5 Field-effect transistors

5.1 Introduction

In this chapter we introduce the second major type of transistor: the *field-effect transistor*. Like the bipolar junction transistors (BJTs) we studied in Chapter 4, field-effect transistors (FETs) allow the user to control a current with another signal. The key difference is that the FET control signal is a voltage while the BJT control signal is a current. Also, the FET control input (called the gate) has a much higher input impedance than the base of a BJT. Indeed, the DC gate impedance for FETs varies from a few megaohms to astounding values in excess of 10^{14} Ω. High input impedance is a highly desirable feature that greatly simplifies circuit analysis.

The BJT has three connections: the collector, base, and emitter. The corresponding connections on an FET are called the *drain*, *gate*, and *source*. Some versions of the FET have a fourth connection called the *bulk* connection. Bipolar transistors come in just two types with opposite polarities: the npn and the pnp. Field-effect transistors have greater variety. In addition to the polarity pairs (termed *n-channel* and *p-channel*), there are differences in gate construction (*junction* and *metal oxide*), and doping (*depletion* and *enhancement*). In terms of analysis, however, they are all very similar, so we will not have to consider each variety separately. Also, as we did with the bipolar transistor, we will focus on one of the polarities (n-channel) since the other polarity simply involves swapping the labels for n and p and changing the sign of the voltages and the direction of the currents.

FETs have both advantages and disadvantages when compared with BJTs. As noted above, FETs have extremely high gate impedance which makes them ideal as a buffer or input stage for a complex circuit. They are also generally less sensitive to temperature variations and more suitable for the large scale integration of modern micro-circuits. On the other hand, FET amplifiers tend to have lower gain than their BJT counterparts. The metal oxide gate construction of some FETs is highly susceptible to damage from static electricity, which means that you can destroy the transistor simply by touching it. Finally, the manufacturing spread in FET

parameters is larger than for BJTs. This makes it trickier to design circuits for mass production.

5.2 Field-effect transistor fundamentals

5.2.1 Junction field-effect transistors

We first consider the *junction field-effect transistor* or JFET. The n-channel version of the JFET is shown with typical external bias voltages in Fig. 5.1. The gate of the device is a piece of p-type semiconductor placed in a larger piece of n-type semiconductor. The two ends of the n-type semiconductor are called the source and the drain, and the region surrounding the p-n junction is called the *channel*. Current enters the drain, flows through the channel and exits from the source. Since the current is flowing through a single material, the I–V characteristic between the drain current I_d and the drain-source voltage V_{ds} is that of a resistor (i.e., linear). However, this simple behavior is modified when a gate voltage is applied or when V_{ds} gets too large.

As shown in the figure, the polarity of the gate-source voltage V_{gs} is such as to reverse bias the p-n junction of this device. As noted in Chapter 3, a reverse biased p-n junction passes very little current, and this is the reason for the high gate input impedance of this device. Recall also that there is a depletion region in the vicinity of the p-n junction where the density of the charge carriers is markedly reduced, and the size of this depletion region increases with reverse bias. Thus, as V_{gs} becomes more negative the depletion region extends further into the channel, effectively reducing its cross-sectional area. Since the resistance of a material varies inversely

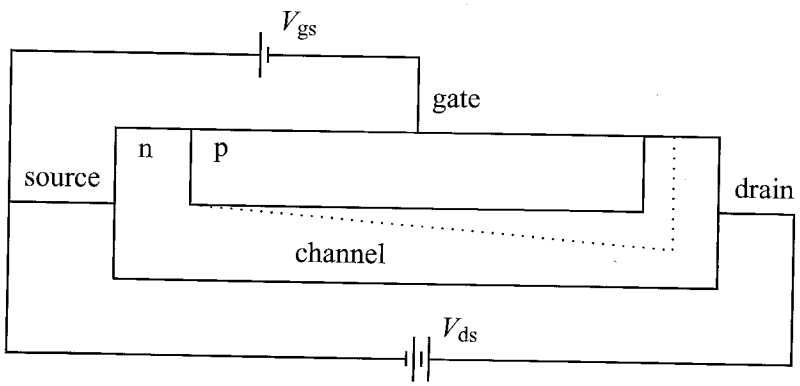


Figure 5.1 An n-channel JFET with typical biasing voltages.

with cross-sectional area (cf. Eq. (1.6) in Chapter 1), this reduces the drain current and the slope of the I–V characteristic curve.

This, however, is not the whole story. The potential of the channel relative to the source varies from zero to V_{ds} as we move thorough the channel from the source to the drain. Thus, the reverse bias of the p-n junction and the extent of the depletion region also vary with position. This is indicated in Fig. 5.1 by the dotted line representing the boundary of the depletion region.

If the reverse bias of the p-n junction at some location is sufficient, the depletion region will extend across the entire channel and the boundary will touch the bottom of the channel. This is called the *pinchoff* or *saturation* point. If V_{ds} is further increased after saturation, the drain current remains essentially constant, reflecting a balance between the increased voltage and the reduced conductivity of the channel.

A set of representative I–V characteristics showing these features is given in Fig. 5.2, which also serves to introduce some additional nomenclature. The different curves represent different values of V_{gs} which are all taken to be above the *threshold value* V_t, the value of V_{gs} that reduces the drain current to zero for all V_{ds}. For small V_{ds} the curves are linear, but eventually curve over and reach their maximum value $I_{d(sat)}$ at $V_{ds(sat)}$. The boundary of this *linear region* (also called the *resistance*, *ohmic*, or *non-saturation* region) is indicated by the dotted line. For values of $V_{ds} > V_{ds(sat)}$, the drain current is essentially constant. This is called the *saturation*

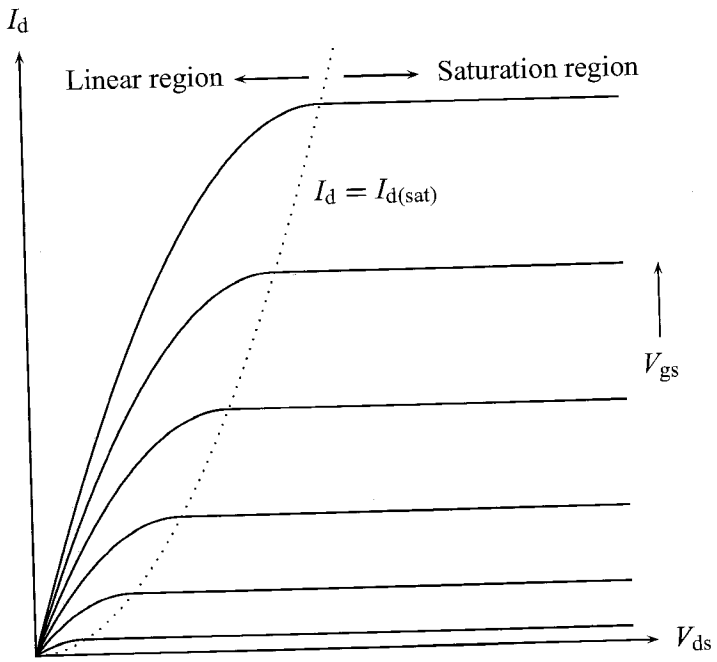


Figure 5.2 I–V characteristics for an FET.

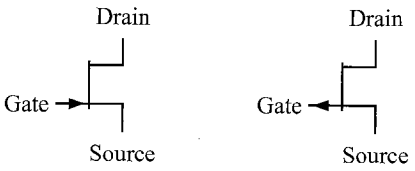


Figure 5.3 Circuit symbols for the n-channel (left) and p-channel (right) JFETs.

region (also known as the *pinchoff* or *active* region). Although not shown in Fig. 5.2, for large enough V_{ds} we enter a *breakdown region* where the drain current rises steeply.

The characteristic curves can also be represented by a set of model equations.[1] The saturation value of the drain-source voltage for a given gate-source voltage is given by

$$V_{ds(sat)} = V_{gs} - V_t. \tag{5.1}$$

Here V_t is the threshold voltage previously defined. In the linear region where $V_{ds} < V_{ds(sat)}$, the drain current is given by

$$I_d = K[2(V_{gs} - V_t) - V_{ds}]V_{ds} = K[2V_{ds(sat)} - V_{ds}]V_{ds} \tag{5.2}$$

where K is a constant. From this we can see that an approximately linear relationship between I_d and V_{ds} requires $V_{ds} \ll 2V_{ds(sat)}$. Finally, in the saturation region where $V_{ds} > V_{ds(sat)}$, we have

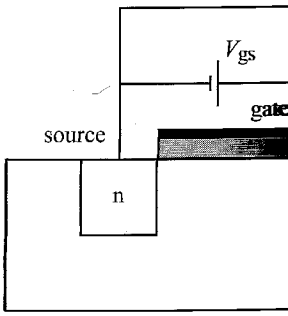
$$I_d = I_{d(sat)} = K(V_{gs} - V_t)^2 = KV_{ds(sat)}^2. \tag{5.3}$$

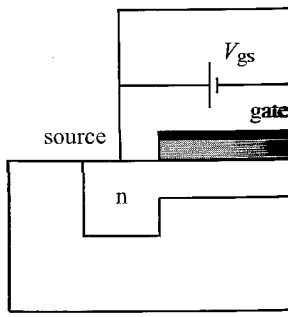
The circuit symbols for both the n-channel and p-channel versions of the JFET are shown in Fig. 5.3. The arrow represents the diode formed by the transistor junction. One must keep in mind, however, that in the JFET this junction is normally reverse biased.

5.2.2 Metal oxide field-effect transistors

The second type of FET we will consider is the *metal oxide semiconductor field-effect transistor* or MOSFET.[2] These transistors are further divided into *enhancement* and *depletion* versions. The n-channel versions of the enhancement and depletion MOSFETs are shown with typical external bias voltages in Figs. 5.4 and 5.5, respectively.[3] As with the JFET, the name of this device refers to the

[1] We refer the reader to the references for further discussion of the physical basis of these equations.
[2] This device is sometimes called an IGFET for *insulated gate field-effect transistor*.
[3] In some advanced applications, the bulk connection would have its own bias, but we will assume it is connected to the source





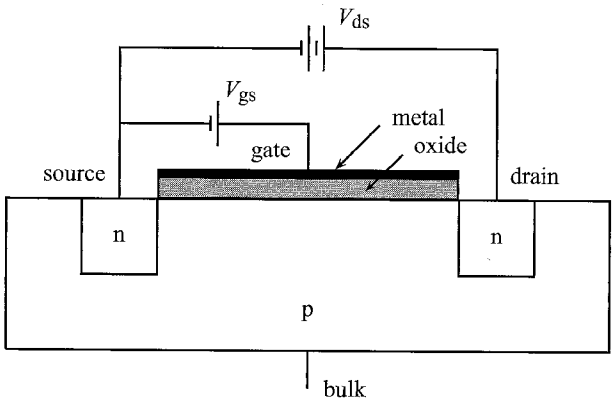


Figure 5.4 An n-channel enhancement-mode MOSFET with typical biasing voltages.

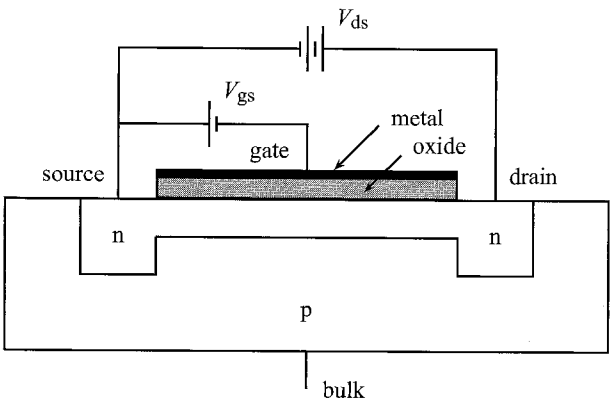


Figure 5.5 An n-channel depletion-mode MOSFET with typical biasing voltages.

structure of the gate. For the MOSFET, the gate pin of the device connects to a layer of metal or other high conductivity material. The metal is deposited on a layer of silicon oxide or other insulating material, which in turn is deposited on the semiconductor. Because of the insulating material, the gate is electrically isolated from the semiconductor and has an extremely high DC input impedance.

To understand the operation of MOSFETs, consider Fig. 5.4. The source and drain are connected to two pieces of n-semiconductor embedded in a p-semiconductor substrate (bulk). Without a gate voltage, the source-bulk and drain-bulk junctions form two back-to-back diodes and no current can flow from drain to source. When a positive voltage is applied to the gate, minority electrons in the p-semiconductor are attracted to the gate region and form an *inversion layer*. The electrons in this inversion layer allow current to flow between the drain and source. The larger the gate voltage, the more electrons in the inversion layer and the larger the drain current. This type of MOSFET, where a conduction channel is created by the action of the gate voltage, is called an *enhancement-mode* MOSFET.

In contrast, the *depletion-mode* MOSFET shown in Fig. 5.5 already has a permanent channel between the source and drain, so drain current will flow even with zero gate voltage. When a negative gate voltage is applied, electrons in this channel are repelled, depleting the channel of charge carriers and reducing the drain current. Adding to the versatility of this device is the fact that a positive gate voltage will attract electrons to the channel and increase the drain current. This versatility comes at the price of somewhat confusing terminology: the depletion-mode MOSFET can be operated in depletion-mode (negative gate voltage) or enhancement-mode (positive gate voltage), while the enhancement-mode MOSFET can only be operated in enhancement-mode.

The circuit symbols for MOSFETs are shown in Figs. 5.6 and 5.7. There is some variety in the way MOSFETs are represented, and in each figure we have shown two versions, one that is more representative (top row) and one that is simpler (bottom row). For the enhancement-mode MOSFET symbols in Fig. 5.6, the top-row versions have a gap between the drain and source that represents the lack of a permanent channel. In line with this scheme, the depletion-mode MOSFET

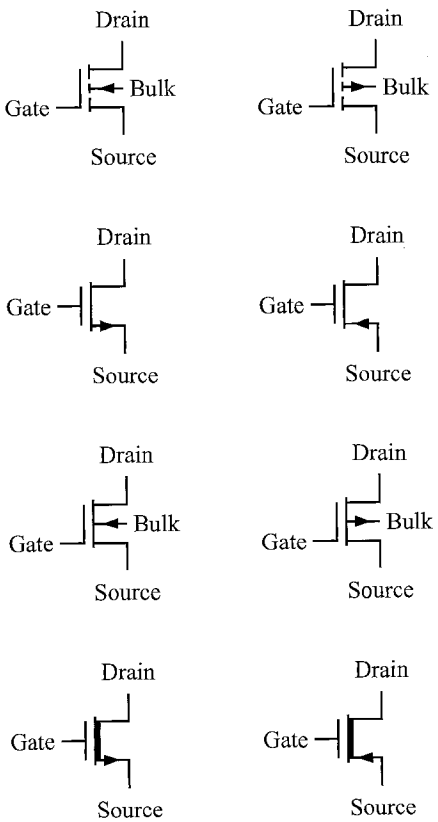


Figure 5.6 Circuit symbols for the n-channel (left) and p-channel (right) enhancement-mode MOSFETs. The second row shows simplified versions of the symbols.

Figure 5.7 Circuit symbols for the n-channel (left) and p-channel (right) depletion-mode MOSFETs. The second row shows simplified versions of the symbols.

symbols in Fig. 5.7 have a solid line between drain and source. Some MOSFETS have a separate pin for the bulk connection (also called the body or substrate connection) and this is shown in the top-row symbols. The bottom-row symbols have the advantage of being similar to those for bipolar transistors and reflect the fact that the bulk connection is often simply tied to the source connection.

The variety of field-effect transistors can be daunting, but the various types are actually quite similar in terms of circuit analysis. The I–V characteristics in Fig. 5.2 could apply to any of the FETs we have discussed, the only difference being the values of V_{gs} assigned to the curves. For the n-channel JFET, the threshold voltage is always negative and there is not much to gain by making the gate voltage greater than zero. For the n-channel enhancement-mode MOSFET, the threshold voltage is zero and the higher curves would correspond to increasingly positive values of V_{gs}. For the n-channel depletion-mode MOSFET, the threshold voltage is negative and a zero value of V_{gs} would correspond to one of the mid-level curves. Higher curves would be positive values of V_{gs}.

A graphical way to see this is to plot the *transfer curve* for the various devices. The transfer curve is a plot of I_d versus V_{gs} for a fixed value of V_{ds} in the saturation region. In Fig. 5.8 we have plotted representative transfer curves for the three types

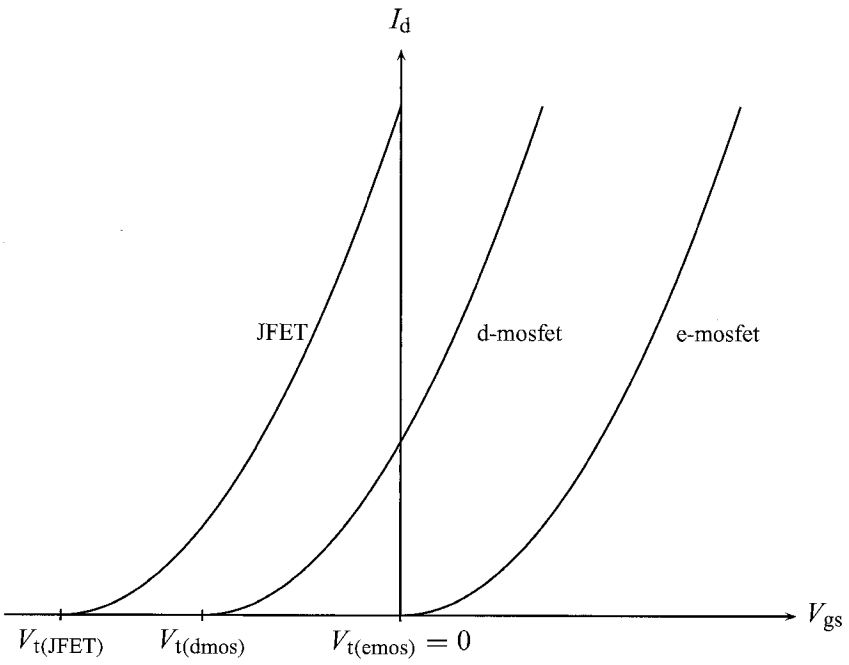


Figure 5.8 Transfer curves for (from left to right) a JFET, a depletion-mode MOSFET, and an enhancement-mode MOSFET (all n-channel).

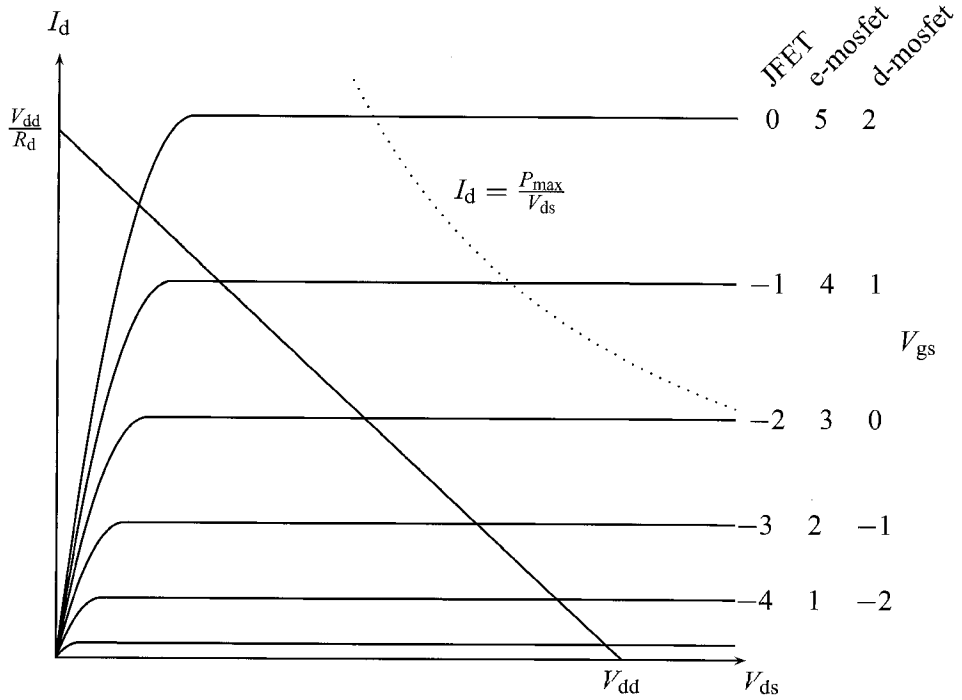


Figure 5.10 Graphical solution for the circuit of Fig. 5.9.

only have positive voltages available for V_g. The circuit can be adapted to this situation by simply using an e-mosfet rather than the JFET. Now +5 V would turn I_d on and 0 V would turn it off. Finally, we note that, as always, the circuit design is constrained by the power rating of the FET used and our operating point must be to the left of the maximum power curve shown by the dotted line.

5.4 Amplifiers

We now turn to the use of FETs in amplifier circuits. Again, the development is very similar to that used for BJT amplifier circuits in Chapter 4. The first task is to set the DC operating point of the transistor in the central part of the saturation region. We will use the universal DC bias circuit to accomplish this. Then we will develop an AC equivalent circuit to describe variations around the operating point. This equivalent circuit will then be used to analyze typical amplifier configurations and obtain the four quantities (a, g, Z_{in}, and Z_{out}) required for the black box model of the amplifiers.

5.4.1 The universal DC bias circuit

Our DC bias circuit is shown in Fig. 5.11. In this example we have used an e-mosfet, but the analysis applies for any FET. Other bias circuits can be obtained from this one by adjusting the resistor values ($R = 0$ for a shorted resistor and $R = \infty$ for a missing resistor).

The analysis of the left part of the circuit is simplified by the very high input resistance of the FET. We can thus ignore the extremely small current into the gate and obtain the gate bias V_g from the voltage divider equation

$$V_g = \left(\frac{R_2}{R_1 + R_2}\right) V_{dd}. \tag{5.5}$$

The transistor is controlled, however, by the gate-source voltage V_{gs}, not by V_g. To obtain this, we start at the gate and complete the voltage loop to ground, giving $V_g - V_{gs} - I_d R_s = 0$. Note here that, because of the lack of any gate current, the current exiting the source is the same as that entering the drain, I_d. Solving for I_d gives

$$I_d = \frac{V_g - V_{gs}}{R_s}. \tag{5.6}$$

As usual, this equation is deceptively simple because I_d is a function of V_{gs} and V_{ds}. Since we intend to set our operating point in the saturation region, the dependence on V_{ds} is negligible and the dependence on V_{gs} is given by either Eq. (5.3) or the transfer curve. Using Eq. (5.3) we obtain

$$K(V_{gs} - V_t)^2 = \frac{V_g - V_{gs}}{R_s} \tag{5.7}$$

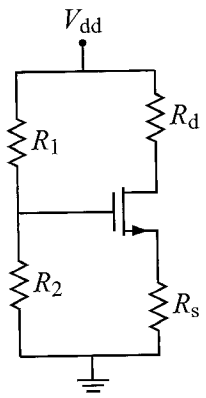


Figure 5.11 DC bias circuit.

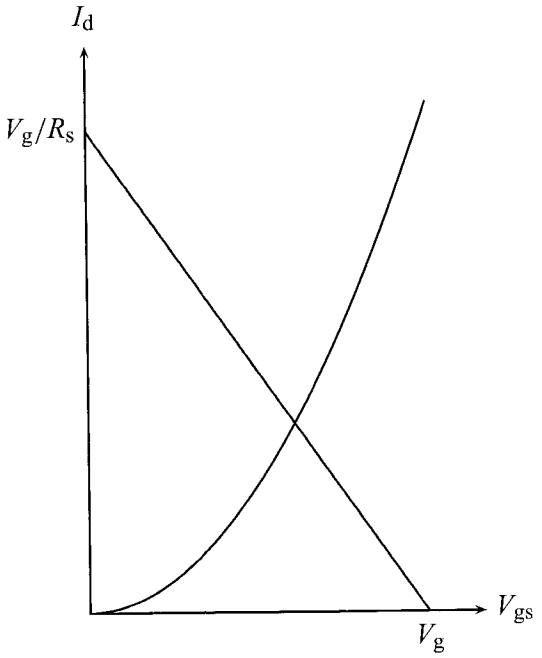


Figure 5.12 Graphical solution to Eq. (5.6).

which can be written

$$V_{gs}^2 + \left(\frac{1}{KR_s} - 2V_t\right) V_{gs} + \left(V_t^2 - \frac{V_g}{KR_s}\right) = 0. \tag{5.8}$$

This quadratic equation can then be solved for V_{gs}. This value is then used in Eq. (5.6) to obtain I_d. Alternatively, a graphical solution can be obtained by plotting the Eq. (5.6) load line and finding its intersection with the transfer curve, as shown in Fig. 5.12.

Turning to the right side of our bias circuit and applying the voltage loop law, we obtain $V_{dd} - I_d R_d - V_{ds} - I_d R_s = 0$ or

$$I_d = \frac{V_{dd} - V_{ds}}{R_s + R_d} \tag{5.9}$$

where V_{ds} is the voltage from the drain to the source of the transistor. Since we know I_d from our solution of Eq. (5.6), we can solve this for V_{ds} and complete our determination of the operating point. Alternatively, we can obtain a graphical solution by the load line method. The resulting plot will be similar to Fig. 5.10 except the y-intercept of the load line will be changed to $V_{dd}/(R_s + R_d)$.