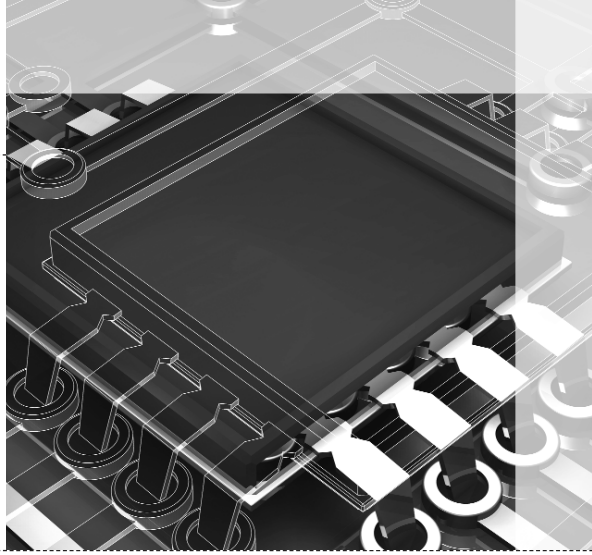
Chapter 3

第 3 章　CMOS 器件模型

CMOS Device Modeling



Before one can design a circuit to be integrated in CMOS technology, one must first have a model describing the behavior of all the components available for use in the design. A model can take the form of mathematical equations, circuit representations, or tables. Most of the modeling used in this text will focus on the active and passive devices discussed in the previous chapter as opposed to higher-level modeling such as macromodeling or behavioral modeling.

It should be stressed at the outset that a model is just that and no more—it is not the real thing! In an ideal world, we would have a model that accurately describes the behavior of a device under all possible conditions. Realistically, we are happy to have a model that predicts simulated performance to within a few percent of measured performance. There is no clear agreement as to which model comes closest to meeting this "ideal" model [1]. This lack of agreement is illustrated by the fact that, at this writing, HSPICE [2] offers the user 43 different MOS transistor models from which to choose!

This text will concentrate on only three of these models. The simplest model, which is appropriate for hand calculations, was described in Section 2.3 and will be further developed here to include capacitance, noise, and ohmic resistance. In SPICE terminology, this simple model is called the LEVEL 1 model. Next, a small-signal model is derived from the LEVEL 1 large-signal model and is presented in Section 3.3.

A far more complex model, the SPICE LEVEL 3 model, is presented in Section 3.4. This model includes many effects that are more evident in modern short-channel technologies as well as subthreshold conduction. It is adequate for device geometries down to about 0.8 μm. Finally, the BSIM3v3 [3] model is presented. This model is the closest to becoming a standard for computer simulation.

Notation

SPICE was originally implemented in FORTRAN where all input was required to be uppercase ASCII characters. Lowercase, Greek, and super-subscripting were not allowed. Modern SPICE implementations generally accept (but do not distinguish between) uppercase and lowercase but the tradition of using uppercase ASCII still lives on. This is particularly evident in the device model parameters. Since Greek characters are not available, these were simply spelled out, for example, γ entered as GAMMA. Superscripts and subscripts were simply not used.

It is inconvenient to adopt the SPICE naming convention throughout the book because equations would appear unruly and would not be familiar to what is commonly seen in the literature. On the other hand, it is necessary to provide the correct notation where application to SPICE is intended. To address this dilemma, we have decided to use SPICE uppercase (nonitalic) notation for all model parameters except those applied to the simple model (SPICE LEVEL 1).

3.1 简单的 MOS 管大信号模型（SPICE LEVEL 1）

Simple MOS Large-Signal Model (SPICE LEVEL 1)

All large-signal models will be developed for the n-channel MOS device with the positive polarities of voltages and currents shown in Fig. 3.1-1(a). The same models can be used for the p-channel MOS device if all voltages and currents are multiplied by -1 and the absolute value of the p-channel threshold is used. This is equivalent to using the voltages and currents defined by Fig. 3.1-1(b), which are all positive quantities. As mentioned in Chapter 1, lowercase variables with capital subscripts will be used for the variables of large-signal models and lowercase variables with lowercase subscripts will be used for the variables of small-signal models. When the voltage or current is a model parameter, such as threshold voltage, it will be designated by an uppercase variable and an uppercase subscript.

When the length and width of the MOS device is greater than about 10 μm, the substrate doping is low, and when a simple model is desired, the model suggested by Sah [4] and used in SPICE by Shichman and Hodges [5] is very appropriate. This model was developed in Eq. (2.3-27) and is given below.

$$i_D = \frac{\mu_0 C_{\text{ox}} W}{L}\left[(v_{GS} - V_T) - \left(\frac{v_{DS}}{2}\right)\right]v_{DS} \tag{3.1-1}$$

The terminal voltages and currents have been defined in the previous chapter. The various parameters of Eq. (3.1-1) are defined as

μ_0 = surface mobility of the channel for the n-channel or p-channel device ($\text{cm}^2/\text{V-s}$)

$C_{\text{ox}} = \dfrac{\varepsilon_{\text{ox}}}{t_{\text{ox}}}$ = capacitance per unit area of the gate oxide (F/cm^2)

W = effective channel width

L = effective channel length

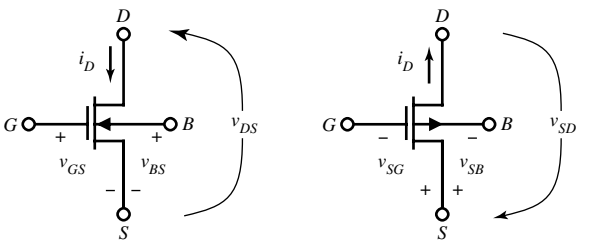


Figure 3.1-1 Positive sign convention for (a) n-channel and (b) p-channel MOS transistor.

The threshold voltage V_T is given by Eq. (2.3-19) for an n-channel transistor:

$$V_T = V_{T0} + \gamma\left(\sqrt{2|\phi_F| + v_{SB}} - \sqrt{2|\phi_F|}\right) \tag{3.1-2}$$

$$V_{T0} = V_T(v_{SB} = 0) = V_{FB} + 2|\phi_F| + \frac{\sqrt{2q\varepsilon_{Si}N_{SUB}2|\phi_F|}}{C_{ox}} \tag{3.1-3}$$

$$\gamma = \text{bulk threshold parameter } (V^{1/2}) = \frac{\sqrt{2\varepsilon_{Si}qN_{SUB}}}{C_{ox}} \tag{3.1-4}$$

$$\phi_F = \text{strong inversion surface potential (V)} = \frac{kT}{q}\ln\left(\frac{N_{SUB}}{n_i}\right) \tag{3.1-5}$$

$$V_{FB} = \text{flatband voltage (V)} = \phi_{MS} - \frac{Q_{ss}}{C_{ox}} \tag{3.1-6}$$

$$\phi_{MS} = \phi_F(\text{substrate}) - \phi_F(\text{gate}) \quad [\text{see Eq. (2.3-17)}] \tag{3.1-7}$$

$$\phi_F(\text{substrate}) = -\frac{kT}{q}\ln\left(\frac{N_{SUB}}{n_i}\right)[\text{n-channel with p-substrate}] \tag{3.1-8}$$

$$\phi_F(\text{gate}) = -\frac{kT}{q}\ln\left(\frac{N_{GATE}}{n_i}\right)[\text{n-channel with } n^+ \text{ polysilicon gate}] \tag{3.1-9}$$

$$Q_{ss} = \text{oxide-charge} = qN_{ss} \tag{3.1-10}$$

k = Boltzmann's constant

T = temperature (K)

n_i = intrinsic carrier concentration

Table 3.1-1 gives some of the pertinent constants for silicon.

A unique aspect of the MOS device is its dependence on the voltage from the source to bulk as shown by Eq. (3.1-2). This dependence means that the MOS device must be treated as a four-terminal element. It will be shown later how this behavior can influence both the large- and small-signal performance of MOS circuits.

Table 3.1-1 Constants for Silicon

Constant Symbol	Constant Description	Value	Units
V_{G0}	Silicon bandgap (27 °C)	1.205	V
k	Boltzmann's constant	1.381×10^{-23}	J/K
n_i	Intrinsic carrier concentration (27 °C)	1.45×10^{10}	cm^{-3}
ε_0	Permittivity of free space	8.854×10^{-14}	F/cm
ε_{Si}	Permittivity of silicon	11.7 ε_0	F/cm
ε_{ox}	Permittivity of SiO_2	3.9 ε_0	F/cm

Table 3.1-2 Model Parameters for a Typical CMOS Bulk Process Suitable for Hand Calculations Using the Simple Model with Values Based on a 0.8 μm Silicon-Gate Bulk CMOS n-Well Process

Parameter Symbol	Parameter Description	Typical Parameter Value		Units
		n-Channel	p-Channel	
V_{T0}	Threshold voltage ($V_{BS} = 0$)	0.7 ± 0.15	−0.7 ± 0.15	V
K'	Transconductance parameter (in saturation)	110.0 ± 10%	50.0 ± 10%	$\mu A/V^2$
γ	Bulk threshold parameter	0.4	0.57	$V^{1/2}$
λ	Channel length modulation parameter	0.04 ($L = 1$ μm) 0.01 ($L = 2$ μm)	0.05 ($L = 1$ μm) 0.01 ($L = 2$ μm)	V^{-1}
$2\lvert\phi_F\rvert$	Surface potential at strong inversion	0.7	0.8	V

In the realm of circuit design, it is more desirable to express the model equations in terms of electrical rather than physical parameters. For this reason, the drain current is often expressed as

$$i_D = \beta\left[(v_{GS} - V_T) - \frac{v_{DS}}{2}\right]v_{DS} \tag{3.1-11}$$

or

$$i_D = K'\frac{W}{L}\left[(v_{GS} - V_T) - \frac{v_{DS}}{2}\right]v_{DS} \tag{3.1-12}$$

where the transconductance parameter β is given in terms of physical parameters as

$$\beta = K'\frac{W}{L} \cong \mu_0 C_{\text{ox}}\frac{W}{L} \quad (\text{A/V}^2) \tag{3.1-13}$$

When devices are characterized in the nonsaturation region with low gate and drain voltages, the value for K' is approximately equal to $\mu_0 C_{\text{ox}}$ in the simple model. This is not the case when devices are characterized with larger voltages introducing effects such as mobility degradation. For these latter cases, K' is usually smaller. Typical values for the model parameters of Eq. (3.1-12) are given in Table 3.1-2.

There are various regions of operation of the MOS transistor based on the model of Eq. (3.1-1). These regions of operation depend on the value of $v_{GS} - V_T$. If $v_{GS} - V_T$ is zero or negative, then the MOS device is in the cutoff* region and Eq. (3.1-1) becomes

$$i_D = 0, \qquad v_{GS} - V_T \leq 0 \tag{3.1-14}$$

In this region, the channel acts like an open circuit.

*We will learn later that MOS transistors can operate in the subthreshold region where the gate–source voltage is less than the threshold voltage.

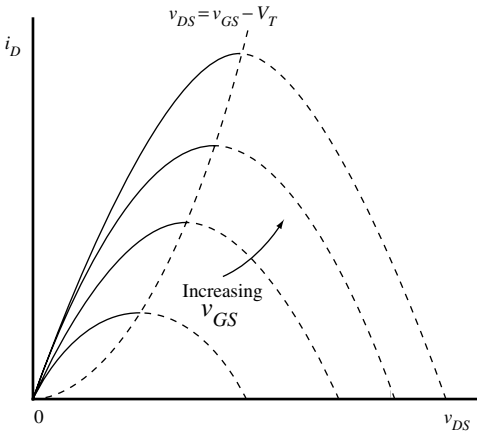


Figure 3.1-2 Graphical illustration of the modified Sah equation.

A plot of Eq. (3.1-1) with $\lambda = 0$ as a function of v_{DS} is shown in Fig. 3.1-2 for various values of $v_{GS} - V_T$. At the maximum of these curves the MOS transistor is said to saturate. The value of v_{DS} at which this occurs is called the saturation voltage and is given as

$$v_{DS}(\text{sat}) = v_{GS} - V_T \tag{3.1-15}$$

Thus, v_{DS}(sat) defines the boundary between the remaining two regions of operation. If v_{DS} is less than v_{DS}(sat), then the MOS transistor is in the nonsaturated region and Eq. (3.1-1) becomes

$$i_D = K' \frac{W}{L}\left[(v_{GS} - V_T) - \frac{v_{DS}}{2}\right]v_{DS}, \qquad 0 < v_{DS} \leq (v_{GS} - V_T) \tag{3.1-16}$$

In Fig. 3.1-2, the nonsaturated region lies between the vertical axis ($v_{DS} = 0$) and the $v_{DS} = v_{GS} - V_T$ curve.

The third region occurs when v_{DS} is greater than v_{DS}(sat) or $v_{GS} - V_T$. At this point the current i_D becomes independent of v_{DS}. Therefore, v_{DS} in Eq. (3.1-1) is replaced by v_{DS}(sat) of Eq. (3.1-15) to get

$$i_D = K' \frac{W}{2L}(v_{GS} - V_T)^2, \qquad 0 < (v_{GS} - V_T) \leq v_{DS} \tag{3.1-17}$$

Equation (3.1-17) indicates that drain current remains constant once v_{DS} is greater than $v_{GS} - V_T$. In reality, this is not true. As drain voltage increases, the channel length is reduced, resulting in increased current. This phenomenon is called *channel length modulation* and is accounted for in the saturation model with the addition of the factor $(1 + \lambda v_{DS})$, where v_{DS} is the actual drain–source voltage and not v_{DS}(sat). The saturation region model modified to include channel length modulation is given in Eq. (3.1-18):

$$i_D = K' \frac{W}{2L}(v_{GS} - V_T)^2(1 + \lambda v_{DS}), \qquad 0 < (v_{GS} - V_T) \leq v_{DS} \tag{3.1-18}$$

Up to this point in time, it has been convention to refer to the various regions of the MOS transistor as cutoff when $(v_{GS} - V_T) \leq 0$; as active, ohmic, triode, or nonsaturation when $0 \leq v_{DS} \leq (v_{GS} - V_T)$; and as saturation when $(v_{GS} - V_T) \leq v_{DS}$. While we will continue to use this convention throughout this text, the reader needs to be aware that a different convention

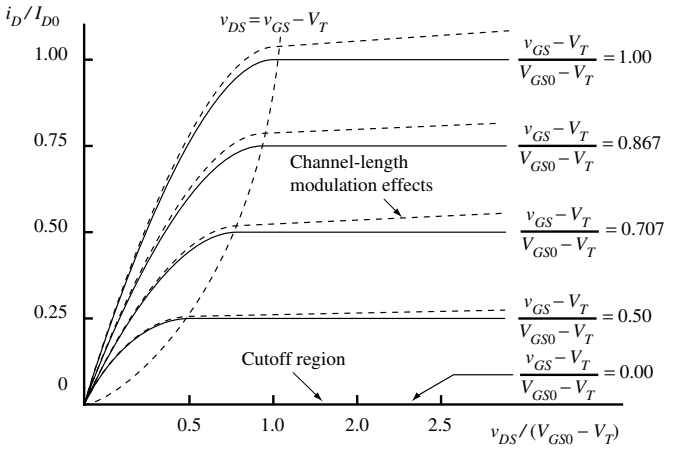


Figure 3.1-3 Output characteristics of the MOS device.

was introduced recently in the fourth edition of Gray, Hurst, Lewis, and Meyer [6] that defines cutoff when $(v_{GS} - V_T) \leq 0$, triode or ohmic when $0 \leq v_{DS} \leq (v_{GS} - V_T)$, and active when $(v_{GS} - V_T) \leq v_{DS}$. The important difference is that "active" has replaced "saturation" when describing the region when $(v_{GS} - V_T) \leq v_{DS}$. This new convention makes the names of the regions of operation for an MOS transistor similar to the bipolar transistor.

The output characteristics of the MOS transistor can be developed from Eqs. (3.1-14), (3.1-16), and (3.1-18). Figure 3.1-3 shows these characteristics plotted on a normalized basis. These curves have been normalized to the upper curve, where V_{GS0} is defined as the value of v_{GS} that causes a drain current of I_{D0} in the saturation region. The entire characteristic is developed by extending the solid curves of Fig. 3.1-2 horizontally to the right from the maximum points. The solid curves of Fig. 3.1-3 correspond to $\lambda = 0$. If $\lambda \neq 0$, then the curves are the dashed lines.

Another important characteristic of the MOS transistor can be obtained by plotting i_D versus v_{GS} using Eq. (3.1-18). Figure 3.1-4 shows this result. This characteristic of the MOS

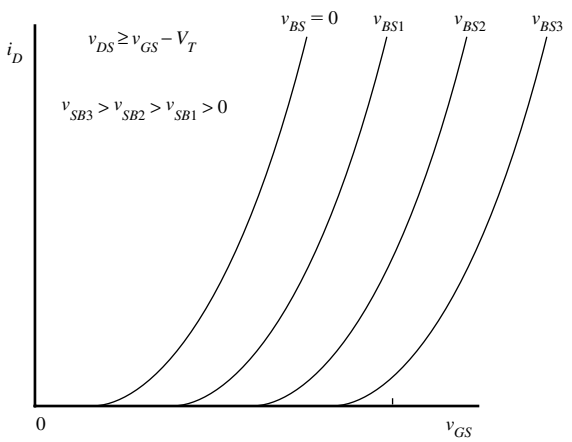


Figure 3.1-4 Transconductance characteristic of the MOS transistor as a function of the source–bulk voltage, v_{SB}.

transistor is called the transconductance characteristic. We note that the transconductance characteristic in the saturation region can be obtained from Fig. 3.1-3 by drawing a vertical line to the right of the parabolic dashed line and plotting values of i_D versus v_{GS}. Figure 3.1-4 is also useful for illustrating the effect of the source–bulk voltage, v_{SB}. As the value of v_{SB} increases, the value of V_T increases for the enhancement, n-channel devices (for a p-channel device, $|V_T|$ increases as v_{BS} increases). V_T also increases positively for the n-channel depletion device, but since V_T is negative, the value of V_T approaches zero from the negative side. If v_{SB} is large enough, V_T will actually become positive and the depletion device becomes an enhancement device.

Since the MOS transistor is a bidirectional device, determining which physical node is the drain and which the source may seem arbitrary. This is not really the case. For an n-channel transistor, the source is always at the lower potential of the two nodes. For the p-channel transistor, the source is always at the higher potential. It is obvious that the drain and source designations are not constrained to a given node of a transistor but can switch back and forth depending on the terminal voltages applied to the transistor.

A circuit version of the large-signal model of the MOS transistor consists of a current source connected between the drain and source terminals, that depends on the drain, source, gate, and bulk terminal voltages defined by the simple model described in this section. This simple model has five electrical and process parameters that completely define it. These parameters are K', V_T, γ, λ, and $2\phi_F$. The subscript n or p will be used when the parameter refers to an n-channel or p-channel device, respectively. They constitute the LEVEL 1 model parameters of SPICE [7]. Typical values for these model parameters are given in Table 3.1-2.

The function of the large-signal model is to solve for the drain current given the terminal voltages of the MOS device. An example will help to illustrate this as well as show how the model is applied to the p-channel device.

Example 3.1-1

Application of the Simple MOS Large-Signal Model

Assume that the transistors in Fig. 3.1-1 have a W/L ratio of 5 μm/1 μm and that the large-signal model parameters are those given in Table 3.1-2. If the drain, gate, source, and bulk voltages of the n-channel transistor are 3 V, 2 V, 0 V, and 0 V, respectively, find the drain current. Repeat for the p-channel transistor if the drain, gate, source, and bulk voltages are −3 V, −2 V, 0 V, and 0 V, respectively.

SOLUTION

We must first determine in which region the transistor is operating. Equation (3.1-15) gives $v_{DS}(\text{sat})$ as 2 V − 0.7 V = 1.3 V. Since v_{DS} is 3 V, the n-channel transistor is in the saturation region. Using Eq. (3.1-18) and the values from Table 3.1-2, we have

$$i_D = \frac{K'_N W}{2L}(v_{GS} - V_{TN})^2(1 + \lambda_N v_{DS})$$

$$= \frac{110 \times 10^{-6}(5\ \mu\text{m})}{2(1\ \mu\text{m})}(2 - 0.7)^2(1 + 0.04 \times 3) = 520\ \mu\text{A}$$

Evaluation of Eq. (3.1-15) for the p-channel transistor is given as

$$v_{SD}(\text{sat}) = v_{SG} - |V_{TP}| = 2\ \text{V} - 0.7\ \text{V} = 1.3\ \text{V}$$

Since v_{SD} is 3 V, the p-channel transistor is also in the saturation region, and Eq. (3.1-17) is applicable. The drain current of Fig. 3.1-1(b) can be found using the values from Table 3.1-2 as

$$\begin{aligned} i_D &= \frac{K_P' W}{2L}(v_{SG} - |V_{TP}|)^2 (1 + \lambda_P v_{SD}) \\ &= \frac{50 \times 10^{-6}(5\ \mu\text{m})}{2(1\ \mu\text{m})}(2 - 0.7)^2(1 + 0.05 \times 3) = 243\ \mu\text{A} \end{aligned}$$

It is often useful to describe v_{GS} in terms of i_D in saturation as shown below:

$$v_{GS} = V_T + \sqrt{2i_D/\beta} \tag{3.1-19}$$

This expression illustrates that there are two components to v_{GS}—an amount to invert the channel plus an additional amount to support the desired drain current. This second component is often referred to in the literature as V_{ON} or overdrive. Thus, V_{ON} can be defined as

$$V_{ON} = \sqrt{2i_D/\beta} \tag{3.1-20}$$

The term V_{ON} should be recognized as the term for saturation voltage V_{DS} (sat). They can be used interchangeably.

As the channel length of the technology decreases, the carriers in the channel will experience velocity saturation and the velocity is no longer proportional to the electric field [8]. An approximation for the carrier velocity is given as

$$v_d \approx \frac{\mu E}{1 + \dfrac{E}{E_c}} \tag{3.1-21}$$

where E_c is the critical electrical field at which velocity saturation of the carrier occurs. Eq. (2.3-25) can be expressed as

$$i_D = WQ_I(y)\mu_n \frac{dv(y)}{dy} = WQ_I(y)\mu_n E(y) = WQ_I(y)\mu_n v_d \tag{3.1-22}$$

If we substitute Eq. (3.1-21) into Eq. (3.1-22) and replace the electric field E by dv/dy, we obtain

$$i_D\left(1 + \frac{1}{E_c}\frac{dv}{dy}\right)dy = WQ_I(y)\mu_n dv(y) \tag{3.1-23}$$

which is the form of Eq. (2.3-25) that includes velocity saturation for electrons. Integrating the distance and voltage along the channel from 0 to L and 0 to v_{DS} as done in Section 2.4 gives

$$i_D = \frac{\mu_n C_{ox}}{2\left(1 + \dfrac{1}{E_c}\dfrac{v_{DS}}{L}\right)}\frac{W}{L}[2(v_{GS} - V_T)v_{DS} - {v_{DS}}^2] \tag{3.1-24}$$

which is equivalent to Eq. (2.3-27) but incorporates the velocity saturation effect. It is convenient to define a constant θ as

$$\theta = \frac{1}{E_c L}\,(V^{-1}) \tag{3.1-25}$$

Using this definition, Eq. (3.1-24) can rewritten as

$$i_D = \frac{\mu_n C_{ox}}{2(1 + \theta v_{DS})}\frac{W}{L}[2(v_{GS} - V_T)v_{DS} - v_{DS}^2] \tag{3.1-26}$$

This expression is also equivalent to Eq. (3.1-1) but includes the influence of velocity saturation.

The next step, which follows the previous development in this section, is to solve for v_{DS}(sat) as was done in Eq. (3.1-15). Unfortunately, this is not very straightforward because of v_{DS} in the denominator of Eq. (3.1-26). An approximation for v_{DS}(sat) is given as [6]

$$\begin{aligned} v'_{DS}(\text{sat}) &\approx (V_{GS} - V_T)\left(1 - \frac{\theta(V_{GS} - V_T)}{2} + \cdots\right) \\ &= v_{DS}(\text{sat})\left(1 - \frac{\theta(V_{GS} - V_T)}{2} + \cdots\right) \end{aligned} \tag{3.1-27}$$

The large-signal model for the saturation region that includes velocity saturation can be developed by assuming that $0.5\theta\ (v_{GS} - V_T) < 1$ so that Eq. (3.1-27) reduces to the definition of v_{DS}(sat). With this assumption, Eq. (3.1-26) becomes

$$i_D = \frac{\mu_n C_{ox}}{2[1 + \theta(v_{GS} - V_T)]}\frac{W}{L}[(v_{GS} - V_T)^2], v_{DS} \geq (V_{GS} - V_T)\left(1 - \frac{\theta(V_{GS} - V_T)}{2} + \cdots\right) \tag{3.1-28}$$

which extends the large-signal model of Eq. (3.1-17) to include the effect of velocity saturation. Figure 3.1-5 shows the influence of the θ parameter on the transconductance characteristics of an n-channel MOSFET with $K' = 110\ \mu A/V^2$ and $W/L = 1$. We see that the effect of velocity saturation (when $\theta \neq 0$) causes the slope of the transconductance curve to decrease. The transconductance characteristic moves from square law to linear as θ increases from zero.

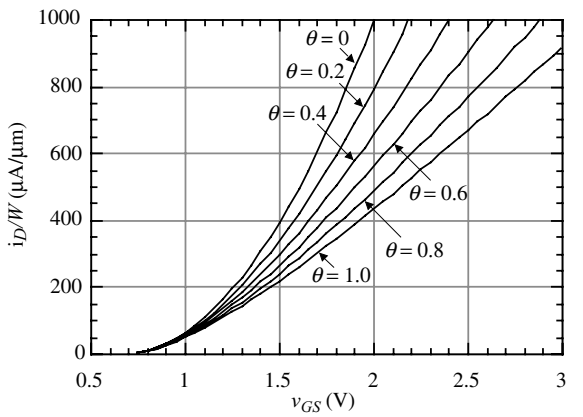


Figure 3.1-5 Influence of velocity saturation on the transconductance characteristics.

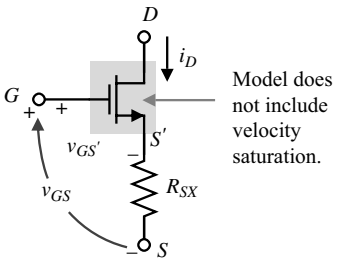


Figure 3.1-6 n-Channel MOSFET with a degeneration resistor R_{SX} to model velocity saturation.

A simple way of modeling velocity saturation is found by considering degeneration of the MOSFET [6]. Degeneration is the insertion of a resistor in series with the source as illustrated in Fig. 3.1-6. In terms of this figure, we can write

$$i_D = \frac{K'W}{2L}(v'_{GS} - V_T)^2 \tag{3.1-29}$$

where

$$v'_{GS} = v_{GS} - i_D R_{XS} \tag{3.1-30}$$

Solving for i_D by substituting Eq. (3.1-30) into Eq. (3.1-29) results in

$$i_D = \frac{K'}{2\left[1 + K'\dfrac{W}{L}R_{SX}(v_{GS} - V_T)\right]}\frac{W}{L}(v_{GS} - V_T)^2 \tag{3.1-31}$$

Comparing Eq. (3.1-31) with Eq. (3.1-28) gives the result that

$$R_{SX} = \frac{\theta L}{K'W} = \frac{1}{E_c K'W} \tag{3.1-32}$$

Therefore, given the critical field, E_c, one can calculate a value of R_{SX} that can be inserted in series with the source of the MOSFET to model velocity saturation. Therefore, given E_c, K', and W, we can calculate a resistor to be placed in series with the source of the MOS transistor to model velocity saturation using the MOS transistor model that does not include velocity saturation.

Example 3.1-2

Application of the Simple MOS Large-Signal Model for Velocity Saturation

Assume that the transistor in Fig. 3.1-6 has a W/L ratio of 1 μm/0.18 μm and that the large-signal model parameters are those given in Table 3.1-2. Assume also that this transistor experiences velocity saturation where E_c is 1.5×10^6 V/m. Find the value of R_{SX}.

SOLUTION

The value of R_{SX} can be calculated as

$$R_{SX} = \frac{1}{E_c K'W} = \frac{1}{1.5 \times 10^6 \cdot 110 \times 10^{-6} \cdot 1 \times 10^{-6}} = 6.06\ \text{k}\Omega$$

3.2　其他 MOS 管大信号模型的参数

Other MOS Large-Signal Model Parameters

The large-signal model also includes several other characteristics such as the source/drain bulk junctions, source/drain ohmic resistances, various capacitors, noise, and temperature dependence. The complete version of the large-signal model is given in Fig. 3.2-1.

The diodes of Fig. 3.2-1 represent the pn junctions between the source and substrate and the drain and substrate. For proper transistor operation, these diodes must always be reverse biased. Their purpose in the dc model is primarily to model leakage currents. These currents are expressed as

$$i_{BD} = I_s\left[\exp\left(\frac{qv_{BD}}{kT}\right) - 1\right] \tag{3.2-1}$$

and

$$i_{BS} = I_s\left[\exp\left(\frac{qv_{BS}}{kT}\right) - 1\right] \tag{3.2-2}$$

where I_s is the reverse saturation current of a pn junction, q is the charge of an electron, k is Boltzmann's constant, and T is temperature in kelvin units.

The resistors r_D and r_S represent the ohmic resistance of the drain and source, respectively. Typically, these resistors may be 50–100 Ω* and can often be ignored at low drain currents.

The capacitors of Fig. 3.2-1 can be separated into three types. The first type includes capacitors C_{BD} and C_{BS}, which are associated with the backbiased depletion region between

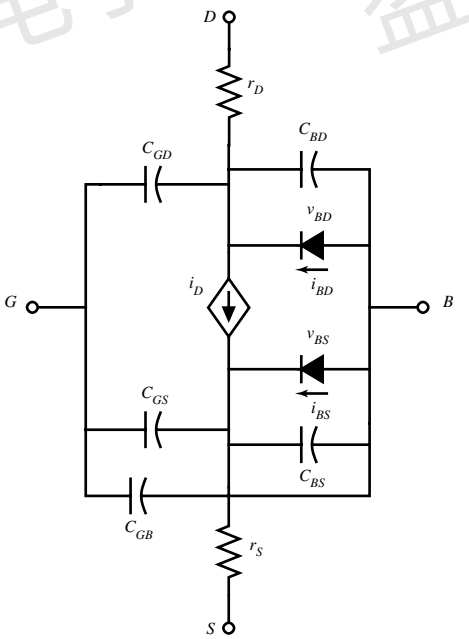


Figure 3.2-1 Complete large-signal model for the MOS transistor.

*For silicide process, these resistances will be much less—on the order of 5–10 Ω.

the drain and substrate and the source and substrate. The second type includes capacitors C_{GD}, C_{GS}, and C_{GB}, which are all common to the gate and are dependent on the operating condition of the transistor. The third type includes parasitic capacitors, which are independent of the operating conditions.

The depletion capacitors are a function of the voltage across the pn junction. The expression of this junction-depletion capacitance is divided into two regions to account for the high injection effects. The first is given as

$$C_{BX} = (\text{CJ})\,(\text{AX})\left[1 - \frac{v_{BX}}{\text{PB}}\right]^{-\text{MJ}}, \qquad v_{BX} \leq (\text{FC})(\text{PB}) \tag{3.2-3}$$

where

$X = D$ for C_{BD} or $X = S$ for C_{BS}

AX = area of the source (X = S) or drain (X = D)

CJ = zero-bias ($v_{BX} = 0$) junction capacitance (per unit area)

$$\text{CJ} \cong \sqrt{\frac{q\varepsilon_{\text{Si}}N_{\text{SUB}}}{2\text{PB}}}$$

PB = bulk junction potential [same as ϕ_0 given in Eq. (2.2-6)]

FC = forward-bias nonideal junction-capacitance coefficient ($\cong$0.5)

MJ = bulk junction grading coefficient ($\frac{1}{2}$ for step junctions and $\frac{1}{3}$ for graded junctions)

The second region is given as

$$C_{BX} = \frac{(\text{CJ})(\text{AX})}{(1 - \text{FC})^{1+\text{MJ}}}\left[1 - (1 + \text{MJ})\text{FC} + \text{MJ}\,\frac{v_{BX}}{\text{PB}}\right], \qquad v_{BX} > (\text{FC})(\text{PB}) \tag{3.2-4}$$

Figure 3.2-2 illustrates how the junction-depletion capacitances of Eqs. (3.2-3) and (3.2-4) are combined to model the large-signal capacitances C_{BD} and C_{BS}. It is seen that Eq. (3.2-4) prevents C_{BX} from approaching infinity as v_{BX} approaches PB.

A closer examination of the depletion capacitors in Fig. 3.2-3 shows that this capacitor is like a tub. It has a bottom with an area equal to the area of the drain or source. However, there are the sides that are also part of the depletion region. This area is called the sidewall. AX in Eqs. (3.2-3) and (3.2-4) should include both the bottom and sidewall assuming the zero-bias capacitances of the two regions are similar. To more closely model the depletion capacitance, it is separated into the bottom and sidewall components, given as follows:

$$C_{BX} = \frac{(\text{CJ})(\text{AX})}{\left[1 - \left(\frac{v_{BX}}{\text{PB}}\right)\right]^{\text{MJ}}} + \frac{(\text{CJSW})(\text{PX})}{\left[1 - \left(\frac{v_{BX}}{\text{PB}}\right)\right]^{\text{MJSW}}}, \qquad v_{BX} \leq (\text{FC})(\text{PB}) \tag{3.2-5}$$

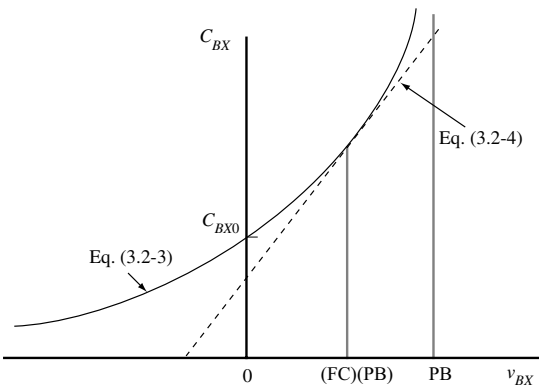


Figure 3.2-2 Example of the method of modeling the voltage dependence of the bulk junction capacitances.

and

$$C_{BX} = \frac{(\text{CJ})(\text{AX})}{(1 - \text{FC})^{1+\text{MJ}}}\left[1 - (1 + \text{MJ})\text{FC} + \text{MJ}\frac{v_{BX}}{\text{PB}}\right] + \frac{(\text{CJSW})(\text{PX})}{(1 - \text{FC})^{1+\text{MJSW}}}\left[1 - (1 + \text{MJSW})\text{FC} + \frac{v_{BX}}{\text{PB}}(\text{MJSW})\right],$$
$$v_{BX} \geq (\text{FC})(\text{PB}) \tag{3.2-6}$$

where

AX = area of the source (X = S) or drain (X = D)

PX = perimeter of the source (X = S) or drain (X = D)

CJSW = zero-bias, bulk–source/drain sidewall capacitance

MJSW = bulk–source/drain sidewall grading coefficient

Table 3.2-1 gives the values for CJ, CJSW, MJ, and MJSW for an MOS device that has an oxide thickness of 140 Å resulting in a $C_{ox} = 24.7 \times 10^{-4}$ F/m^2. It can be seen that the

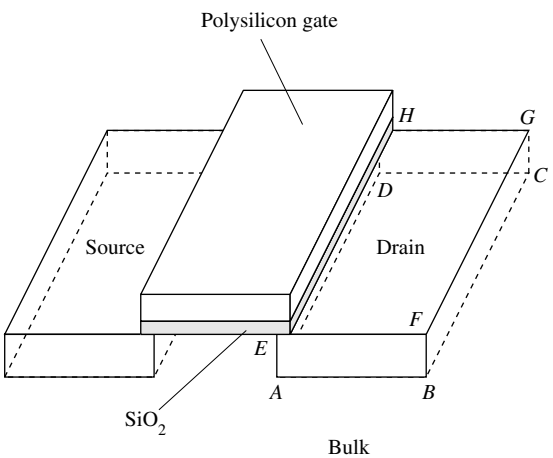


Figure 3.2-3 Illustration showing the bottom (*ABCD*) and sidewall (*ABFE* + *BCGF* + *DCGH* + *ADHE*) components of the bulk junction capacitors.

Table 3.2-1 Capacitance Values and Coefficients for the MOS Model

Type	p-Channel	n-Channel	Units
CGSO	220×10^{-12}	220×10^{-12}	F/m
CGDO	220×10^{-12}	220×10^{-12}	F/m
CGBO	700×10^{-12}	700×10^{-12}	F/m
CJ	560×10^{-6}	770×10^{-6}	F/m^2
CJSW	350×10^{-12}	380×10^{-12}	F/m
MJ	0.5	0.5	
MJSW	0.35	0.38	

Based on an oxide thickness of 140 Å or $C_{ox} = 24.7 \times 10^{-4}$ F/m^2.

depletion capacitors cannot be accurately modeled until the geometry of the device is known, for example, the area and perimeter of the source and drain. However, values can be assumed for the purpose of design. For example, one could consider a typical source or drain to be 1.8 μm by 5 μm. Thus, a value for C_{BX} of 12.1 F and 9.8 F results for n-channel and p-channel devices, respectively, for $V_{BX} = 0$.

The large-signal, charge-storage capacitors of the MOS device consist of the gate-to-source (C_{GS}), gate-to-drain (C_{GD}), and gate-to-bulk (C_{GB}) capacitances. Figure 3.2-4 shows a cross section of the various capacitances that constitute the charge-storage capacitors of the MOS device. C_{BS} and C_{BD} are the bulk-to-source and bulk-to-drain capacitors discussed above. The following discussion represents a heuristic development of a model for the large-signal charge-storage capacitors.

C_1 and C_3 are overlap capacitances and are due to an overlap of two conducting surfaces separated by a dielectric. The overlapping capacitors are shown in more detail in Fig. 3.2-5. The amount of overlap is designated as LD. This overlap is due to the lateral diffusion of the source and drain underneath the polysilicon gate. For example, a 0.8 μm CMOS process might have a lateral diffusion component, LD, of approximately 16 nm. The overlap capacitances can be approximated as

$$C_1 = C_3 \cong (\text{LD})(W_{\text{eff}})C_{\text{ox}} = (\text{CGXO})W_{\text{eff}} \tag{3.2-7}$$

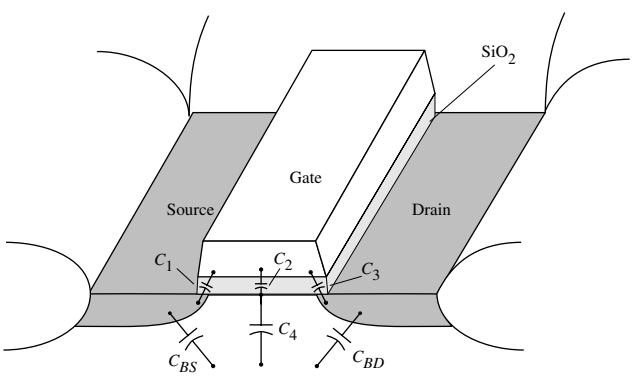


Figure 3.2-4 Large-signal, charge-storage capacitors of the MOS device.

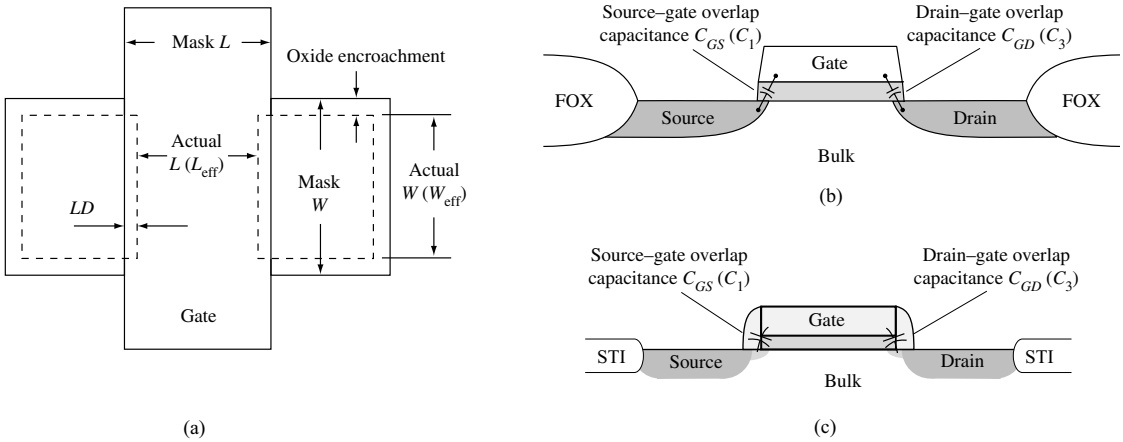


Figure 3.2-5 Overlap capacitances of an MOS transistor. (a) Top view showing the overlap between the source or drain and the gate for LOCOS technology. (b) Side view for LOCOS technology. (c) Side view for STI technology.

where W_{eff} is the effective channel width and CGXO (X = S or D) is the overlap capacitance in F/m for the gate–source or gate–drain overlap. The difference between the mask W and actual W is due to the encroachment of the field oxide under the silicon nitride. Table 3.2-1 gives a value for CGSO and CGDO based on a device with an oxide thickness of 140 Å. A third overlap capacitance that can be significant is the overlap between the gate and the bulk. Figure 3.2-6 shows this overlap capacitor (C_5) in more detail. This is the capacitance that occurs between the gate and bulk at the edges of the channel and is a function of the effective length of the channel, L_{eff}. Table 3.2-1 gives a typical value for CGBO for a device based on an oxide thickness of 140 Å.

If the device illustrated in Fig. 3.2-4 was in the saturated state, the channel would extend almost to the drain and would extend completely to the drain if the MOS device was in the nonsaturated state. C_2 is the gate-to-channel capacitance and is given as

$$C_2 = W_{\text{eff}}(L - 2\text{LD})C_{\text{ox}} = W_{\text{eff}}(L_{\text{eff}})C_{\text{ox}} \tag{3.2-8}$$

The term L_{eff} is the effective channel length resulting from the mask-defined length being reduced by the amount of lateral diffusion (note that up until now, the symbols L and W were

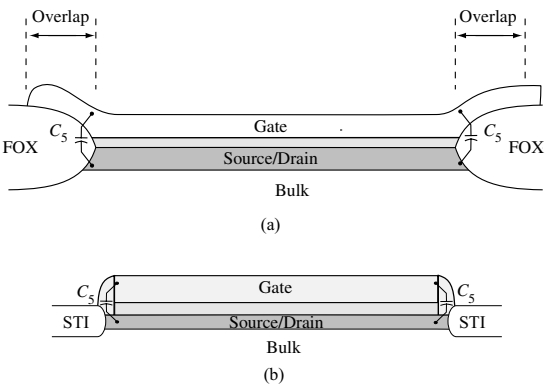


Figure 3.2-6 Gate–bulk overlap capacitances. (a) LOCOS technology. (b) STI technology.

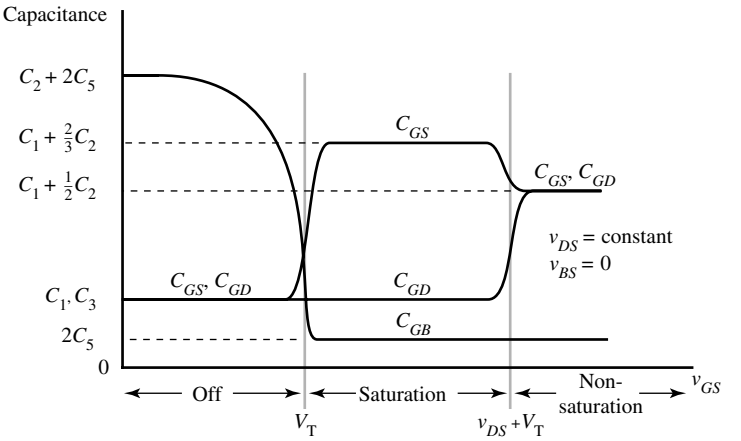


Figure 3.2-7 Voltage dependence of C_{GS}, C_{GD}, and C_{GB} as a function of V_{GS} with V_{DS} constant and $V_{BS} = 0$.

used to refer to "effective" dimensions whereas now these have been changed for added clarification). C_4 is the channel-to-bulk capacitance, which is a depletion capacitance that will vary with voltage like C_{BS} or C_{BD}.

It is of interest to examine C_{GB}, C_{GS}, and C_{GD} as v_{DS} is held constant and v_{GS} is increased from zero. To understand the results, one can imagine following a vertical line on Fig. 3.1-3 at, say, $v_{DS} = 0.5(V_{GS0} - V_T)$, as v_{GS} increases from zero. The MOS device will first be off until v_{GS} reaches V_T. Next, it will be in the saturated region until v_{GS} becomes equal to $v_{DS}(\text{sat}) + V_T$. Finally, the MOS device will be in the nonsaturated region. The approximate variation of C_{GB}, C_{GS}, and C_{GD} under these conditions is shown in Fig. 3.2-7. In cutoff, there is no channel and C_{GB} is approximately equal to $C_2 + 2C_5$. As v_{GS} approaches V_T from the off region, a thin depletion layer is formed, creating a large value of C_4. Since C_4 is in series with C_2, little effect is observed. As v_{GS} increases, this depletion region widens, causing C_4 to decrease and reducing C_{GB}. When $v_{GS} = V_T$, an inversion layer is formed that prevents further decreases of C_4 (and thus C_{GB}).

C_1, C_2, and C_3 constitute C_{GS} and C_{GD}. The problem is how to allocate C_2 to C_{GS} and C_{GD}. The approach used is to assume in saturation that approximately two-thirds of C_2 belongs to C_{GS} and none to C_{GD}. This is, of course, an approximation. However, it has been found to give reasonably good results. Figure 3.2-7 shows how C_{GS} and C_{GD} change values in going from the off to the saturation region. Finally, when v_{GS} is greater than $v_{DS} + V_T$, the MOS device enters the nonsaturated region. In this case, the channel extends from the drain to the source and C_2 is simply divided evenly between C_{GD} and C_{GS} as shown in Fig. 3.2-7.

As a consequence of the above considerations, we shall use the following formulas for the charge-storage capacitances of the MOS device in the indicated regions.

Off

$$C_{GB} = C_2 + 2C_5 = C_{\text{ox}}(W_{\text{eff}})(L_{\text{eff}}) + \text{CGBO}(L_{\text{eff}}) \tag{3.2-9a}$$

$$C_{GS} = C_1 \cong C_{\text{ox}}(\text{LD})(W_{\text{eff}}) = \text{CGSO}(W_{\text{eff}}) \tag{3.2-9b}$$

$$C_{GD} = C_3 \cong C_{\text{ox}}(\text{LD})(W_{\text{eff}}) = \text{CGDO}(W_{\text{eff}}) \tag{3.2-9c}$$

Saturation

$$C_{GB} = 2C_5 = \text{CGBO}\,(L_{\text{eff}}) \tag{3.2-10a}$$

$$\begin{aligned} C_{GS} &= C_1 + \tfrac{2}{3}C_2 = C_{\text{ox}}(\text{LD} + 0.67L_{\text{eff}})(W_{\text{eff}}) \\ &= \text{CGSO}(W_{\text{eff}}) + 0.67C_{\text{ox}}(W_{\text{eff}})(L_{\text{eff}}) \end{aligned} \tag{3.2-10b}$$

$$C_{GD} = C_3 \cong C_{\text{ox}}(\text{LD})(W_{\text{eff}}) = \text{CGDO}(W_{\text{eff}}) \tag{3.2-10c}$$

Nonsaturated

$$C_{GB} = 2C_5 = \text{CGBO}\,(L_{\text{eff}}) \tag{3.2-11a}$$

$$\begin{aligned} C_{GS} &= C_1 + 0.5C_2 = C_{\text{ox}}(\text{LD} + 0.5L_{\text{eff}})(W_{\text{eff}}) \\ &= (\text{CGSO} + 0.5C_{\text{ox}}L_{\text{eff}})W_{\text{eff}} \end{aligned} \tag{3.2-11b}$$

$$\begin{aligned} C_{GD} &= C_3 + 0.5C_2 = C_{\text{ox}}(\text{LD} + 0.5L_{\text{eff}})(W_{\text{eff}}) \\ &= (\text{CGDO} + 0.5C_{\text{ox}}L_{\text{eff}})W_{\text{eff}} \end{aligned} \tag{3.2-11c}$$

Equations that provide a smooth transition between the three regions can be found in the literature [9].

Other capacitor parasitics associated with transistors are due to interconnect to the transistor, for example, polysilicon over field (substrate). This type of capacitance typically constitutes the major portion of C_{GB} in the nonsaturated and saturated regions and thus is very important and should be considered in the design of CMOS circuits.

Another important aspect of modeling the CMOS device is noise. The existence of noise is due to the fact that electrical charge is not continuous but is carried in discrete amounts equal to the charge of an electron. In electronic circuits, noise manifests itself by representing a lower limit below which electrical signals cannot be amplified without significant deterioration in the quality of the signal. Noise can be modeled by a current source connected in parallel with i_D of Fig. 3.2-1. This current source represents two sources of noise, called thermal noise and flicker noise [10,11]. These sources of noise were discussed in Section 2.5. The mean-square current-noise source is defined as

$$i_n^2 = \left[\frac{8kTg_m(1+\eta)}{3} + \frac{(\text{KF})I_D}{f\,C_{\text{ox}}\,L^2}\right]\Delta f \quad (A^2) \tag{3.2-12}$$

where

Δf = a small bandwidth (typically 1 Hz) at a frequency f

η = g_{mbs}/g_m [see Eq. (3.3-8)]

k = Boltzmann's constant

T = temperature (K)

g_m = small-signal transconductance from gate to channel [see Eq. (3.3-6)]

KF = flicker noise coefficient (F-A)

f = frequency (Hz)

KF has a typical value of 10^{-28} F-A. Both sources of noise are process dependent and the values are usually different for enhancement and depletion mode field effect transistors (FETs).

The mean-square current noise can be reflected to the gate of the MOS device by dividing Eq. (3.2-12) by g_m^2, if the source is on ac ground, giving

$$e_n^2 = \frac{i_n^2}{g_m^2} = \left[\frac{8kT(1+\eta)}{3\,g_m} + \frac{\text{KF}}{2f\,C_{\text{ox}}\,WLK'}\right]\Delta f \quad (\text{V}^2) \tag{3.2-13}$$

The equivalent input-mean-square voltage-noise form of Eq. (3.2-13) will be useful for analyzing the noise performance of CMOS circuits in later chapters.

The experimental noise characteristics of n-channel and p-channel devices are shown in Figs. 3.2-8(a) and 3.2-8(b). These devices were fabricated using a submicron, silicon-gate, n-well, CMOS process. The data in Figs. 3.2-8(a) and 3.2-8(b) are typical for MOS devices and show that the 1/*f* noise is the dominant source of noise for frequencies below 100 kHz (at the given bias conditions).* Consequently, in many practical cases, the equivalent input-mean-square voltage noise of Eq. (3.2-13) is simplified to

$$e_{\text{eq}}^2 = \left[\frac{\text{KF}}{2f\,C_{\text{ox}}\,WLK'}\right]\Delta f \quad (\text{V}^2) \tag{3.2-14}$$

or in terms of the input-voltage-noise spectral density we can rewrite Eq. (3.2-14) as

$$e_{\text{eq}}^2 = \frac{e_{\text{eq}}^2}{\Delta f} = \frac{\text{KF}}{2f\,C_{\text{ox}}\,WLK'} = \frac{B}{f\,WL} \quad (\text{V}^2/\text{Hz}) \tag{3.2-15}$$

where B is a constant for an n-channel or a p-channel device of a given process.** The right-hand expression of Eq. (3.2-15) will be important in optimizing the design with respect to noise performance.

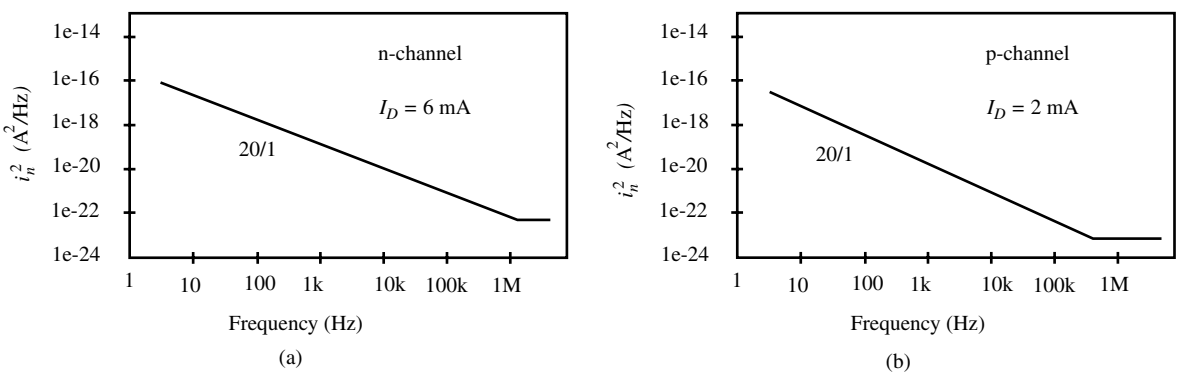


Figure 3.2-8 Drain-current noise for (a) an n-channel and (b) a p-channel MOSFET measured on a silicon-gate submicron process.

*If the bias current is reduced, the thermal noise floor increases, thus moving the 1/*f* noise corner to a lower frequency. Therefore, the 1/*f* noise corner is a function of the thermal noise floor.

**Since the same symbol is used for voltage (current) noise and voltage (current) spectral density, the units are generally used to distinguish the difference if it is not clear in the text.

The temperature behavior of the MOS was given in Eqs. (2.5-8) and (2.5-9). Equation (2.5-8) can be expressed as

$$K'(T) = K'(T_0)\,(T/T_0)^{-1.5} \tag{3.2-16}$$

Equation (2.5-9) is a series approximation of the threshold voltage temperature dependence and can be written as

$$V_T(T) = V_T(T_0) + \alpha(T - T_0) + \cdots \tag{3.2-17}$$

where a plus sign is used instead of the minus sign in Eq. (2.5-9). This means that α will be negative for NMOS and positive for PMOS. Substituting Eqs. (3.2-16) and (3.2-17) into the simple, large-signal model of the MOSFET in Eq. (3.1-17) gives the temperature dependence of the transistor:

$$I_D(T) = \frac{\mu_o C_{ox} W}{2L}\left(\frac{T}{T_0}\right)^{-1.5}[V_{GS} - V_{T0} - \alpha(T - T_0)]^2 \tag{3.2-18}$$

where V_{T0} is the value of the threshold voltage at the reference temperature, T_0.

While Eq. (3.2-18) can be used to determine the influence of temperature on the transistor current, there is an interesting characteristic called the *zero temperature coefficient* point that we want to examine. The differentiation of Eq. (3.2-18) is given as

$$\begin{aligned}\frac{dI_D}{d_T} = {} & \frac{-1.5\mu_o C_{ox} W}{2LT_0}\left(\frac{T}{T_0}\right)^{-2.5}[V_{GS} - V_{T0} - \alpha(T - T_0)]^2 \\ & - \alpha\,\frac{\mu_o C_{ox} W}{L}\left(\frac{T}{T_0}\right)^{-1.5}[V_{GS} - V_{T0} - \alpha(T - T_0)]\end{aligned} \tag{3.2-19}$$

If Eq. (3.2-19) is set to zero, the value of V_{GS} that gives a zero temperature coefficient is given as

$$V_{GS}\,(\text{ZTC}) = V_{T0} - \alpha T_0 - \frac{\alpha T}{3} \tag{3.2-20}$$

The zero temperature characteristics of an NMOS transistor are illustrated in Fig. 3.2-9. Note that in general, the ZTC point is well defined only for temperatures less than 150 °C. For higher temperatures the ZTC drifts from the low-temperature value. Also, note that for values of gate–source voltage above the ZTC value, the drain current *decreases* as the temperature *increases*. This reduces the thermal runaway characteristics of the MOSFET compared to the BJT. The following example illustrates the application of the ZTC concept.

Example 3.2-1

Example 3.2-1 Application of the ZTC Point to an NMOS Transistor

Assume that $K' = 110\ \mu\text{A/V}^2$, $V_{T0} = 0.7$ V, and $\alpha = -2.3$ mV/°C for an NMOS transistor. Find the ZTC point if the reference temperature is 300 K and the new temperature is 400 K. Find the drain current under these conditions if $W/L = 10$.

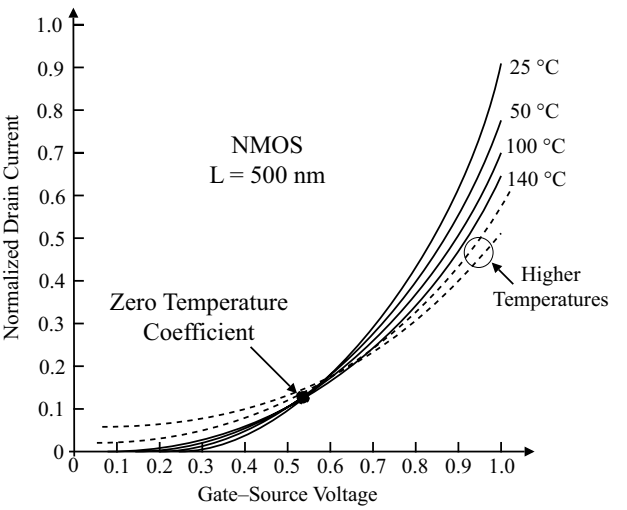


Figure 3.2-9 Zero temperature coefficient characteristics of a modern CMOS technology (L_{min} = 50 nm).

SOLUTION

Evaluating Eq. (3.2-20) gives

$$V_{GS}(\text{ZTC}) = 0.7\text{ V} + 0.0023(\text{V/°C})(300\text{ K}) + \frac{0.0023(\text{V/°C})(400\text{ K})}{3}$$

$$= 1.697\ V$$

At this value of VGS the drain current is

$$I_D(400\text{ °K}) = \frac{110\ \mu\text{A/V}^2(10)}{2}\left(\frac{400}{300}\right)^{-1.5}[1.697 - 0.7 - 0.0023(100)]^2$$

$$= 176\ \mu\text{A}$$

3.3　MOS 管小信号模型

Small-Signal Model for the MOS Transistor

Up to this point, we have been considering the large-signal model of the MOS transistor shown in Fig. 3.2-1. However, after the large-signal model has been used to find the dc conditions, the small-signal model becomes important. The small-signal model is a linear model that helps to simplify calculations. It is valid only over voltage or current regions where the large-signal voltage and currents can adequately be represented by a straight line.

Figure 3.3-1 shows a linearized small-signal model for the MOS transistor. The parameters of the small-signal model will be designated by lowercase subscripts. The various parameters of this small-signal model are all related to the large-signal model parameters and dc variables. The normal relationship between these two models assumes that the small-signal parameters are defined in terms of the ratio of small perturbations of the large-signal variables or as the partial differentiation of one large-signal variable with respect to another.

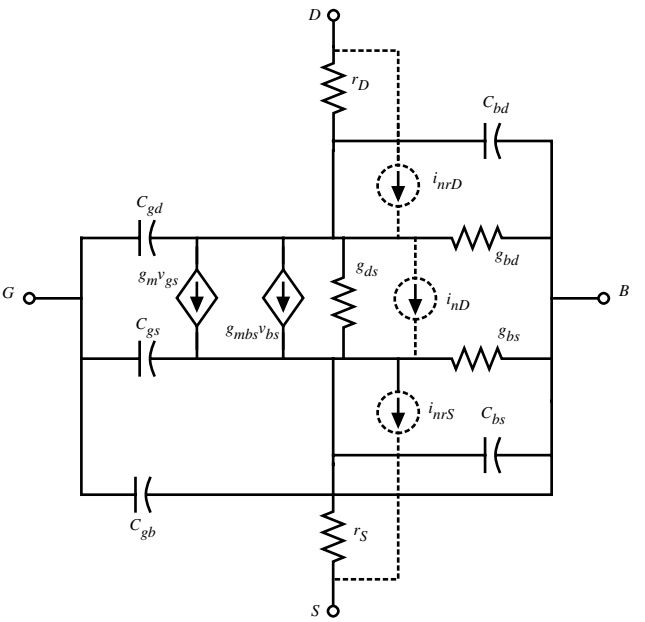


Figure 3.3-1 Small-signal model of the MOS transistor.

The conductances g_{bd} and g_{bs} are the equivalent conductances of the bulk-to-drain and bulk-to-source junctions. Since these junctions are normally reverse biased, the conductances are very small. They are defined as

$$g_{bd} = \frac{\partial i_{BD}}{\partial v_{BD}} \text{ (evaluated at the quiescent point)} \cong 0 \tag{3.3-1}$$

and

$$g_{bs} = \frac{\partial i_{BS}}{\partial v_{BS}} \text{ (evaluated at the quiescent point)} \cong 0 \tag{3.3-2}$$

The channel transconductances g_m and g_{mbs} and the channel conductance g_{ds} are defined as

$$g_m = \frac{\partial i_D}{\partial v_{GS}} \text{ (evaluated at the quiescent point)} \tag{3.3-3}$$

$$g_{mbs} = \frac{\partial i_D}{\partial v_{BS}} \text{ (evaluated at the quiescent point)} \tag{3.3-4}$$

and

$$g_{ds} = \frac{\partial i_D}{\partial v_{DS}} \text{ (evaluated at the quiescent point)} \tag{3.3-5}$$

The values of these small-signal parameters depend on which region the quiescent point occurs in. For example, in the saturated region g_m can be found from Eq. (3.1-18) as

$$g_m = \sqrt{(2K'W/L)|I_D|(1 + \lambda V_{DS})} \cong \sqrt{(2K'W/L)|I_D|} \tag{3.3-6}$$

which emphasizes the dependence of the small-signal parameters on the large-signal operating conditions. The small-signal channel transconductance due to v_{SB} is found by rewriting Eq. (3.3-4) as

$$g_{mbs} = \frac{-\partial i_D}{\partial v_{SB}} = -\left(\frac{\partial i_D}{\partial V_T}\right)\left(\frac{\partial V_T}{\partial v_{SB}}\right) \tag{3.3-7}$$

Using Eq. (3.1-2) and noting that $\partial i_D/\partial V_T = -\partial i_D/\partial v_{GS}$, we get*

$$g_{mbs} = g_m \frac{\gamma}{2(2|\phi_F| + |V_{SB}|)^{1/2}} = \eta g_m \tag{3.3-8}$$

This transconductance will become important in our small-signal analysis of the MOS transistor when the ac value of the source–bulk potential v_{sb} is not zero.

The small-signal channel conductance, g_{ds} (g_0), is given as

$$g_{ds} = g_0 = \frac{I_D \lambda}{1 + \lambda V_{DS}} \cong I_D \lambda \tag{3.3-9}$$

The channel conductance will be dependent on L through λ, which is inversely proportional to L. We have assumed the MOS transistor is in saturation for the results given by Eqs. (3.3-6), (3.3-8), and (3.3-9).

The important dependence of the small-signal parameters on the large-signal model parameters and dc voltages and currents is illustrated in Table 3.3-1. In this table we see that the three small-signal model parameters of g_m, g_{mbs}, and g_{ds} have several alternate forms. An example of the typical values of the small-signal model parameters follows.

Example 3.3-1

Typical Values of Small-Signal Model Parameters

Find the values of g_m, g_{mbs}, and g_{ds} using the large-signal model parameters in Table 3.1-2 for both an n-channel and a p-channel device if the dc value of the magnitude of the drain current is 50 μA and the magnitude of the dc value of the source–bulk voltage is 2 V. Assume that the W/L ratio is 1 μm/1 μm.

SOLUTION

Using the values of Table 3.1-2 and Eqs. (3.3-6), (3.3-8), and (3.3-9) gives $g_m = 105$ μA/V, $g_{mbs} = 12.8$ μA/V, and $g_{ds} = 2.0$ μA/V for the n-channel device and $g_m = 70.7$ μA/V, $g_{mbs} = 12.0$ μA/V, and $g_{ds} = 2.5$ μA/V for the p-channel device.

*Note that absolute signs are used for V_{SB} in order to prevent g_{mbs} from becoming infinite. However, in a few rare cases the source–bulk junction is forward biased and in this case the absolute signs must be removed and V_{SB} becomes negative (for an n-channel transistor).

Table 3.3-1 Dependence of the Small-Signal Model Parameters on the dc Values of Voltage and Current in the Saturation Region

Small-Signal Model Parameters	dc Current	dc Current and Voltage	dc Voltage
g_m	$\cong (2K'I_D W/L)^{1/2}$	—	$\cong \dfrac{K'W}{L}(V_{GS} - V_T)$
g_{mbs}	—	$\dfrac{\gamma(2I_D\beta)^{1/2}}{2(2\lvert\phi_F\rvert + \lvert V_{SB}\rvert)^{1/2}}$	$\dfrac{\gamma[\beta(V_{GS} - V_T)]^{1/2}}{2(2\lvert\phi_F\rvert + \lvert V_{SB}\rvert)^{1/2}}$
g_{ds}	$\cong \lambda I_D$	—	—

Although MOS devices are not often used in the nonsaturation region in analog circuit design, the relationships of the small-signal model parameters in the nonsaturation region are given as

$$g_m = \frac{\partial i_D}{\partial v_{GS}} \cong \beta V_{DS} \tag{3.3-10}$$

$$g_{mbs} = \frac{\partial i_D}{\partial v_{BS}} = \frac{\beta\gamma V_{DS}}{2(2|\phi_F| + |V_{SB}|)^{1/2}} \tag{3.3-11}$$

and

$$g_{ds} \cong \beta(V_{GS} - V_T - V_{DS}) \tag{3.3-12}$$

Table 3.3-2 summarizes the dependence of the small-signal model parameters on the large-signal model parameters and dc voltages and currents for the nonsaturated region. The typical values of the small-signal model parameters for the nonsaturated region are illustrated in the following example.

Example 3.2-1

Typical Values of the Small-Signal Model Parameters in the Nonsaturated Region

Find the values of the small-signal model parameters in the nonsaturation region for an n-channel and a p-channel transistor if $V_{GS} = 5$ V, $V_{DS} = 1$ V, and $|V_{BS}| = 2$ V. Assume that the W/L ratio for both transistors is 1 μm/1 μm. Also assume that the value for K' in the nonsaturation region is the same as that for the saturation (generally a poor assumption).

SOLUTION

First, it is necessary to calculate the threshold voltage of each transistor using Eq. (3.1-2). The results are a V_T of 1.02 V for the n-channel and −1.14 V for the p-channel. This gives a dc current of 383 μA and 168 μA, respectively. Using Eqs. (3.3-10), (3.3-11), and (3.3-12), we get $g_m = 110$ μA/V, $g_{mbs} = 13.4$ μA/V, and $r_{ds} = 3.05$ kΩ for the n-channel transistor and $g_m = 50$ μA/V, $g_{mbs} = 8.52$ μA/V, and $r_{ds} = 6.99$ kΩ for the p-channel transistor.

Table 3.3-2 Dependence of the Small-Signal Model Parameters on the dc Values of Voltage and Current in the Nonsaturation Region

Small-Signal Model Parameters	dc Voltage and/or Current Dependence
g_m	$\cong \beta V_{DS}$
g_{mbs}	$\dfrac{\beta\gamma V_{DS}}{2(2\lvert\phi_F\rvert + \lvert V_{SB}\rvert)^{1/2}}$
g_{ds}	$\cong \beta (V_{GS} - V_T - V_{DS})$

The values of r_d and r_s are assumed to be the same as r_D and r_S of Fig. 3.2-1. Likewise, for small-signal conditions C_{gs}, C_{gd}, C_{gb}, C_{bd}, and C_{bs} are evaluated for C_{gs}, C_{gd}, and C_{gb} by knowing the region of operation (cutoff, saturation or nonsaturation) and for C_{bd} and C_{bs} by knowing the value of V_{BD} and V_{BS}. With this information, C_{gs}, C_{gd}, C_{gb}, C_{bd}, and C_{bs} can be found from C_{GS}, C_{GD}, C_{GB}, C_{BD}, and C_{BS}, respectively.

If the noise of the MOS transistor is to be modeled, then three additional current sources are added to Fig. 3.3-1 as indicated by the dashed lines. The values of the mean-square noise-current sources are given as

$$i_{nrD}^2 = \left(\frac{4kT}{r_D}\right)\Delta f \quad (\text{A}^2) \tag{3.3-13}$$

$$i_{nrS}^2 = \left(\frac{4kT}{r_S}\right)\Delta f \quad (\text{A}^2) \tag{3.3-14}$$

and

$$i_{nD}^2 = \left[\frac{8kT\, g_m(1 + \eta)}{3} + \frac{(\text{KF})I_D}{f\, C_{\text{ox}}\, L^2}\right]\Delta f \quad (\text{A}^2) \tag{3.3-15}$$

The various parameters for these equations have previously been defined. With the noise modeling capability, the small-signal model of Fig. 3.3-1 is a very general model.

It will be important to be familiar with the small-signal model for the saturation region developed in this section. This model, along with the circuit simplification techniques given in Appendix A, will be the key element in analyzing the circuits in the following chapters.

3.4 计算机仿真模型　Computer Simulation Models

The large-signal model of the MOS device previously discussed is simple to use for hand calculations but neglects many important second-order effects. While a simple model for hand calculation and design intuition is critical, a more accurate model is required for computer simulation. There are many model choices available for the designer when choosing a device model to use for computer simulation. At one time, HSPICE* supported 43 different MOSFET

*HSPICE is now owned by Avant! Inc. and has been renamed Star-Hspice.

models [2] (many of which were company proprietary) while SmartSpice publishes support for 14 [12]. Which model is the right one to use? In the fabless semiconductor environment, the user must use the model provided by the wafer foundry. In companies where the foundry is captive (i.e., the company owns its own wafer fabrication facility) a modeling group provides the model to circuit designers. It is seldom that a designer chooses a model and performs parameter extraction to get the terms for the model chosen.

The SPICE LEVEL 3 dc model will be covered in some detail because it is a relatively straightforward extension of the LEVEL 2 model. The BSIM3v3 model will be introduced but the detailed equations will not be presented because of the volume of equations required to describe it—there are other good texts that deal with the subject of modeling exclusively [13,14], and there is little additional design intuition derived from covering the details.

Models developed for computer simulation have improved over the years but no model has yet been developed that, with a single set of parameters, covers device operation for all possible geometries. Therefore, many SPICE simulators offer a feature called "model binning." Parameters are derived for transistors of different geometry (W's and L's) and the simulator determines which set of parameters to use based on the particular W and L called out in the device instantiation line in the circuit description. The circuit designer need only be aware of this since the binning is done by the model provider.

SPICE LEVEL 3 Model

The large-signal model of the MOS device previously discussed is simple to use for hand calculations but neglects many important second-order effects. Most of these second-order effects are due to narrow or short channel dimensions (less than about 3 μm). In this section, we will consider a more complex model that is suitable for computer-based analysis (circuit simulation, i.e., SPICE simulation). In particular, the SPICE LEVEL 3 model will be covered (see Table 3.4-1). This model is typically good for MOS technologies down to about 0.8 μm. We will also consider the effects of temperature on the parameters of the MOS large-signal model.

We first consider second-order effects due to small geometries (Fig. 3.4-1). When v_{GS} is greater than V_T, the drain current for a small device can be given as [2] follows:

Drain Current

$$i_{DS} = \text{BETA}\left[v_{GS} - V_T - \left(\frac{1+f_b}{2}\right)v_{DE}\right]v_{DE} \tag{3.4-1}$$

$$\text{BETA} = \text{KP}\frac{W_{\text{eff}}}{L_{\text{eff}}} = \mu_{\text{eff}}\text{COX}\frac{W_{\text{eff}}}{L_{\text{eff}}} \tag{3.4-2}$$

$$L_{\text{eff}} = L - 2(\text{LD}) \tag{3.4-3}$$

$$W_{\text{eff}} = W - 2(\text{WD}) \tag{3.4-4}$$

$$v_{DE} = \min(v_{DS}, v_{DS}(\text{sat})) \tag{3.4-5}$$

$$f_b = f_n + \frac{\text{GAMMA}\cdot f_s}{4(\text{PHI} + v_{SB})^{1/2}} \tag{3.4-6}$$

Table 3.4-1 Typical Model Parameters Suitable for SPICE Simulations Using LEVEL-3 Model (Extended Model)*

Parameter Symbol	Parameter Description	Typical Parameter Value		
		n-Channel	p-Channel	Units
VTO	Threshold	0.7 ± 0.15	−0.7 ± 0.15	V
UO	Mobility	660	210	$cm^2/V\text{-}s$
DELTA	Narrow-width threshold adjustment factor	2.4	1.25	—
ETA	Static-feedback threshold adjustment factor	0.1	0.1	—
KAPPA	Saturation field factor in channel length modulation	0.15	2.5	1/V
THETA	Mobility degradation factor	0.1	0.1	1/V
NSUB	Substrate doping	3×10^{16}	6×10^{16}	cm^{-3}
TOX	Oxide thickness	140	140	A
XJ	Metallurgical junction depth	0.2	0.2	μm
WD	Delta width			μm
LD	Lateral diffusion	0.016	0.015	μm
NFS	Parameter for weak inversion modeling	7×10^{11}	6×10^{11}	cm^{-2}
CGSO		220×10^{-12}	220×10^{-12}	F/m
CGDO		220×10^{-12}	220×10^{-12}	F/m
CGBO		700×10^{-12}	700×10^{-12}	F/m
CJ		770×10^{-6}	560×10^{-6}	F/m^2
CJSW		380×10^{-12}	350×10^{-12}	F/m
MJ		0.5	0.5	
MJSW		0.38	0.35	

*These values are based on a 0.8 μm silicon-gate bulk CMOS n-well process and include capacitance parameters from Table 3.2-1.

Note that PHI is the SPICE model term for the quantity $2\phi_F$. Also be aware that PHI is always positive in SPICE regardless of the transistor type (p- or n-channel). In this text, the term PHI will always be positive while the term $2\phi_F$ will have a polarity determined by the transistor type as shown in Table 2.3-1.

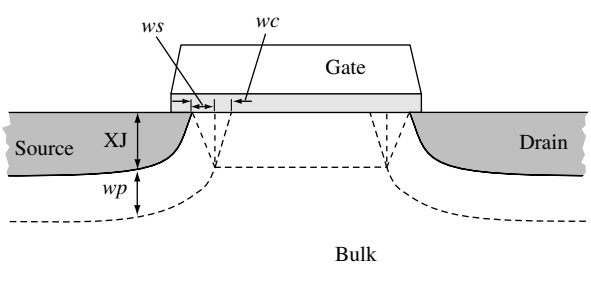


Figure 3.4-1 Illustration of the short-channel effects in the MOS transistor.

$$f_n = \frac{\text{DELTA}}{W_{\text{eff}}} \frac{\pi \varepsilon_{\text{Si}}}{2 \cdot \text{C}_{\text{ox}}} \tag{3.4-7}$$

$$f_s = 1 - \frac{\text{XJ}}{L_{\text{eff}}} \left\{ \frac{\text{LD} + wc}{\text{XJ}} \left[1 - \left(\frac{wp}{\text{XJ} + wp} \right)^2 \right]^{1/2} - \frac{\text{LD}}{\text{XJ}} \right\} \tag{3.4-8}$$

$$wp = xd(\text{PHI} + v_{SB})^{1/2} \tag{3.4-9}$$

$$xd = \left(\frac{2 \cdot \varepsilon_{\text{Si}}}{q \cdot \text{NSUB}} \right)^{1/2} \tag{3.4-10}$$

$$wc = \text{XJ} \left[k_1 + k_2 \left(\frac{wp}{\text{XJ}} \right) - k_3 \left(\frac{wp}{\text{XJ}} \right)^2 \right] \tag{3.4-11}$$

$$k_1 = 0.0631353, \qquad k_2 = 0.08013292, \qquad k_3 = 0.01110777$$

Threshold Voltage

$$V_T = V_{bi} - \left(\frac{\text{ETA} - 8.14 \times 10^{-22}}{C_{\text{ox}} L_{\text{eff}}^3} \right) v_{DS} + \text{GAMMA} \cdot f_s(\text{PHI} + v_{SB})^{1/2} + f_n(\text{PHI} + v_{SB}) \tag{3.4-12}$$

$$v_{bi} = v_{fb} + \text{PHI} \tag{3.4-13}$$

or

$$v_{bi} = \text{VTO} - \text{GAMMA} \cdot \sqrt{\text{PHI}} \tag{3.4-14}$$

Saturation Voltage

$$v_{\text{sat}} = \frac{v_{gs} - V_T}{1 + f_b} \tag{3.4-15}$$

$$v_{DS}\,(\text{sat}) = v_{\text{sat}} + v_C - \left(v_{\text{sat}}^2 + v_C^2 \right)^{1/2} \tag{3.4-16}$$

$$v_C = \frac{\text{VMAX} \cdot L_{\text{eff}}}{\mu_{\text{s}}} \tag{3.4-17}$$

If VMAX is not given, then $v_{DS}\,(\text{sat}) = v_{\text{sat}}$.

Effective Mobility

$$\mu_s = \frac{\text{U0}}{1 + \text{THETA}\,(v_{Gs} - V_T)}, \quad \text{when VMAX} = 0 \tag{3.4-18}$$

$$\mu_{\text{eff}} = \frac{\mu_s}{1 + \dfrac{v_{DE}}{v_C}}, \quad \text{when VMAX} > 0; \quad \text{otherwise } \mu_{\text{eff}} = \mu_s \tag{3.4-19}$$

Channel Length Modulation

$$\Delta L = xd\left[\text{KAPPA}\,(v_{DS} - v_{DS}(\text{sat}))\right]^{1/2}, \quad \text{when VMAX} = 0 \tag{3.4-20}$$

$$\Delta L = -\frac{ep \cdot xd^2}{2} + \left[\left(\frac{ep \cdot xd^2}{2}\right)^2 + \text{KAPPA} \cdot xd^2\,(v_{DS} - v_{DS}(\text{sat}))\right]^{1/2}, \tag{3.4-21}$$

$$\text{when VMAX} > 0$$

where

$$ep = \frac{v_C\,(v_C + v_{DS}\,(\text{sat}))}{L_{\text{eff}}\,v_{DS}\,(\text{sat})} \tag{3.4-22}$$

$$i_{DS} = \frac{i_{DS}}{1 - \Delta L} \tag{3.4-23}$$

The temperature-dependent variables in the models developed so far include the Fermi potential, PHI, EG, bulk junction potential of the source–bulk and drain–bulk junctions, PB, the reverse currents of the pn junctions, I_S, and the dependence of mobility on temperature. The temperature dependence of most of these variables is found in the equations given previously or from well-known expressions. The dependence of mobility on temperature is given as

$$\text{U0}(T) = \text{U0}(T_0)\left(\frac{T}{T_0}\right)^{\text{BEX}} \tag{3.4-24}$$

where BEX is the temperature exponent for mobility and is typically −1.5.

$$v_{\text{therm}}\,(T) = \frac{k\,T}{q} \tag{3.4-25}$$

$$\text{EG}(T) = 1.16 - 7.02 \cdot 10^{-4}\left[\frac{T^2}{T + 1108.0}\right] \tag{3.4-26}$$

$$\text{PHI}(T) = \text{PHI}(T_0) \cdot \left(\frac{T}{T_0}\right) - v_{\text{therm}}\,(T)\left[3\ln\left(\frac{T}{T_0}\right) + \frac{\text{EG}(T_0)}{v_{\text{therm}}\,(T_0)} - \frac{\text{EG}(T)}{v_{\text{therm}}\,(T)}\right] \tag{3.4-27}$$

$$v_{bi}(T) = v_{bi}(T_0) + \frac{\text{PHI}(T) - \text{PHI}(T_0)}{2} + \frac{\text{EG}(T_0) - \text{EG}(T)}{2} \tag{3.4-28}$$

$$\text{VT0}(T) = v_{bi}\,(T) + \text{GAMMA}\left[\sqrt{\text{PHI}(T)}\right] \tag{3.4-29}$$

$$\text{PHI}(T) = 2v_{\text{therm}}\ln\left(\frac{\text{NSUB}}{n_i(T)}\right) \tag{3.4-30}$$

$$n_i(T) = 1.45 \cdot 10^{16}\left(\frac{T}{T_0}\right)^{3/2}\exp\left[\text{EG} \cdot \left(\frac{T}{T_0} - 1\right)\left(\frac{1}{2 \cdot v_{\text{therm}}\,(T_0)}\right)\right] \tag{3.4-31}$$

For drain and source junction diodes, the following relationships apply:

$$\mathrm{PB}(T) = \mathrm{PB} \cdot \left(\frac{T}{T_0}\right) - v_{\mathrm{therm}}(T)\left[3\ln\left(\frac{T}{T_0}\right) + \frac{\mathrm{EG}(T_0)}{v_{\mathrm{therm}}(T_0)} - \frac{\mathrm{EG}(T)}{v_{\mathrm{therm}}(T)}\right] \tag{3.4-32}$$

and

$$I_S(T) = \frac{I_S(T_0)}{\mathrm{N}} \cdot \exp\left[\frac{\mathrm{EG}(T_0)}{v_{\mathrm{therm}}(T_0)} - \frac{\mathrm{EG}(T)}{v_{\mathrm{therm}}(T)} + 3\ln\left(\frac{T}{T_0}\right)\right] \tag{3.4-33}$$

where N is the diode emission coefficient. The nominal temperature, T_0, is 300 K.

An alternate form of the temperature dependence of the MOS model can be found elsewhere [15].

BSIM 3v3 Model

MOS transistor models introduced thus far in this chapter have been used successfully when applied to 0.8 μm technologies and above. As geometries shrink below 0.8 μm, better models are required. Researchers in the Electrical Engineering and Computer Sciences Department at the University of California at Berkeley have been leaders in the development of SPICE and the models used in it. In 1984 they introduced the BSIM1 model [16] to address the need for a better submicron MOS transistor model. The BSIM1 model approached the modeling problem as a multiparameter curve-fitting exercise. The model contained 60 parameters covering the dc performance of the MOS transistor. There was some relationship to device physics, but in large part, it was a nonphysical model. Later, in 1991, UC Berkeley released the BSIM2 model that improved performance related to the modeling of output resistance changes due to hot-electron effects, source/drain parasitic resistance, and inversion-layer capacitance. This model contained 99 dc parameters, making it more unwieldy than the 60-parameter (dc parameters) BSIM1 model. In 1994, UC Berkeley introduced the BSIM3 model (version 2), which, unlike the earlier BSIM models, returned to a more device-physics-based modeling approach. The model is simpler to use and has only 40 dc parameters. Moreover, the BSIM3 model provides good performance when applied to analog as well as digital circuit simulation. In its third version, BSIM3v3 [3], it has become the industry standard MOS transistor model.

The BSIM3 model addresses the following important effects seen in deep-submicron MOSFET operation:

- Threshold voltage reduction
- Mobility degradation due to a vertical field
- Velocity saturation effects
- Drain-induced barrier lowering (DIBL)
- Channel length modulation
- Subthreshold (weak inversion) conduction
- Parasitic resistance in the source and drain
- Hot-electron effects on output resistance

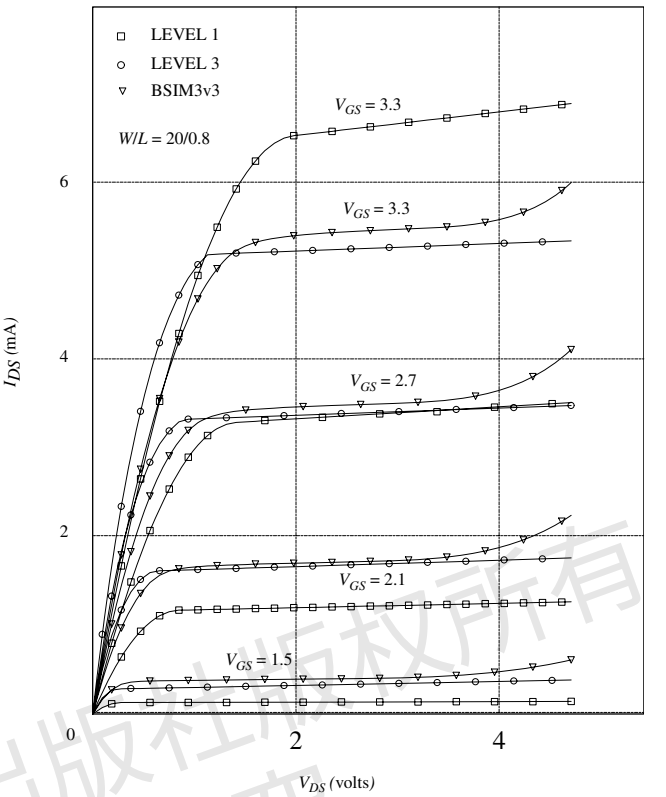


Figure 3.4-2 Simulation of MOSFET transconductance characteristic using LEVEL = 1, LEVEL = 3 and the BSIM3v3 models.

The plot shown in Fig. 3.4-2 shows a comparison of a 20/0.8 device using the LEVEL 1, LEVEL 3, and BSIM3v3 models. The model parameters were adjusted to provide similar characteristics (given the limitations of each model). Assuming that the BSIM3v3 model closely approximates actual transistor performance, this figure indicates that the LEVEL 1 model is grossly in error, while the LEVEL 3 model shows a significant difference in modeling the transition from the nonsaturation to linear region.

3.5　亚阈区 MOS 模型　　Subthreshold MOS Model

The models discussed in previous sections predict that no current will flow in a device when the gate–source voltage is at or below the threshold voltage. In reality, this is not the case. As v_{GS} approaches V_T, the $i_D - v_{GS}$ characteristics change from square-law to exponential. Whereas the region where v_{GS} is above the threshold is called the *strong inversion* region, the region below (actually, the transition between the two regions is not well defined as will be explained later) is called the *subthreshold,* or *weak inversion* region. This is illustrated in Fig. 3.5-1 where the transconductance characteristic of a MOSFET in saturation is shown with the square root of current plotted as a function of the gate–source voltage. When the gate–source voltage reaches the value designated as V_{ON} (this relates to the SPICE model