

10

Field Effect Transistors

Bogdan M.
Wilamowski
Auburn University
J. David Irwin
Auburn University

10.1	Introduction	10-1
10.2	MOS Transistor	10-1
	MOS Structure and Threshold Voltage • MOS Transistor Current Characteristics • Second-Order Effects on a MOS Transistor	
10.3	Junction Field Effect Transistor	10-10
10.4	Static Induction Transistor	10-13
	Theory of SIT Operation for Small Currents • Theory of SIT for Large Currents • Bipolar Mode of Operation of the SIT	
10.5	Lateral Punch-through Transistor	10-17
10.6	Power MOS Transistors	10-19
	References	10-21

10.1 Introduction

There are several different types of field effect transistors (FETs), each of which has a different operational principle. For example, there are metal oxide semiconductor (MOS) transistors, junction field effect transistors (JFETs), static induction transistors (SITs), the punch-through transistors (PHTs), and others. All of these devices employ the flow of majority carriers. The most popular one among this group is the MOS transistor, which is primarily used in integrated circuits [T99, N06, S05]. In contrast, the JFET is not suitable for integration and so it is primarily fabricated as an individual device [E97, R99]. All FETs have very large input resistance on the order of $10^{12} \Omega$. The MOS transistor typically operates with very small currents [N02] and thus for power electronics applications thousands of MOS transistors are connected in parallel. A JFET usually operates with larger currents. Both JFET and MOS transistors have relatively small transconductances, and this means that they cannot control current flow as effectively as bipolar junction transistors (BJTs). Since the parasitic capacitors are of the same order of magnitude, BJTs can charge and discharge these capacitors much faster and so BJTs are more suitable for high-frequency operations. Because current flow in MOS transistors is very close to the silicon surface where surface states can fluctuate with time, MOS devices have a relatively higher noise level, especially at low frequencies.

10.2 MOS Transistor

The MOS transistor can be considered a capacitor in which the applied voltage to the gate G would attract carriers (electrons in NMOS and holes in PMOS) from the semiconductor substrate. This layer of accumulated carriers near the surface conducts current between source and drain [T99]. If the voltage on the gate is increased, then more carriers (electrons or holes) will be accumulated near the surface, causing a larger current to flow, as indicated in Figure 10.1. In order to better understand the process of carrier accumulation under the gate, the MOS structure must be analyzed in detail.

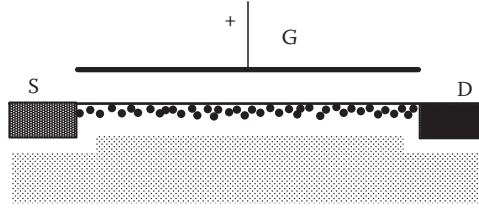


FIGURE 10.1 The principle of operation for a NMOS transistor, where electrons are accumulated by the positive gate voltage.

10.2.1 MOS Structure and Threshold Voltage

Figure 10.2 shows the cross section of the MOS band structure with a p-type silicon substrate. Note that the Fermi level has a different location in every material. In metals, there is no forbidden energy gap and the Fermi level E_{Fm} is on the edge of the conduction band. Albert Einstein received his Nobel Prize for the photoelectric effect, in which he was able to measure the energy required to free an electron from metal to the vacuum. This energy is now known as the work function ϕ_m . Work functions for various materials are shown in Table 10.1.

It is important to note that the Fermi levels in semiconductors may depend on the doping level N and on the type of impurities present. In the n-type material, the Fermi level E_{Fs} is above the center of the energy gap E_i and in the p-type material, the Fermi level E_{Fs} is below E_i , as shown in Figure 10.2. The work function ϕ_s for intrinsic, i.e., undoped, silicon is 3.8 eV, as listed in Table 10.1, and the energy needed to free an electron from the p-type silicon is

$$\phi_s = 3.8 + \phi_F \quad \text{where } \phi_F = V_T \ln \left(\frac{N_A}{n_i} \right) \quad (10.1)$$

and this energy is dependent on the acceptor doping level N_A and intrinsic carrier concentration n_i . At room temperature in silicon $n_i = 10^{10} \text{ cm}^{-3}$. Similarly, in the n-type silicon:

$$\phi_s = 3.8 - \phi_F \quad \text{where } \phi_F = V_T \ln \left(\frac{N_D}{n_i} \right) \quad (10.2)$$

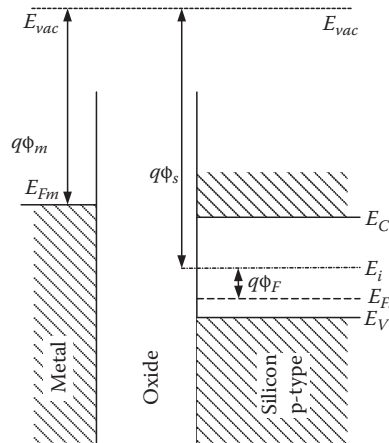


FIGURE 10.2 The location of Fermi levels in metal and in the silicon.

TABLE 10.1 Work Functions ϕ_m for Various Materials

	Si	p ⁺ Si	n ⁺ Si	Al	Mo	Au	Cu
ϕ_m (eV)	3.8	4.5	3.05	3.2	3.95	4.1	3.8

and N_D is the donor doping level. When positive voltage V_G is applied to the gate, the metal's band structure will move down by qV_G , as shown in Figure 10.3, and the depletion layer in the silicon will be formed.

Figure 10.3 shows the case with the gate (metal) biasing exactly at the threshold as the accumulation layer of carriers at the silicon surface is just being formed. α is particular state, called strong inversion, is one in which the silicon surface is now a n-type level with the same electron concentration as the hole concentration in the bulk p-type silicon. It also means the voltage drop on the depletion layer is

$$V_d = 2\phi_F \quad (10.3)$$

Knowing the voltage drop V_d , the thickness w of the depletion layer can be found from

$$w = \sqrt{\frac{2\epsilon_o\epsilon_{Si}V_d}{qN}} = \sqrt{\frac{4\epsilon_o\epsilon_{Si}\phi_F}{qN}} \quad (10.4)$$

where

$$\epsilon_{Si} = 11.8 \quad \epsilon_o = 8.85 \cdot 10^{-14} \text{ F/cm} \quad (10.5)$$

and N is the impurity concentration in the silicon substrate. q is the charge of ionized impurities in the depletion layer is

$$Q_d = qNw = \sqrt{4\epsilon_{Si}\epsilon_o\phi_FqN} = \sqrt{6.68 \cdot 10^{-31}\phi_FN} \quad (10.6)$$

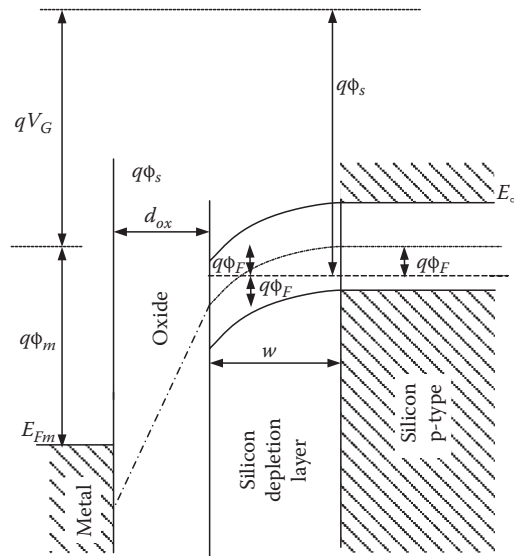
**FIGURE 10.3** MOS band structure with positive voltage on the gate V_G in p-type silicon.

TABLE 10.2 Surface Charges in Silicon for Various Crystal Orientations

Crystal Orientation	$\langle 111 \rangle$	$\langle 110 \rangle$	$\langle 100 \rangle$
Number of surface states Q_{ss}/q (cm^{-2})	$1.5 \cdot 10^{11}$	10^{11}	$2 \cdot 10^{10}$
Surface charge Q_{ss} (C)	$2.4 \cdot 10^{-8}$	$1.6 \cdot 10^{-8}$	$3.2 \cdot 10^{-9}$

On top of the depletion layer charge, there exists a silicon surface charge Q_{ss} that depends on the silicon crystal orientation as indicated in Table 10.2.

Since the MOS structure can be considered as a capacitor, knowledge of the charge in silicon $Q_d + Q_{ss}$ yields the corresponding voltage

$$V = \frac{Q_d + Q_{ss}}{C_{ox}} \quad (10.7)$$

the MOS structure unit capacitance is

$$C_{ox} = \frac{\epsilon_{ox}\epsilon_o}{d_{ox}} = \frac{3.45 \cdot 10^{-9}}{d_{ox} (\mu\text{m})} \text{ - } > (\text{F}/\text{cm}^2) \quad (10.8)$$

where d_{ox} is the thickness of the oxide and $\epsilon_{ox} = 3.9$. In addition to the electrical charges, the voltage drop on the depletion layer $V_d = 2\phi_F$ and the difference in the work functions should also be considered in determining the threshold voltage

$$\phi_{ms} = \phi_m - \phi_s = \phi_m - (3.8 + s\phi_F) = \phi_m - 3.8 - s\phi_F \quad (10.9)$$

where the symbol s indicates the sign, which is

$$s = \begin{cases} 1 & \text{for n channel (p-type impurities)} \\ -1 & \text{for p channel (n-type impurities)} \end{cases} \quad (10.10)$$

In conclusion, the threshold voltage for a MOS structure is given by

$$V_{th} = \frac{sQ_d - Q_{ss} + sqF_{imp}}{C_{ox}} + s2\phi_F + \phi_{ms} \quad (10.11)$$

Note that the effective charge can be controlled by ion implantation using F_{imp} dose (cm^{-2}), which can be made from p-type or n-type impurities, and as a result, the threshold voltage can be properly adjusted.

Example 10.1

Calculate the threshold voltage V_{th} for a MOS transistor with a p^+ polysilicon gate and a p-type substrate with $N_A = 10^{16} \text{ cm}^{-3}$. Assume a $\langle 100 \rangle$ crystal orientation and a $d_{ox} = 0.1 \mu\text{m}$. Find the implantation dose required for adjusting the threshold voltage to $V_{th} = +2 \text{ V}$. Specify if boron or phosphor should be used for implantation. The calculations required for this example are

$$\phi_F = V_T \ln \frac{N}{n_i} = 0.0258 \ln \frac{10^{16}}{1.5 \cdot 10^{10}} = 0.347 \quad 2\phi_F = 0.694$$

$$s = \begin{cases} 1 & \text{for n channel (p-type substrate)} \\ -1 & \text{for p channel (n-type substrate)} \end{cases} = +1$$

$$\phi_{ms} = \phi_m - \phi_s \quad \phi_s = 3.8 + s\phi_F$$

$$\phi_{ms} = 4.5 - (3.8 + s\phi_F) = 4.5 - (3.8 + 0.347) = 0.353$$

$$Q_d = qNw = \sqrt{4\epsilon_{Si}\epsilon_o\phi_FqN} = \sqrt{6.68 \cdot 10^{-31}\phi_FN} = \sqrt{6.68 \cdot 10^{-31} \cdot 0.347 \cdot 10^{16}} = 4.8145 \cdot 10^{-8}$$

$$Q_{ss} \xrightarrow{<100>} 3.3 \cdot 10^{-9}$$

$$C_{ox} = \frac{\epsilon_{ox}\epsilon_o}{d_{ox}} = \frac{3.45 \cdot 10^{-9}}{d_{ox}(\mu m)} = \frac{3.45 \cdot 10^{-9}}{0.1} = 3.45 \cdot 10^{-8} \text{ (F/cm}^2\text{)}$$

$$V_{th} = s \frac{Q_d}{C_{ox}} + s2\phi_F + \phi_{ms} - \frac{Q_{ss}}{C_{ox}} = \frac{4.8145 - 0.32}{3.45} + 0.694 + 0.353 = 2.35$$

Since the requirement specifies a $V_{th} = 2$ V, this value must be lowered by 0.35 V:

$$\Delta V_{th} = \frac{qF_{imp}}{C_{ox}} \quad F_{imp} = \frac{\Delta V_{th}C_{ox}}{q} = \frac{0.35 \cdot 3.45 \cdot 10^{-8}}{1.6 \cdot 10^{-19}} = 7.54 \cdot 10^{10} \text{ atm/cm}^2$$

The threshold voltage obtained from Equation 10.11 is valid for the case in which the substrate has the same potential as the source for the MOS transistor. If the substrate is biased with an additional voltage V_{SB} , then due to the substrate biasing, the threshold voltage will change by

$$\Delta V_{th} = \frac{\sqrt{2\epsilon_{Si}\epsilon_oqN}}{C_{ox}} \left(\sqrt{2\phi_F + |V_{SB}|} - \sqrt{2\phi_F} \right) = \gamma \left(\sqrt{2\phi_F + |V_{SB}|} - \sqrt{2\phi_F} \right) \quad (10.12)$$

10.2.2 MOS Transistor Current Characteristics

Depending on the value of the drain-source voltage V_{DS} , the MOS transistor characteristics are described by different formulas. For small values of V_{DS} , known as the “linear” or “triode” region, the current is a strong function of the drain voltage, as shown in Figure 10.4. For large values of V_{DS} known as the “current saturation” or “pentode” region, the current is almost independent of the drain voltage. The current-voltage characteristics are often approximated by the formula:

$$I_D = \begin{cases} K[(\Delta)V_{DS} - 0.5V_{DS}^2](1 + \lambda V_{DS}) & \text{for } V_{DS} \leq \Delta \\ 0.5K(\Delta)^2(1 + \lambda V_{DS}) & \text{for } V_{DS} \geq \Delta \end{cases} \quad (10.13)$$

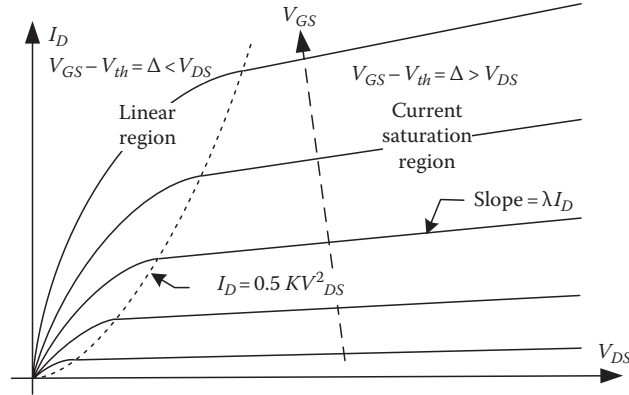


FIGURE 10.4 Output characteristics for MOS transistors.

where Δ shows how much the gate voltage V_{GS} exceeds the threshold voltage V_{th} , i.e.,

$$\Delta = V_{GS} - V_{th} \quad (10.14)$$

$$K = K' \frac{W}{L} = \infty C_{ox} \frac{W}{L} \quad (10.15)$$

The λ parameter describes the slope of the output characteristics in the current saturation region. Typical values of the λ parameter are 0.02 to 0.04 [V^{-1}]. For small signal analysis, a MOS transistor in the current saturation region can be described by two parameters r_m and r_o :

$$r_m = \frac{1}{g_m} = \frac{\Delta}{2I_D} = \frac{1}{K\Delta} = \frac{1}{\sqrt{2KI_D}} \quad (10.16)$$

$$r_o = \frac{1}{\lambda I_D} \quad (10.17)$$

Figure 10.5 shows a small-signal equivalent model of the MOS transistor. For the voltage-controlled circuit, the input capacitances need not to be included, i.e., input capacitance is a part of the previous stage. Assuming that the loading capacitance C is the capacitance of the identical transistor of the next stage $C = C_{ox}WL$, the maximum frequency of operation is

$$f_{max} = \frac{1}{2\pi(r_m \parallel r_o)C} \approx \sqrt{\frac{2\mu C_{ox} \frac{W}{L} I_D}{2\pi C_{ox} WL}} = \sqrt{\frac{2\mu I_D}{C_{ox} W}} \quad (10.18)$$

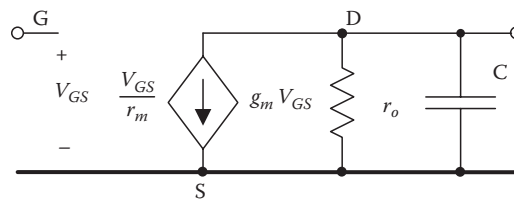


FIGURE 10.5 Small-signal equivalent model for a MOS transistor.

Example 10.2

Consider the NMOS transistor described in Example 10.1 with $V_{th} = +2$ V. Neglecting the channel length modulation ($\lambda = 0$), and assuming the following parameters: electron mobility $\mu = 600$ cm²/Vs, $L = 2$ μ m and $W = 20$ μ m, calculate

- Drain current for $V_{GS} = 4$ V and $V_{DS} = 1$ V
- Drain current for $V_{GS} = 4$ V and $V_{DS} = 10$ V
- Maximum frequency for $V_{GS} = 4$ V and $V_{DS} = 10$ V

$$a. \quad K = \mu C_{ox} \frac{W}{L} = 600 \cdot 3.45 \cdot 10^{-8} \frac{20}{2} = 0.2 \cdot 10^{-3} = 200 \mu A / V^2$$

$$I_D = K \left[(V_{GS} - V_{th})V_{DS} - \frac{V_{DS}^2}{2} \right] (1 + \lambda V_{DS}) = 200 \cdot 10^{-6} \left[2 \cdot 1 - \frac{1}{2} \right] (1 + 0.03) = 309 \mu A$$

$$b. \quad I_D = \frac{K}{2} (V_{GS} - V_{th})^2 (1 + \lambda V_{DS}) = \frac{200 \cdot 10^{-6}}{2} (4 - 2)^2 (1 + 0.3) = 520 \mu A$$

$$c. \quad r_m = \frac{1}{\sqrt{2KI_D}} = \frac{1}{\sqrt{2 \cdot 200 \mu A \cdot 520 \mu A}} = 2.19 k\Omega$$

$$r_o = \frac{1}{\lambda I_D} = \frac{1}{0.03 \cdot 520 \mu A} = 64.1 k\Omega$$

$$C = C_{ox}WL = 3.45 \cdot 10^{-8} \cdot 2 \cdot 10^{-4} \cdot 20 \cdot 10^{-4} = 13.8 fF$$

$$f_{max} = \frac{1}{2\pi(r_m \parallel r_o)C} = \frac{1}{2\pi(2.19 k\Omega \parallel 64.1 k\Omega)13.8 fF} = 5.45 GHz$$

Note that only the gate oxide capacitance was included. Since all other junction parasitic capacitances were ignored, the calculated maximum frequency f_{max} is significantly larger than the actual one.

10.2.3 Second-Order Effects on a MOS Transistor

There are a number of second-order effects that significantly affect the operation of a MOS transistor, such as channel length modulation, carrier velocity limitation, surface mobility degradation, subthreshold conduction, etc. [T99]. These effects will now be described with some detail.

10.2.3.1 Channel Length Modulation

The effect of channel length modulation is shown in Figure 10.6. The thickness of the depletion layer d depends on the drain-substrate voltage and is described by an equation that is similar to Equation 10.4:

$$d = \sqrt{2 \frac{\epsilon_o \epsilon_{Si} V_D}{qN}} \quad (10.19)$$

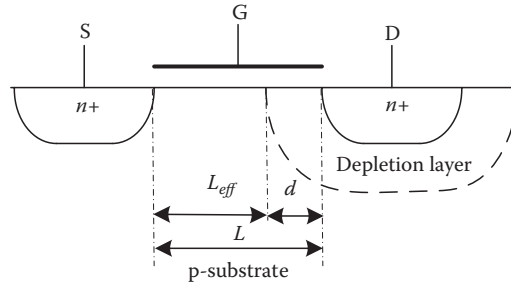


FIGURE 10.6 Channel-length modulation.

As a consequence, the effective channel length L_{eff} is shorter than the distance L between the implanted source and drain regions, and, therefore, instead of Equation 10.15, the transconductance coefficient K should be expressed as

$$K = K' \frac{W}{L_{eff}} = K' \frac{W}{L - \sqrt{2 \frac{\epsilon_o \epsilon_{si} V_D}{qN}}} = K' \frac{W}{L} \frac{1}{\left(1 - \frac{\eta}{L} \sqrt{V_D}\right)} \approx K' \frac{W}{L} \left(1 + \frac{\eta}{L} \sqrt{V_D}\right) \quad (10.20)$$

As indicated in Figure 10.6, the effect of channel length modulation becomes more significant with the reduction in channel length. As us, the textbook formula, i.e., Equation 10.13, using the λ parameter to describe the effect of channel length modulation, and also employed in the basic Spice MOS transistor models (Level 2 and Level 3), is not correct. Equation 10.13 implies that for a given drain voltage, the ratio between d and L_{eff} is always the same; however, clearly this cannot be true for different channel lengths. With longer transistor channels, the effect of channel length modulation is less significant and the output resistance actually increases with the channel length L .

As Equation 10.19 indicates, the channel length modulation can be reduced only by increasing the impurity concentration N in the silicon. Without a significant increase in impurities, there would be a punch-through effect in a short channel transistor. Punch-through occurs when the depletion layer thickness d becomes equal to the channel length L , as illustrated in Figure 10.6. Unfortunately, a larger impurity concentration in the substrate leads to large parasitic capacitances, and therefore a reduction in the transistor size does not result in a similar reduction of the parasitic capacitances. For example, it is interesting to note that a significant reduction in the size of transistors in the last decade from $0.3 \mu\text{m}$ to $0.05 \mu\text{m}$ did not result in any noticeable increase in the computer clock frequencies. Of course, there are also other factors that are limiting clock frequencies. One of the most important limitations is an ability to dissipate heat, since power dissipation in computer chips is proportional to clock frequency.

10.2.3.2 Effect of Carrier Velocity Saturation

Most of the textbook equations are derived with an assumption that the mobility of carriers is constant and independent of the electrical field. Actually, in silicon, the maximum carrier velocity for both electrons and holes is $v_{sat} = 10^7 \text{ cm/s} = 10^{11} \mu\text{m/s}$. As indicated in Figure 10.7, the critical field in which the velocity saturates is about $1 \text{ V}/\mu\text{m}$. Note that the channel length in modern MOS transistors is significantly smaller than $1 \mu\text{m}$. Therefore, even reducing the supply voltage below 2 V produces an average electrical field in MOS transistors, which is significantly larger than $1 \text{ V}/\mu\text{m}$. Fortunately, large electrical fields exist near the drain and significantly smaller electrical fields are near the source, and these smaller electric fields shape the transistor current characteristics, so Equations 10.13 are

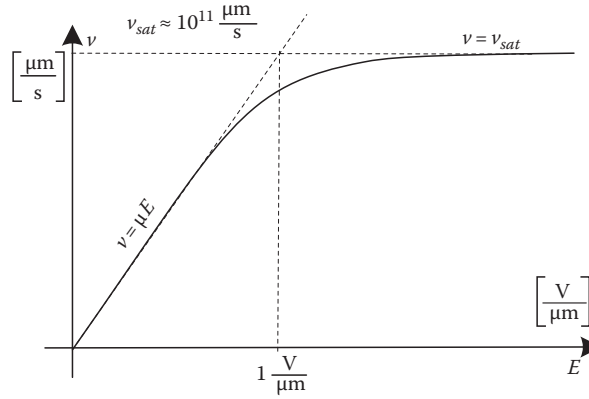


FIGURE 10.7 Carriers' velocity as a function of the electrical field in silicon.

still in use even if the drain current in the current saturation region no longer is described by a quadratic equation:

$$I_D \sim (V_{GS} - V_{th})^n \quad (10.21)$$

where n has value between 1 and 2.

10.2.3.3 Carrier Mobility Degradation near the Surface

A key feature of MOS transistor operation is the fact that most of the current flows near the silicon surface and, as a result of crystal imperfections, carrier mobilities near the surface are reduced. This effect becomes even more significant with increased gate voltage when a large electrical field is created perpendicular to the direction of current flow. As a consequence, with larger gate voltages, a larger number of carriers are accumulated near the surface. However, these carriers are moving slower due to surface mobility degradation and the fact that the drain current is not increasing as fast as would be predicted by Equation 10.13 with a quadratic relationship. In addition to mobility degradation in the transverse electric field, i.e., the gate voltage, there is a strong degradation due to the longitudinal electric field, i.e., the drain voltage. As a result, experimental characteristics for short-channel MOS transistors exhibit almost linear dependence with gate voltage (Figure 10.8).

AQ1

10.2.3.4 Subthreshold Conduction

As illustrated in Section 2.1, as the gate voltage increases, carriers gradually accumulate near the surface. The assumption, that suddenly there is a strong surface inversion where the concentration of minority carriers near the surface is exactly the same as the concentration of carriers in the substrate, is very artificial. Below and near the threshold, the drain current in a MOS transistor is described by an exponential relationship:

$$I_D = I_{ON} \exp\left(\frac{V_{GS} - V_{th} - \eta V_T}{\eta V_T}\right) \quad \text{for } V_{GS} < V_{th} + \eta V_T \quad (10.22)$$

where

$$I_{ON} = \frac{K}{2} (\eta V_T)^2 \quad (10.23)$$

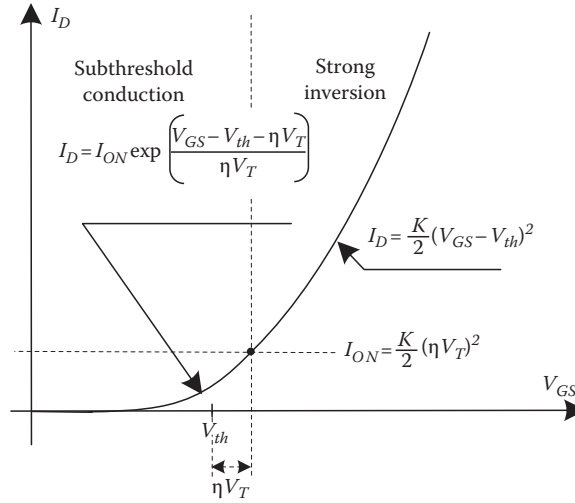


FIGURE 10.8 Drain current versus the gate-source voltage in subthreshold conduction and the strong inversion regions.

$V_T = kT/q$ is the thermal potential and η depends on the device geometry and lies between 1.5 and 2.5. The subthreshold conduction is the reason why MOS transistors are actually never completely turned OFF and there is always some current leak through MOS transistors. When the popular CMOS technology was first developed, one of the underlying assumptions was that along the power path, there would never be even one MOS transistor in the OFF state, so power would not be taken from power supply. While a transistor can be in the OFF state, there is a leakage current caused by the subthreshold conduction in CMOS VLSI circuits, which can be very significant in situations where the number of MOS transistors exceeds one billion.

10.3 Junction Field Effect Transistor

The principle of operation for a JFET is quite different than that of a MOS transistor [S05, N06]. The current flows through a thin semiconductor layer that is surrounded by a gate made of semiconductor material of the opposite type, as shown in Figure 10.9. The gate-channel (source) junction is biased in the reverse direction, so there is no gate current. The thickness of the depletion regions controls current flow between source S and drain D. The thickness of the depletion regions is the function of the gate voltage:

$$d = \sqrt{2 \frac{\epsilon_o \epsilon_{Si} V_{channel}}{qN}} \quad (10.24)$$

As indicated in Figure 10.9a, the thickness of the depletion layer d is constant only if there is no voltage applied between source and drain. With applied drain voltage, this issue becomes much more complicated because the channel width is a nonlinear function of distance, and there is a nonlinear voltage drop along the channel length. As a consequence, there is a different gate-channel voltage and a different depletion layer thickness d along the channel, as shown in Figure 10.9b. With any further increase in the gate voltage, the channel is pinched off and only a depletion layer exists between point x in Figure 10.9c and the drain D. In that region, the operation of a JFET is very interesting. In the n-type JFET case shown in Figure 10.9, electrons in the channel close to the source are being pushed away by the negative gate voltage. This is why the depletion layer is formed. But the large positive drain voltage creates an electric field between point x and drain, and this electric field swipes all electrons that

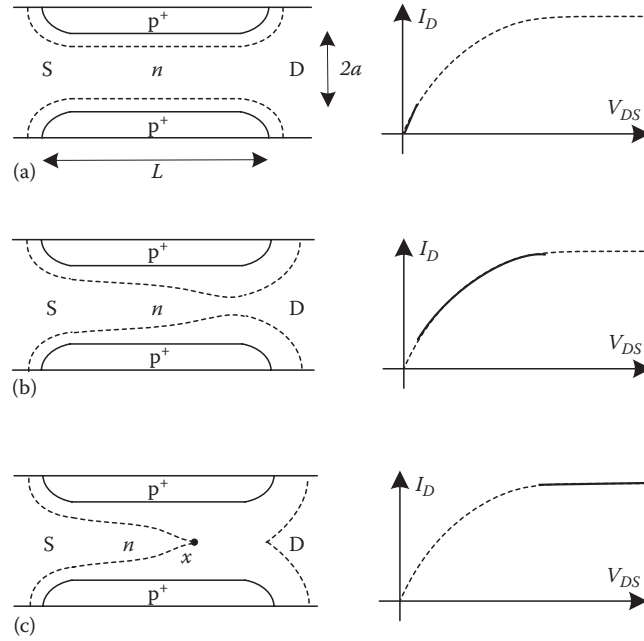


FIGURE 10.9 Cross sections of an n-channel JFET with fixed gate voltage and different drain-source voltages. Characteristics on the right side of the figure show the corresponding modes of operations.

have reached point x through the pinch-off region near the drain. Interestingly, when in this mode of operation, the drain voltage does not have a direct effect on the drain current, and the drain current is determined by the triangular shape of the channel region to the left of point x . Of course, in a manner analogous to that of the MOS transistor, the JFET also experiences second-order effects from the channel-length modulation. With an increase in drain voltage, the point x is moved to the left, which reduces the effective resistance of the “triangular-like channel region” and results in a slight increase in drain current.

The geometry of the JFET is usually more complicated than the one shown in Figure 10.9. Also, in most JFET devices, there is a nonlinear impurity distribution in the channel. Therefore, a derivation of the current-voltage characteristics would be either too simplistic or too complicated. As a consequence, various approximation formulas are being used. The most popular equations for determining JFET drain currents are

$$\text{if } V_{GS} < V_P \quad I_D = 0 \quad (10.25)$$

$$\text{if } V_{GS} \geq V_P \quad I_D = \begin{cases} 2I_{DSS} \left[\left(\frac{V_{GS}}{V_P} - 1 \right) \frac{V_{DS}}{V_P} - 0.5 \left(\frac{V_{DS}}{V_P} \right)^2 \right] (1 + \lambda V_{DS}) & \text{for } V_{DS} \leq V_{GS} - V_P \\ I_{DSS} \left(\frac{V_{GS}}{V_P} - 1 \right)^2 (1 + \lambda V_{DS}) & \text{for } V_{DS} \geq V_{GS} - V_P \end{cases} \quad (10.26)$$

where V_P is the pinch-off voltage and I_{DSS} is the drain current for the case in which the gate is connected with source and a relatively large drain voltage is applied. Both V_P and I_{DSS} can be related to the JFET geometry as shown in Figure 10.9.

the pinch-off voltage V_p can be calculated as

$$V_p = \frac{qa^2N}{2\epsilon\epsilon_0} \quad (10.27)$$

and the drain current I_{DSS} can be calculated from Ohms law by dividing V_p by the resistance of the channel between source and drain. the nonlinear distribution of the resistance is usually assumed to be many times larger than the resistance without the applied drain voltage, as shown in Figure 10.9a:

$$R_{channel} = 3 \frac{\rho L}{2aW} \quad (10.28)$$

where a is half of the thickness of the channel, L is the channel length, and W is the channel width:

$$I_{DSS} = \frac{V_p}{3} \frac{2aW}{\rho L} = V_p \frac{2aW}{3L} \sigma = V_p \frac{2aW}{3L} q\mu N \quad (10.29)$$

Example 10.3

An n-channel junction JFET has a uniformly doped channel 2 μm thick, 20 μm wide, and 20 μm long with $N_D = 0.5 \cdot 10^{16} \text{ cm}^{-3}$. Determine the pinch-off voltage as well as the drain current for $V_{GS} = -1 \text{ V}$ and $V_{DS} = 10 \text{ V}$. Assume $\mu_n = 500 \text{ cm}^2/\text{Vs}$. Neglect the channel-length modulation effects.

Since the gate concentration is not given, the increase in potential is neglected:

$$V_p = \frac{qa^2N_D}{2\epsilon\epsilon_0} = \frac{1.6 \cdot 10^{-19} (10^{-4})^2 0.5 \cdot 10^{16}}{2 \cdot 11.8 \cdot 8.85 \cdot 10^{-14}} = 3.83 \text{ V}$$

$$I_{DSS} = V_p \frac{2aW}{3L} q\mu N = \frac{2 \cdot 10^{-4} \cdot 20 \cdot 10^{-4}}{20 \cdot 10^{-4}} (1.6 \cdot 10^{-19} \cdot 500 \cdot 0.5 \cdot 10^{16}) \frac{3.83}{3} = 1.0 \text{ mA}$$

For $V_{GS} = -1 \text{ V}$ in a n-channel JFET, the pinch-off voltage must be negative, e.g., $V_p = -3.83 \text{ V}$, and neglecting the channel length modulation effect $\lambda = 0$:

$$I_D = I_{DSS} \left(\frac{V_{GS}}{V_p} - 1 \right)^2 = 1.0 \text{ mA} \left(\frac{-1 - (-3.83)}{-3.83} \right)^2 = 0.55 \text{ mA}$$

Note the similarities in Equations 10.13 for the MOS transistor and Equation 10.26 for the JFET. Actually, if the pinch-off voltage V_p and the drain initial current I_{DSS} are replaced by

$$V_{th} = V_p \quad \text{and} \quad K = \frac{2I_{DSS}}{V_p^2} \quad (10.30) \quad \text{AQ2}$$

then, to calculate the drain current in a JFET, we need not employ the JFET equations (10.26) and use instead the well-known equations for MOS transistors working in the depletion mode:

$$I_D = \begin{cases} K[(V_{GS} - V_{th})V_{DS} - 0.5V_{DS}^2](1 + \lambda V_{DS}) & \text{for } V_{DS} \leq V_{GS} - V_{th} \\ 0.5K(V_{GS} - V_{th})^2(1 + \lambda V_{DS}) & \text{for } V_{DS} \geq V_{GS} - V_{th} \end{cases} \quad (10.31)$$

Because current flow in JFETs is far from the surface, JFETs have a significantly smaller $1/f$ noise level and they are a preferred choice for low-noise amplifiers. Another advantage of the JFET is that it is

relatively safe to exceed gate-source or gate-drain break voltages. In the case of MOS transistors, large gate voltages may result in a permanent break-through of the oxide, leading to the destruction of the transistor. It is however very difficult to integrate several JFETs into one chip. Another disadvantage of the JFET is that the gate voltage should always have a polarity opposite that of the drain voltage, and this makes it almost impossible to fabricate digital circuits using JFETs.

10.4 Static Induction Transistor

Static induction devices were invented by J. Nishizawa [NTS75]. The device has characteristics similar to that of the vacuum triode. Its fabrication is relatively difficult and Japan is actually the only country where a family of static induction devices was successfully fabricated [W99].

The cross section of the static induction transistor (SIT) is shown in Figure 10.10. In this n-channel structure, the gate is biased with a negative potential and the drain has a significantly large positive potential. There are two reverse-biased junctions; one between gate and source and second between gate and drain. Because n^- regions have a very low concentration of impurities (10^{14} cm^{-3} or below) with very small applied voltages or simply the junctions' built-in potential, these n^- regions are depleted of carriers. As a consequence, gate-drain voltages form a relatively complex potential surface. Samples of such surfaces for a gate voltage equal -10 V and a drain voltage equal $+50 \text{ V}$ are shown in Figure 10.11. Note that gate and drain voltages create opposite electric fields near the source. With an increase in gate voltage, the height of the potential barrier increases, as shown in Figure 10.11, while the larger drain

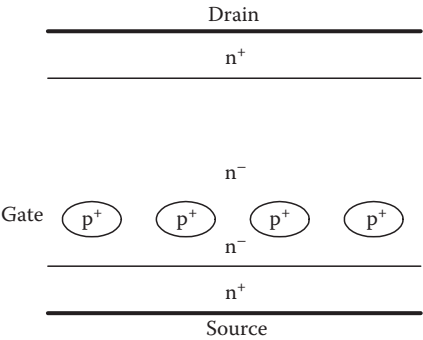


FIGURE 10.10 Cross section of the static induction transistor.

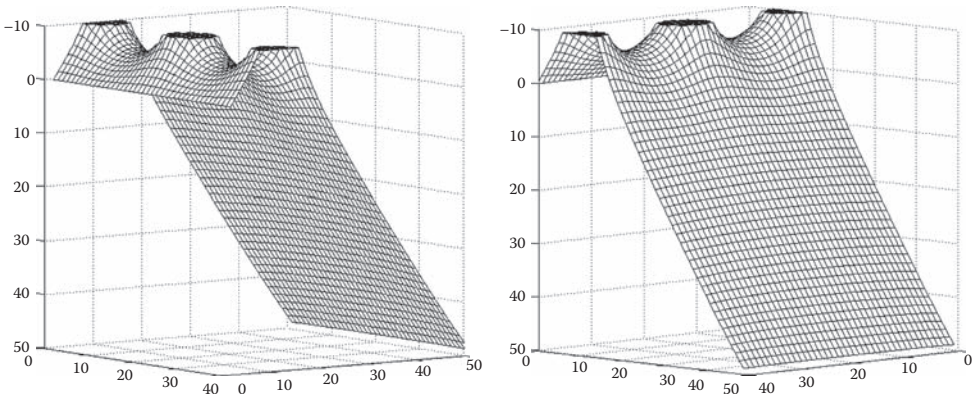


FIGURE 10.11 Potential distribution in the static induction transistor with a gate voltage of -10 V and a drain voltage of $+50 \text{ V}$.

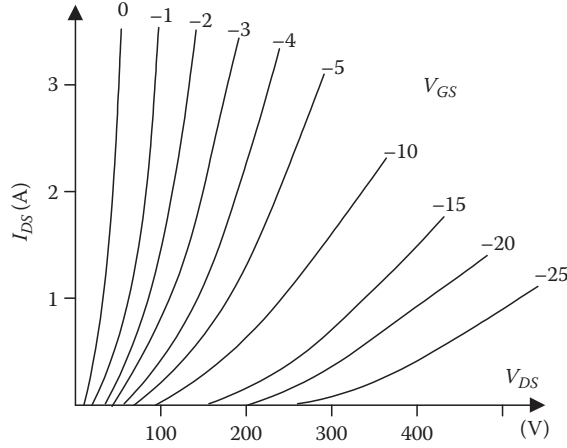


FIGURE 10.12 Typical current-voltage characteristics for the SIT.

voltage leads to a lowering of the potential barrier. Typical current-voltage characteristics for the SIT are shown in Figure 10.12.

10.4.1 Theory of SIT Operation for Small Currents

Consider first a derivation of the formula for the one-dimensional electron current flow through a potential with a parabolic shape. In the n -channel device, the electron current is described by a differential equation that includes both drift and diffusion of carriers [N02, WSM92]:

$$J_n = -qn(x)\mu_n \frac{d\phi(x)}{dx} + qD_n \frac{dn(x)}{dx} \quad (10.32) \quad \text{AQ3}$$

where $D_n = \mu_n V_T$ and $V_T = \frac{kT}{q}$. By multiplying both sides of the equation by [PW80]

$$\exp\left(-\frac{\phi(x)}{V_T}\right) \quad (10.33)$$

and rearranging

$$J_n \exp\left(-\frac{\phi(x)}{V_T}\right) = qD_n \frac{d}{dx} \left[n(x) \exp\left(-\frac{\phi(x)}{V_T}\right) \right] \quad (10.34)$$

After integration from source x_s to drain x_D :

$$J_n = qD_n \frac{n(x_D) \exp\left(-\frac{\phi(x_D)}{V_T}\right) - n(x_s) \exp\left(-\frac{\phi(x_s)}{V_T}\right)}{\int_{x_s}^{x_D} \exp\left(-\frac{\phi(x)}{V_T}\right) dx} \quad (10.35)$$

By inserting

$$\begin{aligned}\varphi(x_S) &= 0 & n(x_S) &= N_S \\ \varphi(x_D) &= V_D & n(x_D) &= N_D\end{aligned}\quad (10.36)$$

Equation 10.35 reduces to

$$J_n = \frac{qD_n N_S}{\int_{x_S}^{x_D} \exp\left(-\frac{\varphi(x)}{V_T}\right) dx} \quad (10.37)$$

Equation 10.37 is very general and it describes the current flow over a potential barrier with a shape given by $\varphi(x)$. This equation can be used not only in SIT devices but also in bipolar transistors or it can be used to calculate the subthreshold conduction in a MOS transistor.

Note that because of this exponential relationship, only the shape of the potential distribution near the top of the potential barrier is important, and in the SIT, this shape can be approximated (see Figure 10.11) by quadratic equations along the x direction and across the y direction of the channel:

$$\varphi(y, x) = \Phi \left(\left(\frac{x}{L} \right)^2 - \left(\frac{y}{W} \right)^2 \right) \quad (10.38)$$

where L is the effective channel length, W is the effective channel width, and Φ is the height of the potential barrier in the center of the channel. Using (10.38) and integrating (10.37) first along the channel and then across it, leads to a simple formula for the drain current of a SIT as a function of the height of potential barrier:

$$I_D = qD_p N_S Z \frac{W}{L} \exp\left(\frac{\Phi}{V_T}\right) \quad (10.39)$$

where Φ is the potential barrier height in reference to the source potential, and N_S is the electron concentration at the source, the W/L ratio describes the shape of the potential saddle in the vicinity of the barrier, and Z is the length of the source strip.

Since the barrier height Φ can be a linear function of gate and drain voltages:

$$I_D = qD_p N_S Z \frac{W}{L} \exp\left(\frac{aV_{GS} + bV_{DS} + \Phi_0}{V_T}\right) \quad (10.40)$$

The actual characteristics of the SIT device on a logarithmic scale are shown in Figure 10.13. Indeed, for a small current range, the characteristics have an exponential character, but this is not true for large currents where the space charge for moving electrons affects the potential distribution.

10.4.2 Theory of SIT for Large Currents

For large current levels, the SIT current is controlled by the space charge of the moving carriers [PW81, MW83]. In the one-dimensional case, the potential distribution is described by the Poisson equation:

$$\frac{d^2\varphi}{dx^2} = -\frac{\rho(x)}{\epsilon_{Si}\epsilon_0} = \frac{I_{DS}}{A v(x)} \quad (10.41)$$

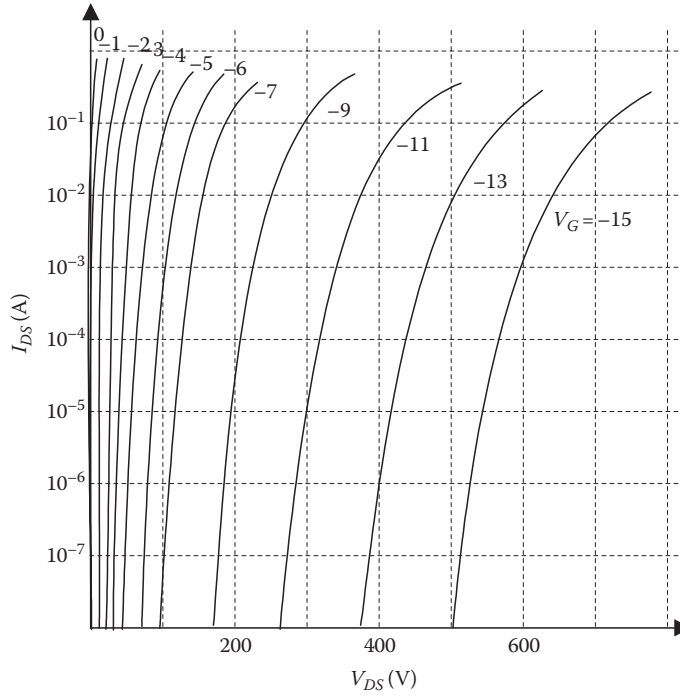


FIGURE 10.13 Characteristics of the static induction transistor for a small current range.

where A is the effective device cross section and $v(x)$ is the carrier velocity. For a small electrical field, $v(x) = \mu E(x)$ and the solution of (10.41) is

$$I_D = \frac{9}{8} V_{DS}^2 \epsilon_{Si} \epsilon_o \frac{A}{L^3} \quad (10.42)$$

For a large electrical field, $v(x) = v_{sat}$ and

$$I_D = 2 V_{DS} v_{sat} \epsilon_{Si} \epsilon_o \frac{A}{L^2} \quad (10.43)$$

where L is the channel length and $v_{sat} \approx 10^{11}$ $\mu\text{m/s}$ is the carrier saturation velocity. In practical devices, the current–voltage relationship is described by an exponential relationship (10.9) for small currents, a quadratic relationship (10.11), and finally for large voltages, by an almost linear relationship (10.12). SIT characteristics drawn in linear and logarithmic scales are shown in Figures 10.12 and 10.13, respectively.

10.4.3 Bipolar Mode of Operation of the SIT

The bipolar mode of operation for the SIT was first reported in 1977 [NW77a]. Several complex theories for the bipolar mode of operation were developed [NTT86, NOC82], but actually the simple formula (10.37) works well not only for the typical mode of SIT operation, but also for the bipolar mode as well. Furthermore, the same formula works very well for classical bipolar transistors. Typical characteristics of the SIT, operating in normal and in bipolar modes, are shown in Figure 10.14.

In a SIT, a virtual base is formed, not by impurity doping but rather by a potential barrier that is induced by the gate voltage. As a consequence, in the bipolar mode of operation, the SIT may have a

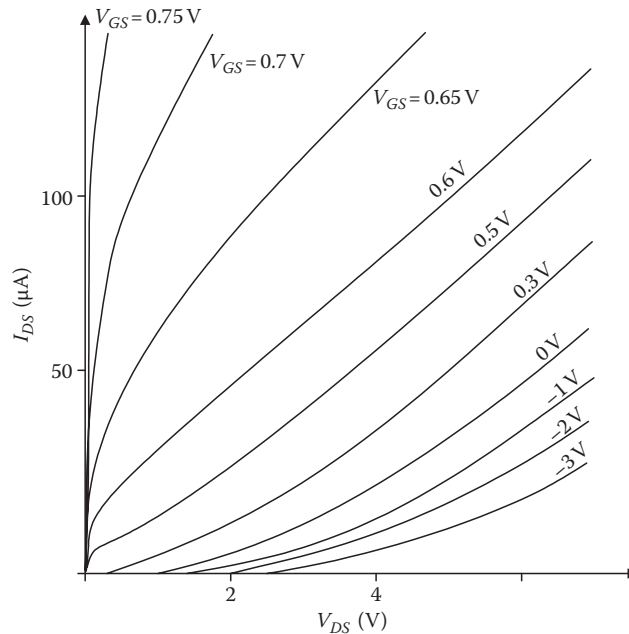


FIGURE 10.14 SIT transistor characteristic operating in both normal and the bipolar modes, $I_D = f(V_{DS})$ with V_{GS} as parameter.

very large current gain β . Also, the SIT operates with a very low level of impurity concentration and its parasitic capacitances are very low. When a bipolar transistor in integrated injection logic (I^2L) was replaced by a SIT, the time-delay product of such a device was reduced almost 100 times [NW77]. Such a drastic improvement in the power-delay product is possible because the SITL structure has a significantly smaller junction parasitic capacitance, and, furthermore, the voltage swing is reduced.

Another interesting application of the SIT is a replacement for Schottky diodes in the protection of a bipolar transistor against saturation, leading to a much faster switching time. The use of a SIT [WMS84] is more advantageous than that obtained with a Schottky diode since it does not require additional area on a chip and it does not introduce additional capacitance between the base and the collector.

10.5 Lateral Punch-Through Transistor

The fabrication of SITs is a very challenging endeavor. The channel area requires very low impurity concentration ($N < 10^{14} \text{ cm}^{-3}$), and, at least, a part of the channel near the source has to be made using an epitaxial layer, as shown in Figure 10.10, which should be about 100 times more pure than that which is considered an epitaxial layer with low impurity concentration ($N = 10^{16} \text{ cm}^{-3}$). The second difficult issue is the creation of a buried gate region. With high temperature epitaxial growth and subsequent diffusion processes, it is extremely difficult to concentrate gate impurity in one place without spreading it into the channel area and actually closing the channel. Only a couple of Japanese companies (Yamaha and Sony) were able to develop a fabrication process for SIT devices.

The lateral punch-through transistor (LPTT), which has characteristics that are similar to the SIT, can be fabricated with a very simple process [WJ82]. The cross section of the LPTT device is shown in Figure 10.15, and its characteristics are shown in Figure 10.16. The LPTT device, in contrast to SIT device, must use the same type of the impurity in the channel as is used in the gate. For the n-type channel, p^- must be used instead of n^- . With an increase in positive drain voltage, the thickness of the drain depletion region increases. Once the depletion layer reaches the source, the punch-through current will start to flow between source and drain. The current can be stopped by applying a negative voltage on the

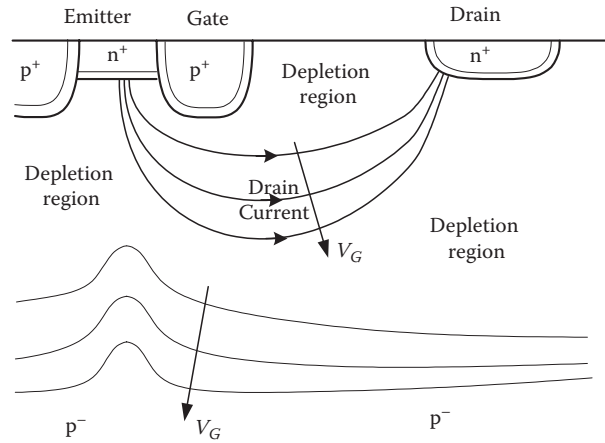


FIGURE 10.15 Cross section of a lateral punch-through transistor.

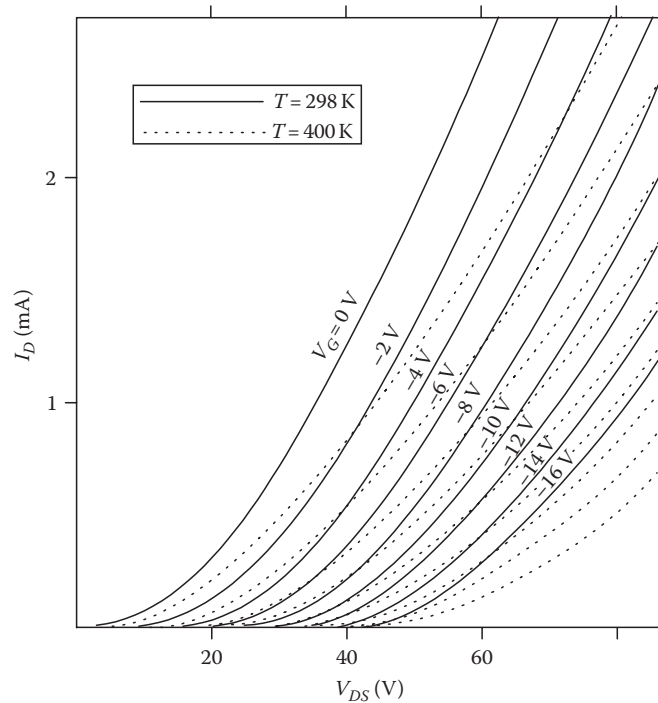


FIGURE 10.16 Characteristics of the LPT transistor.

gate. Eventually, the electrical field near the source will be affected by both the gate and source voltages. Because of its proximity to the source, the gate voltage can be much more effective in controlling the device current. Typical current–voltage characteristics are shown in Figure 10.16.

The concept of the LPTT can be extended to a punch-through MOS transistor (PTMOS) where the current is controlled by the MOS gate [W83a, WJF84]. The principle of operation of such a device is shown in Figure 10.17. In this device, instead of the implanted p-type gate, the gate is formed by an accumulation p-type layer under the MOS gate, as shown in Figure 10.17a, and the punch-through current can be controlled by a potential applied to the adjacent p-type region, which is normally

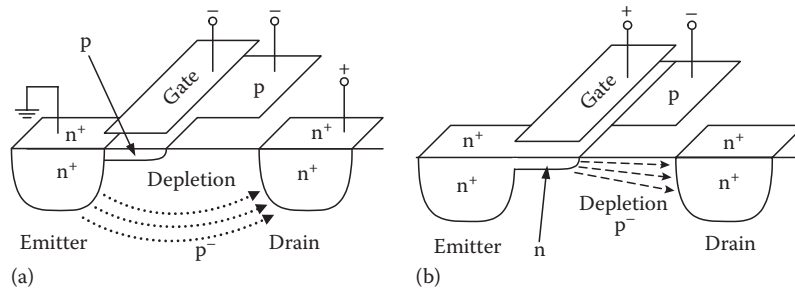


FIGURE 10.17 Punch-through MOS transistor: (a) transistor in the punch-through mode for a negative gate potential; and (b) transistor in the on-state for a positive gate potential.

biased with a large negative potential. With positive voltage on the MOS gate, the n-type accumulation layer is formed under the gate and the current may flow easily from source to drain, as indicated in Figure 10.17b. Note that in typical CMOS technology, in order to prevent a punch-through phenomenon between source and drain, a large impurity doping in the substrate must be used, which in turn leads to larger parasitic capacitances. In a MOS transistor with shorter channels, a larger impurity concentration must be used to prevent the punch-through phenomenon, and parasitic capacitances are also larger. The PTMOS takes advantage of the punch-through phenomenon and the substrate impurity concentration can be very low, which leads to very small parasitic capacitances and a significant reduction in power consumption for digital circuits operating with very large clock frequencies.

The PTMOS transistor has several advantages over the traditional MOS transistor:

1. The gate capacitance is very small.
2. Carriers are moving with a velocity close to saturation velocity.
3. The substrate doping is much lower and the existing depletion layer leads to a much smaller drain capacitance.

The device operates in a fashion that is similar to that of the MOS transistor in subthreshold conditions, but this process occurs at much higher current levels. Such a “bipolar mode” of operation may have many advantages in VLSI applications.

10.6 Power MOS Transistors

MOS transistors have a relatively small transconductance in comparison to bipolar transistors. Therefore, in power electronic applications, the integrated device structures usually should consist of thousands of transistors connected in parallel. There are two types of power MOS transistors: VMOS, shown in Figure 10.18a, and DMOS, shown in Figure 10.18b. In the VMOS structure, MOS gates and channels are formed on etched surfaces. This way many transistors can be efficiently connected together. VMOS uses the silicon surface very efficiently.

The DMOS transistor does not use the chip area as efficiently as the VMOS transistor, but it can be fabricated with much larger breakdown voltages. In DMOS, a fragile MOS structure is protected from a large electric field by a concept borrowed from the SIT [NTS75] and the high-voltage Schottky diode [W83]. The n^- area under the gate, as illustrated in Figure 10.18b, is depleted from carriers, and neighboring p-type regions work as electrostatic screens as is done in SIT devices. As a result, this transistor may withstand much larger drain voltages and also the effect of channel-length modulation is significantly reduced. The latter effect leads to larger output resistances of the transistor. Therefore, the drain current is less sensitive to drain voltage variations. In fact, the DMOS structure can be considered as a composition of the MOS transistor and the SIT, as is shown in Figure 10.19.

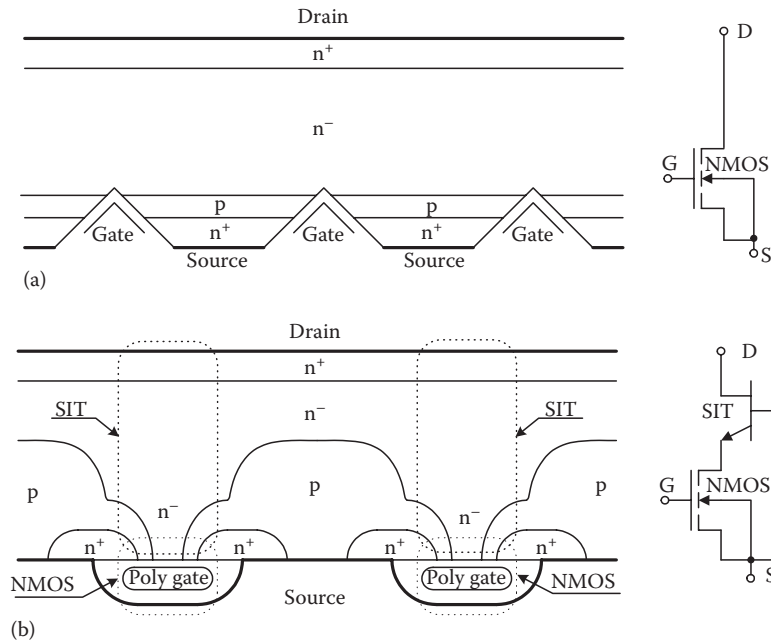


FIGURE 10.18 Cross section of power MOS transistors: (a) VMOS and (b) DMOS.

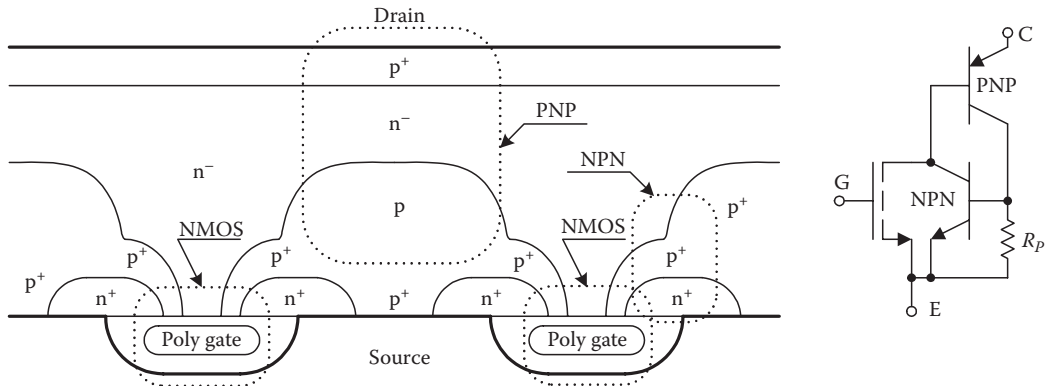


FIGURE 10.19 Cross section of an IGBT and its internal equivalent circuit.

A major disadvantage of power MOS transistors is the relatively larger drain series resistance and much smaller transconductance in comparison to bipolar transistors. Both of these parameters can be improved dramatically if the n^+ layer near the drain is replaced by p^+ layer as is shown in Figure 10.19. A way an integrated structure is being built where its equivalent diagram consists of a MOS transistor integrated with a bipolar transistor, as shown in Figure 10.19. Such a structure has a transconductance that is β times larger, where β is the current gain of the PNP bipolar transistor, and a much smaller series resistance due to the conductivity modulation effect caused by holes injected into the lightly doped drain region. Such device is known as insulated gate bipolar transistor (IGBT). An IGBT can work with large currents and voltages. Its main disadvantage is a large switching time that is limited primarily by the poor switching performance of the bipolar transistor. Another difficulty is related to a possible latch-up action of four layer $n^+pn^-p^+$ structure. This undesired effect could be suppressed by using a heavily doped p^+ region in the base of the NPN structure, which leads to a significant reduction in the current

gain of this parasitic transistor, shown in Figure 10.19. The gain of the PNP transistor must be kept large so the transconductance of the entire device can be large. IGBT transistors may have breakdown voltages over 1000 V, with turn-off times in the range from 0.1 to 0.5 μs . In addition, they may operate with currents above 100 A with a forward voltage drop of about 3 V.

AQ4 References

- [N02] Kwok K. Ng, *The Complete Guide to Semiconductor Devices*, Wiley-IEEE Press, 2002.
- [N06] D. Neamen, *An Introduction to Semiconductor Devices*, McGraw-Hill, 2006.
- [E97] R. Enderlein, *Fundamentals of Semiconductor Physics and Devices*, World Scientifics, 1997.
- [WJF84] B. M. Wilamowski, R. C. Jaeger, and J. N. Fordemwalt, Buried MOS transistor with punch-through, *Solid State Electronics*, **27**, 8/9, 811–815, 1984.
- [WMS84] B. M. Wilamowski, R. H. Mattson, and Z. J. Staszak, The SIT saturation protected bipolar transistor, *IEEE Electron Device Letters*, **5**, 263–265, 1984.
- [T99] Y. Tsividis, *Operation and Modeling the MOS Transistor*, McGraw-Hill, 1999.
- [R99] D. J. Roulston, *Introduction to the Physics of Semiconductor Devices*, Oxford University Press, 1999.
- [S05] B. Streetman, *Solid State Electronic Devices*, Prentice Hall, 2005.
- [NTS75] J. Nishizawa, T. Terasaki, and J. Shibata, Field-effect transistor versus analog transistor (static induction transistor), *IEEE Transactions on Electron Devices*, **22**, 4, 185–197, 1975.
- [NW77] J. Nishizawa and B. M. Wilamowski, Integrated logic—State induction transistor logic, *International Solid State Circuit Conference*, Philadelphia, USA, pp. 222–223, 1977.
- [NW77a] J. Nishizawa and B. M. Wilamowski, Static induction logic—A simple structure with very low switching energy and very high packing density, *International Conference on Solid State Devices*, Tokyo, Japan, pp. 53–54, 1976; and *Journal of Japanese Society of Applied Physics*, **16-1**, 158–162, 1977.
- [NTT86] Y. Nakamura, H. Tadano, M. Takigawa, I. Igarashi, and J. Nishizawa, Experimental study on current gain of BSIT, *IEEE Transactions on Electron Devices*, **33**, 6, 810–815, 1986.
- [MW83] R. H. Mattson and B. M. Wilamowski, Punch-through devices operating in space-charge-limited modes, *IEEE International Workshop on the Physics of Semiconductor Devices*, Delhi, India, December 5–10, 1983.
- [NOC82] J. Nishizawa, T. Ohmi, and H. L. Chen, Analysis of static characteristics of a bipolar-mode SIT (BSIT), *IEEE Transactions on Electron Devices*, **29**, 8, 1233–1244, 1982.
- [WSM92] B. M. Wilamowski, Z. J. Staszak, and R. H. Mattson, An electrical network approach to the analyses of semiconductor devices, *IEEE Transactions on Education*, **35**, 2, 144–152, 1992.
- [W99] B. M. Wilamowski, High speed, high voltage, and energy efficient static induction devices, *12 Symposium of Static Induction Devices—SSID'99*, Tokyo, Japan, pp. 23–28, April 23, 1999.
- AQ5 [O82] T. Ohmi, Punching through device and its integration—Static induction transistor, *IEEE Transactions on Electron Devices*, **27**, 536–545, 1980.
- [WJ82] B. M. Wilamowski and R. C. Jaeger, The lateral punch-through transistor, *IEEE Electron Device Letters*, **3**, 10, 277–280, 1982.
- [PW80] P. Plotka and B. M. Wilamowski, Interpretation of exponential type drain characteristics of the SIT, *Solid-State Electronics*, **23**, 693–694, 1980.
- [PW81] P. Plotka and B. M. Wilamowski, Temperature properties of the static induction transistor, *Solid-State Electronics*, **24**, 105–107, 1981.
- [W83] B. M. Wilamowski, Schottky diodes with high breakdown voltage, *Solid-State Electronics*, **26**, 5, 491–493, 1983.
- [W83a] B. M. Wilamowski, The punch-through transistor with MOS controlled gate, *Physica Status Solidi (a)*, **79**, 631–637, 1983.

