure 7.19 shows the MOSFET model augmented to include the controlled source $g_{mb}v_{bs}$ that models the body effect. Ideally, this is the model to be used whenever the source is not connected to the substrate. It has been found, however, that except in some very particular

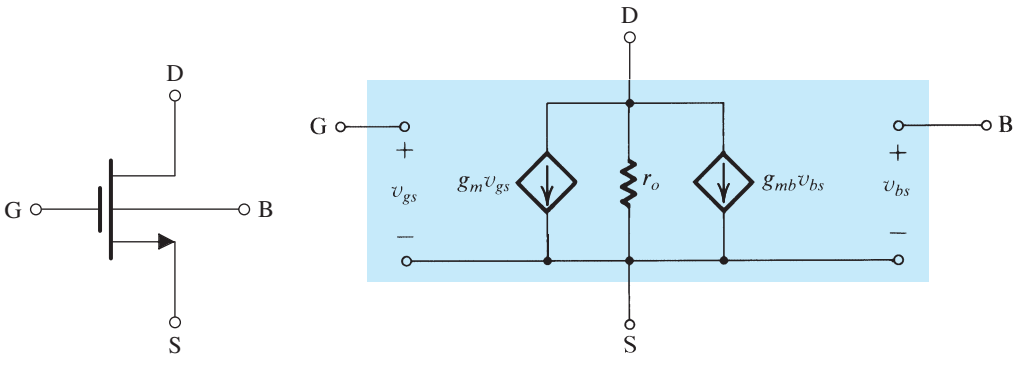


Figure 7.19 Small-signal, equivalent-circuit model of a MOSFET in which the source is not connected to the body.

situations, the body effect can generally be ignored in the initial, pencil-and-paper design of MOSFET amplifiers.

Finally, although the analysis above was performed on an NMOS transistor, the results and the equivalent circuit of Fig. 7.19 apply equally well to PMOS transistors, except for using $|V_{GS}|$, $|V_t|$, $|V_{OV}|$, $|V_A|$, $|V_{SB}|$, $|\gamma|$, and $|\lambda|$ and replacing k_n' with k_p' in the appropriate formula.

EXERCISES

7.6 For the amplifier in Fig. 7.4, let $V_{DD} = 5$ V, $R_D = 10$ kΩ, $V_t = 1$ V, $k_n' = 20\ \mu$A/V^2, $W/L = 20$, $V_{GS} = 2$ V, and $\lambda = 0$.

(a) Find the dc current I_D and the dc voltage V_{DS}.

(b) Find g_m.

(c) Find the voltage gain.

(d) If $v_{gs} = 0.2 \sin \omega t$ volts, find v_{ds} assuming that the small-signal approximation holds. What are the minimum and maximum values of v_{DS}?

(e) Use Eq. (7.28) to determine the various components of i_D. Using the identity ($\sin^2 \omega t = \frac{1}{2} - \frac{1}{2} \cos 2\,\omega t$), show that there is a slight shift in I_D (by how much?) and that there is a second-harmonic component (i.e., a component with frequency $2\,\omega$). Express the amplitude of the second-harmonic component as a percentage of the amplitude of the fundamental. (This value is known as the second-harmonic distortion.)

Ans. (a) 200 μA, 3 V; (b) 0.4 mA/V; (c) –4 V/V; (d) $v_{ds} = -0.8 \sin \omega t$ volts, 2.2 V, 3.8 V; (e) $i_D = (204 + 80 \sin \omega t - 4 \cos 2\,\omega t)$ μA, 5%

7.7 An NMOS transistor has $\mu_n C_{ox} = 60\ \mu$A/V^2, $W/L = 40$, $V_t = 1$ V, and $V_A = 15$ V. Find g_m and r_o when (a) the bias voltage $V_{GS} = 1.5$ V, (b) the bias current $I_D = 0.5$ mA.

Ans. (a) 1.2 mA/V, 50 kΩ; (b) 1.55 mA/V, 30 kΩ

7.8 A MOSFET is to operate at $I_D = 0.1$ mA and is to have $g_m = 1$ mA/V. If $k_n' = 50\ \mu A/V^2$, find the required W/L ratio and the overdrive voltage.
Ans. 100; 0.2 V

7.9 For a fabrication process for which $\mu_p \simeq 0.4\,\mu_n$, find the ratio of the width of a PMOS transistor to the width of an NMOS transistor so that the two devices have equal g_m for the same bias conditions. The two devices have equal channel lengths.
Ans. 2.5

7.10 A PMOS transistor has $V_t = -1$ V, $k_p' = 60\ \mu A/V^2$, and $W/L = 16\ \mu m/0.8\ \mu m$. Find I_D and g_m when the device is biased at $V_{GS} = -1.6$ V. Also, find the value of r_o if λ (at $L = 1\ \mu m$) = –0.04 V^{-1}.
Ans. 216 μA; 0.72 mA/V; 92.6 kΩ

7.11 Derive an expression for $(g_m r_o)$ in terms of V_A and V_{OV}. As we shall see in Chapter 8, this is an important transistor parameter and is known as the intrinsic gain. Evaluate the value of $g_m r_o$ for an NMOS transistor fabricated in a 0.8-μm CMOS process for which $V_A' = 12.5$ V/μm of channel length. Let the device have minimum channel length and be operated at an overdrive voltage of 0.2 V.
Ans. $g_m r_o = 2V_A/V_{OV}$; 100 V/V

7.2.2 The BJT Case

We next consider the small-signal operation of the BJT and develop small-signal equivalent-circuit models that represent its operation at a given bias point. The following development parallels what we used for the MOSFET except that here we have an added complication: The BJT draws a finite base current. As will be seen shortly, this phenomenon (finite β) manifests itself as a finite input resistance looking into the base of the BJT (as compared to the infinite input resistance looking into the gate of the MOSFET).

Consider the *conceptual* amplifier circuit shown in Fig. 7.20(a). Here the base–emitter junction is forward biased by a dc voltage V_{BE}. The reverse bias of the collector–base junction is established by connecting the collector to another power supply of voltage V_{CC} through a resistor R_C. The input signal to be amplified is represented by the voltage source v_{be} that is superimposed on V_{BE}.

The DC Bias Point We consider first the dc bias conditions by setting the signal v_{be} to zero. The circuit reduces to that in Fig. 7.20(b), and we can write the following relationships for the dc currents and voltages:

$$I_C = I_S e^{V_{BE}/V_T} \tag{7.52}$$

$$I_E = I_C/\alpha \tag{7.53}$$

$$I_B = I_C/\beta \tag{7.54}$$

$$V_{CE} = V_{CC} - I_C R_C \tag{7.55}$$

For active-mode operation, V_{CE} should be greater than $(V_{BE} - 0.4)$ by an amount that allows for the required negative signal swing at the collector.

Table 7.1 Systematic Procedure for the Analysis of Transistor Amplifier Circuits

1. Eliminate the signal source and determine the dc operating point of the transistor.
2. Calculate the values of the parameters of the small-signal model.
3. Eliminate the dc sources by replacing each dc voltage source by a short circuit and each dc current source by an open circuit.
4. Replace the transistor with one of its small-signal, equivalent-circuit models. Although any of the models can be used, one might be more convenient than the others for the particular circuit being analyzed. This point will be made clearer in the next section.
5. Analyze the resulting circuit to determine the required quantities (e.g., voltage gain, input resistance).

Table 7.2 Small-Signal Models of the MOSFET

Small-Signal Parameters

NMOS transistors

- **Transconductance:**

$$g_m = \mu_n C_{ox} \frac{W}{L} V_{OV} = \sqrt{2\mu_n C_{ox} \frac{W}{L} I_D} = \frac{2I_D}{V_{OV}}$$

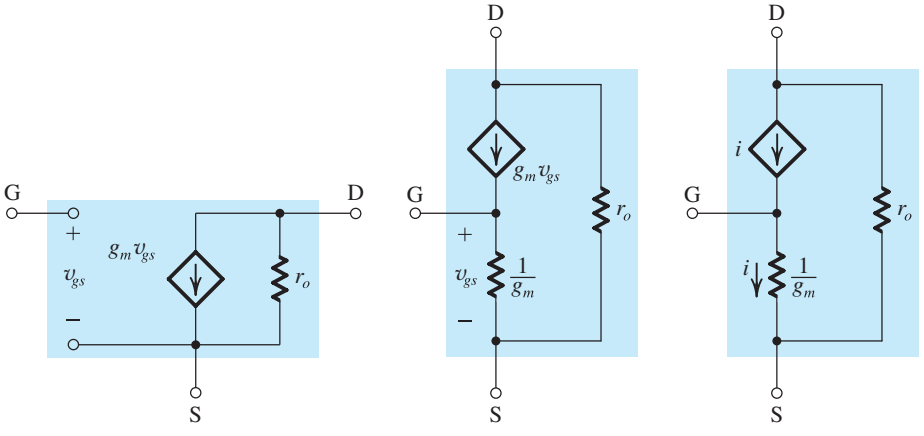
- Output resistance:

$$r_o = V_A / I_D = 1/\lambda I_D$$

PMOS transistors

Same formulas as for NMOS *except* using $|V_{OV}|$, $|V_A|$, $|\lambda|$ and replacing μ_n with μ_p.

Small-Signal, Equivalent-Circuit Models



Hybrid-π model T models

Chapter 3

Single-Stage Amplifiers

Amplification is an essential function in most analog (and many digital) circuits. We amplify an analog or digital signal because it may be too small to drive a load, overcome the noise of a subsequent stage, or provide logical levels to a digital circuit. Amplification also plays a critical role in feedback systems (Chapter 8).

In this chapter, we study the low-frequency behavior of single-stage CMOS amplifiers. Analyzing both the large-signal and the small-signal characteristics of each circuit, we develop intuitive techniques and models that prove useful in understanding more complex systems. An important part of a designer's job is to use proper approximations so as to create a simple mental picture of a complicated circuit. The intuition thus gained makes it possible to formulate the behavior of most circuits by inspection rather than by lengthy calculations.

Following a brief review of basic concepts, we describe in this chapter four types of amplifiers: common-source and common-gate topologies, source followers, and cascode configurations. In each case, we begin with a simple model and gradually add second-order phenomena such as channel-length modulation and body effect.

3.1 Basic Concepts

The input-output characteristic of an amplifier is generally a nonlinear function (Fig. 3.1) that can be approximated by a polynomial over some signal range:

$$y(t) \approx \alpha_0 + \alpha_1 x(t) + \alpha_2 x^2(t) + \cdots + \alpha_n x^n(t) \quad x_1 \leq x \leq x_2. \tag{3.1}$$

The input and output may be current or voltage quantities. For a sufficiently narrow range of x,

$$y(t) \approx \alpha_0 + \alpha_1 x(t), \tag{3.2}$$

where α_0 can be considered the operating (bias) point and α_1 the small-signal gain. So long as $\alpha_1 x(t) \ll \alpha_0$, the bias point is disturbed negligibly, (3.2) provides a reasonable

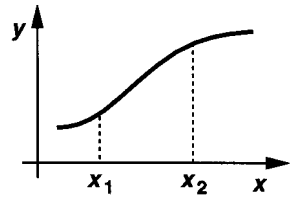


Figure 3.1 Input-output characteristic of a nonlinear system.

approximation, and higher order terms are insignificant. In other words, $\Delta y = \alpha_1 \Delta x$, indicating a linear relationship between the *increments* at the input and output. As $x(t)$ increases in magnitude, higher order terms manifest themselves, leading to nonlinearity and necessitating large-signal analysis. From another point of view, if the slope of the characteristic (the incremental gain) *varies* with the signal level, then the system is nonlinear. These concepts are described in detail in Chapter 13.

What aspects of the performance of an amplifier are important? In addition to gain and speed, such parameters as power dissipation, supply voltage, linearity, noise, or maximum voltage swings may be important. Furthermore, the input and output impedances determine how the circuit interacts with preceding and subsequent stages. In practice, most of these parameters trade with each other, making the design a multi-dimensional optimization problem. Illustrated in the "analog design octagon" of Fig. 3.2, such trade-offs present many challenges in the design of high-performance amplifiers, requiring intuition and experience to arrive at an acceptable compromise.

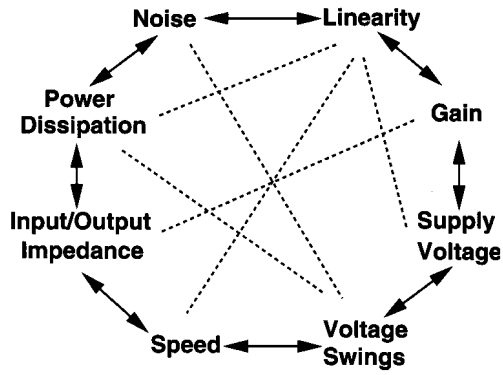


Figure 3.2 Analog design octagon.

3.2 Common-Source Stage

3.2.1 Common-Source Stage with Resistive Load

By virtue of its transconductance, a MOSFET converts variations in its gate-source voltage to a small-signal drain current, which can pass through a resistor to generate an output voltage. Shown in Fig. 3.3(a), the common-source (CS) stage performs such an operation.

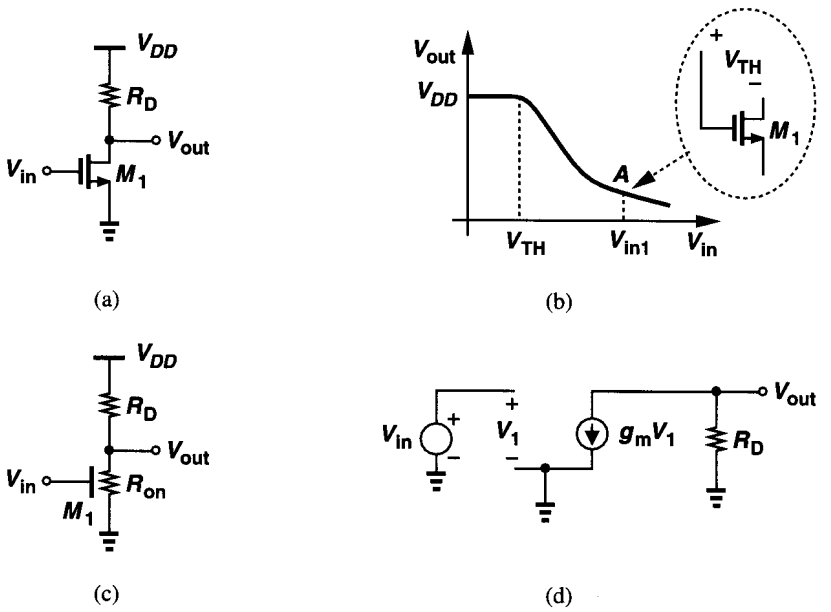


Figure 3.3 (a) Common-source stage, (b) input-output characteristic, (c) equivalent circuit in deep triode region, (d) small-signal model for the saturation region.

We study both the large-signal and the small-signal behavior of the circuit. Note that the input impedance of the circuit is very high at low frequencies.

If the input voltage increases from zero, M_1 is off and $V_{out} = V_{DD}$ [Fig. 3.3(b)]. As V_{in} approaches V_{TH}, M_1 begins to turn on, drawing current from R_D and lowering V_{out}. If V_{DD} is not excessively low, M_1 turns on in saturation, and we have

$$V_{out} = V_{DD} - R_D \frac{1}{2}\mu_n C_{ox} \frac{W}{L}(V_{in} - V_{TH})^2, \tag{3.3}$$

where channel-length modulation is neglected. With further increase in V_{in}, V_{out} drops more and the transistor continues to operate in saturation until V_{in} exceeds V_{out} by V_{TH} [point A in Fig. 3.3(b)]. At this point,

$$V_{in1} - V_{TH} = V_{DD} - R_D \frac{1}{2}\mu_n C_{ox} \frac{W}{L}(V_{in1} - V_{TH})^2, \tag{3.4}$$

from which $V_{in1} - V_{TH}$ and hence V_{out} can be calculated.

For $V_{in} > V_{in1}$, M_1 is in the triode region:

$$V_{out} = V_{DD} - R_D \frac{1}{2}\mu_n C_{ox} \frac{W}{L}\left[2(V_{in} - V_{TH})V_{out} - V_{out}^2\right]. \tag{3.5}$$

If V_{in} is high enough to drive M_1 into deep triode region, $V_{out} \ll 2(V_{in} - V_{TH})$, and, from the equivalent circuit of Fig. 3.3(c),

$$V_{out} = V_{DD}\frac{R_{on}}{R_{on} + R_D} \tag{3.6}$$

$$= \frac{V_{DD}}{1 + \mu_n C_{ox}\dfrac{W}{L}R_D(V_{in} - V_{TH})}. \tag{3.7}$$

Since the transconductance drops in the triode region, we usually ensure that $V_{out} > V_{in} - V_{TH}$, operating to the left of point A in Fig. 3.3(b). Using (3.3) as the input-output characteristic and viewing its slope as the small-signal gain, we have:

$$A_v = \frac{\partial V_{out}}{\partial V_{in}} \tag{3.8}$$

$$= -R_D\mu_n C_{ox}\frac{W}{L}(V_{in} - V_{TH}) \tag{3.9}$$

$$= -g_m R_D. \tag{3.10}$$

This result can be directly derived from the observation that M_1 converts an input voltage change ΔV_{in} to a drain current change $g_m \Delta V_{in}$, and hence an output voltage change $-g_m R_D \Delta V_{in}$. The small-signal model of Fig. 3.3(d) yields the same result.

Even though derived for small-signal operation, the equation $A_v = -g_m R_D$ predicts certain effects if the circuit senses a *large* signal swing. Since g_m itself varies with the input signal according to $g_m = \mu_n C_{ox}(W/L)(V_{GS} - V_{TH})$, the gain of the circuit changes substantially if the signal is large. In other words, if the gain of the circuit *varies* significantly with the signal swing, then the circuit operates in the large-signal mode. The dependence of the gain upon the signal level leads to nonlinearity (Chapter 13), usually an undesirable effect.

A key result here is that to minimize the nonlinearity, the gain equation must be a weak function of signal-dependent parameters such as g_m. We present several examples of this concept in this chapter and in Chapter 13.

Example 3.1

Sketch the drain current and transconductance of M_1 in Fig. 3.3(a) as a function of the input voltage.

Solution

The drain current becomes significant for $V_{in} > V_{TH}$, eventually approaching V_{DD}/R_D if $R_{on1} \ll R_D$ [Fig. 3.4(a)]. Since in saturation, $g_m = \mu_n C_{ox}(W/L)(V_{in} - V_{TH})$, the transconductance begins to rise for $V_{in} > V_{TH}$. In the triode region, $g_m = \mu_n C_{ox}(W/L)V_{DS}$, falling as V_{in} exceeds V_{in1} [Fig. 3.4(b)].

How do we maximize the voltage gain of a common-source stage? Writing (3.10) as

$$A_v = -\sqrt{2\mu_n C_{ox}\frac{W}{L}I_D}\frac{V_{RD}}{I_D}, \tag{3.11}$$

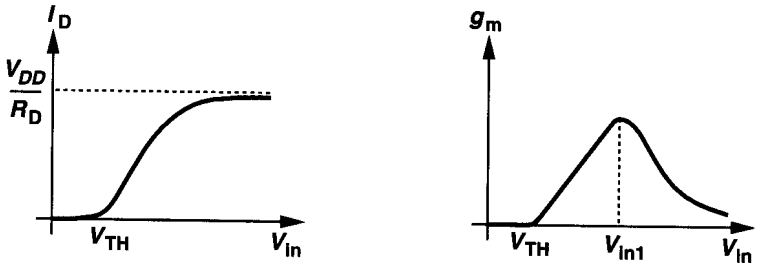


Figure 3.4

where V_{RD} denotes the voltage drop across R_D, we have

$$A_v = -\sqrt{2\mu_n C_{ox}\frac{W}{L}}\frac{V_{RD}}{\sqrt{I_D}}. \tag{3.12}$$

Thus, the magnitude of A_v can be increased by increasing W/L or V_{RD} or decreasing I_D if other parameters are constant. It is important to understand the trade-offs resulting from this equation. A larger device size leads to greater device capacitances, and a higher V_{RD} limits the maximum voltage swings. For example, if $V_{DD} - V_{RD} = V_{in} - V_{TH}$, then M_1 is at the edge of the triode region, allowing only very small swings at the output (and input). If V_{RD} remains constant and I_D is reduced, then R_D must increase, thereby leading to a greater time constant at the output node. In other words, as noted in the analog design octagon, the circuit exhibits trade-offs between gain, bandwidth, and voltage swings. Lower supply voltages further tighten these trade-offs.

For large values of R_D, the effect of channel length modulation in M_1 becomes significant. Modifying (3.4) to include this effect,

$$V_{out} = V_{DD} - R_D\frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{in} - V_{TH})^2(1 + \lambda V_{out}), \tag{3.13}$$

we have

$$\begin{aligned}\frac{\partial V_{out}}{\partial V_{in}} &= -R_D\mu_n C_{ox}\frac{W}{L}(V_{in} - V_{TH})(1 + \lambda V_{out}) \\ &\quad -R_D\frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{in} - V_{TH})^2\lambda\frac{\partial V_{out}}{\partial V_{in}}.\end{aligned} \tag{3.14}$$

Using the approximation $I_D \approx (1/2)\mu_n C_{ox}(W/L)(V_{in} - V_{TH})^2$, we obtain:

$$A_v = -R_D g_m - R_D I_D \lambda A_v \tag{3.15}$$

and hence

$$A_v = -\frac{g_m R_D}{1 + R_D \lambda I_D}. \tag{3.16}$$

Since $\lambda I_D = 1/r_O$,

$$A_v = -g_m \frac{r_O R_D}{r_O + R_D}. \tag{3.17}$$

The small-signal model of Fig. 3.5 gives the same result with much less effort. That is, since

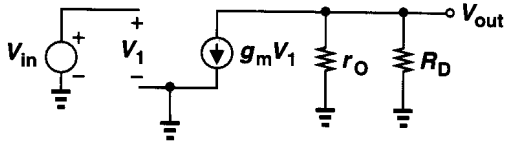


Figure 3.5 Small-signal model of CS stage including the transistor output resistance.

$g_m V_1(r_O \| R_D) = -V_{out}$ and $V_1 = V_{in}$, we have $V_{out}/V_{in} = -g_m(r_O \| R_D)$. Note that, as mentioned in Chapter 1, V_{in}, V_1, and V_{out} in this figure denote small-signal quantities.

Example 3.2

Assuming M_1 in Fig. 3.6 is biased in saturation, calculate the small-signal voltage gain of the circuit.

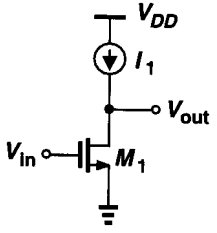


Figure 3.6

Solution

Since I_1 introduces an infinite impedance, the gain is limited by the output resistance of M_1:

$$A_v = -g_m r_O. \tag{3.18}$$

Called the "intrinsic gain" of a transistor, this quantity represents the maximum voltage gain that can be achieved using a single device. In today's CMOS technology, $g_m r_O$ of short-channel devices is between roughly 10 and 30. Thus, we usually assume $1/g_m \ll r_O$.

In Fig. 3.6, Kirchhoff's current law (KCL) requires that $I_{D1} = I_1$. Then, how can V_{in} change the current of M_1 if I_1 is constant? Writing the total drain current of M_1 as

$$I_{D1} = \frac{1}{2}\mu_n C_{ox}(V_{in} - V_{TH})^2(1 + \lambda V_{out}) \tag{3.19}$$

$$= I_1, \tag{3.20}$$

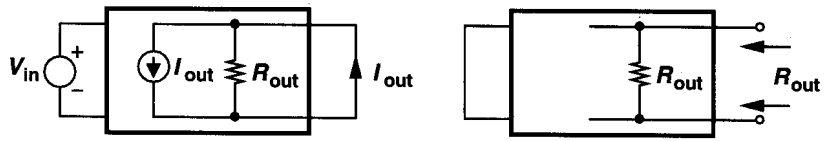


Figure 3.25 Modeling output port of an amplifier by a Norton equivalent.

Defining $G_m = I_{out}/V_{in}$, we have $V_{out} = -G_m V_{in} R_{out}$. This lemma proves useful if G_m and R_{out} can be determined by inspection.

Example 3.6

Calculate the voltage gain of the circuit shown in Fig. 3.26. Assume I_0 is ideal.

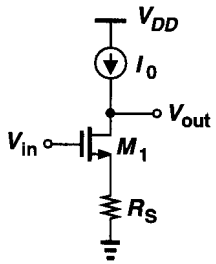


Figure 3.26

Solution

The transconductance and output resistance of the stage are given by Eqs. (3.55) and (3.60), respectively. Thus,

$$A_v = -\frac{g_m r_O}{R_S + [1 + (g_m + g_{mb})R_S]r_O}\{[1 + (g_m + g_{mb})r_O]R_S + r_O\} \tag{3.74}$$

$$= -g_m r_O. \tag{3.75}$$

Interestingly, the voltage gain is equal to the intrinsic gain of the transistor and independent of R_S. This is because, if I_0 is ideal, the current through R_S cannot change and hence the small-signal voltage drop across R_S is zero—as if R_S were zero itself.

3.3 Source Follower

Our analysis of the common-source stage indicates that, to achieve a high voltage gain with limited supply voltage, the load impedance must be as large as possible. If such a stage is to drive a low-impedance load, then a "buffer" must be placed after the amplifier so as to drive the load with negligible loss of the signal level. The source follower (also called the "common-drain" stage) can operate as a voltage buffer.

Illustrated in Fig. 3.27(a), the source follower senses the signal at the gate and drives

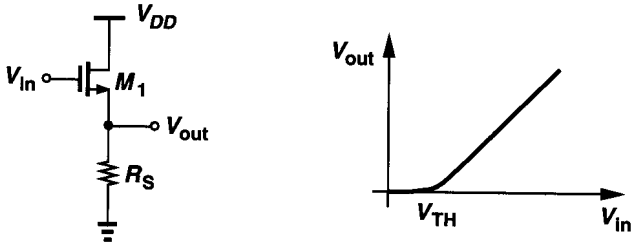


Figure 3.27 (a) Source follower, and (b) its input-output characteristic.

the load at the source, allowing the source potential to "follow" the gate voltage. Beginning with the large-signal behavior, we note that for $V_{in} < V_{TH}$, M_1 is off and $V_{out} = 0$. As V_{in} exceeds V_{TH}, M_1 turns on in saturation (for typical values of V_{DD}) and I_{D1} flows through R_S [Fig. 3.27(b)]. As V_{in} increases further, V_{out} follows the input with a difference (level shift) equal to V_{GS}. We can express the input-output characteristic as:

$$\frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{in} - V_{TH} - V_{out})^2 R_S = V_{out}. \tag{3.76}$$

Let us calculate the small-signal gain of the circuit by differentiating both sides of (3.76) with respect to V_{in}:

$$\frac{1}{2}\mu_n C_{ox}\frac{W}{L}2(V_{in} - V_{TH} - V_{out})\left(1 - \frac{\partial V_{TH}}{\partial V_{in}} - \frac{\partial V_{out}}{\partial V_{in}}\right)R_S = \frac{\partial V_{out}}{\partial V_{in}}. \tag{3.77}$$

Since $\partial V_{TH}/\partial V_{in} = \eta \partial V_{out}/\partial V_{in}$,

$$\frac{\partial V_{out}}{\partial V_{in}} = \frac{\mu_n C_{ox}\dfrac{W}{L}(V_{in} - V_{TH} - V_{out})R_S}{1 + \mu_n C_{ox}\dfrac{W}{L}(V_{in} - V_{TH} - V_{out})R_S(1+\eta)}. \tag{3.78}$$

Also, note that

$$g_m = \mu_n C_{ox}\frac{W}{L}(V_{in} - V_{TH} - V_{out}). \tag{3.79}$$

Consequently,

$$A_v = \frac{g_m R_S}{1 + (g_m + g_{mb})R_S}. \tag{3.80}$$

The same result is more easily obtained with the aid of a small-signal equivalent circuit. From Fig. 3.28, we have $V_{in} - V_1 = V_{out}$, $V_{bs} = -V_{out}$, and $g_m V_1 - g_{mb} V_{out} = V_{out}/R_S$.

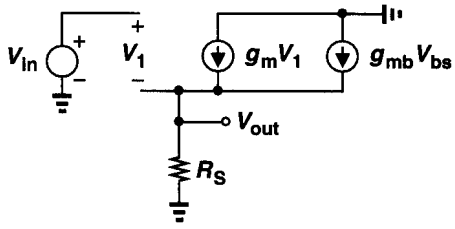


Figure 3.28 Small-signal equivalent circuit of source follower.

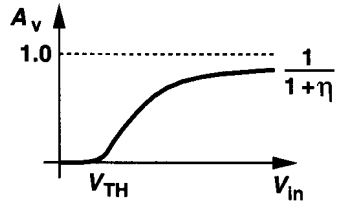


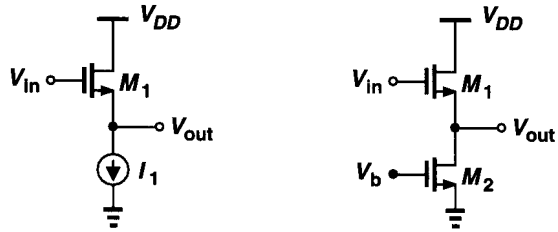
Figure 3.29 Voltage gain of source follower versus input voltage.

Thus, $V_{out}/V_{in} = g_m R_S/[1 + (g_m + g_{mb})R_S]$.

Sketched in Fig. 3.29 vs. V_{in}, the voltage gain begins from zero for $V_{in} \approx V_{TH}$ (that is, $g_m \approx 0$) and monotonically increases. As the drain current and g_m increase, A_v approaches $g_m/(g_m + g_{mb}) = 1/(1+\eta)$. Since η itself slowly decreases with V_{out}, A_v would eventually become equal to unity, but for typical allowable source-bulk voltages, η remains greater than roughly 0.2.

An important result of (3.80) is that even if $R_S = \infty$, the voltage gain of a source follower is not equal to one. We return to this point later. Note that M_1 in Fig. 3.27 does not enter the triode region if V_{in} remains below V_{DD}.

In the source follower of Fig. 3.27, the drain current of M_1 heavily depends on the input dc level. For example, if V_{in} changes from 1.5 V to 2 V, I_D may increase by a factor of 2 and hence $V_{GS} - V_{TH}$ by $\sqrt{2}$, thereby introducing substantial nonlinearity in the input-output characteristic. To alleviate this issue, the resistor can be replaced by a current source as shown in Fig. 3.30(a). The current source itself is implemented as an NMOS transistor operating in the saturation region [Fig. 3.30(b)].



(a) (b)

Figure 3.30 Source follower using an NMOS transistor as current source.

Example 3.7

Suppose in the source follower of Fig. 3.30(a), $(W/L)_1 = 20/0.5$, $I_1 = 200\ \mu A$, $V_{TH0} = 0.6$ V, $2\Phi_F = 0.7$ V, $\mu_n C_{ox} = 50\ \mu A/V^2$, and $\gamma = 0.4\ V^2$.

(a) Calculate V_{out} for $V_{in} = 1.2$ V.

(b) If I_1 is implemented as M_2 in Fig. 3.30(b), find the minimum value of $(W/L)_2$ for which M_2 remains saturated.

Solution

(a) Since the threshold voltage of M_1 depends on V_{out}, we perform a simple iteration. Noting that

$$(V_{in} - V_{TH} - V_{out})^2 = \frac{2I_D}{\mu_n C_{ox}\left(\frac{W}{L}\right)_1}, \tag{3.81}$$

we first assume $V_{TH} \approx 0.6$ V, obtaining $V_{out} = 0.153$ V. Now we calculate a new V_{TH} as

$$V_{TH} = V_{TH0} + \gamma(\sqrt{2\Phi_F + V_{SB}} - \sqrt{2\Phi_F}) \tag{3.82}$$

$$= 0.635\ \text{V}. \tag{3.83}$$

This indicates that V_{out} is approximately 35 mV less than that calculated above, i.e., $V_{out} \approx 0.119$ V.

(b) Since the drain-source voltage of M_2 is equal to 0.119 V, the device is saturated only if $(V_{GS} - V_{TH})_2 \leq 0.119$ V. With $I_D = 200\ \mu A$, this gives $(W/L)_2 \geq 283/0.5$. Note the substantial drain junction and overlap capacitance contributed by M_2 to the output node.

To gain a better understanding of source followers, let us calculate the small-signal output resistance of the circuit in Fig. 3.31(a). Using the equivalent circuit of Fig. 3.31(b) and noting that $V_1 = -V_X$, we write

$$I_X - g_m V_X - g_{mb} V_X = 0. \tag{3.84}$$

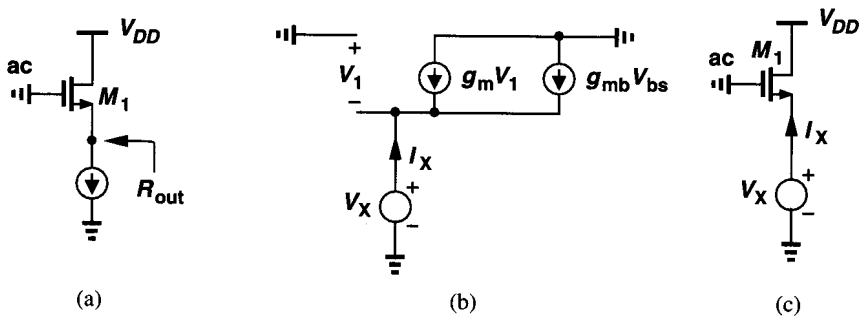


Figure 3.31 Calculation of the output impedance of a source follower.

It follows that

$$R_{out} = \frac{1}{g_m + g_{mb}}. \tag{3.85}$$

Interestingly, body effect decreases the output resistance of source followers. To understand why, suppose in Fig. 3.31(c), V_X decreases by ΔV so that the drain current increases. With no body effect, only the gate-source voltage of M_1 would increase by ΔV. With body effect, on the other hand, the threshold voltage of the device decreases as well. Thus, in $(V_{GS} - V_{TH})^2$ the first term increases and the second decreases, resulting in a greater change in the drain current and hence a lower output impedance.

The above phenomenon can also be studied with the aid of the small-signal model shown in Fig. 3.32(a). It is important to note that the magnitude of the current source $g_{mb}V_{bs}$ is

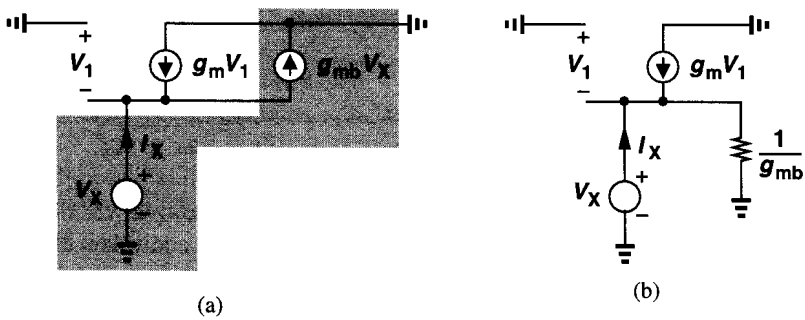


Figure 3.32 Source follower including body effect.

linearly proportional to the voltage across it. Such behavior is that of a simple resistor equal to $1/g_{mb}$, yielding the small-signal model shown in Fig. 3.32(b). The equivalent resistor simply appears in parallel with the output, thereby lowering the overall output resistance. The reader can show that, without $1/g_{mb}$, the output resistance equals $1/g_m$, concluding that

$$R_{out} = \frac{1}{g_m} \| \frac{1}{g_{mb}} \tag{3.86}$$

$$= \frac{1}{g_m + g_{mb}}. \tag{3.87}$$

Modeling the effect of g_{mb} by a resistor—which is only valid for source followers—also helps explain the less-than-unity voltage gain implied by (3.80) for $R_S = \infty$. As shown in

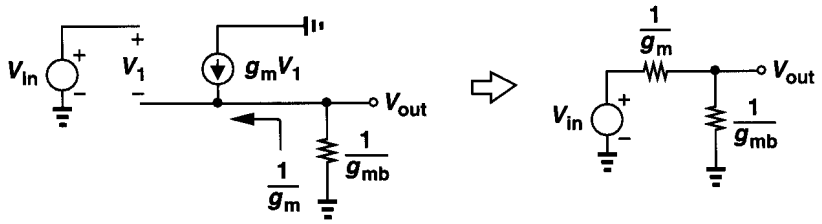


Figure 3.33 Representation of intrinsic source follower by a Thevenin equivalent.

the Thevenin equivalent of Fig. 3.33,

$$A_v = \frac{\dfrac{1}{g_{mb}}}{\dfrac{1}{g_m} + \dfrac{1}{g_{mb}}} \tag{3.88}$$

$$= \frac{g_m}{g_m + g_{mb}}. \tag{3.89}$$

For completeness, we also study the source follower of Fig. 3.34(a) with finite channel-length modulation in M_1 and M_2. From the equivalent circuit in Fig. 3.34(b), we have

$$A_v = \frac{\dfrac{1}{g_{mb}} \| r_{O1} \| r_{o2} \| R_L}{\dfrac{1}{g_{mb}} \| r_{O1} \| r_{o2} \| R_L + \dfrac{1}{g_m}}. \tag{3.90}$$

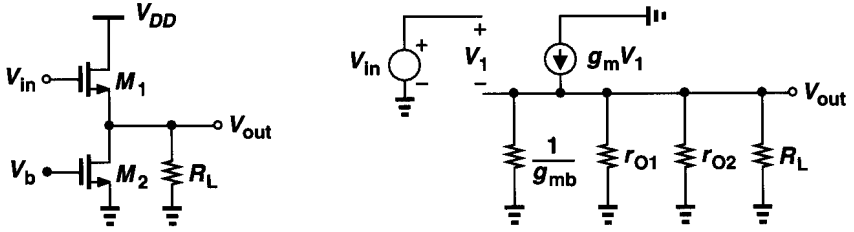


Figure 3.34 (a) Source follower driving load resistance, (b) small-signal equivalent circuit.

Example 3.8

Calculate the voltage gain of the circuit shown in Fig. 3.35.

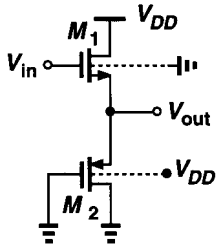


Figure 3.35

Solution

The impedance seen looking into the source of M_2 is equal to $[1/(g_{m2} + g_{mb2})] \| r_{O2}$. Thus,

$$A_v = \frac{\dfrac{1}{g_{m2}+g_{mb2}} \| r_{O2} \| r_{O1} \| \dfrac{1}{g_{mb1}}}{\dfrac{1}{g_{m2}+g_{mb2}} \| r_{O2} \| r_{O1} \| \dfrac{1}{g_{mb1}} + \dfrac{1}{g_{m1}}}. \tag{3.91}$$

Source followers exhibit a high input impedance and a moderate output impedance, but at the cost of two drawbacks: nonlinearity and voltage headroom limitation. We consider these issues in detail.

As mentioned in relation to Fig. 3.27(a), even if a source follower is biased by an ideal current source, its input-output characteristic displays some nonlinearity due to the nonlinear dependence of V_{TH} upon the source potential. In submicron technologies, r_O of the transistor also changes substantially with V_{DS}, thus introducing additional variation in the small-signal gain of the circuit (Chapter 16). For this reason, typical source followers suffer from several percent of nonlinearity.

The nonlinearity due to body effect can be eliminated if the bulk is tied to the source. This is usually possible only for PFETs because all NFETs share the same substrate. Fig. 3.36 shows a PMOS source follower employing two separate n-wells so as to eliminate the body effect of M_1. The lower mobility of PFETs, however, yields a higher output impedance in this case than that available in an NMOS counterpart.

Source followers also shift the dc level of the signal by V_{GS}, thereby consuming voltage headroom and limiting the voltage swings. To understand this point, consider the example illustrated in Fig. 3.37, a cascade of a common-source stage and a source follower. Without the source follower, the minimum allowable value of V_X would be equal to $V_{GS1} - V_{TH1}$ (for M_1 to remain in saturation). With the source follower, on the other hand, V_X must be greater than $V_{GS2} + (V_{GS3} - V_{TH3})$ so that M_3 is saturated. For comparable overdrive voltages in M_1 and M_3, this means the allowable swing at X is reduced by V_{GS2}, a substantial amount.

It is also instructive to compare the gain of source followers and common-source stages when the load impedance is relatively low. A practical example is the need to drive an external 50-Ω termination in a high-frequency setup. As shown in Fig. 3.38(a), the load can

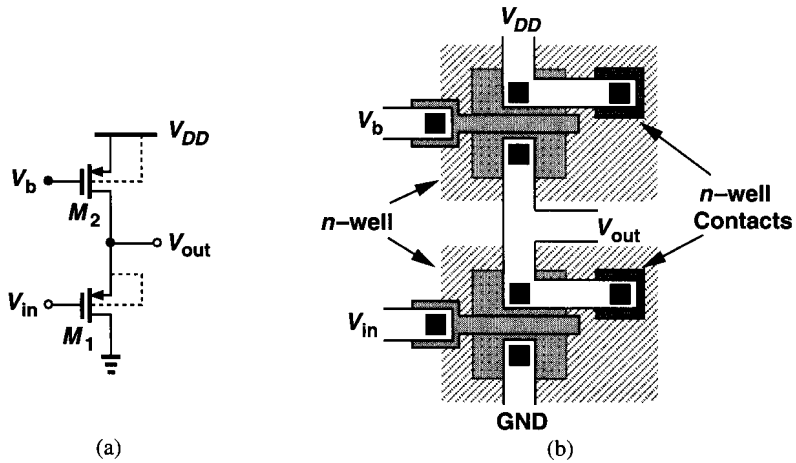


Figure 3.36 PMOS source follower with no body effect.

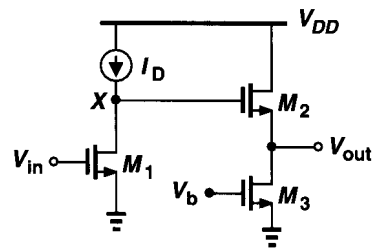


Figure 3.37 Cascade of source follower and CS stage.

be driven by a source follower with an overall voltage gain of

$$\frac{V_{out}}{V_{in}}|_{SF} \approx \frac{R_L}{R_L + 1/g_{m1}}. \tag{3.92}$$

On the other hand, as depicted in Fig. 3.38(b), the load can be included as part of a common-

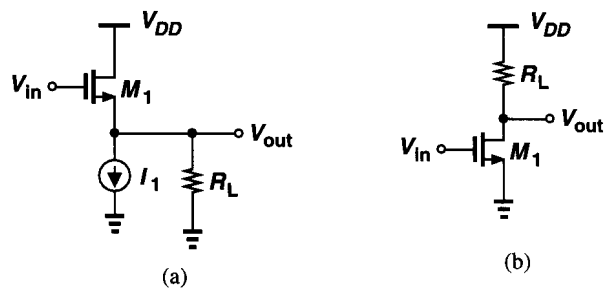


Figure 3.38 (a) Source follower and (b) CS stage driving a load resistance.

source stage, providing a gain of

$$\frac{V_{out}}{V_{in}}|_{CS} \approx -g_{m1}R_L. \tag{3.93}$$

The key difference between these two topologies is the achievable voltage gain for a given bias current. For example, if $1/g_{m1} \approx R_L$, then the source follower exhibits a gain of at most 0.5 whereas the common-source stage provides a gain close to unity. Thus, source followers are not necessarily efficient drivers.

The drawbacks of source followers, namely, nonlinearity due to body effect, voltage headroom consumption due to level shift, and poor driving capability, limit the use of this topology. Perhaps the most common application of source followers is in performing voltage level shift.

Example 3.9

(a) In the circuit of Fig. 3.39(a), calculate the voltage gain if C_1 acts as an ac short at the frequency of interest. What is the maximum dc level of the input signal for which M_1 remains saturated?

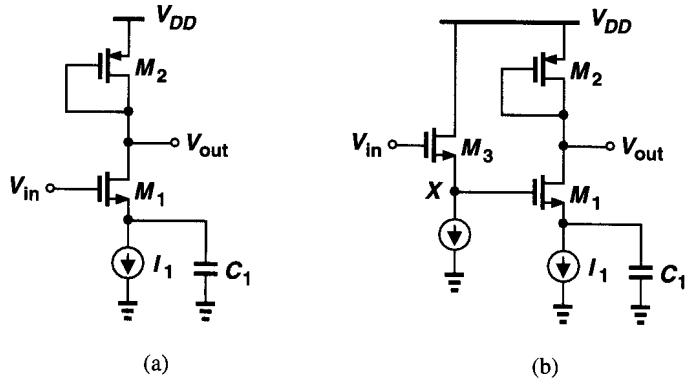


Figure 3.39

(b) To accommodate an input dc level close to V_{DD}, the circuit is modified as shown in Fig. 3.39(b). What relationship among the gate-source voltages of M_1-M_3 guarantees that M_1 is saturated?

Solution

(a) The gain is given by

$$A_v = -g_{m1}[r_{O1} \| r_{O2} \| (1/g_{m2})]. \tag{3.94}$$

Since $V_{out} = V_{DD} - |V_{GS2}|$, the maximum allowable dc level of V_{in} is equal to $V_{DD} - |V_{GS2}| + V_{TH1}$.

(b) If $V_{in} = V_{DD}$, then $V_X = V_{DD} - V_{GS3}$. For M_1 to be saturated, $V_{DD} - V_{GS3} - V_{TH1} \leq V_{DD} - |V_{GS2}|$ and hence $V_{GS3} + V_{TH1} \geq |V_{GS2}|$.

As explained in Chapter 7, source followers also introduce substantial noise. For this reason, the circuit of Fig. 3.39(b) is ill-suited to low-noise applications.

3.4 Common-Gate Stage

In common-source amplifiers and source followers, the input signal is applied to the gate of a MOSFET. It is also possible to apply the signal to the source terminal. Shown in Fig. 3.40(a), a common-gate (CG) stage senses the input at the source and produces the output at the drain. The gate is connected to a dc voltage to establish proper operating conditions. Note that the bias current of M_1 flows through the input signal source. Alternatively, as depicted in Fig. 3.40(b), M_1 can be biased by a constant current source, with the signal capacitively coupled to the circuit.

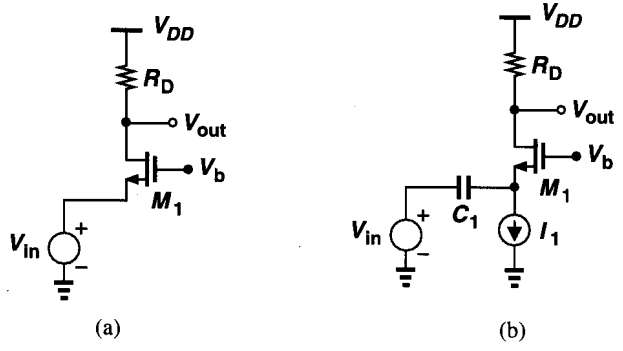


Figure 3.40 (a) Common-gate stage with direct coupling at input, (b) CG stage with capacitive coupling at input.

We first study the large-signal behavior of the circuit in Fig. 3.40(a). For simplicity, let us assume that V_{in} decreases from a large positive value. For $V_{in} \geq V_b - V_{TH}$, M_1 is off and $V_{out} = V_{DD}$. For lower values of V_{in}, we can write

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_b - V_{in} - V_{TH})^2, \tag{3.95}$$

if M_1 is in saturation. As V_{in} decreases, so does V_{out}, eventually driving M_1 into the triode region if

$$V_{DD} - \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_b - V_{in} - V_{TH})^2 R_D = V_b - V_{TH}. \tag{3.96}$$

The input-output characteristic is shown in Fig. 3.41. If M_1 is saturated, we can express the output voltage as

$$V_{out} = V_{DD} - \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_b - V_{in} - V_{TH})^2 R_D, \tag{3.97}$$

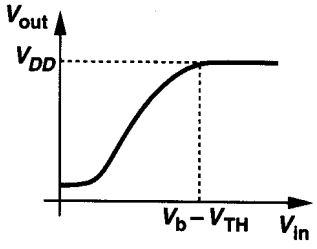


Figure 3.41 Common-gate input-output characteristic.

obtaining a small-signal gain of

$$\frac{\partial V_{out}}{\partial V_{in}} = -\mu_n C_{ox} \frac{W}{L}(V_b - V_{in} - V_{TH}) \left(-1 - \frac{\partial V_{TH}}{\partial V_{in}}\right) R_D. \tag{3.98}$$

Since $\partial V_{TH}/\partial V_{in} = \partial V_{TH}/\partial V_{SB} = \eta$, we have

$$\frac{\partial V_{out}}{\partial V_{in}} = \mu_n C_{ox} \frac{W}{L} R_D (V_b - V_{in} - V_{TH})(1 + \eta) \tag{3.99}$$

$$= g_m(1 + \eta) R_D. \tag{3.100}$$

Note that the gain is positive. Interestingly, body effect increases the equivalent transconductance of the stage.

The input impedance of the circuit is also important. We note that, for $\lambda = 0$, the impedance seen at the source of M_1 in Fig. 3.40(a) is the same as that at the source of M_1 in Fig. 3.31, namely, $1/(g_m + g_{mb}) = 1/[g_m(1 + \eta)]$. Thus, the body effect decreases the input impedance of the common-gate stage. The relatively low input impedance of the common-gate stage proves useful in some applications.

Example 3.10

In Fig. 3.42, transistor M_1 senses ΔV and delivers a proportional current to a 50-Ω transmission line. The other end of the line is terminated by a 50-Ω resistor in Fig. 3.42(a) and a common-gate stage in Fig. 3.42(b). Assume $\lambda = \gamma = 0$.

(a) Calculate V_{out}/V_{in} at low frequencies for both arrangements.

(b) What condition is necessary to minimize wave reflection at node X?

Solution

(a) For small signals applied to the gate of M_1, the drain current experiences a change equal to $g_{m1}\Delta V_X$. This current is drawn from R_D in Fig. 3.42(a) and M_2 in Fig. 3.42(b), producing an output voltage swing equal to $-g_{m1}\Delta V_X R_D$. Thus, $A_v = -g_m R_D$ for both cases.

(b) To minimize reflection at node X, the resistance seen at the source of M_2 must equal 50 Ω and the reactance must be small. Thus, $1/(g_m + g_{mb}) = 50\ \Omega$, which can be ensured by proper sizing and biasing of M_2. To minimize the capacitances of the transistor, it is desirable to use a small device biased at a large current. (Recall that $g_m = \sqrt{2\mu_n C_{ox}(W/L)I_D}$.) In addition to higher power dissipation, this remedy also requires a large V_{GS} for M_2.

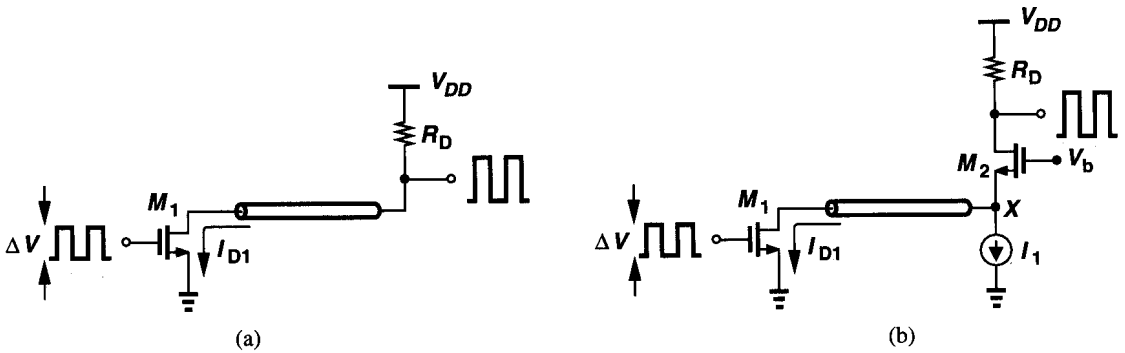


Figure 3.42

The key point in this example is that, while the overall voltage gain in both arrangements equals $-g_{m1}R_D$, the value of R_D in Fig. 3.42(b) can be much greater than 50 Ω without introducing reflections at point X. Thus, the common-gate circuit can provide a much higher voltage gain than that in Fig. 3.42(a).

Now let us study the common-gate topology in a more general case, taking into account both the output impedance of the transistor and the impedance of the signal source. Depicted in Fig. 3.43(a), the circuit can be analyzed with the aid of its equivalent shown

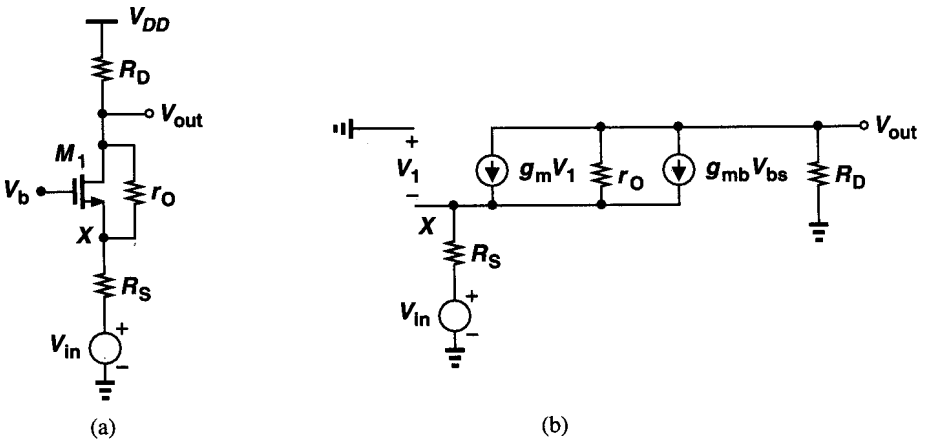


Figure 3.43 (a) CG stage with finite output resistance, (b) small-signal equivalent circuit.

in Fig. 3.43(b). Noting that the current flowing through R_S is equal to $-V_{out}/R_D$, we have:

$$V_1 - \frac{V_{out}}{R_D}R_S + V_{in} = 0. \tag{3.101}$$

Moreover, since the current through r_O is equal to $-V_{out}/R_D - g_m V_1 - g_{mb} V_1$, we can write

$$r_O\left(\frac{-V_{out}}{R_D} - g_m V_1 - g_{mb} V_1\right) - \frac{V_{out}}{R_D} R_S + V_{in} = V_{out}. \tag{3.102}$$

Upon substitution for V_1 from (3.102), (3.101) reduces to

$$r_O\left[\frac{-V_{out}}{R_D} - (g_m + g_{mb})\left(V_{out}\frac{R_S}{R_D} - V_{in}\right)\right] - \frac{V_{out} R_S}{R_D} + V_{in} = V_{out}. \tag{3.103}$$

It follows that

$$\frac{V_{out}}{V_{in}} = \frac{(g_m + g_{mb})r_O + 1}{r_O + (g_m + g_{mb})r_O R_S + R_S + R_D} R_D. \tag{3.104}$$

Note the similarity between (3.104) and (3.71). The gain of the common-gate stage is slightly higher due to body effect.

Example 3.11

Calculate the voltage gain of the circuit shown in Fig. 3.44(a) if $\lambda \neq 0$ and $\gamma \neq 0$.

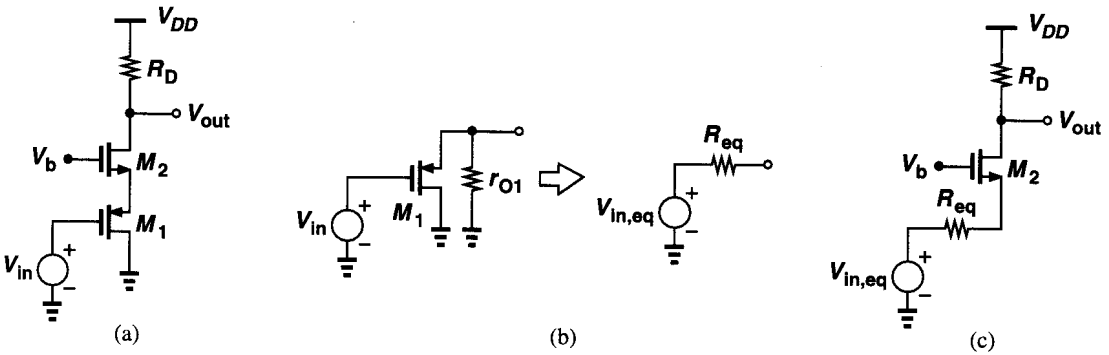


Figure 3.44

Solution

We first find the Thevenin equivalent of M_1. As shown in Fig. 3.44(b), M_1 operates as a source follower and the equivalent Thevenin voltage is given by

$$V_{in,eq} = \frac{r_{O1} \left\| \frac{1}{g_{mb1}} \right.}{r_{O1} \left\| \frac{1}{g_{mb1}} + \frac{1}{g_{m1}} \right.} V_{in}, \tag{3.105}$$

and the equivalent Thevenin resistance is

$$R_{eq} = r_{O1} \left\| \frac{1}{g_{mb1}} \right\| \frac{1}{g_{m1}}. \tag{3.106}$$

Redrawing the circuit as in Fig. 3.44(c), we use (3.104) to write

$$\frac{V_{out}}{V_{in}} = \frac{(g_{m2} + g_{mb2})r_{O2} + 1}{r_{O2} + [1 + (g_{m2} + g_{mb2})r_{O2}]\left(r_{O1} \left\| \frac{1}{g_{mb1}} \right\| \frac{1}{g_{m1}}\right) + R_D} R_D \frac{r_{O1} \left\| \frac{1}{g_{mb1}} \right.}{r_{O1} \left\| \frac{1}{g_{mb1}} + \frac{1}{g_{m2}} \right.}. \tag{3.107}$$

The input and output impedances of the common-gate topology are also of interest. To obtain the impedance seen at the source [Fig. 3.45(a)], we use the equivalent circuit in

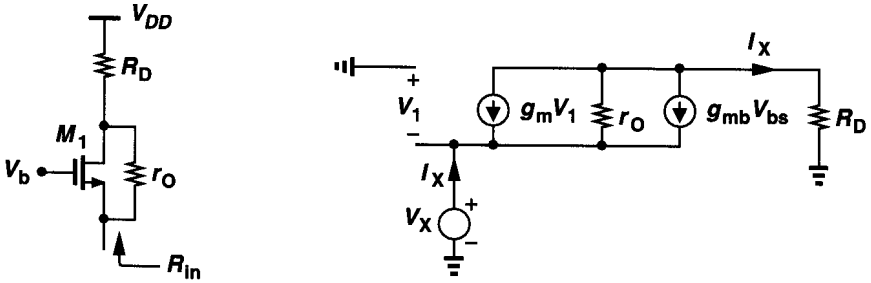


Figure 3.45 (a) Input resistance of a CG stage, (b) small-signal equivalent circuit.

Fig. 3.45(b). Since $V_1 = -V_X$ and the current through r_O is equal to $I_X + g_m V_1 + g_{mb} V_1 = I_X - (g_m + g_{mb})V_X$, we can add up the voltages across r_O and R_D as

$$R_D I_X + r_O[I_X - (g_m + g_{mb})V_X] = V_X. \tag{3.108}$$

Thus,

$$\frac{V_X}{I_X} = \frac{R_D + r_O}{1 + (g_m + g_{mb})r_O} \tag{3.109}$$

$$\approx \frac{R_D}{(g_m + g_{mb})r_O} + \frac{1}{g_m + g_{mb}}, \tag{3.110}$$

if $(g_m + g_{mb})r_O \gg 1$. This result reveals that the drain impedance is divided by $(g_m + g_{mb})r_O$ when seen at the source. This is particularly important in short-channel devices because of their low intrinsic gain. Two special cases of (3.109) are worth studying. First, suppose

$R_D = 0$. Then,

$$\frac{V_X}{I_X} = \frac{r_O}{1 + (g_m + g_{mb})r_O} \tag{3.111}$$

$$= \frac{1}{\frac{1}{r_O} + g_m + g_{mb}}, \tag{3.112}$$

which is simply the impedance seen at the source of a source follower, a predictable result because if $R_D = 0$, the circuit configuration is the same as in Fig. 3.31(a).

Second, let us replace R_D with an ideal current source. Equation (3.110) predicts that the input impedance approaches *infinity*. While somewhat surprising, this result can be explained with the aid of Fig. 3.46. Since the total current through the transistor is fixed and equal to I_1, a change in the source potential cannot change the device current, and hence $I_X = 0$. In other words, the input impedance of a common-gate stage is relatively low *only* if the load impedance connected to the drain is small.

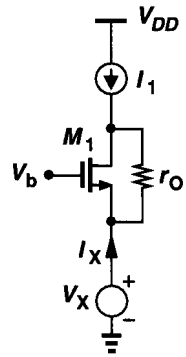


Figure 3.46 Input resistance of a CG stage with ideal current source load.

Example 3.12

Calculate the voltage gain of a common-gate stage with a current-source load [Fig. 3.47(a)].

Solution

Letting R_D approach infinity in (3.104), we have

$$A_v = (g_m + g_{mb})r_O + 1. \tag{3.113}$$

Interestingly, the gain does not depend on R_S. From our foregoing discussion, we recognize that if $R_D \to \infty$, so does the impedance seen at the source of M_1, and the small-signal voltage at node X becomes *equal* to V_{in}. We can therefore simplify the circuit as shown in Fig. 3.47(b), readily arriving at (3.113).

In order to calculate the output impedance of the common-gate stage, we use the circuit

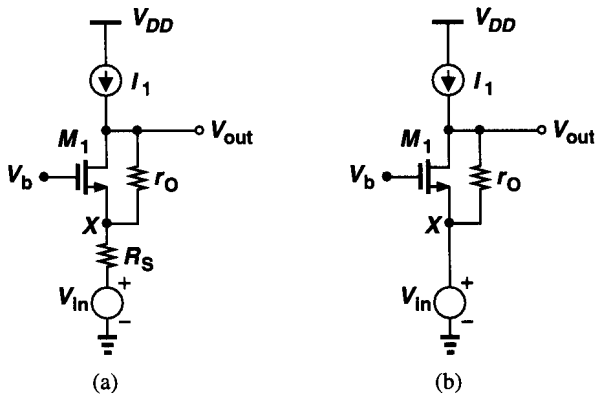


Figure 3.47

in Fig. 3.48. We note that the result is similar to that in Fig. 3.22 and hence

$$R_{out} = \{[1 + (g_m + g_{mb})r_O]R_S + r_O\} \| R_D. \tag{3.114}$$

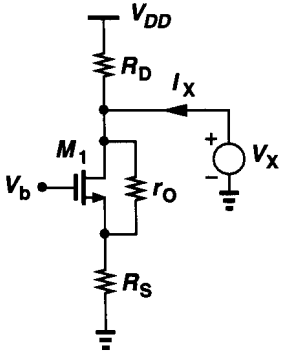


Figure 3.48 Calculation of output resistance of a CG stage.

Example 3.13

As seen in Example 3.10 the input signal of a common-gate stage may be a current rather than a voltage. Shown in Fig. 3.49 is such an arrangement. Calculate V_{out}/I_{in} and the output impedance of the circuit if the input current source exhibits an output impedance equal to R_P.

Solution

To find V_{out}/I_{in}, we replace I_{in} and R_P with a Thevenin equivalent and use (3.104) to write

$$\frac{V_{out}}{I_{in}} = \frac{(g_m + g_{mb})r_O + 1}{r_O + (g_m + g_{mb})r_O R_P + R_P + R_D} R_D R_P. \tag{3.115}$$

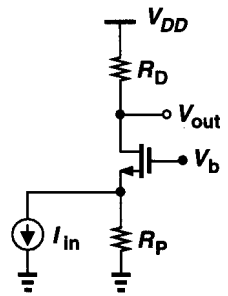


Figure 3.49

The output impedance is simply equal to

$$R_{out} = \{[1 + (g_m + g_{mb})r_O]R_P + r_O\} \| R_D. \tag{3.116}$$

3.5 Cascode Stage

As mentioned in Example 3.10 the input signal of a common-gate stage may be a current. We also know that a transistor in a common-source arrangement converts a voltage signal to a current signal. The cascade of a CS stage and a CG stage is called a "cascode"[1] topology, providing many useful properties. Fig. 3.50 shows the basic configuration: M_1 generates a small-signal drain current proportional to V_{in} and M_2 simply routes the current to R_D.

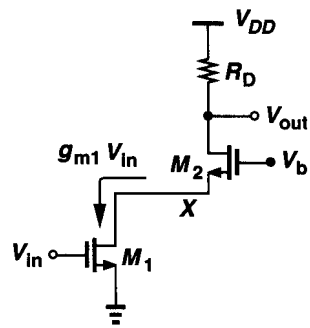


Figure 3.50 Cascode stage.

We call M_1 the input device and M_2 the cascode device. Note that in this example, M_1 and M_2 carry equal currents. As we describe the attributes of the circuit in this section, many advantages of the cascode topology over a simple common-source stage become evident.

First, let us study the bias conditions of the cascode. For M_1 to operate in saturation, $V_X \geq V_{in} - V_{TH1}$. If M_1 and M_2 are both in saturation, then V_X is determined primarily by

[1]The term *cascode* is believed to be the acronym for "cascaded triodes," possibly invented in vacuum tube days.

Chapter 5

Passive and Active Current Mirrors

Our study of single-stage and differential amplifiers in Chapters 3 and 4 points to the wide usage of current sources. In these circuits current sources act as a large resistor without consuming excessive voltage headroom. We also noted that MOS devices operating in saturation can act as a current source.

Current sources find other applications in analog design as well. For example, some digital-to-analog (D/A) converters employ an array of current sources to produce an analog output proportional to the digital input. Also, current sources, in conjunction with "current mirrors," can perform useful functions on analog signals.

This chapter deals with the design of current mirrors as both bias elements and signal processing components. Following a review of basic current mirrors, we study cascode mirror operation. Next, we analyze active current mirrors and describe the properties of differential pairs using such circuits as loads.

5.1 Basic Current Mirrors

Fig. 5.1 illustrates two examples where a current source proves useful. From our study in Chapter 2, recall that the output resistance and capacitance and the voltage headroom of a current source trade with the magnitude of the output current. In addition to these issues, several other aspects of current sources are important: supply, process, and temperature dependence, output noise current, and matching with other current sources. We postpone noise and matching considerations to Chapters 7 and 13, respectively.

How should a MOSFET be biased so as to operate as a stable current source? To gain a better view of the issues, let us consider the simple resistive biasing shown in Fig. 5.2. Assuming M_1 is in saturation, we can write

$$I_{out} \approx \frac{1}{2}\mu_n C_{ox}\frac{W}{L}\left(\frac{R_2}{R_1+R_2}V_{DD} - V_{TH}\right)^2. \tag{5.1}$$

This expression reveals various dependencies of I_{out} upon the supply, process, and temperature. The overdrive voltage is a function of V_{DD} and V_{TH}; the threshold voltage may

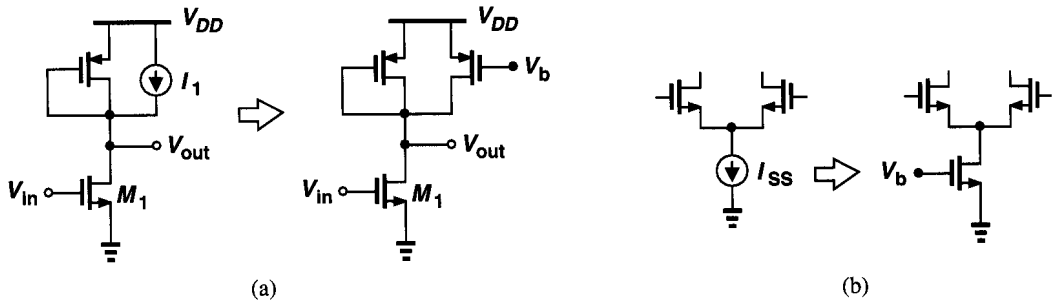


Figure 5.1 Applications of current sources.

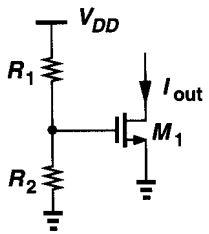


Figure 5.2 Definition of current by resistive divider.

vary by 100 mV from wafer to wafer. Furthermore, both μ_n and V_{TH} exhibit temperature dependence. Thus, I_{out} is poorly defined. The issue becomes more severe as the device is biased with a smaller overdrive voltage, e.g., to consume less headroom. With a nominal overdrive of, say, 200 mV, a 50-mV error in V_{TH} results in a 44% error in the output current.

It is important to note that the above process and temperature dependencies exist even if the gate voltage is not a function of the supply voltage. In other words, if the gate-source *voltage* of a MOSFET is precisely defined, then its drain *current* is not! For this reason, we must seek other methods of biasing MOS current sources.

The design of current sources in analog circuits is based on "copying" currents from a reference, with the assumption that *one* precisely-defined current source is already available. While this method may appear to entail an endless cycle, it is carried out as illustrated in Fig. 5.3. A relatively complex circuit—sometimes requiring external adjustments—is used

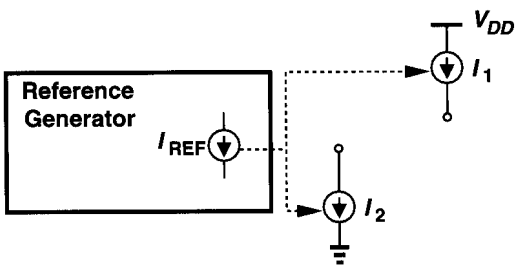


Figure 5.3 Use of a reference to generate various currents.

to generate a stable reference current, I_{REF}, which is then copied to many current sources in the system. We study the copying operation here and the reference generator circuit in Chapter 11.

How do we generate copies of a reference current? For example, in Fig. 5.4, how do we guarantee $I_{out} = I_{REF}$? For a MOSFET, if $I_D = f(V_{GS})$, where $f(\cdot)$ denotes the

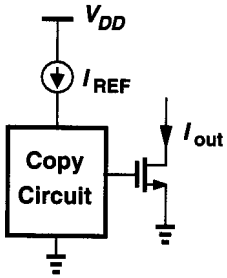


Figure 5.4 Conceptual means of copying currents.

functionality of I_D versus V_{GS}, then $V_{GS} = f^{-1}(I_D)$. That is, if a transistor is biased at I_{REF}, then it produces $V_{GS} = f^{-1}(I_{REF})$ [Fig. 5.5(a)]. Thus, if this voltage is applied to the gate and source terminals of a second MOSFET, the resulting current is $I_{out} = ff^{-1}(I_{REF}) = I_{REF}$ [Fig. 5.5(b)]. From another point of view, two identical MOS devices that have equal gate-source voltages and operate in saturation carry equal currents (if $\lambda = 0$).

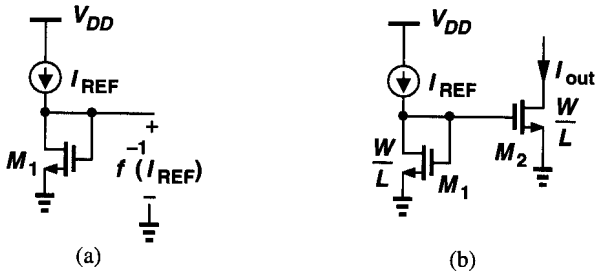


Figure 5.5 (a) Diode-connected device providing inverse function, (b) basic current mirror.

The structure consisting of M_1 and M_2 in Fig. 5.5(b) is called a "current mirror." In the general case, the devices need not be identical. Neglecting channel-length modulation, we can write

$$I_{REF} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_1 (V_{GS} - V_{TH})^2 \tag{5.2}$$

$$I_{out} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_2 (V_{GS} - V_{TH})^2, \tag{5.3}$$

obtaining

$$I_{out} = \frac{(W/L)_2}{(W/L)_1} I_{REF}. \tag{5.4}$$

The key property of this topology is that it allows precise copying of the current with no dependence on process and temperature. The ratio of I_{out} and I_{REF} is given by the *ratio* of device dimensions, a quantity that can be controlled with reasonable accuracy.

Example 5.1

In Fig. 5.6, find the drain current of M_4 if all of the transistors are in saturation.

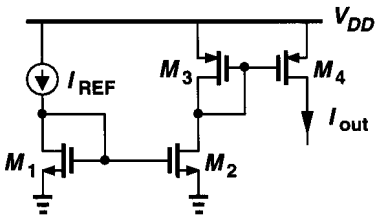


Figure 5.6

Solution

We have $I_{D2} = I_{REF}[(W/L)_2/(W/L)_1]$. Also, $|I_{D3}| = |I_{D2}|$ and $I_{D4} = I_{D3}[(W/L)_4/(W/L)_3]$. Thus, $|I_{D4}| = \alpha\beta I_{REF}$, where $\alpha = (W/L)_2/(W/L)_1$ and $\beta = (W/L)_4/(W/L)_3$. Proper choice of α and β can establish large or small ratios between I_{D4} and I_{REF}. For example, $\alpha = \beta = 5$ yields a magnification factor of 25. Similarly, $\alpha = \beta = 0.2$ can be utilized to generate a small, well-defined current.

Current mirrors find wide application in analog circuits. Fig. 5.7 illustrates a typical case, where a differential pair is biased by means of an NMOS mirror for the tail current source and a PMOS mirror for the load current sources. The device dimensions shown establish a

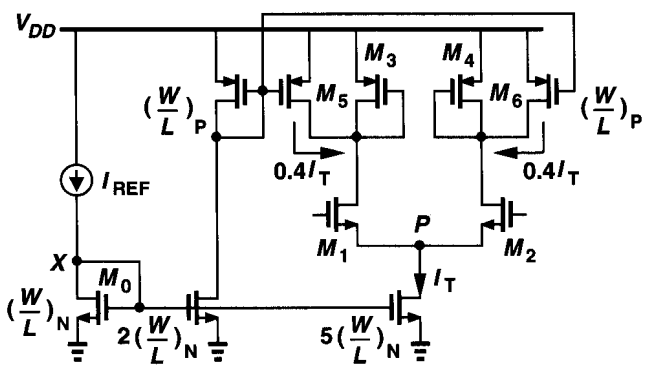


Figure 5.7 Current mirrors used to bias a differential amplifier.

drain current of $0.4I_T$ in M_5 and M_6, reducing the drain current of M_3 and M_4 and hence increasing the gain.

Current mirrors usually employ the same *length* for all of the transistors so as to minimize errors due to the side-diffusion of the source and drain areas (L_D). For example, in Fig. 5.7, the NMOS current sources must have the same channel length as M_0. This is because if, L_{drawn} is, say, doubled, then $L_{eff} = L_{drawn} - 2L_D$ is not. Furthermore, the threshold voltage of short-channel devices exhibits some dependence on the channel length (Chapter 16). Thus, current ratioing is achieved by only scaling the width of transistors.[1]

We should also mention that current mirrors can process *signals* as well. In Fig. 5.5(b), for example, if I_{REF} increases by ΔI, then I_{out} increases by $\Delta I(W/L)_2/(W/L)_1$. That is, the circuit *amplifies* the small-signal current if $(W/L)_2/(W/L)_1 > 1$ (but at the cost of proportional multiplication of the bias current).

Example 5.2

Calculate the small-signal voltage gain of the circuit shown in Fig. 5.8.

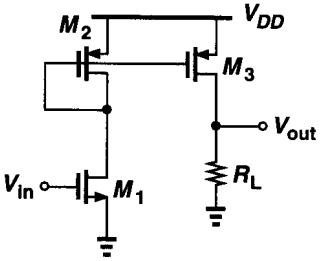


Figure 5.8

Solution

The small-signal drain current of M_1 is equal to $g_{m1}V_{in}$. Since $I_{D2} = I_{D1}$ and $I_{D3} = I_{D2}(W/L)_3/(W/L)_2$, the small-signal drain current of M_3 is equal to $g_{m1}V_{in}(W/L)_3/(W/L)_2$, yielding a voltage gain of $g_{m1}R_L(W/L)_3/(W/L)_2$.

5.2 Cascode Current Mirrors

In our discussion of current mirrors thus far, we have neglected channel length modulation. In practice, this effect results in significant error in copying currents, especially if minimum-length transistors are used so as to minimize the width and hence the output capacitance of the current source. For the simple mirror of Fig. 5.5(b), we can write

$$I_{D1} = \frac{1}{2}\mu_n C_{ox}\left(\frac{W}{L}\right)_1 (V_{GS} - V_{TH})^2(1 + \lambda V_{DS1}) \tag{5.5}$$

[1] As explained in Chapter 18, the widths are actually scaled by placing multiple unit transistors in parallel rather than making a device wider.