Chapter

6

MOSFET Operation

In this chapter we discuss MOSFET operation. Figure 6.1 shows how we define the voltages, currents, and terminal designations for a MOSFET. When the substrate is connected to ground and the well is tied to *VDD*, we use the simplified models shown at the bottom of the figure. It is important to keep in mind that the MOSFET is a four-terminal device and that the source and drain of the MOSFET are interchangeable. Note that **all voltages and currents are positive** using the naming convention seen in the figure. If we say "the V_{SG} of the MOSFET is...," we know that we are talking about a PMOS device. Also note, the drain current flows from the top of the symbol to the bottom. For the NMOS, the drain is at the top of the symbol, while for the PMOS the source is at the top of the symbol. The devices are complementary.

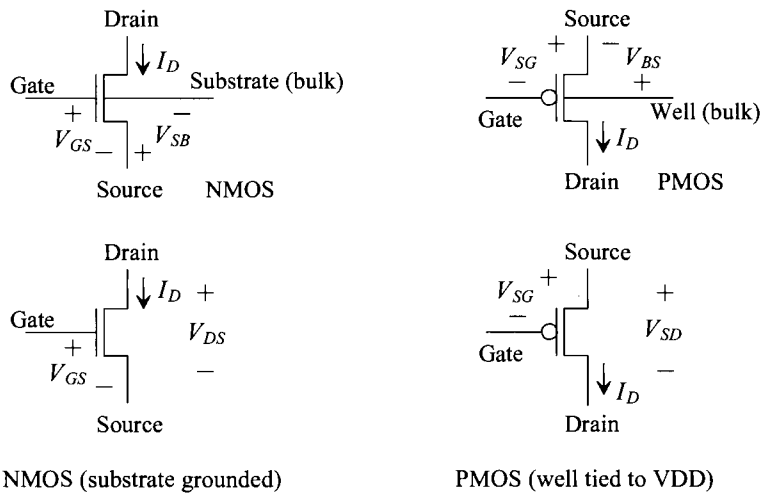

Figure 6.1 Voltage and current designations for MOSFETs in this chapter.

6.1 MOSFET Capacitance Overview/Review

In this section we'll discuss and review the capacitances of a MOSFET operating in the accumulation, depletion (weak inversion), and strong inversion regions.

Case I: Accumulation

Examine the cross-sectional view seen in Fig. 6.2. When $V_{GS} < 0$, mobile holes from the substrate are attracted (or *accumulated*) under the gate oxide. Remembering from Eq. (5.8) that

$$C'_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{6.1}$$

the capacitance between the *gate* electrode and the *substrate* electrode is given by

$$C_{gb} = \frac{\varepsilon_{ox}}{t_{ox}} \cdot (L_{drawn} - 2 \cdot L_{diff}) \cdot W_{drawn} \cdot (scale)^2 = C'_{ox} \cdot L_{eff} \cdot W_{drawn} \cdot (scale)^2 \tag{6.2}$$

where ε_{ox} (= 3.97 · 8.85 aF/μm) is the dielectric constant of the gate oxide, W_{drawn} is the drawn width (neglecting oxide encroachment), and $L_{drawn} - 2 \cdot L_{diff}$ is the effective channel length, L_{eff}. The capacitance between the gate and drain/source (the overlap capacitances, see Eq. [5.23]) is given by

$$C_{gs} = C'_{ox} \cdot L_{diff} \cdot W_{drawn} \cdot (scale)^2 = C_{gd} \tag{6.3}$$

neglecting oxide encroachment on the width of the MOSFET. The gate-drain overlap capacitance is present in a MOSFET regardless of the biasing conditions.

The total capacitance between the gate and ground in the circuit of Fig. 6.2 is the sum of C_{gd}, C_{gs}, and C_{gb} and is given by

$$C_{gs} + C_{gb} + C_{gd} = C_{ox} = C'_{ox} \cdot L_{drawn} \cdot W_{drawn} \cdot (scale)^2 \tag{6.4}$$

There is a significant resistance in series with C_{gb}. The resistance comes from the physical distance between the substrate connection and the area under the gate oxide. The resistivity of the n+ source and drain regions in series with C_{gs} and C_{gd} tends to be small enough to neglect in most circuit design applications.

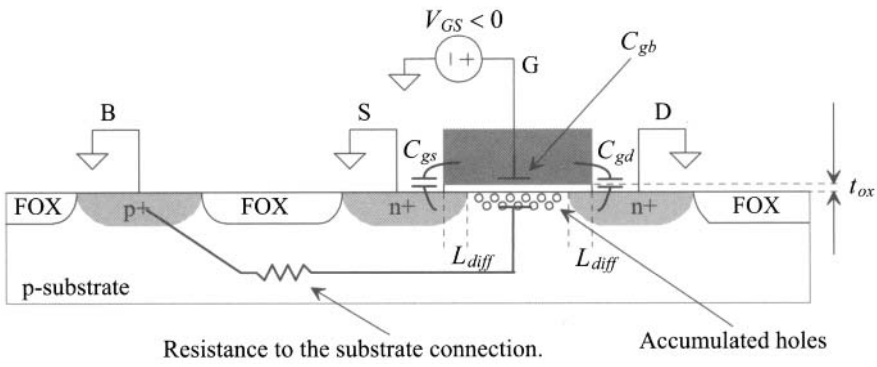

Figure 6.2 Cross-sectional view of a MOSFET operating in accumulation.

Case II: Depletion

Referring again to Fig. 6.2, let's consider the case when V_{GS} is not negative enough to attract a large number of holes under the oxide and not positive enough to attract a large number of electrons. Under these conditions, the surface under the gate is said to be nearly depleted (depeleted of free electrons and holes). Consider the cross-sectional view seen in Fig. 6.3. As V_{GS} is increased from some negative voltage, holes will be displaced under the gate, leaving immobile acceptor ions that contribute a negative charge. We see that as we increase V_{GS} a capacitance between the gate and the induced (n) channel under the oxide exists. Also, a depletion capacitance between the depleted channel and the substrate is formed. The capacitance between the gate and the source/drain is simply the overlap capacitance, while the capacitance between the gate and the substrate is the oxide capacitance *in series* with the depletion capacitance. The depletion layer shown in Fig. 6.3 is formed between the substrate and the induced channel. The MOSFET operated in this region is said to be in *weak inversion* or the *subthreshold region* because the surface under the oxide is not heavily n+.

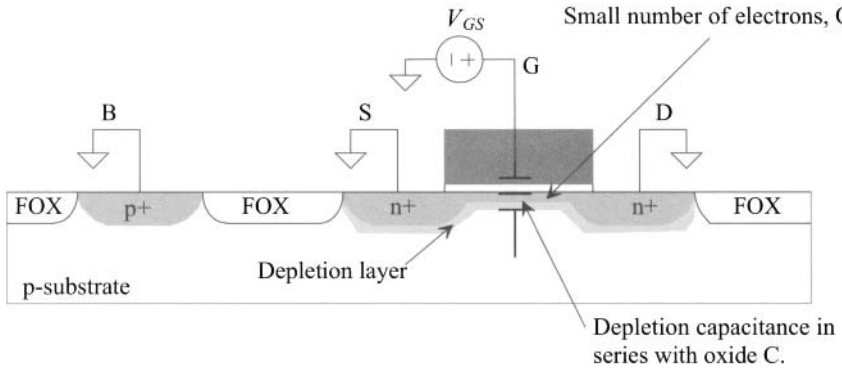

Figure 6.3 Cross-sectional view of a MOSFET operating in depletion.

Case III: Strong Inversion

When V_{GS} is sufficiently large ($>> V_{THN}$, the threshold voltage of the NMOS device) so that a large number of electrons are attracted under the gate, the surface is said to be inverted, that is, no longer p-type. Figure 6.4 shows how the capacitance at the gate changes as V_{GS} is varied, for an NMOS device, when the source, drain, and bulk are grounded. This figure can be misleading. It may appear that we can operate the MOSFET in accumulation if we need a good capacitor. Remembering that when the MOSFET is in the accumulation region the majority of the capacitance to ground, C_{gb}, runs through the large parasitic resistance of the substrate, we see that to operate the MOSFET in this region we need plentiful substrate connections around the gate oxide (to reduce this parasitic substrate resistance). It's preferable to operate the MOSFET in strong inversion when we need a capacitor. The attracted electrons under the gate oxide short the drain and source together forming a low-resistance bottom plate for the capacitor. We will make a capacitor in this fashion many times when designing circuits.

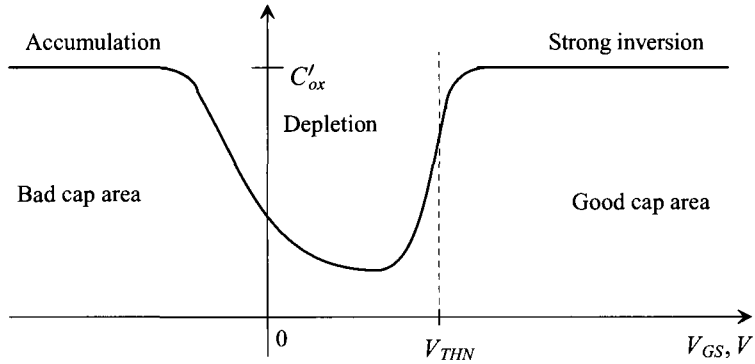

Figure 6.4 The variation of the gate capacitance with DC gate-source voltage.

Example 6.1

Suppose the MOSFET configuration seen in Fig. 6.5 is to be used as a capacitor. If the width and length of the MOSFET are both 100, estimate the capacitance between the gate and the source/drain terminals. Use the long-channel process oxide capacitance listed in Table 5.1. Are there any restrictions on the voltages we can use across the capacitor?

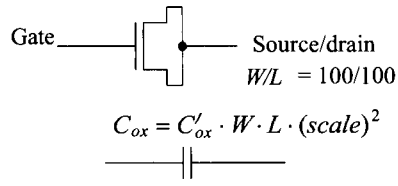

Figure 6.5 Using the MOSFET as a capacitor.

Since the MOSFET is to be used as a capacitor, we require operation in the strong inversion region, that is, $V_{GS} >> V_{THN}$ (the gate potential at least a threshold voltage plus 5% of *VDD* above the source/drain potentials). The capacitance between the gate and the source/drain is then $C_{tot} = C_{ox} = C'_{ox} \cdot W \cdot L$ or from Table 5.1

$$C_{tot} = (1.75\,fF/\mu m^2)(100\ \mu m)(100\ \mu m) = 17.5\,pF$$

Note that we did not concern ourselves with the substrate connection. Since we are assuming strong inversion, the bulk (substrate) connection only affects the capacitances from the drain/source to substrate (those of the source/drain implant regions). We see, however, that the connection of the substrate significantly affects the threshold voltage of the devices and thus the point we label strong inversion. ■

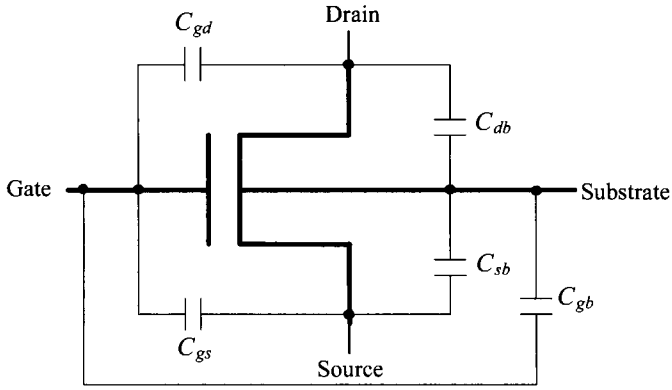

Figure 6.6 MOSFET capacitances.

Summary

Figure 6.6 shows our MOSFET symbol with capacitances. Table 6.1 lists the capacitances based on the region of operation (without a scale factor). The capacitance *CGBO* is the capacitance associated with the gate poly extension over the field region (Fig. 5.13). The gate-drain capacitance, C_{gd}, and the gate-source capacitance, C_{gs}, are determined by the region of operation. For example, as we'll see later, when the MOSFET operates in the triode region, the inverted channel extends between the source and drain implants (the channel resistively connects the source and drain together). The capacitance between the gate and this channel is the oxide capacitance, C_{ox}. We assume that half of this capacitance is between the drain and the other half is between the source.

Table 6.1 MOSFET capacitances.

Name	Off	Triode	Saturation
C_{gd}	$CGDO \cdot W$	$\frac{1}{2} \cdot W \cdot L \cdot C'_{ox}$	$CGDO \cdot W$
C_{db}	C_{jd}	C_{jd}	C_{jd}
C_{gb}	$C'_{ox} WL_{eff} + CGBO \cdot L$	$CGBO \cdot L$	$CGBO \cdot L$
C_{gs}	$CGSO \cdot W$	$\frac{1}{2} \cdot W \cdot L \cdot C'_{ox}$	$\frac{2}{3} \cdot W \cdot L \cdot C'_{ox}$
C_{sb}	C_{js}	C_{js}	C_{js}

6.2 The Threshold Voltage

In the last section we said that the semiconductor/oxide surface is inverted when V_{GS} is greater than the threshold voltage V_{THN}. Under these conditions a channel of electrons is formed under the gate oxide. Below this channel, electrons fill the holes in the substrate

giving rise to a depletion region (depleted of free carriers). The thickness of the depletion region (Fig. 6.7) is given from pn junction theory by

$$X_d = \sqrt{\frac{2\varepsilon_{si}|V_s - V_{fp}|}{qN_A}} \tag{6.5}$$

where N_A is the number of acceptor atoms in the substrate, V_s is the electrostatic potential at the oxide-silicon interface (the channel), and the electrostatic potential of the p-type substrate is given by (see Eq. [2.11])

$$V_{fp} = -\frac{E_i - E_{fp}}{q} = -\frac{kT}{q}\ln\frac{N_A}{n_i} \tag{6.6}$$

noting that this is a negative number. As seen in Fig. 6.7, one edge of the depletion region is the MOSFET's gate oxide, while the other edge is the p-substrate (holes). The positive potential on the gate attracts electrons under the gate oxide. This charge is equal and opposite to the charge in the polysilicon gate material. The charge/unit area is given by

$$Q'_b = qN_A X_d = \sqrt{2\varepsilon_{si} qN_A |V_s - V_{fp}|} \tag{6.7}$$

If the surface electrostatic potential at the oxide interface, V_s, is the same as the bulk electrostatic potential V_{fp} (i.e., $V_s = V_{fp}$ and then $Q'_b = 0$), the MOSFET is operating in the accumulation mode, or the MOSFET is OFF in circuit terms. At this point the number of holes at the oxide-semiconductor surface is N_A, the same concentration as the bulk.

As V_{GS} is increased, the surface potential becomes more positive. When $V_s = 0$, the surface under the oxide has become depleted (the carrier concentration is n_i). When $V_s = -V_{fp}$ (a positive number), the channel is inverted (electrons are pulled under the oxide forming a channel), and the electron concentration at the semiconductor-oxide interface is equal to the substrate doping concentration. The value of V_{GS} when $V_s = -V_{fp}$ is arbitrarily defined as the threshold voltage, V_{THN}. Note that the surface potential changed a total of $2|V_{fp}|$ between the strong inversion and accumulation cases.

For $V_{GS} = V_{THN}$ ($V_s = -V_{fp}$), the negative charge under the gate oxide is given by

$$Q'_{bo} = \sqrt{2qN_A\varepsilon_{si}|-2V_{fp}|} \tag{6.8}$$

with units of Coulombs/m^2. Up to this point we have assumed that the substrate and source were tied together to ground. If the source of the NMOS device is at a higher potential than the substrate, the potential difference is given by V_{SB}; the negative charge under the gate oxide becomes

$$Q'_b = \sqrt{2qN_A\varepsilon_{si}|-2V_{fp}+V_{SB}|} \tag{6.9}$$

Example 6.2

For a substrate doping of 10^{15} atoms/cm^3, $V_{GS} = V_{THN}$ and $V_{SB} = 0$, estimate the electrostatic potential in the substrate region, V_{fp}, and at the oxide-semiconductor interface, V_s, the depletion layer width, X_d, and the charge contained in the depletion region, Q'_b, and thus the inverted region under the gate.

The electrostatic potential of the substrate is

$$V_{fp} = -\frac{kT}{q}\ln\frac{N_A}{n_i} = -26\text{ mV}\cdot\ln\frac{10^{15}}{14.5\times10^{9}} = -290\text{ mV}$$

and therefore the electrostatic potential at the oxide semiconductor interface ($V_{GS} = V_{THN}$), V_s, is 290 mV. The depletion layer thickness is given by

$$X_d = \sqrt{\frac{2\cdot11.7\cdot(8.85\times10^{-18}F/\mu m)(2\cdot0.29V)}{(1.6\times10^{-19}\frac{C}{atom})(10^{15}\frac{atoms}{cm^3})(\frac{cm^3}{10^{12}\mu m^3})}} = 0.866\ \mu m$$

and the charge contained in this region, from Eq. (6.7) or (6.8) with $V_s = -V_{fp}$, by

$$Q'_{bo} = qN_AX_d = \left(1.6\times10^{-19}\frac{C}{atom}\right)\left(10^{15}\frac{atoms}{cm^3}\right)\left(\frac{cm^3}{10^{12}\mu m^3}\right)(0.866\,\mu m)$$

$$= 139\ \frac{aC}{\mu m^2}$$

Note that this is true only when $V_{GS} = V_{THN}$. ■

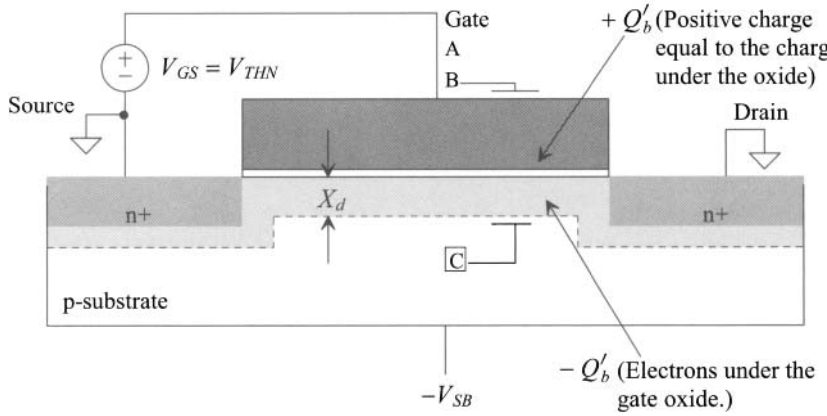

Figure 6.7 Calculation of the threshold voltage.

Contact Potentials

Again, consider the MOSFET shown in Fig. 6.7. We assume that the applied $V_{GS} = V_{THN}$ so that the preceding discussions and assumptions hold. The potential across the gate-oxide capacitance, C'_{ox}, is simply

$$V_{BC} = \frac{Q'_b}{C'_{ox}} \tag{6.10}$$

The surface potential *change*, V_C (= ΔV_s), from the equilibrium case is $|2V_{fp}|$. (The absolute voltage of the channel is 0 V; that is, the source, drain, and channel are at ground.) The potential needed to change the surface potential and fill the fixed holes in the substrate is

$$V_B = \frac{Q'_b}{C'_{ox}} - 2V_{fp} \tag{6.11}$$

An additional charge, Q'_{ss} (columbs/area), can be used to model surface states (aka interface trapped charges, Q'_{it}) that may exist because of dangling bonds at the oxide-silicon interface. Here we assume electrons are attracted to the surface of the semiconductor (and trapped directly under the oxide) causing the threshold voltage to decrease. Equation (6.11) may be rewritten to include these surface-state charges as

$$V_B = \frac{Q'_b - Q'_{ss}}{C'_{ox}} - 2V_{fp} \tag{6.12}$$

The final component needed to determine the threshold voltage is the contact potential between point C (the bulk) and point A (the gate material) in Fig. 6.7. The potential difference between the gate and bulk (p-substrate) can be determined by summing the difference between the materials in the MOS system shown in Fig. 6.8. Adding the contact potentials, we get $(V_G - V_{ox}) + (V_{ox} - V_{fp}) = V_G - V_{fp}$. Note that if we were to include the surface electrostatic potential, V_s, the result would be the same, that is, only the two outer materials are of concern when calculating the contact potential difference. The contact potential between the bulk and the gate poly, we will assume n+ poly with doping concentration $N_{D,poly}$, is given by

$$V_{ms} = V_G - V_{fp} = \frac{kT}{q} \ln\left(\frac{N_{D,poly}}{n_i}\right) + \frac{kT}{q} \ln\frac{N_A}{n_i} \tag{6.13}$$

The threshold voltage, V_{THN}, is given by

$$V_{THN} = \frac{Q'_b - Q'_{ss}}{C'_{ox}} - 2V_{fp} - V_{ms} \tag{6.14}$$

$$= -V_{ms} - 2V_{fp} + \frac{Q'_{bo} - Q'_{ss}}{C'_{ox}} - \frac{Q'_{bo} - Q'_b}{C'_{ox}} \tag{6.15}$$

$$= -V_{ms} - 2V_{fp} + \frac{Q'_{bo} - Q'_{ss}}{C'_{ox}} + \frac{\sqrt{2q\varepsilon_{si}N_A}}{C'_{ox}}\left[\sqrt{|2V_{fp}| + V_{SB}} - \sqrt{|2V_{fp}|}\right] \tag{6.16}$$

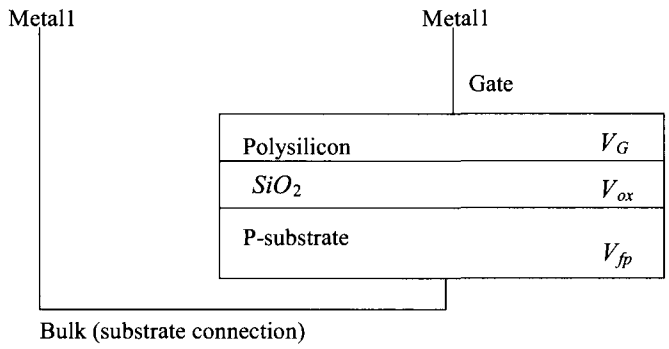

Figure 6.8 Determining the contact potential between poly and substrate.

When the source is shorted to the substrate, $V_{SB} = 0$, we can define the zero-bias threshold voltage as

$$V_{THN0} = -V_{ms} - 2V_{fp} + \frac{Q'_{bo} - Q'_{ss}}{C'_{ox}} \tag{6.17}$$

We can define a body effect coefficient or body factor by

$$\gamma = \frac{\sqrt{2q\varepsilon_{si}N_A}}{C'_{ox}} \tag{6.18}$$

Equation (6.16) can now be written as

$$V_{THN} = V_{THN0} + \gamma\left(\sqrt{|2V_{fp}| + V_{SB}} - \sqrt{|2V_{fp}|}\right) \tag{6.19}$$

It is interesting to note that a voltage, called the flatband voltage V_{FB}, must be applied for the oxide-semiconductor interface surface potential, V_s, to become the same potential as the bulk surface potential, V_{fp}. The flatband voltage is given by

$$V_{FB} = -V_{ms} - \frac{Q'_{ss}}{C'_{ox}} \tag{6.20}$$

The zero-bias threshold voltage may then be written in terms of the flatband voltage as

$$V_{THN0} = V_{FB} - 2V_{fp} + \frac{Q'_{bo}}{C'_{ox}} \tag{6.21}$$

These equations describe how the threshold voltage of the MOSFET is affected by substrate doping, oxide thickness, source/substrate bias, gate material, and surface charge density.

Example 6.3

Assuming $N_A = 10^{16}$ atoms/cm^3 and $C'_{ox} = 1.75$ fF/μm^2, estimate γ (GAMMA, the body effect coefficient).

From Eq. (6.18) the calculated γ is

$$\gamma = \frac{\sqrt{2 \cdot 1.6 \times 10^{-19} \frac{columbs}{atom} \cdot 11.7 \cdot 8.85 \frac{aF}{\mu m} \cdot 10^{16} \frac{atoms}{cm^3} \cdot \frac{cm^3}{10^{12} \mu m^3}}}{1.75 \frac{fF}{\mu m^2}} = 0.330\ V^{1/2} \ \blacksquare$$

Example 6.4

Estimate the zero-bias threshold voltage for the MOSFET of Ex. 6.2. Assume that the poly doping level is 10^{20} atoms/cm^3. What happens to the threshold voltage if sodium contamination causes an impurity of 40 aC/μm^2 at the oxide-semiconductor interface with $C'_{ox} = 1.75\,fF/\mu m^2$?

The electrostatic potential between the gate and substrate is given by

$$-V_{ms} = V_{fp} - V_G = -290\ mV - 26\ mV \cdot \ln\frac{10^{20}}{14.5 \times 10^9} = -879\ mV$$

$$-2V_{fp} = 580\ mV$$

$$\frac{Q'_{bo}}{C'_{ox}} = \frac{139\ aC/\mu m^2}{1.75\ fF/\mu m^2} = 79\ mV$$

$$\frac{Q'_{ss}}{C'_{ox}} = 23\ mV$$

The threshold voltage, from Eq. (6.17) without the sodium contamination, is –220 mV; with the sodium contamination the threshold voltage is –243 mV. ■

Threshold Voltage Adjust

These threshold voltages would correspond to *depletion devices* (a negative threshold voltage), that is, MOSFETs that conduct when the $V_{GS} = 0$. In CMOS applications, this is highly undesirable. We normally use *enhancement devices* (devices with positive threshold voltages that are off with $V_{GS} = V_{SG} = 0$). To compensate or adjust the value of the threshold voltage (the channel, the area under the gate poly) can be implanted with p+ ions. This effectively increases the value of the threshold voltage by Q'_c/C'_{ox}, where Q'_c is the charge density/unit area due to the implant. If N_I is the ion implant dose in atoms/unit area, then we can write

$$Q'_c = q \cdot N_I \tag{6.22}$$

and the threshold voltage by

$$V_{THN0} = -V_{ms} - 2V_{fp} + \frac{Q'_{bo} - Q'_{ss} + Q'_c}{C'_{ox}} \tag{6.23}$$

Example 6.5
Estimate the ion implant dose required to change the threshold voltage in Ex. 6.4 without sodium contamination, to 1 V.

From Eqs. (6.22) and (6.23) and the results of Ex. 6.4

$$V_{THN0} = -220\ mV + \frac{qN_I}{C'_{ox}} = 1\ V$$

This gives $N_I = 1.3 \times 10^{12}$ atoms/cm^2. ■

These calculations lend some insight into how the threshold voltage is affected by the different process parameters. In practice, the results of these calculations do not exactly match the measured threshold voltage. From a circuit design engineer's point of view, the threshold voltage and the body factor are measured in the laboratory when the SPICE models are extracted.

6.3 IV Characteristics of MOSFETs

Now that we have some familiarity with the factors influencing the threshold voltage of a MOSFET, let's derive the large-signal IV (current/voltage) characteristics of the MOSFET, namely operation in the triode and the saturation regions. The following derivation is sometimes referred to as the *gradual-channel approximation.* The electric field variation in the channel between the source and drain (the y-direction) doesn't vary significantly when compared to the variation in the direction perpendicular to the channel (the x-direction).

6.3.1 MOSFET Operation in the Triode Region

Consider Fig. 6.9, where $V_{GS} > V_{THN}$, so that the surface under the oxide is inverted and $V_{DS} > 0$, causing a drift current to flow from the drain to the source. In our initial analysis, we assume that V_{DS} is sufficiently small so that the threshold voltage and the depletion layer width are approximately constant.

Initially, we must find the charge stored on the oxide capacitance C'_{ox}. The voltage, with respect to the source of the MOSFET, of the channel a distance y away from the source is labeled $V(y)$. The potential difference between the gate electrode and the channel is then $V_{GS} - V(y)$. The charge/unit area in the inversion layer is given by

$$Q'_{ch} = C'_{ox} \cdot [V_{GS} - V(y)] \tag{6.24}$$

However, we know that a charge Q'_b is present in the inversion layer from the application of the threshold voltage, V_{THN}, necessary for conduction between the drain and the source. This charge is given by

$$Q'_b = C'_{ox} \cdot V_{THN} \tag{6.25}$$

The total charge available in the inverted channel, for conduction of a current between the drain and the source, is given by the difference in these two equations, or

$$Q'_I(y) = C'_{ox} \cdot (V_{GS} - V(y) - V_{THN}) \tag{6.26}$$

The differential resistance of the channel region with a length dy and a width W is given by

$$dR = \overbrace{\frac{1}{\mu_n Q'_I(y)}}^{\text{eff. sheet Res.}} \cdot \frac{dy}{W} \tag{6.27}$$

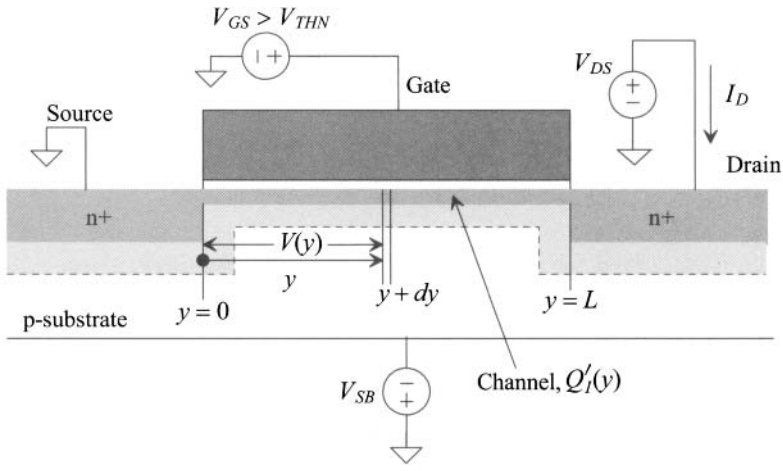

Figure 6.9 Calculation of the large-signal behavior of the MOSFET in the triode (ohmic) region.

where μ_n is the average electron mobility through the channel with units of $cm^2/V \cdot sec$ (see Eq. [5.4]). The mobility is simply a ratio of the electron (or hole) velocity cm/sec to the electric field, V/cm. For short-channel devices, the mobility decreases when the velocity of the carriers starts to saturate. This causes the effective sheet resistance in Eq. (6.27) to increase, resulting in a lowering of the drain current. This (velocity saturation) is discussed in more detail later in the chapter.

The differential voltage drop across this differential resistance is given by

$$dV(y) = I_D \cdot dR = \frac{I_D}{W\mu_n Q_I'(y)} \cdot dy \tag{6.28}$$

or substituting Eq. (6.26) and rearranging

$$I_D \cdot dy = W\mu_n C_{ox}'(V_{GS} - V(y) - V_{THN}) \cdot dV(y) \tag{6.29}$$

At this point, let's define the transconductance parameter, *KP*, for a MOSFET. For an n-channel MOSFET, this parameter is given by

$$KP_n = \mu_n \cdot C_{ox}' = \mu_n \cdot \frac{\varepsilon_{ox}}{t_{ox}} \tag{6.30}$$

and for a p-channel MOSFET, it is given by

$$KP_p = \mu_p \cdot C_{ox}' = \mu_p \cdot \frac{\varepsilon_{ox}}{t_{ox}} \tag{6.31}$$

where μ_p is the mobility of the holes in a PMOS transistor. Typical values of *KP* in the *long-channel* process (with a minimum length of 1 μm) used in this book are 120 $\mu A/V^2$ and 40 $\mu A/V^2$ for n- and p-channel transistors, respectively.

The current can be obtained by integrating the left side of Eq. (6.29) from the source to the drain, that is, from 0 to *L* and the right side from 0 to V_{DS}. This is shown below:

$$I_D \int_0^L dy = W \cdot KP_n \cdot \int_0^{V_{DS}} (V_{GS} - V(y) - V_{THN}) \cdot dV(y) \tag{6.32}$$

or

$$I_D = KP_n \cdot \frac{W}{L} \cdot \left[(V_{GS} - V_{THN})V_{DS} - \frac{V_{DS}^2}{2} \right] \text{ for } V_{GS} \geq V_{THN} \text{ and } V_{DS} \leq V_{GS} - V_{THN} \tag{6.33}$$

This equation is valid when the MOSFET is operating in the triode (aka **linear** or **ohmic**) region. This is the case when the induced channel extends from the source to the drain. Furthermore, we can rewrite Eq. (6.33) defining the transconductance parameter as

$$\beta = KP_n \cdot \frac{W}{L} \tag{6.34}$$

or

$$I_D = \beta \cdot \left[(V_{GS} - V_{THN})V_{DS} - \frac{V_{DS}^2}{2} \right] \tag{6.35}$$

The equivalent equation for the PMOS device operating in the triode region is

$$I_D = KP_p \cdot \frac{W}{L} \cdot \left[(V_{SG} - V_{THP})V_{SD} - \frac{V_{SD}^2}{2} \right] \text{ for } V_{SG} \geq V_{THP} \text{ and } V_{SD} \leq V_{SG} - V_{THP} \quad (6.36)$$

where the threshold voltage of the p-channel MOSFET is positive (noting, again, that from our sign convention in Fig. 6.1 all voltages and currents are positive.)

6.3.2 The Saturation Region

The voltage at $V(y)$ when $y = L$, that is, $V(L)$, in Eq. (6.26) is simply V_{DS}. In the previous subsection, we said that V_{DS} is always less than $V_{GS} - V_{THN}$ so that at no point along the channel is the inversion charge zero. When $V_{DS} = V_{GS} - V_{THN}$, the inversion charge under the gate at $y = L$ (the drain-channel junction) is zero, Eq. (6.26). This drain-source voltage is called $V_{DS,sat}$ $(= V_{GS} - V_{THN})$, and indicates when the channel charge becomes *pinched off* at the drain-channel interface. Increases in V_{DS} beyond $V_{DS,sat}$ attract the fixed channel charge to the drain terminal depleting the charge in the channel directly adjacent to the drain. Further increases in V_{DS} do not cause an increase in the drain current[1]. In other words the *current saturates* and thus stops increasing.

Figure 6.10 shows the depletion region, with a thickness of X_{dl}, between the drain and channel. An increase in V_{DS} results in an increase in X_{dl}. If V_{DS} is increased until X_{dl} extends from the drain to the source, the device is said to be *punched through*. Large currents can flow under these conditions, causing device failure. The maximum voltage, for near minimum-size channel lengths, that can be applied between the drain and source of a MOSFET is set by the "punchthrough" voltage. For long-channel lengths, the maximum voltage is set by the breakdown voltage of the drain (n+) to substrate diode.

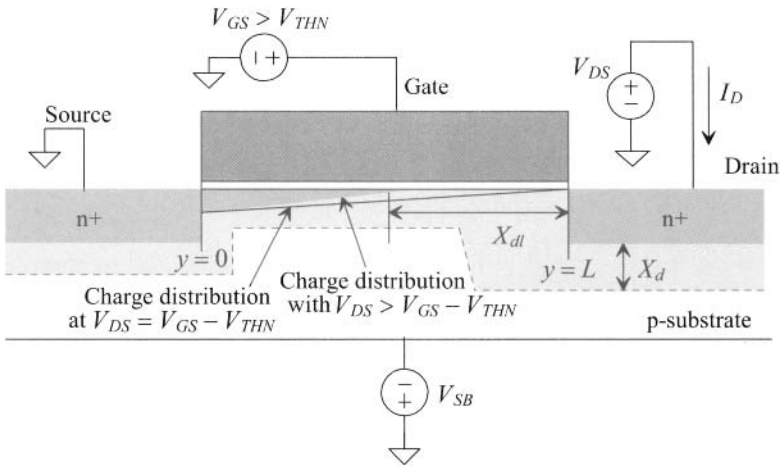

Figure 6.10 The MOSFET in saturation (pinched off).

[1] We will see on the next page that this is not entirely true. An effect called *channel length modulation* causes the drain current to increase slightly with increasing drain-source voltage.

When a MOSFET's channel is pinched off, that is, $V_{DS} \geq V_{GS} - V_{THN}$ and $V_{GS} \geq V_{THN}$, it is operating in the saturation region. Substitution of $V_{DS,sat}$ into Eq. (6.33) yields

$$I_{D,sat} = \frac{KP_n}{2} \cdot \frac{W}{L} \cdot (V_{GS} - V_{THN})^2 = \frac{\beta}{2}(V_{GS} - V_{THN})^2$$

$$\text{for } V_{DS} \geq V_{GS} - V_{THN} \text{ and } V_{GS} \geq V_{THN} \tag{6.37}$$

We can define an electrical channel length, L_{elec}, of the MOSFET as the difference between the drawn channel length, L, neglecting lateral diffusion, and the depletion layer width, X_{dl}

$$L_{elec} = L - X_{dl} \tag{6.38}$$

Substituting this into Eq. (6.37), we obtain a better representation of the drain current

$$I_D = \frac{KP_n}{2} \cdot \frac{W}{L_{elec}}(V_{GS} - V_{THN})^2 \tag{6.39}$$

Qualitatively, this means that since the depletion layer width increases with increasing V_{DS}, the drain current increases as well. This effect is called *channel length modulation* (CLM). As L is increased the effects of X_{dl} changing (CLM) become negligible.

To determine the change in output current with drain-source voltage, we take the derivative of Eq. (6.39) with respect to V_{DS} around $V_{DS,sat}$ (when $L \approx L_{elec}$)

$$\frac{dI_D}{dV_{DS}} = -\frac{KP_n}{2}\frac{W}{L_{elec}^2}(V_{GS} - V_{THN})^2 \cdot \frac{dL_{elec}}{dV_{DS}} = I_{D,sat} \cdot \left[\frac{1}{L}\frac{dX_{dl}}{dV_{DS}}\right] \tag{6.40}$$

where it's common to define λ, the channel length modulation parameter, as

$$\lambda = \frac{1}{L} \cdot \frac{dX_{dl}}{dV_{DS}} \tag{6.41}$$

Typical values for λ range from greater than 0.1 V^{-1} to less than 0.01 V^{-1} (ideally $\lambda = 0$). Note that the units of dV_{DS}/dX_{dl} are V/m and that this term grows as process technology shrinks. The smaller devices are designed to drop larger voltages over smaller distances. This last point is why *a nanometer device can't be made to behave like a long-channel device simply by increasing its length.*

Equation (6.37) can be rewritten to account for CLM as

$$I_D = \frac{KP_n}{2} \cdot \frac{W}{L}(V_{GS} - V_{THN})^2[1 + \lambda(V_{DS} - V_{DS,sat})] = I_{D,sat} \cdot [1 + \lambda(V_{DS} - V_{DS,sat})]$$

$$\text{for } V_{DS} > V_{DS,sat} = V_{GS} - V_{THN} \text{ and } V_{GS} > V_{THN} \tag{6.42}$$

The drain current at the triode/saturation region border occurs when

$$I_D = I_{D,sat} \text{ and } V_{DS} = V_{DS,sat} = V_{GS} - V_{THN} \tag{6.43}$$

In these equations we assumed that the mobility does not vary with V_{DS}. Later in the chapter (in the short-channel MOSFET discussion, Sec. 6.5.2) we'll see that the mobility does indeed vary with V_{DS} making characterizing I_D considerably more challenging. Figure 6.11 shows typical curves for an n-channel MOSFET. Notice how the device *appears to go into saturation earlier* than predicted by $V_{DS,sat} = V_{GS} - V_{THN}$. The bold line in the figure separates the actual triode and saturation regions (and also indicates $I_{D,sat}$).

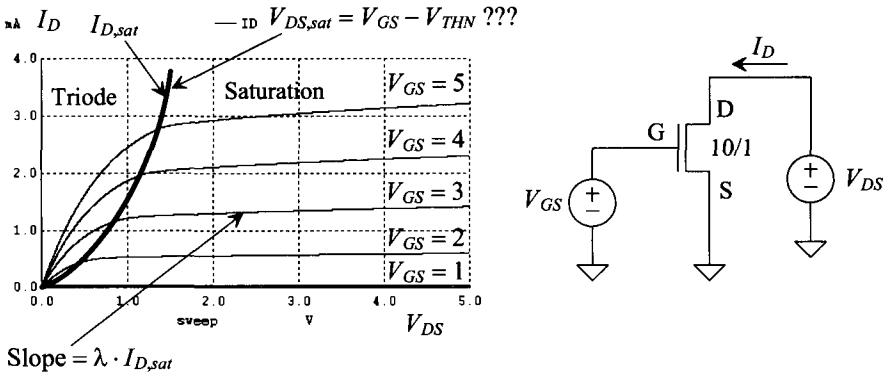

Figure 6.11 Characterisitics of a long-channel NMOS device.

For example, in the simulation results at $V_{GS} = 5$ V (with V_{THN} roughly equal to 1 V), we see that $V_{DS,sat}$ is 1.4 V (not the 4 V we calculate using $V_{GS} - V_{THN}$) The actual charge distribution in the channel is not constant but rather a function of V_{DS}. $Q'_I(y)$ decreases as we move away from the source of the MOSFET, causing $Q'_I(L)$ to become zero earlier, see Fig. 6.10.

C_{gs} Calculation in the Saturation Region

The gate-to-source capacitance of a MOSFET operating in the saturation region can be determined by solving Eq. (6.26) for the total charge in the inverted channel,

$$Q_I = \int_0^L W \cdot Q'_I(y) \cdot dy = WC'_{ox} \int_0^L (V_{GS} - V(y) - V_{THN})\, dy \tag{6.44}$$

or solving Eq. (6.29) for dy and substituting yields

$$Q_I = \frac{(W \cdot C'_{ox})^2 \cdot \mu_n}{I_D} \int_0^{V_{GS}-V_{THN}} (V_{GS} - V(y) - V_{THN})^2 \cdot dV(y) \tag{6.45}$$

where it was used that Q_I goes to zero when $y = L$ occurs when $V_{DS} = V_{GS} - V_{THN}$. Solving this equation using Eqs. (6.30) and (6.37) yields

$$Q_I = \frac{2}{3} \cdot W \cdot L \cdot C'_{ox} \cdot (V_{GS} - V_{THN}) \tag{6.46}$$

We can determine the gate-to-source capacitance while in the saturation region by,

$$C_{gs} = \frac{\partial Q_I}{\partial V_{GS}} = \frac{2}{3} \cdot W \cdot L \cdot C'_{ox} \tag{6.47}$$

See the entry in Table 6.1 (note the discontinuity between saturation and triode).

6.4 SPICE Modeling of the MOSFET

In this section we list the level 1, 2, and 3 SPICE model parameters and their relationship to the equations derived in the last section. The level 1 model is a subset of the level 2 and 3 models which have more elaborate mobility modeling. All three models are based on Eq. (6.42). A level 3 model for a long-channel CMOS process is also presented.

Model Parameters Related to V_{THN}

The following SPICE model parameters are related to the calculation of V_{THN},

Symbol	Name	Description	Default	Typ.	Units
V_{THN0}	VTO	Zero-bias threshold voltage	1.0	0.8	Volts
γ	GAMMA	Body-effect parameter	0	0.4	$V^{1/2}$
$2\lvert V_{fp}\rvert$	PHI	Surface to bulk potential	0.65	0.58	V
N_A	NSUB	Substrate doping	0	1E15	cm^{-3}
Q'_{ss}/q	NSS	Surface state density	0	1E10	cm^{-2}
	TPG	Type of gate material	1	1	

Using Eq. (6.19), we can calculate the threshold voltage, V_{THN}, given the above parameters. If V_{THN0} or γ are not given, then SPICE calculates them using the above information and Eqs. (6.19) – (6.21). TPG specifies the type of gate material: 1 opposite to substrate, –1 same as substrate, and 0 for aluminum gate.

Long-Channel MOSFET Models

The SPICE models used in this book for the "long-channel CMOS process" follow. The scale factor is 1 μm (= minimum drawn channel length).

```
* 1 um Level 3 models
* Don't forget the .options scale=1u if using an Lmin of 1
* 1<L<200 and 10<W<10000 Vdd=5V

.MODEL NMOS NMOS LEVEL = 3
+ TOX  = 200E-10      NSUB  = 1E17       GAMMA = 0.5
+ PHI  = 0.7          VTO   = 0.8        DELTA = 3.0
+ UO   = 650          ETA   = 3.0E-6     THETA = 0.1
+ KP   = 120E-6       VMAX  = 1E5        KAPPA = 0.3
+ RSH  = 0            NFS   = 1E12       TPG   = 1
+ XJ   = 500E-9       LD    = 100E-9
+ CGDO = 200E-12      CGSO  = 200E-12    CGBO  = 1E-10
+ CJ   = 400E-6       PB    = 1          MJ    = 0.5
+ CJSW = 300E-12      MJSW  = 0.5
*
.MODEL PMOS PMOS LEVEL = 3
+ TOX  = 200E-10      NSUB  = 1E17       GAMMA = 0.6
+ PHI  = 0.7          VTO   = -0.9       DELTA = 0.1
+ UO   = 250          ETA   = 0          THETA = 0.1
+ KP   = 40E-6        VMAX  = 5E4        KAPPA = 1
+ RSH  = 0            NFS   = 1E12       TPG   = -1
+ XJ   = 500E-9       LD    = 100E-9
+ CGDO = 200E-12      CGSO  = 200E-12    CGBO  = 1E-10
+ CJ   = 400E-6       PB    = 1          MJ    = 0.5
+ CJSW = 300E-12      MJSW  = 0.5
```

Model Parameters Related to the Drain Current

The SPICE implemention of the square-law equations, Eqs. (6.35) and (6.42), is slightly different [8] than our derivations. Equation [6.42] is implemented in SPICE with $V_{DS,sat} = 0$. To avoid a discontinuity then Eq. (6.35) is multiplied by $(1 + \lambda \cdot V_{DS})$.

Symbol	Name	Description	Default	Typ.	Units
KP	KP	Transconductance parameter	20E-6	50E-6	A/V^2
t_{ox}	TOX	Gate-oxide thickness	1E-7	40E-10	m
λ	Lambda	Channel-length modulation	0	0.01	V^{-1}
L_{diff}	LD	Lateral diffusion	0	2.5E-7	m
$\mu_{n,p}$	UO	Surface mobility	600	580	cm^2/Vs

SPICE Modeling of the Source and Drain Implants

Name	Description	Default	Typical	Units
RD	Drain contact resistance	0	40	Ω
RS	Source contact resistance	0	40	Ω
RSH	Source/drain sheet resistance	0	50	Ω/sq.
CGBO	Gate-bulk overlap capacitance	0	4E-10	F/m
CGDO	Gate-drain overlap capacitance	0	4E-10	F/m
CGSO	Gate-source overlap capacitance	0	4E-10	F/m
PB, PBSW	Bottom, sidewall built-in potential	0.8	0.8	V
MJ, MJSW	Bottom, sidewall grading coefficient	0.6	0.6	
CJ	Bottom zero-bias depletion capacitance	0	3E-4	F/m^2
CJSW	Sidewall zero-bias depletion capacitance	0	2.5E-10	F/m
IS	Bulk-junction saturation current	1E-14	1E-14	A
JS	Bulk-junction saturation current density	0	1E-8	A/m^2
FC	Bulk-junction forward bias coefficient	0.6	0.6	

Summary

Table 6.2 lists the characteristics of the long-channel CMOS process used in this book.

Table 6.2 Summary of device characteristics for the long-channel CMOS process.

Long-channel MOSFET parameters used in this book. The VDD = 5 V and the scale factor is **1 μm** ($scale$ = 1e–6)			
Parameter	**NMOS**	**PMOS**	**Comments**
V_{THN} and V_{THP}	800 mV	900 mV	Typical
KP_n and KP_p	120 μA/V²	40 μA/V²	t_{ox} = 200 Å
$C'_{ox} = \varepsilon_{ox}/t_{ox}$	1.75 fF/μm²	1.75 fF/μm²	$C_{ox} = C'_{ox}WL \cdot (scale)^2$
λ_n and λ_p	0.01 V^{-1}	0.0125 V^{-1}	at L = 2
γ_n and γ_p	0.5 $V^{-1/2}$	0.6 $V^{-1/2}$	Body factor

6.4.1 Some SPICE Simulation Examples

Figure 6.11 shows the I_D-V_{DS} characteristics of an NMOS device in the long-channel process. Figure 6.12 shows the equivalent PMOS device. Notice that the devices are the same size, 10/1, in the two simulations; however, the drain current of the PMOS is less than half the drain current of the NMOS. This is related to the mobility of the holes being two to three times lower than the mobility of the electrons. Electrons in the valence band are more tightly coupled to the nucleus of an atom than are electrons in the conduction band. Because apparent movement of holes is actually the result of electrons moving in the valence band, the hole mobility is lower than electron (in conduction band) mobility.

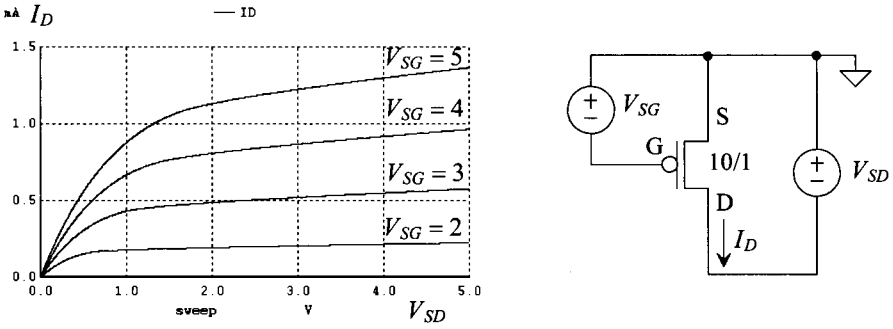

Figure 6.12 Characterisitics of a long-channel PMOS device.

Threshold Voltage and Body Effect

In simple terms, the threshold voltage is the voltage that turns the device on and allows drain current to flow from the drain to the source. Figure 6.13 shows the I_D-V_{GS} curves for an NMOS device. When the source and substrate are at the same potential, $V_{SB} = 0$, the threshold voltage is labeled V_{THN0} (see Eq. [6.17]). When V_{SB} starts to increase, the threshold voltage goes up. This is called the body effect. Figure 6.14 shows two MOSFETs: one with and one without body effect (substrate, or body, is grounded).

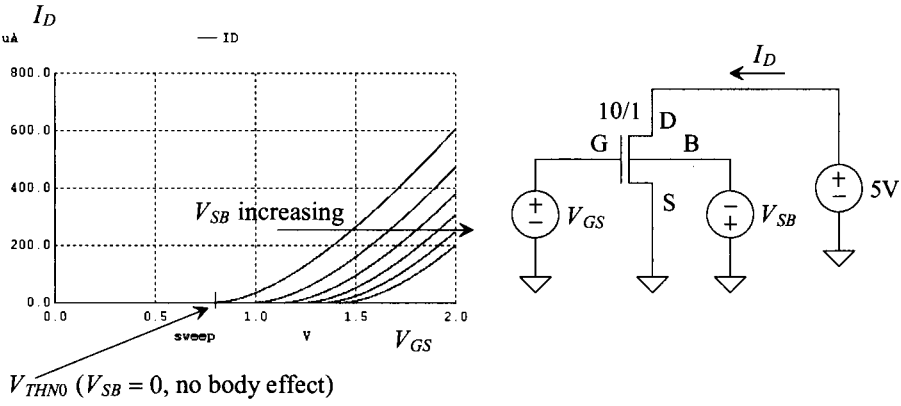

Figure 6.13 Threshold voltage and body effect.

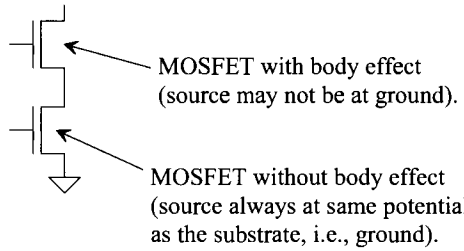

Figure 6.14 How an NMOS can have body effect.

To qualitatively understand the origin of the body effect, consider the MOSFET cross-sectional view seen in Fig. 6.15. As the source potential rises above the bulk (substrate) potential (represented by V_{SB} in the figure), electrons are attracted towards the positive terminal of V_{SB} from the MOSFET's channel. To keep the surface inverted, a larger V_{GS} must be applied to the MOSFET. Thus the effect of the body stealing charge from the channel is an increase in the MOSFET's threshold voltage.

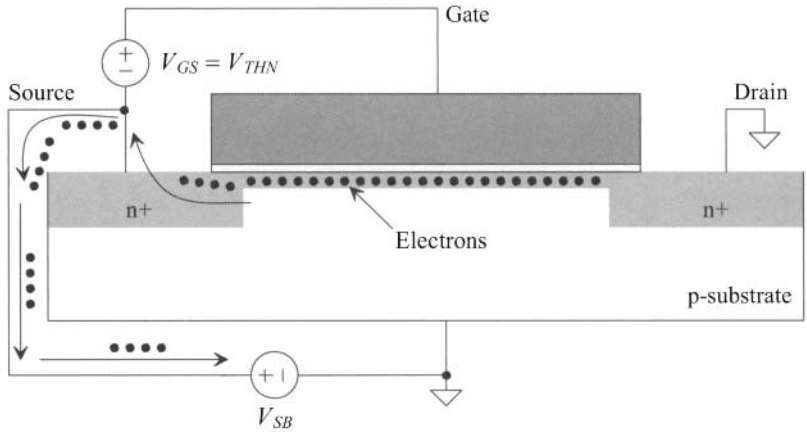

Figure 6.15 Qualitative description of body effect.

6.4.2 The Subthreshold Current

In the last section we said that the MOSFET starts to conduct a current when $V_{GS} = V_{THN}$. In reality there is a drain current, albeit small, when $V_{GS} < V_{THN}$. This current is called subthreshold current. When the MOSFET is operating in the weak inversion region it can also be said to be operating in the subthreshold region. Subthreshold operation can be very useful for low-power operation. Solar-powered calculators, CMOS imagers, or battery- operated watches are examples of devices using CMOS ICs operating in the subthreshold region. The main problems that plague circuits designed to operate in the subthreshold region are matching, noise, and bandwidth. For example, since the drain current is exponentially related to the gate-source voltage (as we'll soon see), any mismatch in these voltages can cause significant differences in the drain current.

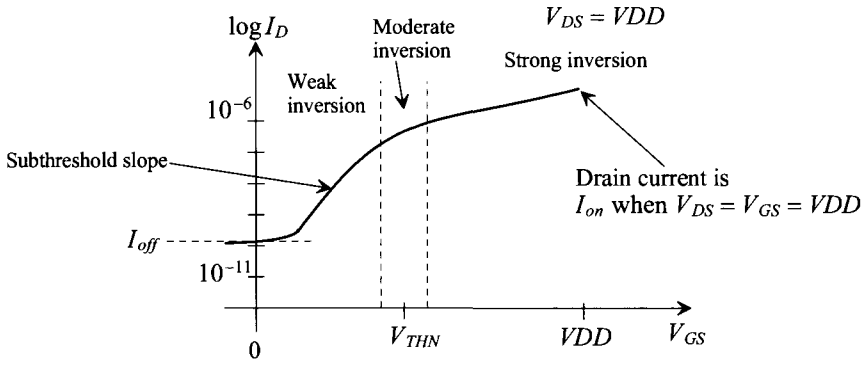

Figure 6.16 Drain current plotted from weak to strong inversion.

The subthreshold region is often characterized using the log I_D plotted against V_{GS} (see Fig. 6.16). The current transport from the drain to source in Sec. 6.3 was via drift. An applied electric field, when the MOSFET is operating in the strong inversion region, causes carriers to drift from the channel to the drain across the depletion region (and hence why we talk about mobility). In the weak inversion, or subthreshold region, the carriers diffuse from the source to the drain just like carrier movement in a bipolar junction transistor, BJT. In a BJT the carriers are emitted from the emitter, diffuse across the base, and are collected at the collector. In a MOSFET operating in subthreshold, the carriers are emitted by the source, diffuse across the body of the device (under the gate oxide) and are collected at the drain. We can write the drain current of the MOSFET in the subthreshold region as

$$I_D = I_{D0} \cdot \frac{W}{L} \cdot e^{q(V_{GS} - V_{THN})/(n \cdot kT)} \tag{6.48}$$

Taking the log of both sides with $V_T = kT/q$ (the thermal voltage), we get

$$\log I_D = \log \frac{W}{L} + \log I_{D0} + -\frac{V_{THN}}{nV_T} \cdot \log e + \overbrace{\left[\frac{1}{V_T \cdot n} \cdot \log e\right]}^{\text{subthreshold slope}} \cdot V_{GS} \tag{6.49}$$

The reciprocal of the subthreshold slope is given by

$$\text{Subthreshold slope}^{-1} = \frac{V_T \cdot n}{\log e} \text{ (mV/decade)} \tag{6.50}$$

If kT/q = 0.026 V = V_T and n (the slope parameter) = 1, the reciprocal of the subthreshold slope is 60 mV/decade (it can be said the subthreshold slope is 60 mV/decade and it is understood it is actually one over the slope). In bulk CMOS n is around 1.6 and the subthreshold slope is 100 mV/decade at room temperature. For the ideal MOSFET used as a switch when V_{GS} is less than the threshold voltage, the drain current goes to zero. The slope of the curve below V_{THN} in Fig. 6.16 is then infinite (corresponding to zero subthreshold slope^{-1}). The subthreshold slope can be a very important MOSFET parameter in many applications (the design of dynamic circuits). Notice that the drain current that flows with V_{GS} = 0 is called I_{off} (with $V_{DS} = VDD$). The drain current that flows when $V_{GS} = V_{DS} = VDD$ (in the strong inversion region) is called I_{on}.

6.5 Short-Channel MOSFETs

The long-channel CMOS process used in the first part of this chapter is useful for illustrating the fundamentals of MOSFET operation. However, modern CMOS transistors have channel lengths that are well below the 1 μm minimum length of this process. The gradual channel approximation used earlier to develop the square-law current-voltage characteristics of the MOSFET falls apart for modern short-channel devices. The electric field under the gate oxide can no longer be treated in a single dimension. In addition, the velocity of the carriers drifting between the channel and the drain of the MOSFET can saturate, Fig. 6.17, an effect called *carrier velocity saturation*, v_{sat}. Typical values for the low electric field mobilities (electric fields, E, less than the critical electric field, E_{crit}, where the velocity saturates) for electrons and holes are 600 cm^2/Vs (μ_n) and 250 cm^2/Vs (μ_p). See text associated with Eq. (5.4) for additional information.

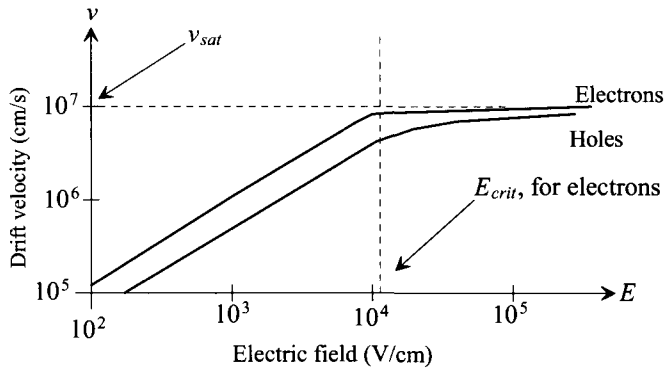

Figure 6.17 Drift velocity plotted against electric field. The slope of these curves is the mobility of the carriers, see Eq. (5.4)

Hot Carriers

In very small devices some of the carriers drifting near the drain can obtain energies much larger than the thermal energy of carriers under equilibrium conditions. These carriers are termed *hot carriers*. The carrier scattering events are no longer localized allowing the carriers to have higher, than the values mentioned above, mobilities. The velocity of these carriers can exceed the saturation velocity indicated in Fig. 6.17. This effect is called *velocity overshoot* and it can enhance the speed and transconductance of the MOSFETs (which is good). Unfortunately, hot carriers can also tunnel through the gate oxide and cause gate current or become trapped in the gate oxide, having the effect of changing the MOSFET's threshold voltage. Hot carriers can also cause impact ionization (avalanche multiplication).

Lightly-Doped Drain (LDD)

A cross-sectional view of an NMOS device using a lightly doped drain (LDD) structure is shown in Fig. 6.18 (and Fig. 4.7). The addition of the lightly doped n- provides a resistive buffer between the channel and the higher-doped source/drain. The effect of the LDD structure is to increase the voltage dropped across the drain-channel interface, dV_{DS}/dX_{dl}.

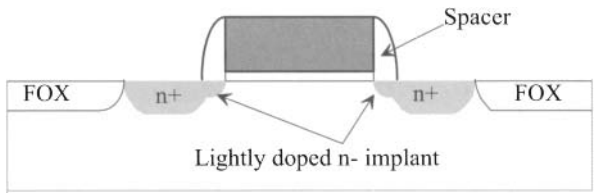

Figure 6.18 Lightly doped drain (LDD) implant.

6.5.1 MOSFET Scaling

Reducing the channel length of a MOSFET can be described in terms of scaling theory. A scaling parameter S ($S < 1$) is used to scale the dimensions of a MOSFET. The value of S is typically in the neighborhood of 0.7 from one CMOS technology generation to the next. For example, if a process uses a *VDD* of 2 V, a next generation process would use a *VDD* of 1.4 V. In other words

$$VDD' = VDD \cdot S \tag{6.51}$$

The channel length of the scaled process is reduced to

$$L' = L \cdot S \tag{6.52}$$

while the width is reduced to

$$W' = W \cdot S \tag{6.53}$$

Table 6.3 describes how S affects the MOSFET parameters. The main benefits of scaling are (1) smaller device sizes and thus reduced chip size (increased yield and more parts per wafer), (2) lower gate delays, allowing higher frequency operation, and (3) reduction in power dissipation. Associated with these benefits are some unwanted side effects referred to as short-channel effects. These unwanted effects are discussed in the next section.

Table 6.3 CMOS scaling relationships.

Parameter	Scaling
Supply voltage (*VDD*)	S
Channel length (L_{min})	S
Channel width (W_{min})	S
Gate-oxide thickness (t_{ox})	S
Substrate doping (N_A)	S^{-1}
On current (I_{on})	S
Gate capacitance (C_{ox})	S
Gate delay	S
Active power	S^3

6.5.2 Short-Channel Effects

The average drift velocity, v, of an electron plotted against electric field, E, was shown in Fig. 6.17. When the electric field reaches a critical value, labeled E_{crit}, the velocity saturates at a value v_{sat}, that is, the velocity ceases to increase with increasing electric field (note that here we are neglecting the potential for velocity overshoot). The ratio of electron drift velocity to applied electric field is the electron mobility (Eq. 5.4), or, again

$$\mu_n = \frac{v}{E} \tag{6.54}$$

Above the critical electric field, the mobility starts to decrease, whereas below E_{crit}, the mobility is essentially constant. Rewriting Eq. (6.29) to determine how the mobility changes with $V(y)$ results in

$$I_D = \mu_n \cdot \frac{dV(y)}{dy} \cdot W \cdot C'_{ox}[V_{GS} - V_{THN} - V(y)] \tag{6.55}$$

We are interested in determining how the drain current of a short-channel MOSFET changes with V_{GS} when operating in the saturation region. (The charge under the gate oxide at the drain channel interface is zero, and the channel is pinched off.) The MOSFET enters the saturation region when $V(L) = V_{DS,sat}$. At high electric fields, the mobility can be approximated by

$$\mu_n = \frac{v_{sat}}{E} = \frac{v_{sat}}{dV(y)/dy} \tag{6.56}$$

so that Eq. (6.55) can be written as

$$I_D = W \cdot v_{sat} \cdot C'_{ox}(V_{GS} - V_{THN} - V_{DS,sat}) \tag{6.57}$$

The drain current of a short-channel MOSFET operating in the saturation region increases linearly with V_{GS}. The long-channel theory, Eq. (6.37), shows the drain current increasing with the square of the gate-source voltage. This result also presents a practical relative figure of merit for the modern CMOS process, the drive current per width of a MOSFET. The on or drive current, I_{on} or I_{drive} (μA/μm), is given by

$$I_{on} = I_{drive} = v_{sat} \cdot C'_{ox}(V_{GS} - V_{THN} - V_{DS,sat}) \tag{6.58}$$

and therefore, see Fig. (6.16),

$$I_D = I_{on} \cdot W = I_{drive} \cdot W \text{ with } V_{GS} = V_{DS} = VDD \tag{6.59}$$

The on (drive) current can be estimated using these equations; however, it is normally measured.

Negative Bias Temperature Instability (NBTI)

In a modern PMOS device when the gate voltage is driven below its source voltage ($V_{SG} > 0$) crucial device parameters, such as the threshold voltage, are observed to shift over time. Historically, both the trapping of holes in oxide defects and the creation of interface states have been suspected to be the cause of the shift. NBTI can be a significant reliability concern in SiO_2 gate dielectrics due to time and temperature-dependent fluctuations in device parameters during both on and off states of operation. NBTI is also present in NMOS devices but it is considerably more pronounced in the PMOS transistor.

Oxide Breakdown

For reliable device operation, the maximum electric field across a device gate oxide should be limited to 10 MV/cm. This translates into 1V / 10 Å of gate oxide. A device with t_{ox} of 20 Å should limit the applied gate voltages to 2 V for reliable long-term operation.

Drain-Induced Barrier Lowering

Drain-induced barrier lowering (DIBL, pronounced "dibble") causes a threshold voltage reduction with the application of a drain-source voltage. The positive potential at the drain terminal helps to attract electrons under the gate oxide and thus increase the surface potential V_s. In other words V_{DS} helps to invert the channel on the drain side of the device, causing a reduction in the threshold voltage. Since V_{THN} decreases with increasing V_{DS}, the result is an increase in drain current and thus a decrease in the MOSFET's output resistance.

Gate-Induced Drain Leakage

Gate-Induced Drain Leakage (GIDL, pronounced "giddle") is a term used to describe a component of the drain to substrate leakage current. When the device is in accumulation (e.g. the gate of an NMOS device is at ground) the surface and substrate potentials are nearly the same. In this situation there can be a dramatic increase in avalanche multiplication or band-to-band tunneling when the drain is at a higher potential. Minority carriers underneath the gate are swept to the substrate creating the leakage current.

Gate Tunnel Current

As the oxide thickness scales downwards, the probability of carriers directly tunneling through the gate oxide increases. For oxide thicknesses less than 15 Å, this gate current can be significant. To reduce the tunnel current, various sandwiches of dielectrics are being explored. Figure 16.67 later in the book presents some results showing values for direct tunnel currents under various operating conditions.

6.5.3 SPICE Models for Our Short-Channel CMOS Process

Section 6.4 presented some SPICE models for the long-channel CMOS process used in this book. In this section we give the BSIM4[2] models for the 50 nm process we use in the book with *VDD* = 1 V, see also Table 5.2. The model listing is given below.

BSIM4 Model Listing (NMOS)

```
* 50nm BSIM4 models
*
* Don't forget the .options scale=50nm if using an Lmin of 1
* 1<Ldrawn<200  10<Wdrawn<10000 Vdd=1V
*
.model      nmos     nmos     level = 54
+binunit = 1         paramchk= 1        mobmod  = 0
+capmod  = 2         igcmod  = 1        igbmod  = 1        geomod  = 1
+diomod  = 1         rdsmod  = 0        rbodymod= 1        rgatemod= 1
```

[2] BSIM4 is a fourth generation MOSFET model developed at the University of California, Berkeley. The acronym stands for Berkeley Short-channel IGFET (insulated gate FET) Model. For more information see: http://www-device.eecs.berkeley.edu

```
+permod  = 1          acnqsmod= 0          trnqsmod= 0

+tnom    = 27         toxe    = 1.4e-009   toxp    = 7e-010    toxm    = 1.4e-009
+epsrox  = 3.9        wint    = 5e-009     lint    = 1.2e-008
+ll      = 0          wl      = 0          lln     = 1         wln     = 1
+lw      = 0          ww      = 0          lwn     = 1         wwn     = 1
+lwl     = 0          wwl     = 0          xpart   = 0         toxref  = 1.4e-009

+vth0    = 0.22       k1      = 0.35       k2      = 0.05      k3      = 0
+k3b     = 0          w0      = 2.5e-006   dvt0    = 2.8       dvt1    = 0.52
+dvt2    = -0.032     dvt0w   = 0          dvt1w   = 0         dvt2w   = 0
+dsub    = 2          minv    = 0.05       voffl   = 0         dvtp0   = 1e-007
+dvtp1   = 0.05       lpe0    = 5.75e-008  lpeb    = 2.3e-010  xj      = 2e-008
+ngate   = 5e+020     ndep    = 2.8e+018   nsd     = 1e+020    phin    = 0
+cdsc    = 0.0002     cdscb   = 0          cdscd   = 0         cit     = 0
+voff    = -0.15      nfactor = 1.2        eta0    = 0.15      etab    = 0
+vfb     = -0.55      u0      = 0.032      ua      = 1.6e-010  ub      = 1.1e-017
+uc      = -3e-011    vsat    = 1.1e+005   a0      = 2         ags     = 1e-020
+a1      = 0          a2      = 1          b0      = -1e-020   b1      = 0
+keta    = 0.04       dwg     = 0          dwb     = 0         pclm    = 0.18
+pdiblc1 = 0.028      pdiblc2 = 0.022      pdiblcb = -0.005    drout   = 0.45
+pvag    = 1e-020     delta   = 0.01       pscbe1  = 8.14e+8   pscbe2  = 1e-007
+fprout  = 0.2        pdits   = 0.2        pditsd  = 0.23      pditsl  = 2.3e+006
+rsh     = 3          rdsw    = 150        rsw     = 150       rdw     = 150
+rdswmin = 0          rdwmin  = 0          rswmin  = 0         prwg    = 0
+prwb    = 6.8e-011   wr      = 1          alpha0  = 0.074     alpha1  = 0.005
+beta0   = 30         agidl   = 0.0002     bgidl   = 2.1e+009  cgidl   = 0.0002
+egidl   = 0.8

+aigbacc = 0.012      bigbacc = 0.0028     cigbacc = 0.002
+nigbacc = 1          aigbinv = 0.014      bigbinv = 0.004     cigbinv = 0.004
+eigbinv = 1.1        nigbinv = 3          aigc    = 0.017     bigc    = 0.0028
+cigc    = 0.002      aigsd   = 0.017      bigsd   = 0.0028    cigsd   = 0.002
+nigc    = 1          poxedge = 1          pigcd   = 1         ntox    = 1

+xrcrg1  = 12         xrcrg2  = 5
+cgso    = 6.238e-010 cgdo    = 6.238e-010 cgbo    = 2.56e-011 cgdl    = 2.495e-10
+cgsl    = 2.495e-10  ckappas = 0.02       ckappad = 0.02      acde    = 1
+moin    = 15         noff    = 0.9        voffcv  = 0.02

+kt1     = -0.21      kt1l    = 0.0        kt2     = -0.042    ute     = -1.5
+ua1     = 1e-009     ub1     = -3.5e-019  uc1     = 0         prt     = 0
+at      = 53000

+fnoimod = 1          tnoimod = 0

+jss     = 0.0001     jsws    = 1e-011     jswgs   = 1e-010    njs     = 1
+ijthsfwd= 0.01       ijthsrev= 0.001      bvs     = 10        xjbvs   = 1
+jsd     = 0.0001     jswd    = 1e-011     jswgd   = 1e-010    njd     = 1
+ijthdfwd= 0.01       ijthdrev= 0.001      bvd     = 10        xjbvd   = 1
+pbs     = 1          cjs     = 0.0005     mjs     = 0.5       pbsws   = 1
+cjsws   = 5e-010     mjsws   = 0.33       pbswgs  = 1         cjswgs  = 3e-010
+mjswgs  = 0.33       pbd     = 1          cjd     = 0.0005    mjd     = 0.5
+pbswd   = 1          cjswd   = 5e-010     mjswd   = 0.33      pbswgd  = 1
+cjswgd  = 5e-010     mjswgd  = 0.33       tpb     = 0.005     tcj     = 0.001
+tpbsw   = 0.005      tcjsw   = 0.001      tpbswg  = 0.005     tcjswg  = 0.001
+xtis    = 3          xtid    = 3


+dmcg    = 0e-006     dmci    = 0e-006     dmdg    = 0e-006    dmcgt   = 0e-007
```

```
+dwj    = 0.0e-008      xgw    = 0e-007      xgl    = 0e-008

+rshg   = 0.4           gbmin  = 1e-010      rbpb   = 5          rbpd   = 15
+rbps   = 15            rbdb   = 15          rbsb   = 15         ngcon  = 1
```

BSIM4 Model Listing (PMOS)

```
.model      pmos        pmos      level = 54

+binunit = 1            paramchk= 1          mobmod  = 0
+capmod  = 2            igcmod  = 1          igbmod  = 1         geomod  = 1
+diomod  = 1            rdsmod  = 0          rbodymod= 1         rgatemod= 1
+permod  = 1            acnqsmod= 0          trnqsmod= 0

+tnom    = 27           toxe    = 1.4e-009   toxp    = 7e-010    toxm    = 1.4e-009
+epsrox  = 3.9          wint    = 5e-009     lint    = 1.2e-008
+ll     = 0             wl      = 0          lln     = 1         wln     = 1
+lw     = 0             ww      = 0          lwn     = 1         wwn     = 1
+lwl    = 0             wwl     = 0          xpart   = 0         toxref  = 1.4e-009

+vth0   = -0.22         k1      = 0.39       k2      = 0.05      k3      = 0
+k3b    = 0             w0      = 2.5e-006   dvt0    = 3.9       dvt1    = 0.635
+dvt2   = -0.032        dvt0w   = 0          dvt1w   = 0         dvt2w   = 0
+dsub   = 0.7           minv    = 0.05       voffl   = 0         dvtp0   = 0.5e-008
+dvtp1  = 0.05          lpe0    = 5.75e-008  lpeb    = 2.3e-010  xj      = 2e-008
+ngate  = 5e+020        ndep    = 2.8e+018   nsd     = 1e+020    phin    = 0
+cdsc   = 0.000258      cdscb   = 0          cdscd   = 6.1e-008  cit     = 0
+voff   = -0.15         nfactor = 2          eta0    = 0.15      etab    = 0
+vfb    = 0.55          u0      = 0.0095     ua      = 1.6e-009  ub      = 8e-018
+uc     = 4.6e-013      vsat    = 90000      a0      = 1.2       ags     = 1e-020
+a1     = 0             a2      = 1          b0      = -1e-020   b1      = 0
+keta   = -0.047        dwg     = 0          dwb     = 0         pclm    = 0.55
+pdiblc1 = 0.03         pdiblc2 = 0.0055     pdiblcb = 3.4e-008  drout   = 0.56
+pvag   = 1e-020        delta   = 0.014      pscbe1  = 8.14e+08  pscbe2  = 9.58e-07
+fprout = 0.2           pdits   = 0.2        pditsd  = 0.23      pditsl  = 2.3e+006
+rsh    = 3             rdsw    = 250        rsw     = 160       rdw     = 160
+rdswmin = 0            rdwmin  = 0          rswmin  = 0         prwg    = 3.22e-008
+prwb   = 6.8e-011      wr      = 1          alpha0  = 0.074     alpha1  = 0.005
+beta0  = 30            agidl   = 0.0002     bgidl   = 2.1e+009  cgidl   = 0.0002
+egidl  = 0.8

+aigbacc = 0.012        bigbacc = 0.0028     cigbacc = 0.002
+nigbacc = 1            aigbinv = 0.014      bigbinv = 0.004     cigbinv = 0.004
+eigbinv = 1.1          nigbinv = 3          aigc    = 0.69      bigc    = 0.0012
+cigc   = 0.0008        aigsd   = 0.0087     bigsd   = 0.0012    cigsd   = 0.0008
+nigc   = 1             poxedge = 1          pigcd   = 1         ntox    = 1

+xrcrg1 = 12            xrcrg2  = 5
+cgso   = 7.43e-010     cgdo    = 7.43e-010  cgbo    = 2.56e-011 cgdl    = 1e-014
+cgsl   = 1e-014        ckappas = 0.5        ckappad = 0.5       acde    = 1
+moin   = 15            noff    = 0.9        voffcv  = 0.02

+kt1    = -0.19         kt1l    = 0          kt2     = -0.052    ute     = -1.5
+ua1    = -1e-009       ub1     = 2e-018     uc1     = 0         prt     = 0
+at     = 33000

+fnoimod = 1            tnoimod = 0

+jss    = 0.0001        jsws    = 1e-011     jswgs   = 1e-010    njs     = 1
```

+ijthsfwd= 0.01	ijthsrev= 0.001	bvs = 10	xjbvs = 1
+jsd = 0.0001	jswd = 1e-011	jswgd = 1e-010	njd = 1
+ijthdfwd= 0.01	ijthdrev= 0.001	bvd = 10	xjbvd = 1
+pbs = 1	cjs = 0.0005	mjs = 0.5	pbsws = 1
+cjsws = 5e-010	mjsws = 0.33	pbswgs = 1	cjswgs = 3e-010
+mjswgs = 0.33	pbd = 1	cjd = 0.0005	mjd = 0.5
+pbswd = 1	cjswd = 5e-010	mjswd = 0.33	pbswgd = 1
+cjswgd = 5e-010	mjswgd = 0.33	tpb = 0.005	tcj = 0.001
+tpbsw = 0.005	tcjsw = 0.001	tpbswg = 0.005	tcjswg = 0.001
+xtis = 3	xtid = 3		
+dmcg = 5e-006	dmci = 5e-006	dmdg = 5e-006	dmcgt = 6e-007
+dwj = 4.5e-008	xgw = 3e-007	xgl = 4e-008	
+rshg = 0.4	gbmin = 1e-010	rbpb = 5	rbpd = 15
+rbps = 15	rbdb = 15	rbsb = 15	ngcon = 1

Simulation Results

Figure 6.19 shows 10/1 PMOS and NMOS device simulation results using the topologies seen in Figs. 6.11–6.13. The actual device sizes are 500 nm (width) by 50 nm (length). From the information in this figure and knowing *VDD* is 1 V, we can estimate the on currents for the MOSFETs. For the NMOS device

$$I_{on,n} \approx 300\ \mu A/(W \cdot scale) = 600\ \mu A/\mu m \tag{6.60}$$

For the PMOS device

$$I_{on,p} \approx 150\ \mu A/(W \cdot scale) = 300\ \mu A/\mu m \tag{6.61}$$

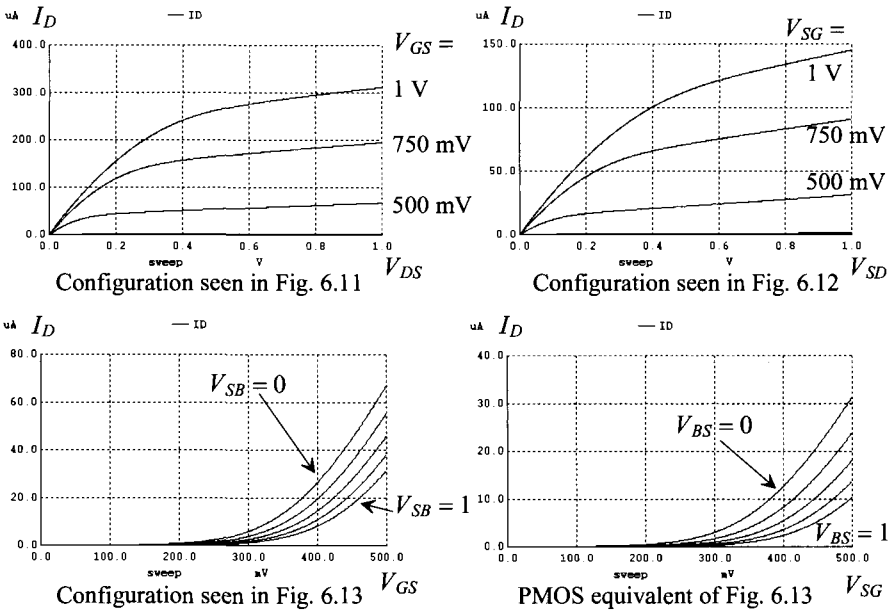

Figure 6.19 Current-voltage characteristics for 50 nm MOSFETs.

The threshold voltages can be estimated as 280 mV. (See Fig. 9.27 and the associated discussion for more information on determining the threshold voltages.) Note that it **doesn't make sense** to try to define a transconductance parameter, *KP*, for a short-channel process (the MOSFETs don't follow the square-law equations, Eqs. [6.33] and [6.37].) Instead we use I_{on}, I_{off}, t_{ox}, *W*, *scale*, $V_{THN,P}$, *VDD*, and plots of measured data. Table 6.4 shows some of the device characteristics for the short-channel CMOS process used in this book. Note that we **don't confuse** I_{on} (Fig. 6.16) with $I_{D,sat}$ (the current at the border between triode and saturation, Figs. 6.11 or 9.4 and Eq. [6.43]).

Table 6.4 Summary of device characteristics for the short-channel CMOS process.

Short-channel MOSFET parameters used in this book. The *VDD* = 1 V and the scale factor is **50 nm** (*scale* = 50e–9)			
Parameter	**NMOS**	**PMOS**	**Comments**
V_{THN} and V_{THP}	280 mV	280 mV	Typical
t_{ox}	14 Å	14 Å	See also Table 5.1
$C'_{ox} = \varepsilon_{ox}/t_{ox}$	25 *f*F/μm^2	25 *f*F/μm^2	$C_{ox} = C'_{ox}WL \cdot (scale)^2$
λ_n and λ_p	0.6 V^{-1}	0.3 V^{-1}	At $L = 2$
$I_{on.n}$ and $I_{on.p}$	600 μA/μm	300 μA/μm	On current
$I_{off.n}$ and $I_{off.p}$	7.1 nA/μm	10 nA/μm	Off current, see Fig. 14.2

ADDITIONAL READING

[1] Y. Taur and T. H. Ning, *Fundamentals of Modern VLSI Devices*, Second Edition, Cambridge University Press, 2010. ISBN 978-0521832946

[2] J.R. Brews, MOSFET Hand Analysis Using BSIM, *IEEE Circuits and Devices Magazine*, vol. 22, No. 1, pp. 28-36, January/February 2006.

[3] R. S. Muller, T. I. Kamins, and M. Chan, *Device Electronics for Integrated Circuits*, John Wiley and Sons Publishers, 2002. ISBN 0-471-59398-2

[4] R. C. Jaeger, *Introduction to Microelectronic Fabrication*, 2nd ed, volume 5 of the Modular Series on Solid State Devices, Prentice-Hall Publishers, 2002. ISBN 0-20-144494-1

[5] W. Liu, *MOSFET Models for Spice Simulation, Including BSIM3v3 and BSIM4*, John Wiley and Sons Publishers, 2001. ISBN 0-471-39697-4

[6] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, 2nd ed., Oxford University Press, 1999. ISBN 978-0195170146.

[7] M. Bohr, "MOS Transistors: Scaling and Performance Trends," *Semiconductor International*, pp. 75–79, June 1995.

[8] G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE, Second Edition*, McGraw-Hill, 1993. Excellent reference for SPICE modeling.

[9] D. A. Neamen, *Semiconductor Physics and Devices-Basic Principles,* Richard D. Irwin, 1992. ISBN 0-256-08405-X.

[10] D. K. Schroder, *Modular Series on Solid State Devices-Advanced MOS Devices,* Addison-Wesley, 1987.

[11] S. M. Sze, *Physics of Semiconductor Devices,* 2nd ed., John-Wiley and Sons, 1981. ISBN 0-471-05661-8.

PROBLEMS

6.1 Plot the magnitude and phase of v_{out} (AC) in the following circuit, Fig. 6.20. Assume that the MOSFET was fabricated using the 50 nm process (see Table 5.1) and is operating in strong inversion. Verify your answer with SPICE.

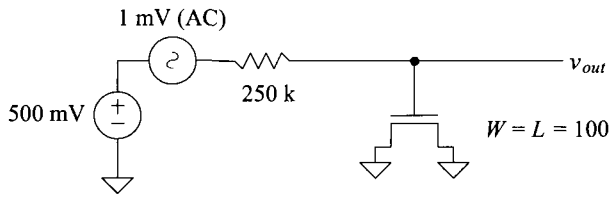

Figure 6.20 Circuit used in Problem 6.1.

6.2 If a MOSFET is used as a capacitor in the strong inversion region where the gate is one electrode and the source/drain is the other electrode, does the gate overlap of the source/drain change the capacitance? Why or not? What is the capacitance?

6.3 Repeat Problem 6.2 when the MOSFET is operating in the accumulation region. Keep in mind that the question is not asking for the capacitance from gate to substrate.

6.4 If the oxide thickness of a MOSFET is 40 Å, what is C'_{ox}?

6.5 Repeat Ex. 6.2 when $V_{SB} = 1$ V.

6.6 Repeat Ex. 6.3 for a p-channel device with a well doping concentration of 10^{16} atoms/cm^3.

6.7 What is the electrostatic potential of the oxide-semiconductor interface when $V_{GS} = V_{THN0}$?

6.8 Repeat Ex. 6.5 to get a threshold voltage of 0.8 V.

6.9 What happens to the threshold voltage in Problem 6.8 if sodium contamination of 100×10^9 sodium ions/cm^2 is present at the oxide-semiconductor interface?

6.10 How much charge (enhanced electrons) is available under the gate for conducting a drain current at the drain-channel interface when $V_{DS} = V_{GS} - V_{THN}$? Why? Assume that the MOSFET is operating in strong inversion, $V_{GS} > V_{THN}$.

6.11 Show the details of the derivation for Eq. (6.33) for the PMOS device.

6.12 Using Eq. (6.35), estimate the small-signal channel resistance (the change in the drain current with changes in the drain-source voltage) of a MOSFET operating in the triode region (the resistance between the drain and source).

6.13 Show, using Eqs. (6.33) and (6.37), that the parallel connection of MOSFETs shown in Fig. 5.18 behave as a single MOSFET with a width equal to the sum of each individual MOSFET width.

6.14 Show that the bottom MOSFET, Fig. 6.21, in a series connection of two MOSFETs cannot operate in the saturation region. Neglect the body effect. *Hint*: Show that M1 is always either in cutoff ($V_{GS1} < V_{THN}$) or triode ($V_{DS1} < V_{GS1} - V_{THN}$).

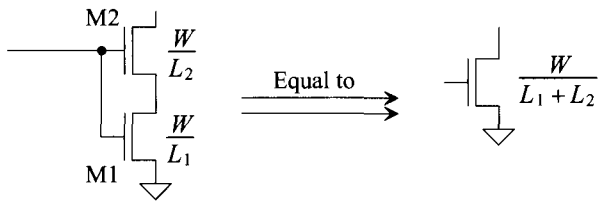

Figure 6.21 MOSFETs operating in series.

6.15 Show that the series connection of MOSFETs shown in Fig. 6.21 behaves as a single MOSFET with twice the length of the individual MOSFETs. Again, neglect the body effect.