ECE 255, MOSFET Small Signal Analysis

6 March 2018

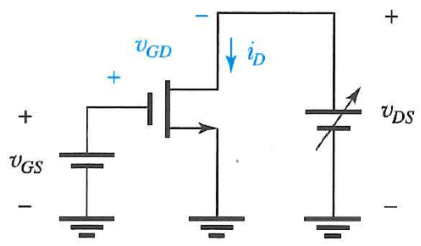
In this lecture, we will introduce small-signal analysis, operation, and models from Section 7.2 of Sedra and Smith. Since the BJT case has been discussed, we will now focus on the MOSFET case. In the small-signal analysis, one assumes that the device is biased at a DC operating point (also called the Q point or the quiescent point), and then, a small signal is super-imposed on the DC biasing point.

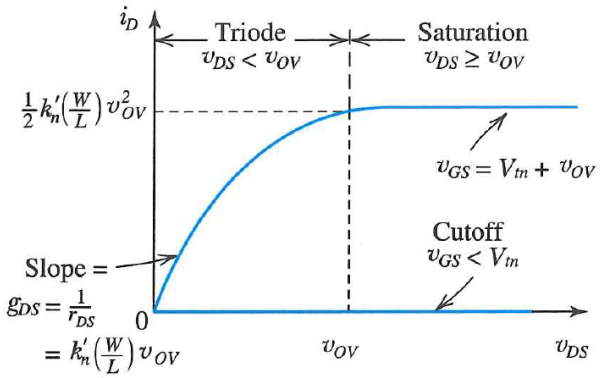
1 The DC Bias Point and Linearization–The MOSFET Case

Before one starts, it will be prudent to refresh our memory on the salient features of the MOSFET from Table 5.1 of Sedra and Smith. Also, Figure 1 is included to remind one of the definition of the relationship between V_{GS}, V_{OV}, V_{GD}, and V_{DS}. When $V_{DS} > V_{OV}$, the channel region is not continuous, and pinching occurs. The device is in the saturation region.

Printed on March 14, 2018 at 10 : 47: W.C. Chew and S.K. Gupta.

Table 5.1 Regions of Operation of the Enhancement NMOS Transistor





- $v_{GS} < V_{tn}$: no channel; transistor in cutoff; $i_D = 0$
- $v_{GS} = V_{tn} + v_{OV}$: a channel is induced; transistor operates in the triode region or the saturation region depending on whether the channel is continuous or pinched off at the drain end;

Triode Region

Continuous channel, obtained by:

$$v_{GD} > V_{tn}$$

or equivalently:

$$v_{DS} < v_{OV}$$

Then,

$$i_D = k_n'\left(\frac{W}{L}\right)\left[(v_{GS} - V_{tn})v_{DS} - \frac{1}{2}v_{DS}^2\right]$$

or equivalently,

$$i_D = k_n'\left(\frac{W}{L}\right)\left(v_{OV} - \frac{1}{2}v_{DS}\right)v_{DS}$$

Saturation Region

Pinched-off channel, obtained by:

$$v_{GD} \leq V_{tn}$$

or equivalently:

$$v_{DS} \geq v_{OV}$$

Then

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)(v_{GS} - V_{tn})^2$$

or equivalently,

$$i_D = \frac{1}{2}k_n'\left(\frac{W}{L}\right)v_{OV}^2$$

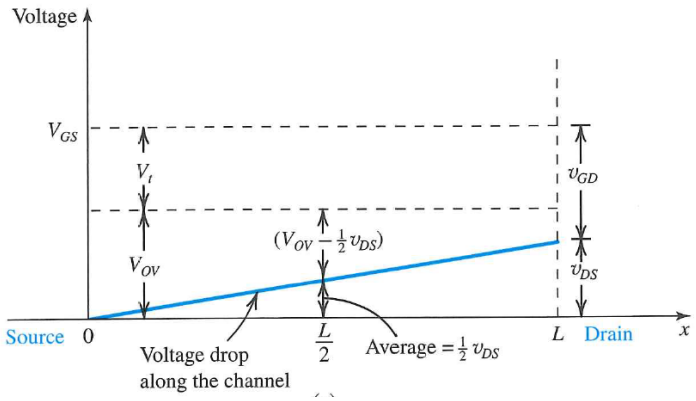


Figure 1: A figure to remind us of the definition of v_{GS}, v_{OV}, v_{DS}, and v_{GD} in the triode region (Courtesy of Sedra and Smith).

Figure 2 illustrates an NMOS operating as an amplifier. It is being biased with a DC voltage, and a small signal is superimposed on top of the DC voltage.

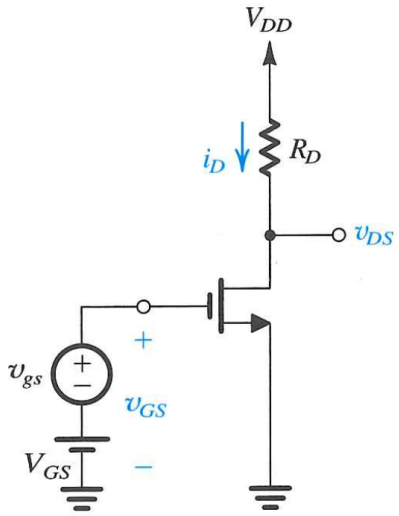


Figure 2: Circuit diagram of a transistor MOSFET (NMOS) amplifier with a small time-varying signal superimposed on top of a DC voltage bias source (Courtesy of Sedra and Smith).

Before proceeding further, one is also reminded of the i-v characteristics of

a MOSFET. Namely, that in the saturation region

$$i_D = \frac{1}{2}k_n(v_{GS} - V_t)^2 = \frac{1}{2}k_n v_{OV}^2 \tag{1.1}$$

where

$$v_{GS} = V_{GS} + v_{gs}, \qquad v_{OV} = V_{OV} + v_{ov} \tag{1.2}$$

where $v_{OV} = v_{GS} - V_t$. Again, the notation will be that the total value is denoted by a lower-case letter with upper-case subscript, the DC value or the Q point is denoted by a upper-case letter with upper-case subscript. The small signal is denoted by a lower-case letter with lower-case subscript. However, the threshold voltage, which is a DC value is denoted as V_t in order not to be confused with V_T the thermal voltage.

It is noted that the overdrive voltage v_{OV} is defined as

$$v_{OV} = v_{GS} - V_t = V_{GS} + v_{gs} - V_t \tag{1.3}$$

If we further define that $v_{OV} = V_{OV} + v_{ov}$, equating the time varying and DC parts, one gets

$$V_{OV} = V_{GS} - V_t, \qquad v_{ov} = v_{gs} \tag{1.4}$$

If the time-varying signals are turned off, then (1.1), keeping only the DC signals, becomes

$$I_D = \frac{1}{2}k_n(V_{GS} - V_t)^2 = \frac{1}{2}k_n V_{OV}^2 \tag{1.5}$$

which is the i-v relation obtained at the Q point, or the operating point.

In (1.1), by letting

$$v_{GS} = V_{GS} + v_{gs} \tag{1.6}$$

then

$$i_D = \frac{1}{2}k_n(V_{GS}+v_{gs}-V_t)^2 = \frac{1}{2}k_n(V_{GS}-V_t)^2 + k_n(V_{GS}-V_t)v_{gs} + \frac{1}{2}k_n v_{gs}^2 \tag{1.7}$$

For small signal analysis, v_{gs} is assumed to be small. Then v_{gs}^2 is even smaller; and hence, the last term can be ignored in the above. By letting the total current here as $i_D = I_D + i_d$, a DC term plus a time-varying term, then from the above, one can see that the time-varying term is approximately given by

$$i_d \approx k_n(V_{GS} - V_t)v_{gs} = g_m v_{gs} \tag{1.8}$$

where quadratic term proportional to v_{gs}^2 has been ignored, and

$$g_m = k_n(V_{GS} - V_t) = k_n V_{OV} = \frac{i_d}{v_{gs}} = \left.\frac{\partial i_D}{\partial v_{GS}}\right|_{v_{GS}=V_{GS}} \tag{1.9}$$

The above has the unit of conductance, and it is called the **MOSFET transconductance**. The last equality follows from that the transconductance is the ratio

between two incrementally small quantities at the operating point. This incremental relationship is shown in Figure 3.

A more detail analysis by comparing the last two terms on the right-hand side, one can show that the last term in (1.7) can be ignored if

$$v_{gs} \ll 2(V_{GS} - V_t) = 2V_{OV} \tag{1.10}$$

Also, it is seen from the Figure 3 that in order for the small signal analysis to be valid, it is required that $v_{gs} \ll 2V_{OV}$.

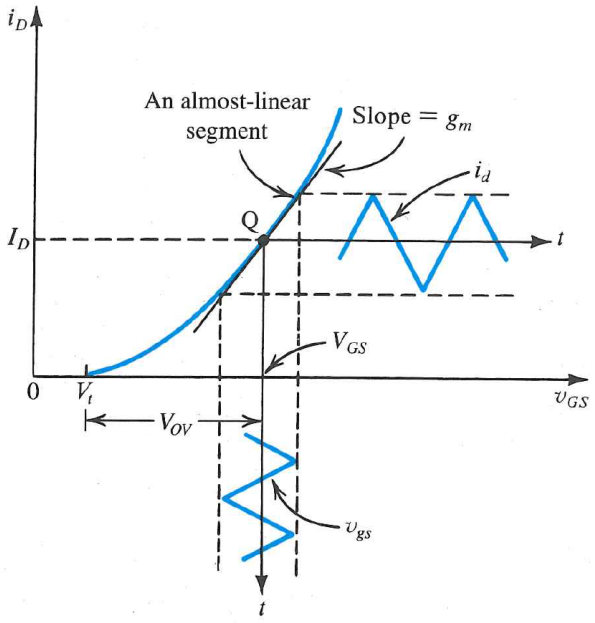


Figure 3: Graphical depiction of the small signal analysis for MOSFET (Courtesy of Sedra and Smith).

Furthermore, it is seen that at the DC operating point, by KVL

$$V_{DS} = V_{DD} - R_D I_D \tag{1.11}$$

If V_{DD} is held fixed, then by KVL again

$$v_{DS} = V_{DD} - R_D i_D = V_{DD} - R_D(I_D + i_d) = V_{DD} - R_D I_D - R_D i_d \tag{1.12}$$

From the above, by first letting $v_{DS} = V_{DS} + v_{ds}$, and then equating the time-varying terms and the DC terms, one concludes that the time varying part of v_{DS}, or v_{ds}, is given by

$$v_{ds} = -i_d R_D = -g_m v_{gs} R_D \tag{1.13}$$

One can define the voltage gain as

$$A_{vo} = \frac{v_{ds}}{v_{gs}} = -g_m R_D \tag{1.14}$$

The negative sign comes about because an increase in v_{gs} causes an increase in i_d, causing the rise in the voltage drop across R_D, and hence, a drop in the voltage v_{gs}. This sign reversal is indicated in Figure 4.

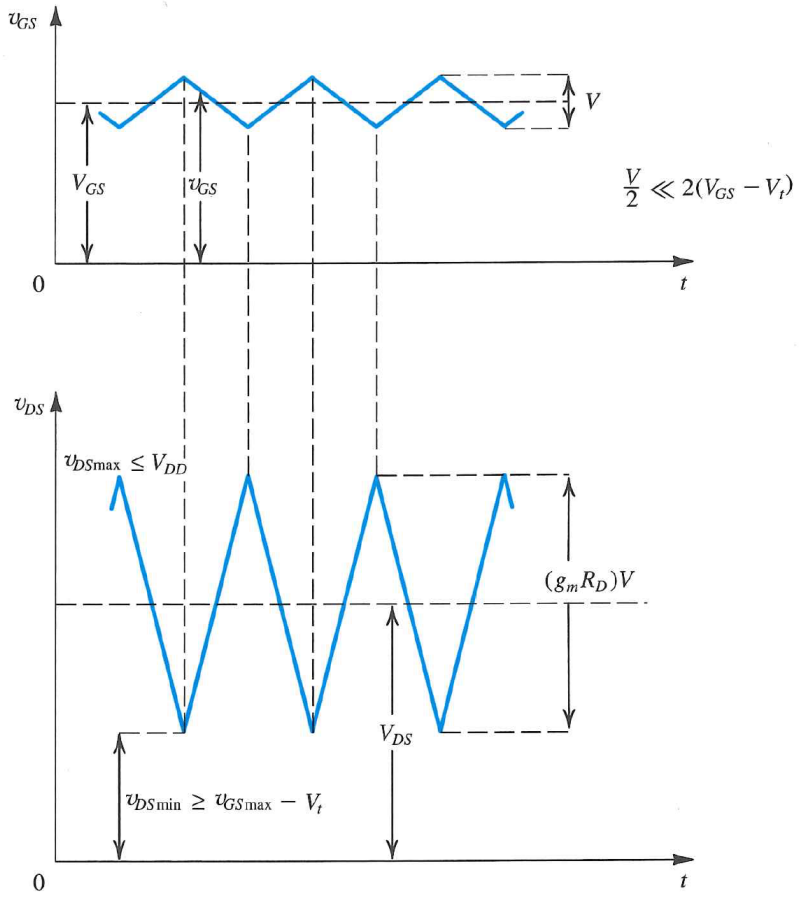


Figure 4: Total instantaneous voltage v_{GS} and v_{DS} for the circuit in Figure 2. Note the sign reversal of the amplified signal (Courtesy of Sedra and Smith).

2 Small Signal Equivalent-Circuit Models

By looking at the i-v characteristic curve of the MOSFET as shown in Table 5.1, it is seen for incremental v_{ds}, the current i_d does not change. This relationship can be modeled by a current source. Moreover, the gate of the MOSFET is essentially an open circuit at DC. Hence, the **small-signal equivalent-circuit model** is presented in Figure 5(a).

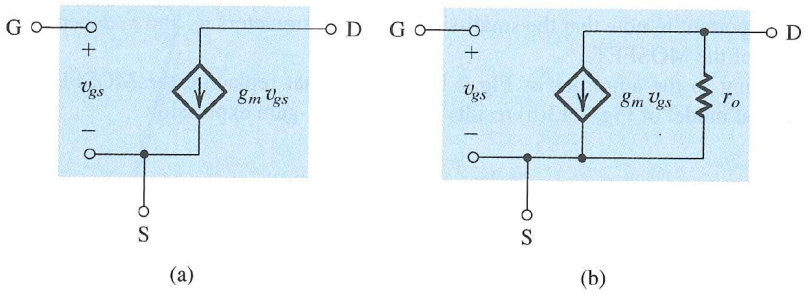


Figure 5: The small-signal model for a MOSFET: (a) no Early effect (channel-length modelulation effect); (b) Early effect is included by adding $r_o = |V_A|/I_D$ (Courtesy of Sedra and Smith).

When the Early effect has to be accounted for, an output resistor r_o can be added as shown in Figure 5(b). The value of r_o is given as

$$r_o = \frac{|V_A|}{I_D}, \text{ where } I_D = \frac{1}{2} k_n V_{OV}^2 \tag{2.1}$$

where $V_A = 1/\lambda$, and $-V_A$ is the negative intercept of the i-v curve.

Here, r_o is typically 10 kΩ to 1000 kΩ. When r_o is included, then the voltage gain becomes

$$A_v = \frac{v_{ds}}{v_{gs}} = -g_m(R_D||r_o) \tag{2.2}$$

In other words, the inclusion of r_o reduces the voltage gain.

The above analysis can be used for PMOS by noting for the sign change in PMOS, and by using $|V_{GS}|$, $|V_t|$, $|V_{OV}|$, and $|V_A|$ in the formulas, and replacing k_n with k_p.

3 Transconductance g_m

The transconductance can be looked at with more details by using $k_n = k_n'(W/L)$, giving

$$g_m = k_n'(W/L)(V_{GS} - V_t) = k_n'(W/L)V_{OV} = \mu_n C_{ox}(W/L)V_{OV} \tag{3.1}$$

The transconductance can be increased by increasing the W/L ratio, and also increasing the overdrive voltage V_{OV}. But increasing V_{OV} implies that the operating point for V_{DS} has to increase in order for the MOSFET to be in the saturation region.

Also, by using the fact that

$$I_D = \frac{1}{2}k_n V_{OV}^2 = \frac{1}{2}k'_n(W/L)V_{OV}^2 \tag{3.2}$$

or that $V_{OV} = \sqrt{2I_D/(k'_n(W/L))}$, then g_m in (3.1) can be alternatively rewritten as

$$g_m = \sqrt{2k'_n(W/L)I_D} \tag{3.3}$$

implying that g_m is proportional to the square roots of the drain current I_D, and W/L. At this point, note that

1. The transconductance g_m of a MOSFET is geometry dependent whereas that of the BJT is not.

2. The transconductance of a BJT is much larger than that of a MOSFET.

For example, with biasing point of $I_D = 50$ mA, with $k'_n = 120\ \mu\text{A/V}^2$, with $W/L = 1$, $g_m = 0.35$ mA/V, whereas when $W/L = 100$, $g_m = 3.5$ mA/V. But typically, the $g_m = 20$ mA/V for BJT when $I_C = 0.5$ mA.

Alternatively, by seeing that $I_D = \frac{1}{2}k_n V_{OV}^2$ from (1.5) and that $g_m = k_n V_{OV}$ from (1.9), an by dividing the two equations, it can be shown that

$$g_m = \frac{2I_D}{V_{OV}} \tag{3.4}$$

Since g_m is an incremental relationship, this relationship can be shown graphically as in Figure 6 where the slope is also shown graphically.

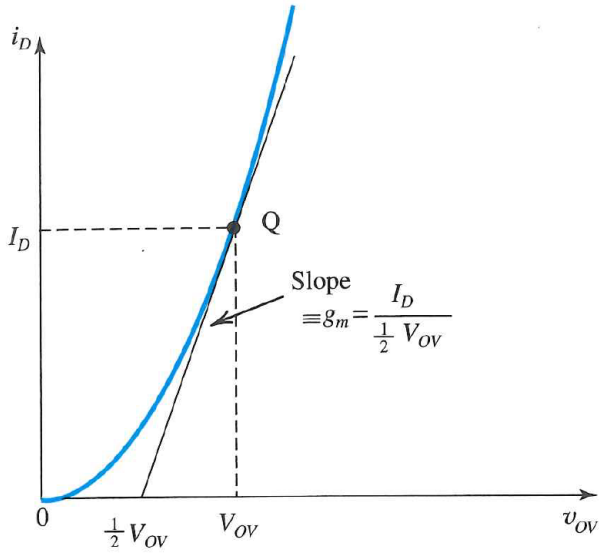


Figure 6: The relationship between I_D, V_{OV}, and g_m as shown in the i_D versus v_{OV} curve (Courtesy of Sedra and Smith).

The above equations indicate that a designer can change g_m by altering W/L, V_{OV}, and I_D.

4 The T Equivalent-Circuit Model

The hybrid-π equivalent-circuit model can be replaced by the T equivalent-circuit model. The morphing of a bybrid-π model to the T model is shown in Figure 7.

1. The current source in the hybrid-π model can be split into two without affecting the branch current. This is seen from the morphing of Figure 7(a) to Figure 7(b).

2. Note that the gate current is zero. The point X can be connected to the gate input, and yet the gate current is zero because of KCL at X. This is indicated in Figure 7(c).

3. But the second voltage-controlled current source is just the voltage and current relation of a resistor whose resistance is $1/g_m$. Hence, it can be replaced by a resistor as shown in Figure 7(d).

Notice that in the T equivalent-circuit model, due to its construction, and KCL, the gate current is always zero, implying that its resistance is infinite.

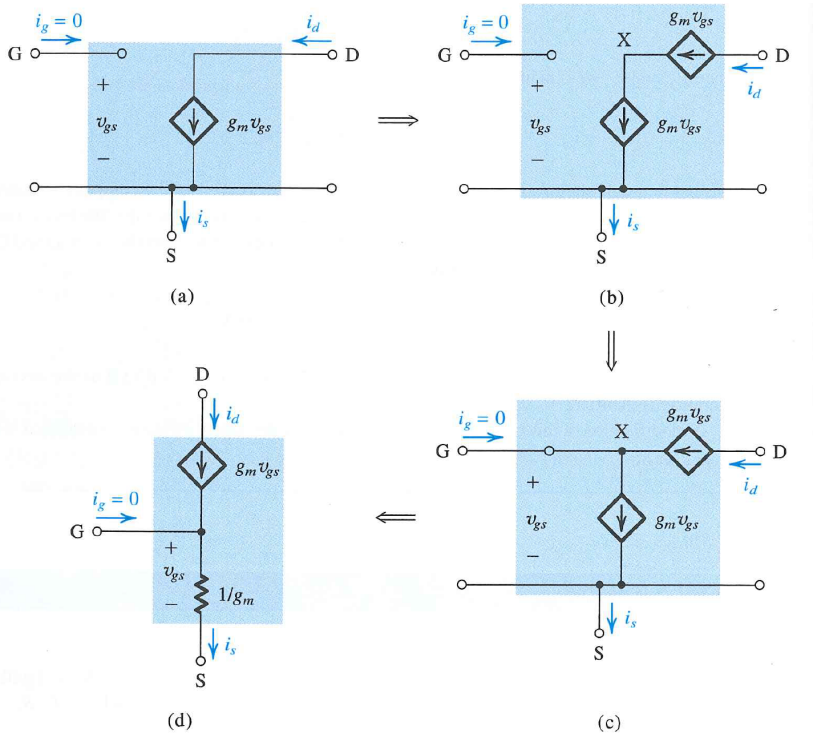


Figure 7: The morphing of the hybrid-π equivalent circuit model in (a) to the T equivalent circuit model in (d) (Courtesy of Sedra and Smith).

Also as the morphing of the hybrid-π equivalent-circuit model to the T equivalent-circuit model is unaffected by connecting a resistor between D and S, an r_o can be thus connected to account for the Early effect or the channel-modulation effect as shown in Figure 8(a). Figure 8(b) is an alternative way of representing the T equivalent-circuit model, so that the gate current is always zero by KCL.

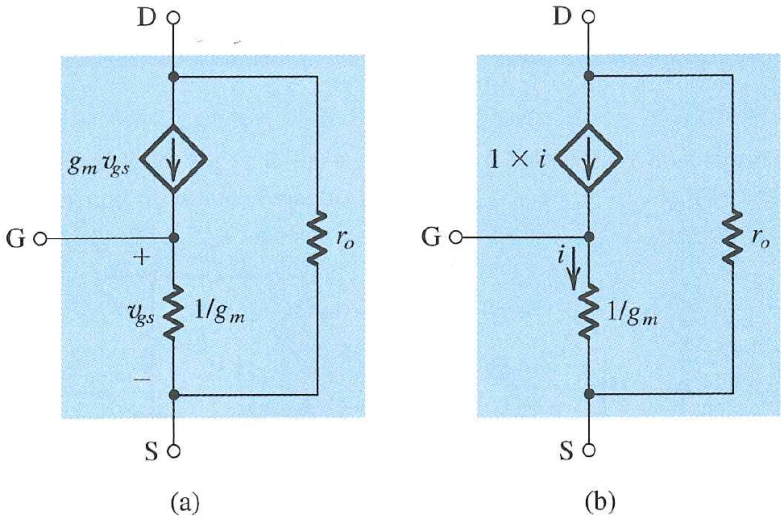
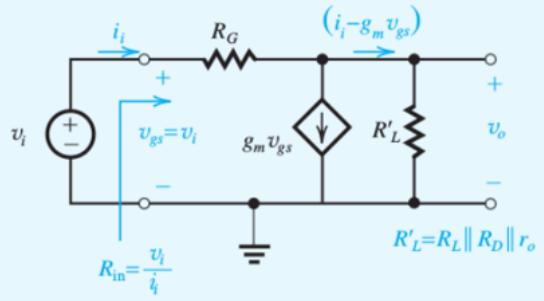


Figure 8: The morphing of the π equivalent-circuit model to the T equivalent circuit model is still valid if a resistor is connected between the drain D and the source S. Hence, in (a), r_o can be connected to account for the Early effect. (b) shows an alternative T model that is equivalent (Courtesy of Sedra and Smith).

Example 7.3

Figure 7.15(a) shows a discrete MOSFET amplifier utilizing a drain-to-gate resistance R_G for biasing purposes. Such a biasing arrangement will be studied in Section 7.4. The input signal v_i is coupled to the gate via a large capacitor, and the output signal at the drain is coupled to the load resistance R_L via another large capacitor. We wish to analyze this amplifier circuit to determine its small-signal voltage gain, its input resistance, and the largest allowable input signal. The transistor has $V_t = 1.5$ V, $k_n'(W/L) = 0.25$ mA/V^2, and $V_A = 50$ V. Assume the coupling capacitors to be sufficiently large so as to act as short circuits at the signal frequencies of interest.



(d)

Figure 7.15 *continued*

Solution

We first determine the dc operating point. For this purpose, we eliminate the input signal v_i, and open-circuit the two coupling capacitors (since they block dc currents). The result is the circuit shown in Fig. 7.14(b). We note that since $I_G = 0$, the dc voltage drop across R_G will be zero, and

$$V_{GS} = V_{DS} = V_{DD} - R_D I_D \tag{7.43}$$

With $V_{DS} = V_{GS}$, the NMOS transistor will be operating in saturation. Thus,

$$I_D = \frac{1}{2} k_n (V_{GS} - V_t)^2 \tag{7.44}$$

where, for simplicity, we have neglected the effect of channel-length modulation on the dc operating point. Substituting $V_{DD} = 15$ V, $R_D = 10\,\text{k}\Omega$, $k_n = 0.25\,\text{mA/V}^2$, and $V_t = 1.5$ V in Eqs. (7.43) and (7.44), and substituting for V_{GS} from Eq. (7.43) into Eq. (7.44) results in a quadratic equation in I_D. Solving the latter and discarding the root that is not physically meaningful yields the solution

$$I_D = 1.06\,\text{mA}$$

which corresponds to

$$V_{GS} = V_{DS} = 4.4\,\text{V}$$

and

$$V_{OV} = 4.4 - 1.5 = 2.9\,\text{V}$$

Next we proceed with the small-signal analysis of the amplifier. Toward that end we replace the MOSFET with its small-signal model to obtain the small-signal equivalent circuit of the amplifier, shown in Fig. 7.15(c). Observe that we have replaced the coupling capacitors with short circuits. The dc voltage supply V_{DD} has also been replaced with a short circuit to ground.

Example 7.3 *continued*

The values of the transistor small-signal parameters g_m and r_o can be determined by using the dc bias quantities found above, as follows:

$$\begin{aligned} g_m &= k_n V_{OV} \\ &= 0.25 \times 2.9 = 0.725 \text{ mA/V} \\ r_o &= \frac{V_A}{I_D} = \frac{50}{1.06} = 47 \text{ k}\Omega \end{aligned}$$

Next we use the equivalent circuit of Fig. 7.15(c) to determine the input resistance $R_{in} \equiv v_i/i_i$ and the voltage gain $A_v = v_o/v_i$. Toward that end we simplify the circuit by combining the three parallel resistances r_o, R_D, and R_L in a single resistance R'_L,

$$\begin{aligned} R'_L &= R_L \| R_D \| r_o \\ &= 10 \| 10 \| 47 = 4.52 \text{ k}\Omega \end{aligned}$$

as shown in Fig. 7.15(d). For the latter circuit we can write the two equations

$$v_o = (i_i - g_m v_{gs}) R'_L \tag{7.45}$$

and

$$i_i = \frac{v_{gs} - v_o}{R_G} \tag{7.46}$$

Substituting for i_i from Eq. (7.46) into Eq. (7.45) results in the following expression for the voltage gain $A_v \equiv v_o/v_i = v_o/v_{gs}$:

$$A_v = -g_m R'_L \frac{1 - (1/g_m R_G)}{1 + (R'_L/R_G)}$$

Since R_G is very large, $g_m R_G \gg 1$ and $R'_L/R_G \ll 1$ (the reader can easily verify this), and the gain expression can be approximated as

$$A_v \simeq -g_m R'_L \tag{7.47}$$

Substituting $g_m = 0.725$ mA/V and $R'_L = 4.52$ kΩ yields

$$A_v = -3.3 \text{ V/V}$$

To obtain the input resistance, we substitute in Eq. (7.46) for $v_o = A_v v_{gs} = -g_m R'_L v_{gs}$, then use $R_{in} \equiv v_i/i_i = v_{gs}/i_i$ to obtain

$$R_{in} = \frac{R_G}{1 + g_m R'_L} \tag{7.48}$$

This is an interesting relationship: The input resistance decreases as the gain $(g_m R_L')$ is increased. The value of R_{in} can now be determined; it is

$$R_{in} = \frac{10\,\text{M}\Omega}{1+3.3} = 2.33\,\text{M}\Omega$$

which is still very large.

The largest allowable input signal $\hat{v}_i$ is constrained by the need to keep the transistor in saturation at all times; that is,

$$v_{DS} \geq v_{GS} - V_t$$

Enforcing this condition with equality at the point v_{GS} is maximum and v_{DS} is minimum, we write

$$v_{DS\,\min} = v_{GS\,\max} - V_t$$

$$V_{DS} - |A_v|\hat{v}_i = V_{GS} + \hat{v}_i - V_t$$

Since $V_{DS} = V_{GS}$, we obtain

$$\hat{v}_i = \frac{V_t}{|A_v| + 1}$$

This is a general relationship that applies to this circuit irrespective of the component values. Observe that it simply states that the maximum signal swing is determined by the fact that the bias arrangement makes $V_D = V_G$ and thus, to keep the MOSFET out of the triode region, the signal between D and G is constrained to be equal to V_t. For our particular design,

$$\hat{v}_i = \frac{1.5}{3.3+1} = 0.35\,\text{V}$$

Since $V_{OV} = 2.9\,\text{V}$, a v_i of 0.35 is indeed much smaller than $2V_{OV} = 5.8\,\text{V}$; thus the assumption of linear operation is justified.

A modification of this circuit that increases the allowable signal swing is investigated in Problem 7.103.

Example 7.4

Figure 7.18(a) shows a MOSFET amplifier biased by a constant-current source I. Assume that the values of I and R_D are such that the MOSFET operates in the saturation region. The input signal v_i is coupled to the source terminal by utilizing a large capacitor C_{C1}. Similarly, the output signal at the drain is taken through a large coupling capacitor C_{C2}. Find the input resistance R_{in} and the voltage gain v_o/v_i. Neglect channel-length modulation.

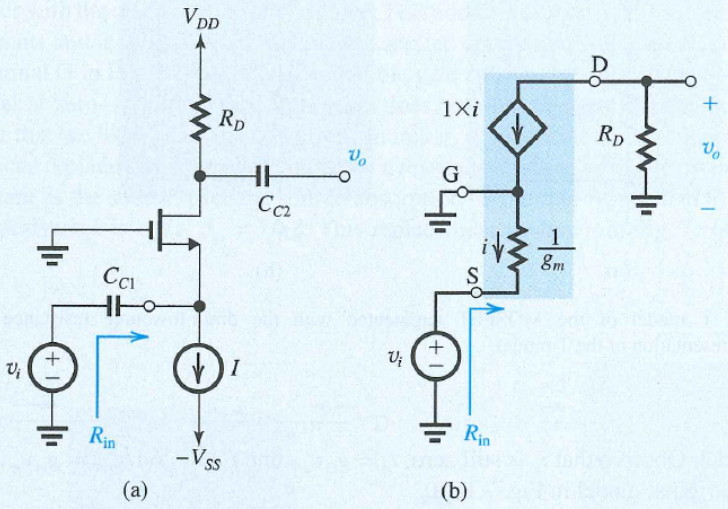


Figure 7.18 (a) Amplifier circuit for Example 7.4. (b) Small-signal equivalent circuit of the amplifier in (a).

Solution

Replacing the MOSFET with its T equivalent-circuit model results in the amplifier equivalent circuit shown in Fig. 7.18(b). Observe that the dc current source I is replaced with an open circuit and the dc voltage source V_{DD} is replaced by a short circuit. The large coupling capacitors have been replaced by short circuits. From the equivalent-circuit model we determine

$$R_{in} = \frac{v_i}{-i} = 1/g_m$$

and

$$v_o = -iR_D = \left(\frac{v_i}{1/g_m}\right)R_D = g_m R_D v_i$$

Thus,

$$A_v \equiv \frac{v_o}{v_i} = g_m R_D$$

We note that this amplifier, known as the common-gate amplifier because the gate at ground potential is common to both the input and output ports, has a low input resistance $(1/g_m)$ and a noninverting gain. We shall study this amplifier type in Section 7.3.5.