Junction Breakdown. For long channel lengths, the drain-depletion region has little effect on the channel, and the I_D-versus-V_{DS} curves closely follow the ideal curves of Fig. 1.29. For increasing V_{DS}, however, eventually the drain-substrate *pn*-junction breakdown voltage is exceeded, and the drain current increases abruptly by avalanche breakdown as described in Section 1.2.2. This phenomenon is not inherently destructive.

Punchthrough. If the depletion region around the drain in an MOS transistor touches the depletion region around the source before junction breakdown occurs, increasing the drain-source voltage increases the drain current by reducing the barrier to electron flow between the source and drain. This phenomenon is called *punchthrough*. Since it depends on the two depletion regions touching, it also depends on the gate length. Punchthrough is not inherently destructive and causes a more gradual increase in the drain current than is caused by avalanche breakdown. Punchthrough normally occurs below the surface of the silicon and is often prevented by an extra ion implantation below the surface to reduce the size of the depletion regions.

Hot Carriers. With sufficient horizontal or vertical electric fields, electrons or holes may reach sufficient velocities to be injected into the oxide, where most of them increase the gate current and some of them become trapped. Such carriers are called *hot* because the required velocity for injection into the oxide is usually greater than the random thermal velocity. Carriers trapped in the oxide shift the threshold voltage and may cause a transistor to remain on when it should turn off or vice versa. In this sense, injection of hot carriers into the oxide is a destructive process. This process is most likely to be problematic in short-channel technologies, where horizontal electric fields are likely to be high.

Oxide Breakdown. In addition to V_{DS} limitations, MOS devices must also be protected against excessive gate voltages. Typical gate oxides break down with an electric field of about 6×10^6 V/cm to 7×10^6 V/cm,[24,25] which corresponds to 6 to 7 V applied from gate to channel with an oxide thickness of 100 angstroms. Since this process depends on the vertical electrical field, it is independent of channel length. However, this process is destructive to the transistor, resulting in resistive connections between the gate and the channel. Oxide breakdown can be caused by static electricity and can be avoided by using *pn* diodes and resistors to limit the voltage range at sensitive nodes internal to the integrated circuit that connect to bonding pads.

1.6 Small-Signal Models of MOS Transistors

As mentioned in Section 1.5, MOS transistors are often used in analog circuits. To simplify the calculation of circuit gain and terminal impedances, *small-signal* models can be used. As in the case for bipolar transistors, a hierarchy of models with increasing complexity can be derived, and choosing the simplest model required to do a given analysis is important in practice.

Consider the MOS transistor in Fig. 1.33 with bias voltages V_{GS} and V_{DD} applied as shown. These bias voltages produce quiescent drain current I_D. If $V_{GS} > V_t$ and $V_{DD} > (V_{GS} - V_t)$, the device operates in the saturation or active region. A small-signal input voltage v_i is applied in series with V_{GS} and produces a small variation in drain current i_d. The total value of the drain current is $I_d = (I_D + i_d)$.

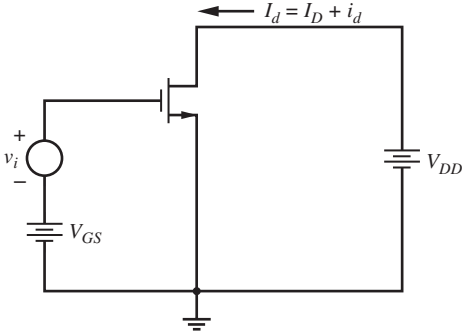

Figure 1.33 Schematic of an MOS transistor with biasing.

1.6.1 Transconductance

Assuming square-law operation, the transconductance from the gate can be determined from (1.165) by differentiating

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = k' \frac{W}{L}(V_{GS} - V_t)(1 + \lambda V_{DS}) \tag{1.179}$$

If $\lambda V_{DS} \ll 1$, (1.179) simplifies to

$$g_m = k' \frac{W}{L}(V_{GS} - V_t) = \sqrt{2k' \frac{W}{L} I_D} \tag{1.180}$$

Unlike the bipolar transistor, the transconductance of the MOS transistor is proportional to the square root of the bias current and depends on device geometry (oxide thickness via k' and W/L). Another key difference between bipolar and MOS transistors can be seen by calculating the ratio of the transconductance to the current. Using (1.157) and (1.180) for MOS transistors shows that

$$\frac{g_m}{I_D} = \frac{2}{V_{GS} - V_t} = \frac{2}{V_{ov}} \tag{1.181}$$

Also, for bipolar transistors, (1.91) shows that

$$\frac{g_m}{I_C} = \frac{q}{kT} = \frac{1}{V_T} \tag{1.182}$$

At room temperature, the thermal voltage V_T is about equal to 26 mV. In contrast, the overdrive V_{ov} for MOS transistors in many applications is chosen to be approximately several hundred mV so that MOS transistors are fast enough for the given application. (Section 1.6.8 shows that the transition frequency f_T of an MOS transistor is proportional to the overdrive.) Under these conditions, the transconductance per given current is much higher for bipolar transistors than for MOS transistors. One of the key challenges in MOS analog circuit design is designing high-quality analog circuits with a low transconductance-to-current ratio.

The transconductance calculated in (1.180) is valid for small-signal analysis. To determine the limitation on the use of small-signal analysis, the change in the drain current resulting from a change in the gate-source voltage will be derived from a large-signal standpoint. The total drain current in Fig. 1.33 can be calculated using (1.157) as

$$I_d = \frac{k'}{2}\frac{W}{L}(V_{GS} + v_i - V_t)^2 = \frac{k'}{2}\frac{W}{L}\left[(V_{GS} - V_t)^2 + 2(V_{GS} - V_t)v_i + v_i^2\right] \tag{1.183}$$

Substituting (1.157) in (1.183) gives

$$I_d = I_D + \frac{k'}{2}\frac{W}{L}\left[2(V_{GS} - V_t)v_i + v_i^2\right] \tag{1.184}$$

Rearranging (1.184) gives

$$i_d = I_d - I_D = k'\frac{W}{L}(V_{GS} - V_t)v_i\left[1 + \frac{v_i}{2(V_{GS} - V_t)}\right] \tag{1.185}$$

If the magnitude of the small-signal input $|v_i|$ is much less than twice the overdrive defined in (1.166), substituting (1.180) into (1.185) gives

$$i_d \simeq g_m v_i \tag{1.186}$$

In particular, if $|v_i| = |\Delta V_{GS}|$ is less than 20 percent of the overdrive, the small-signal analysis is accurate within about 10 percent.

1.6.2 Intrinsic Gate-Source and Gate-Drain Capacitance

If C_{ox} is the oxide capacitance per unit area from gate to channel, then the total capacitance under the gate is $C_{ox}WL$. This capacitance is intrinsic to the device operation and models the gate control of the channel conductance. In the triode region of device operation, the channel exists continuously from source to drain, and the gate-channel capacitance is usually lumped into two equal parts at the drain and source with

$$C_{gs} = C_{gd} = \frac{C_{ox}WL}{2} \tag{1.187}$$

In the saturation or active region, however, the channel pinches off before reaching the drain, and the drain voltage exerts little influence on either the channel or the gate charge. As a consequence, the intrinsic portion of C_{gd} is essentially zero in the saturation region. To calculate the value of the intrinsic part of C_{gs} in the saturation or active region, we must calculate the total charge Q_T stored in the channel. This calculation can be carried out by substituting (1.145) into (1.144) and integrating to obtain

$$Q_T = WC_{ox}\int_0^L [V_{GS} - V(y) - V_t]dy \tag{1.188}$$

Solving (1.150) for dy and substituting into (1.188) gives

$$Q_T = \frac{W^2C_{ox}^2\mu_n}{I_D}\int_0^{V_{GS}-V_t} (V_{GS} - V - V_t)^2 dV \tag{1.189}$$

where the limit $y = L$ corresponds to $V = (V_{GS} - V_t)$ in the saturation or active region. Solution of (1.189) and use of (1.153) and (1.157) gives

$$Q_T = \frac{2}{3}WLC_{ox}(V_{GS} - V_t) \tag{1.190}$$

Therefore, in the saturation or active region,

$$C_{gs} = \frac{\partial Q_T}{\partial V_{GS}} = \frac{2}{3}WLC_{ox} \tag{1.191}$$

and

$$C_{gd} = 0 \tag{1.192}$$

1.6.3 Input Resistance

The gate of an MOS transistor is insulated from the channel by the SiO_2 dielectric. As a result, the low-frequency gate current is essentially zero and the input resistance is essentially infinite. This characteristic is important in some circuits such as sample-and-hold amplifiers, where the gate of an MOS transistor can be connected to a capacitor to sense the voltage on the capacitor without leaking away the charge that causes that voltage. In contrast, bipolar transistors have small but nonzero base current and finite input resistance looking into the base, complicating the design of bipolar sample-and-hold amplifiers.

1.6.4 Output Resistance

In Section 1.5.1, the effect of changes in drain-source voltage on the large-signal characteristics of the MOS transistor was described. Increasing drain-source voltage in an n-channel MOS transistor increases the width of the depletion region around the drain and reduces the effective channel length of the device in the saturation or active region. This effect is called channel-length modulation and causes the drain current to increase when the drain-source voltage is increased. From that treatment, we can calculate the change in the drain current ΔI_D arising from changes in the drain-source voltage ΔV_{DS} as

$$\Delta I_D = \frac{\partial I_D}{\partial V_{DS}} \Delta V_{DS} \tag{1.193}$$

Substitution of (1.161), (1.163), and (1.164) in (1.193) gives

$$\frac{\Delta V_{DS}}{\Delta I_D} = \frac{V_A}{I_D} = \frac{1}{\lambda I_D} = r_o \tag{1.194}$$

where V_A is the Early voltage, λ is the channel-length modulation parameter, I_D is the drain current without channel-length modulation given by (1.157), and r_o is the small-signal output resistance of the transistor.

1.6.5 Basic Small-Signal Model of the MOS Transistor

Combination of the preceding small-signal circuit elements yields the small-signal model of the MOS transistor shown in Fig. 1.34. This model was derived for n-channel transistors in the saturation or active region and is called the *hybrid-π* model. Drain, gate, and source nodes are labeled D, G, and S, respectively. When the gate-source voltage is increased, the model predicts that the incremental current i_d flowing from drain to source increases. Since the dc drain current I_D also flows from drain to source in an n-channel transistor, increasing the gate-source voltage also increases the total drain current I_d. This result is reasonable physically because increasing the gate-source voltage in an n-channel transistor increases the channel conductivity and drain current.

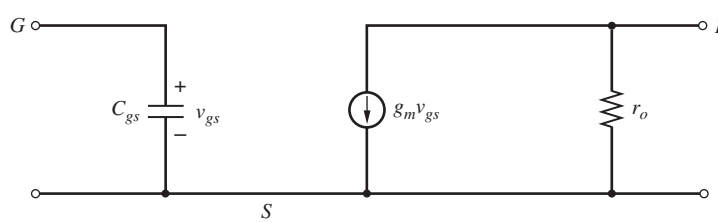

Figure 1.34 Basic small-signal model of an MOS transistor in the saturation or active region.

The model shown in Fig. 1.34 is also valid for p-channel devices. Therefore, the model again shows that increasing the gate-source voltage increases the incremental current i_d flowing from drain to source. Unlike in the n-channel case, however, the dc current I_D in a p-channel transistor flows from source to drain because the source acts as the source of holes. Therefore, the incremental drain current flows in a direction opposite to the dc drain current when the gate-source voltage increases, reducing the total drain current I_d. This result is reasonable physically because increasing the gate-source voltage in a p-channel transistor reduces the channel conductivity and drain current.

1.6.6 Body Transconductance

The drain current is a function of both the gate-source and body-source voltages. On the one hand, the gate-source voltage controls the vertical electric field, which controls the channel conductivity and therefore the drain current. On the other hand, the body-source voltage changes the threshold, which changes the drain current when the gate-source voltage is fixed. This effect stems from the influence of the substrate acting as a second gate and is called the *body effect*. Note that the body of an MOS transistor is usually connected to a constant power-supply voltage, which is a small-signal or ac ground. However, the source connection can have a significant ac voltage impressed on it, which changes the body-source voltage when the body voltage is fixed. Therefore, when the body-source voltage is not constant, two transconductance terms are required to model MOS transistors: one associated with the main gate and the other associated with the body or second gate.

Using (1.165), the transconductance from the body or second gate is

$$g_{mb} = \frac{\partial I_D}{\partial V_{BS}} = -k'\frac{W}{L}(V_{GS} - V_t)(1 + \lambda V_{DS})\frac{\partial V_t}{\partial V_{BS}} \tag{1.195}$$

From (1.140)

$$\frac{\partial V_t}{\partial V_{BS}} = -\frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = -\chi \tag{1.196}$$

This equation defines a factor χ, which is the rate of change of threshold voltage with body bias voltage. Substitution of (1.141) in (1.196) and use of (1.20) gives

$$\chi = \frac{C_{js}}{C_{ox}} \tag{1.197}$$

where C_{js} is the capacitance per unit area of the depletion region under the channel, assuming a one-sided step junction with a built-in potential $\psi_0 = 2\phi_f$. Substitution of (1.196) in (1.195) gives

$$g_{mb} = \frac{\gamma k'(W/L)(V_{GS} - V_t)(1 + \lambda V_{DS})}{2\sqrt{2\phi_f + V_{SB}}} \tag{1.198}$$

If $\lambda V_{DS} \ll 1$, we have

$$g_{mb} = \frac{\gamma k'(W/L)(V_{GS} - V_t)}{2\sqrt{2\phi_f + V_{SB}}} = \gamma\sqrt{\frac{k'(W/L)I_D}{2(2\phi_f + V_{SB})}} \tag{1.199}$$

The ratio g_{mb}/g_m is an important quantity in practice. From (1.179) and (1.198), we find

$$\frac{g_{mb}}{g_m} = \frac{\gamma}{2\sqrt{2\phi_f + V_{SB}}} = \chi \tag{1.200}$$