MOSFETs

19.27 Metal Oxide Semiconductor FET (MOSFET)

The main drawback of *JFET* is that its gate *must* be reverse biased for proper operation of the device *i.e.* it can only have negative gate operation for *n*-channel and positive gate operation for *p*-channel. This means that we can *only* decrease the width of the channel (*i.e.* decrease the *conductivity of the channel) from its zero-bias size. This type of operation is referred to as **_depletion-mode_ operation. Therefore, a *JFET* can only be operated in the depletion-mode. However, there is a field effect transistor (*FET*) that can be operated to enhance (or increase) the width of the channel (with consequent increase in conductivity of the channel) *i.e.* it can have *enhancement-mode* operation. Such a *FET* is called *MOSFET*.

A field effect transistor (FET) that can be operated in the enhancement-mode is called a ***MOSFET***.

A *MOSFET* is an important semiconductor device and can be used in any of the circuits covered for *JFET*. However, a *MOSFET* has several advantages over *JFET* including high input impedance and low cost of production.

19.28 Types of MOSFETs

There are two basic types of *MOSFETs viz.*

1. Depletion-type *MOSFET* or ***D-MOSFET.*** The *D-MOSFET* can be operated in both the depletion-mode and the enhancement-mode. For this reason, a *D-MOSFET* is sometimes called depletion/enhancement *MOSFET*.
2. Enhancement-type *MOSFET* or ***E-MOSFET.*** The *E-MOSFET* can be operated *only* in enhancement-mode.

The manner in which a *MOSFET* is constructed determines whether it is *D-MOSFET* or *E-MOSFET*.

1. D-MOSFET. Fig. 19.43 shows the constructional details of *n*-channel *D-MOSFET*. It is similar to *n*-channel *JFET* except with the following modifications/remarks :

(i) The *n*-channel *D-MOSFET* is a piece of *n*-type material with a *p*-type region (called *substrate*) on the right and an *insulated gate* on the left as shown in Fig. 19.43. The free electrons ($\because$ it is *n*-channel) flowing from source to drain must pass through the narrow channel between the gate and the *p*-type region (*i.e.* substrate).

(ii) Note carefully the gate construction of *D-MOSFET*. A thin layer of metal oxide (usually silicon dioxide, SiO_2) is deposited over a small portion of the channel. A metallic gate is deposited over the oxide layer. As SiO_2 is an insulator, therefore, gate is insulated from the channel. Note that the arrangement forms a capacitor. One plate of this capacitor is the gate and the other plate is the channel with SiO_2 as the dielectric. Recall that we have a gate diode in a *JFET*.

(iii) It is a usual practice to connect the substrate to the source (*S*) internally so that a *MOSFET* has three terminals *viz source* (*S*), *gate* (*G*) and *drain* (*D*).

(iv) Since the gate is insulated from the channel, we can apply either negative or positive voltage to the gate. Therefore, *D-MOSFET* can be operated in both depletion-mode and enhancement-mode. However, *JFET* can be operated only in depletion-mode.

* With the decrease in channel width, the X-sectional area of the channel decreases and hence its resistance increases. This means that conductivity of the channel will decrease. Reverse happens if channel width increases.

** With gate reverse biased, the channel is depleted (*i.e.* emptied) of charge carriers (free electrons for *n*-channel and holes for *p*-channel) and hence the name depletion-mode. Note that depletion means decrease. In this mode of operation, conductivity decreases from the zero-bias level.

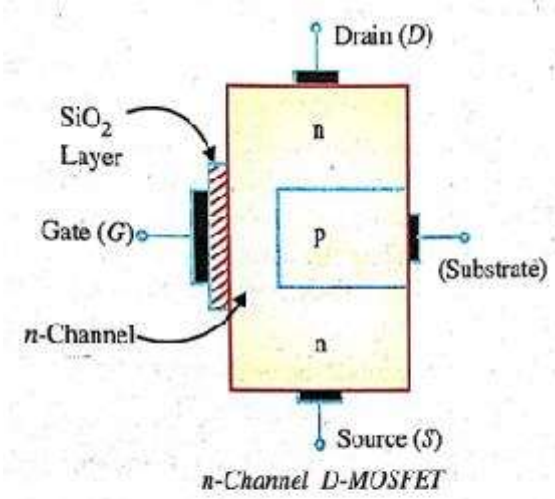


Fig. 19.43

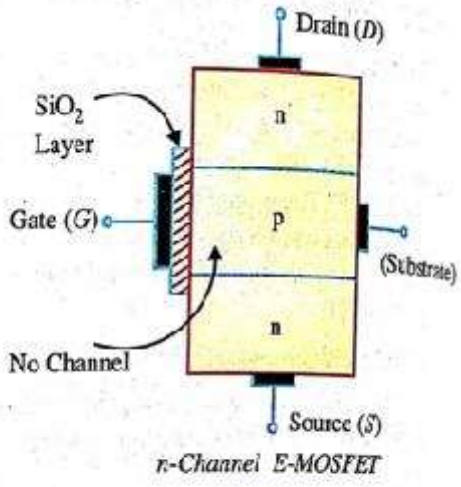


Fig. 19.44

2. E-MOSFET. Fig. 19.44 shows the constructional details of *n*-channel *E-MOSFET*. Its gate construction is similar to that of *D-MOSFET*. The *E-MOSFET* has no channel between source and drain unlike the *D-MOSFET*. Note that the substrate extends completely to the SiO_2 layer so that no channel exists. The *E-MOSFET* requires a proper gate voltage to *form* a channel (called induced channel). It is reminded that *E-MOSFET* can be operated *only* in enhancement mode. In short, the construction of *E-MOSFET* is quite similar to that of the *D-MOSFET* except for the absence of a channel between the drain and source terminals.

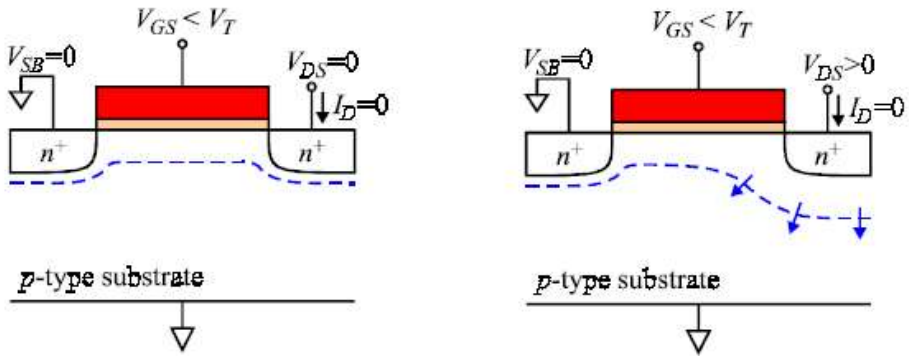
Why the name MOSFET ? The reader may wonder why is the device called *MOSFET*? The answer is simple. The SiO_2 layer is an insulator. The gate terminal is made of a metal conductor. Thus, going from gate to substrate, you have a *metal oxide semiconductor* and hence the name *MOSFET*. Since the gate is insulated from the channel, the *MOSFET* is sometimes called *insulated-gate FET (IGFET)*. However, this term is rarely used in place of the term *MOSFET*.

Operation:

- Assume an n-channel MOSFET, with Source and substrate grounded.

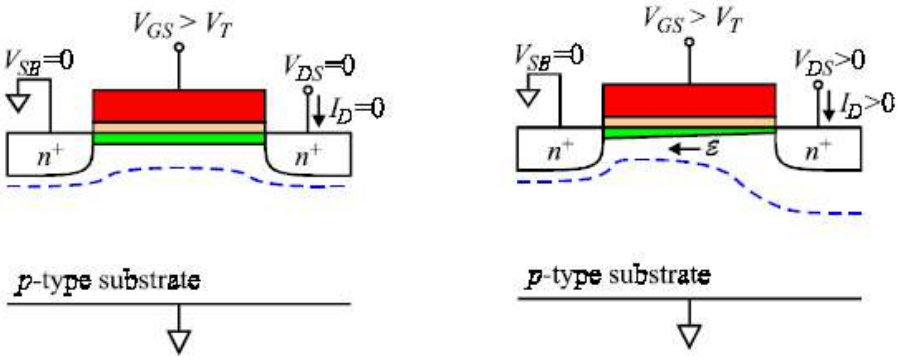
n-Channel MOSFET With $V_{GS} < V_T$

- With $V_{GS} < V_T$, there is no inversion layer present under the surface
- At $V_{DS} = 0$, the source and drain depletion regions are symmetrical
- A positive V_{DS} reverse biases the drain substrate junction, hence the depletion region around the drain widens, and since the drain is adjacent to the gate edge, the depletion region widens in the channel
- No current flows even for $V_{DS} > 0$, since there is no conductive channel between the source and drain for $V_{GS} < V_T$



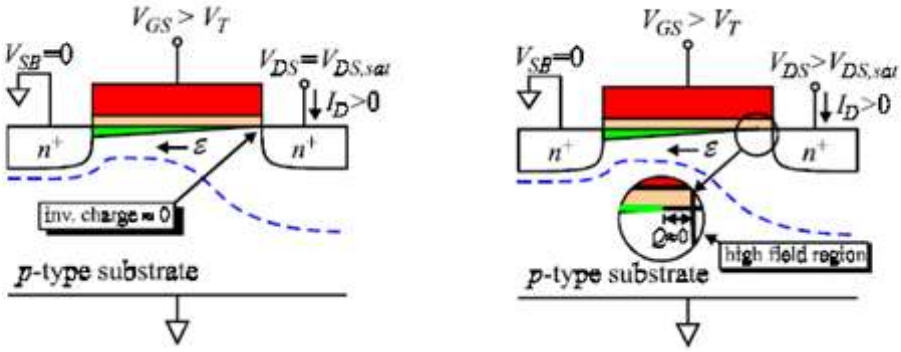
n-Channel MOSFET With $V_{GS} > V_T$, small V_{DS}

- With $V_{GS} > V_T$, a conductive channel forms under the surface - a non-zero transverse field is present
- I_D is zero for $V_{DS} = 0$ since no lateral field is present
- For $V_{DS} > 0$, transverse $\mathcal{E}$ is present and current flows
- The increased reverse bias on the drain substrate junction in contact with the inversion layer causes inversion layer density to decrease



n-Channel MOSFET With $V_{GS} > V_T$, large V_{DS}

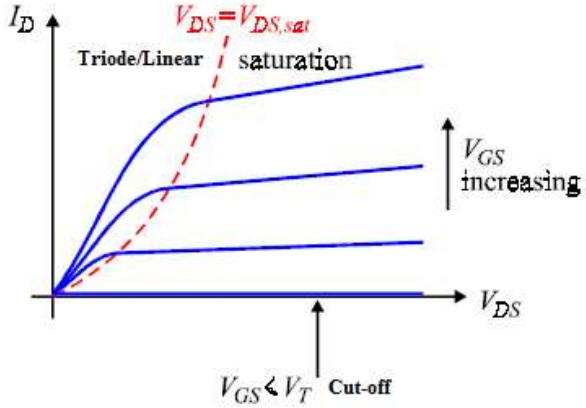
- The point at which the inversion layer density becomes very small (essentially zero) at the drain end is termed pinch-off
- The value of V_{DS} at pinchoff is denoted $V_{DS,sat}$
- Past pinchoff, further increases in lateral electric field are absorbed by the creation of a narrow high field region with low carrier density



Note: The reverse is true for the p-channel enhancement MOS transistor. When $V_{GS} = 0$ the device is "OFF" and the channel is open. The application of a negative (-ve) gate voltage to the p-type eMOSFET enhances the channels conductivity turning it "ON". Then for an p-channel enhancement mode MOSFET: $+V_{GS}$ turns the transistor "OFF", while $-V_{GS}$ turns the transistor "ON".

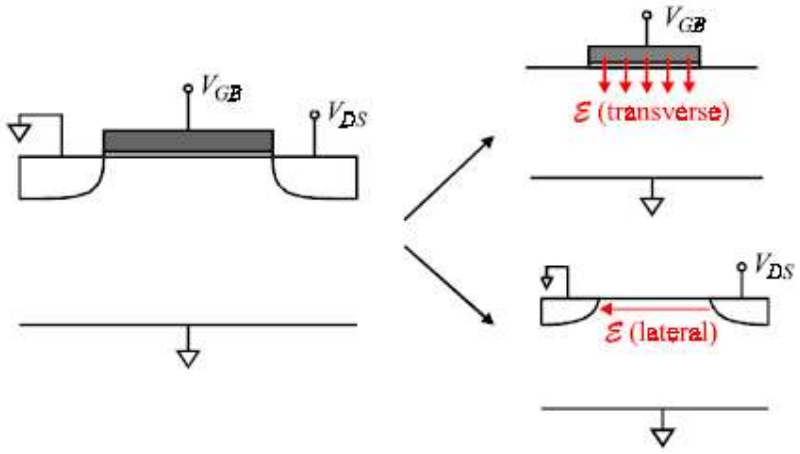
MOSFET I_D-V_{DS} Characteristic

- For $V_{GS} < V_T$, $I_D = 0$
- As V_{DS} increases at a fixed V_{GS}, I_D increases in the triode region due to the increased lateral field, but at a decreasing rate since the inversion layer density is decreasing
- Once pinchoff is reached, further V_{DS} increases only increase I_D due to the formation of the high field region
- The device starts in triode, and moves into saturation at higher V_{DS}

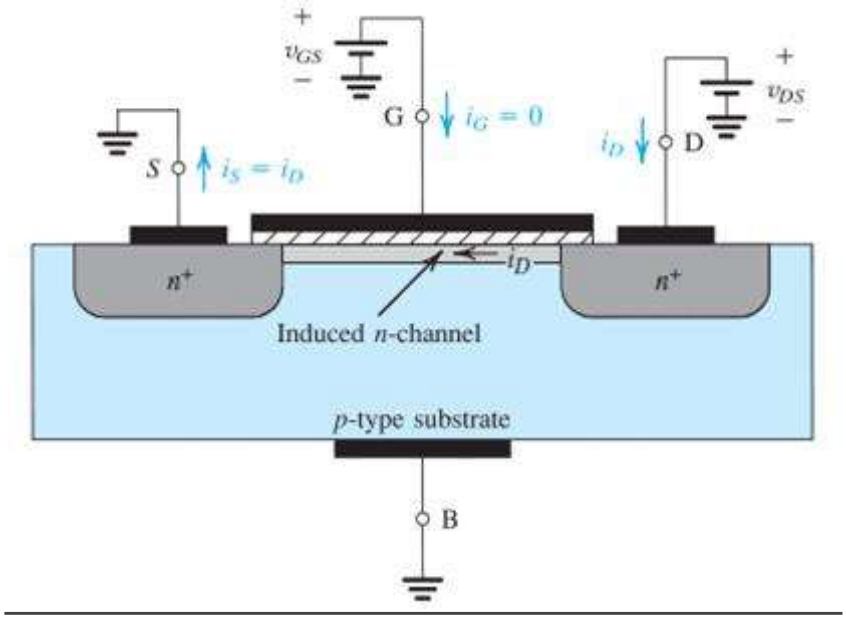


Note:

- The transverse field is caused by the potential difference between the conductive gate and the substrate. This field is supports the substrate depletion region and inversion layer
- The lateral field arises due to a non-zero source to drain potential, and is (in the simple model) the main mechanism for current flow in the MOSFET



Derivation of I_D-V_{DS} relationship:



Increasing this positive gate voltage will cause the channel resistance to decrease further causing an increase in the drain current, I_D through the channel. In other words, for an n-channel enhancement mode MOSFET: $+V_{GS}$ turns the transistor "ON", while a zero or $-V_{GS}$ turns the transistor "OFF". Thus the enhancement-mode MOSFET is equivalent to a "normally-open" switch.

The excess of v_{GS} over V_t **i.e** $(v_{GS} - V_t)$ is termed the **effective voltage** or the **overdrive voltage**

It is denoted by $v_{GS} - V_t \equiv v_{OV}$ (5.1)

We can express the magnitude of the electron charge in the channel by

$$|Q| = C_{ox}(WL)v_{OV} \quad (5.2)$$

where C_{ox}, called the **oxide capacitance**, is the capacitance of the parallel-plate capacitor per unit gate area (in units of F/m^2), W is the width of the channel, and L is the length of the channel. The oxide capacitance C_{ox} is given by

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{5.3}$$

where ε_{ox} is the permittivity of the silicon dioxide,

$$\varepsilon_{ox} = 3.9\varepsilon_0 = 3.9 \times 8.854 \times 10^{-12} = 3.45 \times 10^{-11} \text{ F/m}$$

t_{ox} is oxide thickness

the total capacitance between gate and channel is

$$C = C_{ox}WL$$

the charge per unit channel length, which can be found from Eq. (5.2) as

$$\frac{|Q|}{\text{unit channel length}} = C_{ox}Wv_{OV} \tag{5.4}$$

The voltage v_{DS} establishes an electric field E across the length of the channel,

$$|E| = \frac{v_{DS}}{L} \tag{5.5}$$

This electric field in turn causes the channel electrons to drift toward the drain with a velocity given by

$$\text{Electron drift velocity} = \mu_n |E| = \mu_n \frac{v_{DS}}{L} \tag{5.6}$$

where μ_n is the mobility of the electrons at the surface of the channel. It is a physical parameter whose value depends on the fabrication process technology. The value of i_D can now be found by multiplying the charge per unit channel length (Eq. 5.4) by the electron drift velocity (Eq. 5.6),

$$i_D = \left[(\mu_n C_{ox})\left(\frac{W}{L}\right)v_{OV}\right]v_{DS} \tag{5.7}$$

Thus, for small v_{DS}, the channel behaves as a linear resistance whose value is controlled by the overdrive voltage v_{OV}, which in turn is determined by v_{GS}:

$$i_D = \left[(\mu_n C_{ox})\left(\frac{W}{L}\right)(v_{GS} - V_t)\right]v_{DS} \tag{5.8}$$

Symbols

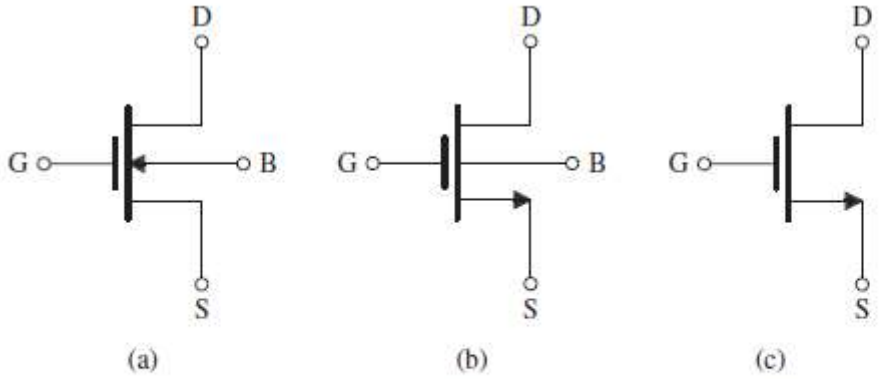
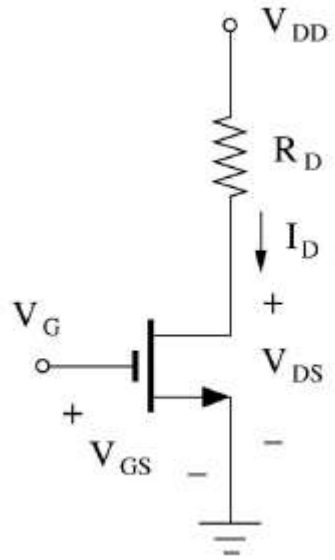


Figure 5.11 (a) Circuit symbol for the *n*-channel enhancement-type MOSFET. (b) Modified circuit symbol with an arrowhead on the source terminal to distinguish it from the drain and to indicate device polarity (i.e., *n* channel). (c) Simplified circuit symbol to be used when the source is connected to the body or when the effect of the body on device operation is unimportant.

Biasing in MOS amplifier circuits:

Biasing by fixing V_{GS}:

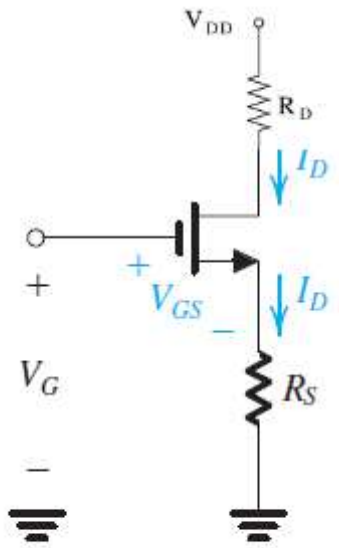


The most straightforward approach to biasing a MOSFET is to fix its gate-to-source voltage *VGS* to the value required to provide the desired *ID*. This voltage value can be derived from the power-supply voltage *VDD* through the use of an appropriate voltage divider.

This method is NOT desirable as $\mu_n C_{ox}\,(W/L)$ and V_t are not "well-defined." Bias point (i.e., I_D and V_{DS}) can change drastically due to temperature and/or manufacturing variability.

Biasing by fixing V_G and connecting a resistance in the Source

An excellent biasing technique for discrete MOSFET circuits consists of fixing the dc voltage at the gate, *VG*, and connecting a resistance in the source lead, as shown in the Figures below.



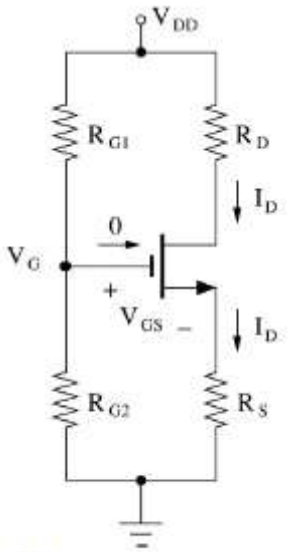
Basic Arrangement

Applying KVL for the input loop we get

$$V_G = V_{GS} + R_S I_D$$

$$V_{GS} = V_G - R_S I_D$$

Now, if V_G is much greater than V_{GS}, I_D will be mostly determined by the values of V_G and R_S. However, even if V_G is not much larger than V_{GS}, resistor R_S provides *negative feedback*, which acts to stabilize the value of the bias current I_D.

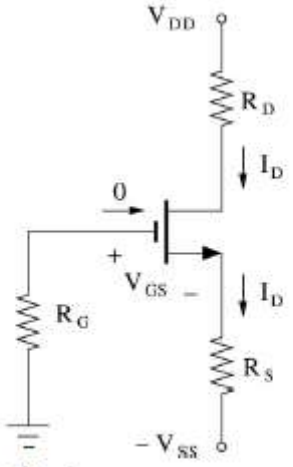


Bias with <u>one</u> power supply

Applying KVL for the input loop we get

$$V_{GS} = V_G - R_S I_D$$

$$V_{GS} = V_{R_{G2}} - R_S I_D$$



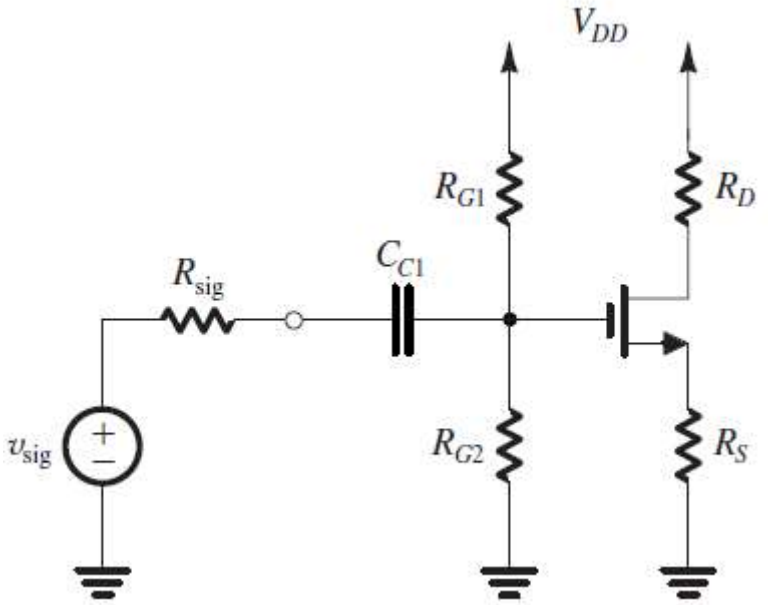
Bias with <u>two</u> power supply

Applying KVL for the input loop we get

$$V_{R_G} - V_{GS} + V_{SS} = R_S I_D$$

Since current through gate is zero, $V_{R_G} = 0\text{V}$

$$\therefore \quad V_{GS} = V_{SS} - R_S I_D$$



coupling of a signal source to the gate using a capacitor C_{C1}

Biasing using a Drain-to-Gate Feedback Resistor

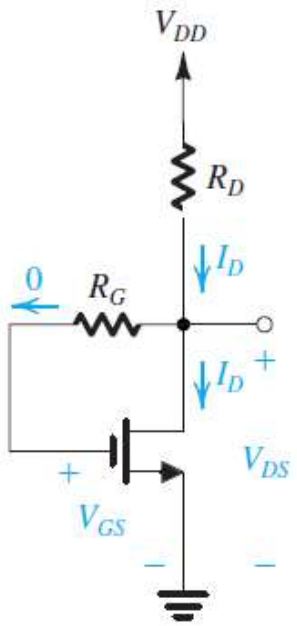


Figure 5.54 Biasing the MOSFET using a large drain-to-gate feedback resistance, R_G.

A simple and effective discrete-circuit biasing arrangement utilizing a feedback resistor connected between the drain and the gate is shown in Fig. 5.54. Here the large feedback resistance R_G (usually in the megohm range) forces the dc voltage at the gate to be equal to that at the drain (because $I_G = 0$). Thus we can write

$$V_{GS} = V_{DS} = V_{DD} - R_D I_D$$

which can be rewritten in the form

$$V_{DD} = V_{GS} + R_D I_D \tag{5.96}$$

if I_D increase , V_{GS} decreases. The decrease in V_{GS} in turn causes a decrease in I_D. Thus the negative feedback or degeneration provided by R_G works to keep the value of I_D as constant as possible.

The circuit of Fig. 5.54 can be utilized as an amplifier by applying the input voltage signal to the gate via a coupling capacitor so as not to disturb the dc bias conditions already established. The amplified output signal at the drain can be coupled to another part of the circuit, again via a capacitor.

Biasing Using a Constant-Current Source

The most effective scheme for biasing a MOSFET amplifier is that using a constant-current source. Figure 5.55(a) shows such an arrangement applied to a discrete MOSFET. Here R_G (usually in the megohm range) establishes a dc ground at the gate and presents a large resistance to an input signal source that can be capacitively coupled to the gate. Resistor R_D establishes an appropriate dc voltage at the drain to allow for the required output signal swing while ensuring that the transistor always remains in the saturation region.

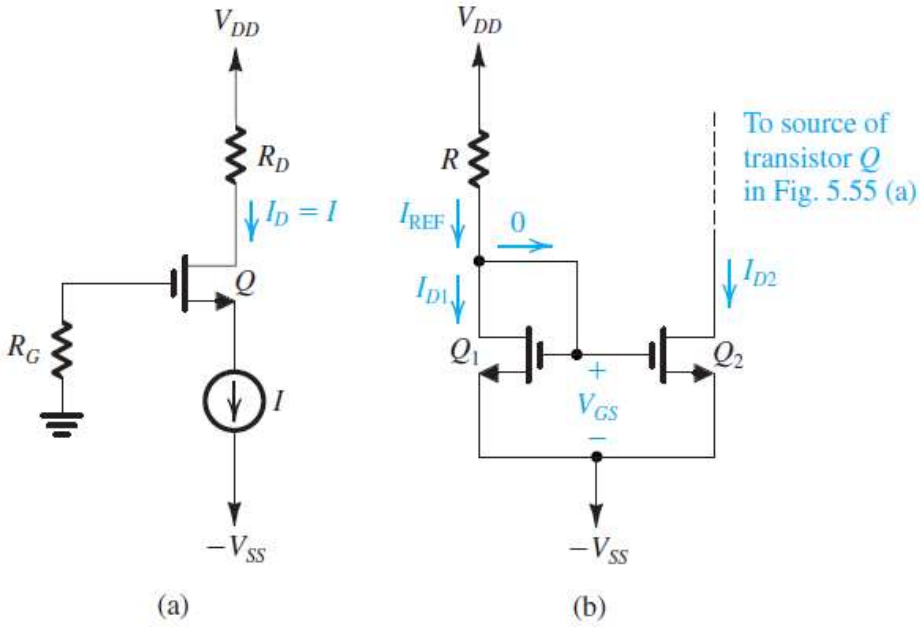


Figure 5.55 **(a)** Biasing the MOSFET using a constant-current source I. **(b)** Implementation of the constant-current source I using a current mirror.

A circuit for implementing the constant-current source I is shown in Fig. 5.55(b). The heart of the circuit is transistor Q_1, whose drain is shorted to its gate, and thus is operating in the saturation region, such that

$$I_{D1} = \frac{1}{2}k_n'\left(\frac{W}{L}\right)_1(V_{GS} - V_t)^2 \tag{5.97}$$

where we have neglected channel-length modulation (i.e., assumed $\lambda = 0$). The drain current of Q_1 is supplied by V_{DD} through resistor R. Since the gate currents are zero,

$$I_{D1} = I_{REF} = \frac{V_{DD} + V_{SS} - V_{GS}}{R} \tag{5.98}$$

where the current through R is considered to be the *reference current* of the current source and is denoted I_{REF}. Given the parameter values of Q_1 and a desired value for I_{REF}, Eqs. (5.97) and (5.98) can be used to determine the value of R. Now consider transistor Q_2: It has the same V_{GS} as Q_1; thus if we assume that it is operating in saturation, its drain current, which is the desired current I of the current source, will be

$$I = I_{D2} = \frac{1}{2}k_n'\left(\frac{W}{L}\right)_2 (V_{GS} - V_t)^2 \tag{5.99}$$

where we have neglected channel-length modulation. Equations (5.98) and (5.99) enable us to relate the current I to the reference current I_{REF},

$$I = I_{\text{REF}}\frac{(W/L)_2}{(W/L)_1} \tag{5.100}$$

Thus I is related to I_{REF} by the ratio of the aspect ratios of Q_1 and Q_2. This circuit, known as a **current mirror**.