

APPENDIX G

COMPARISON OF THE MOSFET AND THE BJT

In this appendix we present a comparison of the characteristics of the two major electronic devices: the MOSFET and the BJT. To facilitate this comparison, typical values for the important parameters of the two devices are first presented. We also discuss the design parameters available with each of the two devices, such as I_C in the BJT, and I_D and V_{OV} in the MOSFET, and the trade-offs encountered in deciding on suitable values for these.

G.1 Typical Values of MOSFET Parameters

Typical values for the important parameters of NMOS and PMOS transistors fabricated in a number of CMOS processes are shown in Table G.1. Each process is characterized by the minimum allowed channel length, L_{\min} ; thus, for example, in a 0.18- μm process, the smallest transistor has a channel length $L = 0.18 \mu\text{m}$. The technologies presented in Table G.1 are in descending order of channel length, with that having the shortest channel length being the most modern. Although the 0.8- μm process is now obsolete, its data are included to show trends in the values of various parameters. It should also be mentioned that although Table G.1 stops at the 65-nm process, by 2014 there were 45-, 32-, and 22-nm processes available, and processes down to 14 nm were in various stages of development. The 0.18- μm and the 0.13- μm processes, however, remained popular in the design of analog ICs. The most recently announced digital ICs utilize 32-nm and 22-nm processes and pack as many as 4.3 billion transistors onto one chip. An important caution is in order regarding the data presented in Table G.1: These data do *not* pertain to any particular commercially available process.

Parameter	0.8 μm		0.5 μm		0.25 μm		0.18 μm		0.13 μm		65 nm	
	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS	NMOS	PMOS
t_{ox} (nm)	15	15	9	9	6	6	4	4	2.7	2.7	1.4	1.4
C_{ox} (fF/ μm^2)	2.3	2.3	3.8	3.8	5.8	5.8	8.6	8.6	12.8	12.8	25	25
μ ($\text{cm}^2/\text{V}\cdot\text{s}$)	550	250	500	180	460	160	450	100	400	100	216	40
μC_{ox} ($\mu\text{A}/\text{V}^2$)	127	58	190	68	267	93	387	86	511	128	540	100
V_{t0} (V)	0.7	-0.7	0.7	-0.8	0.5	-0.6	0.5	-0.5	0.4	-0.4	0.35	-0.35
V_{DD} (V)	5	5	3.3	3.3	2.5	2.5	1.8	1.8	1.3	1.3	1.0	1.0
$ V'_A $ (V/ μm)	25	20	20	10	5	6	5	6	5	6	3	3
C_{ov} (fF/ μm)	0.2	0.2	0.4	0.4	0.3	0.3	0.37	0.33	0.36	0.33	0.33	0.31

Accordingly, these generic data are not intended for use in an actual IC design; rather, they show trends and, as we shall see, help to illustrate design trade-offs as well as enable us to work out design examples and problems with parameter values that are as realistic as possible.

As indicated in Table G.1, the trend has been to reduce the minimum allowable channel length. This trend has been motivated by the desire to pack more transistors on a chip as well as to operate at higher speeds or, in analog terms, over wider bandwidths.

Observe that the oxide thickness, t_{ox} , scales down with the channel length, reaching 1.4 nm for the 65-nm process. Since the oxide capacitance C_{ox} is inversely proportional to t_{ox} , we see that C_{ox} increases as the technology scales down. The surface mobility μ decreases as the technology minimum-feature size is decreased, and μ_p decreases faster than μ_n . As a result, the ratio of μ_p to μ_n has been decreasing with each generation of technology, falling from about 0.5 for older technologies to 0.2 or so for the newer ones. Despite the reduction of μ_n and μ_p , the transconductance parameters $k'_n = \mu_n C_{ox}$ and $k'_p = \mu_p C_{ox}$ have been steadily increasing. As a result, modern short-channel devices achieve required levels of bias currents at lower overdrive voltages. As well, they achieve higher transconductance, a major advantage.

Although the magnitudes of the threshold voltages V_m and V_{tp} have been decreasing with L_{min} from about 0.7–0.8 V to 0.3–0.4 V, the reduction has not been as large as that of the power supply V_{DD} . The latter has been reduced dramatically, from 5 V for older technologies to 1.0 V for the 65-nm process. This reduction has been necessitated by the need to keep the electric fields in the smaller devices from reaching very high values. Another reason for reducing V_{DD} is to keep power dissipation as low as possible given that the IC chip now has a much larger number of transistors.¹

The fact that in modern short-channel CMOS processes $|V_t|$ has become a much larger proportion of the power-supply voltage poses a serious challenge to the circuit design engineer. Recalling that $|V_{GS}| = |V_t| + |V_{OV}|$, where V_{OV} is the overdrive voltage, to keep $|V_{GS}|$ reasonably small, $|V_{OV}|$ for modern technologies is usually in the range of 0.1 V to 0.2 V. To appreciate this point further, recall that to operate a MOSFET in the saturation region, $|V_{DS}|$ must exceed $|V_{OV}|$; thus, to be able to have a number of devices stacked between the power-supply rails in a regime in which V_{DD} is only 1.8 V or lower, we need to keep $|V_{OV}|$ as low as possible. We will shortly see, however, that operating at a low $|V_{OV}|$ has some drawbacks.

Another significant though undesirable feature of modern deep submicron ($L_{min} < 0.25 \mu\text{m}$) CMOS technologies is that the channel-length modulation effect is very pronounced. As a result, V'_A has decreased to about 3 V/ μm , which combined with the decreasing values of L has caused the Early voltage $V_A = V'_A L$ to become very small. Correspondingly, short-channel MOSFETs exhibit low output resistances.

Studying the MOSFET high-frequency² equivalent-circuit model in Section 10.2 and the high-frequency response of the common-source amplifier in Section 10.3 shows that two major MOSFET capacitances are C_{gs} and C_{gd} . While C_{gs} has an overlap component,³ C_{gd} is entirely an overlap capacitance. Both C_{gd} and the overlap component of C_{gs} are almost equal and are denoted C_{ov} . The last line of Table G.1 provides the value of C_{ov} per micron of gate width. Although the normalized C_{ov} has been staying more or less constant with the reduction in L_{min} , we will shortly see that the shorter devices exhibit much higher operating speeds and wider amplifier bandwidths than the longer devices. Specifically, we will, for example, see that f_T for a 0.25- μm NMOS transistor can be as high as 10 GHz.

¹Chip power dissipation is a very serious issue, with some ICs dissipating as much as 100 W. As a result, an important current area of research concerns what is termed “power-aware design.”

²For completeness, this appendix includes material on the high-frequency models and operation of both the MOSFET and the BJT. These topics are covered in Chapter 10. The reader can easily skip the appendix paragraphs dealing with these topics until Chapter 10 has been studied.

³Overlap capacitances result because the gate electrode overlaps the source and drain diffusions (Fig. 5.1).

G.2 Typical Values of IC BJT Parameters

Table G.2 provides typical values for the major parameters that characterize integrated-circuit bipolar transistors. Data are provided for devices fabricated in two different processes: the standard, old process, known as the “high-voltage process,” and an advanced, modern process, referred to as a “low-voltage process.” For each process we show the parameters of the standard *nnp* transistor and those of a special type of *pnp* transistor known as a **lateral *pnp*** (as opposed to **vertical**, as in the *nnp* case) (see Appendix A). In this regard we should mention that a major drawback of standard bipolar integrated-circuit fabrication processes has been the lack of *pnp* transistors of a quality equal to that of the *nnp* devices. Rather, there are a number of *pnp* implementations for which the lateral *pnp* is the most economical to fabricate. Unfortunately, however, as should be evident from Table G.2, the lateral *pnp* has characteristics that are much inferior to those of the vertical *nnp*. Note in particular the lower value of β and the much larger value of the forward transit time τ_F that determines the emitter–base diffusion capacitance C_{de} and, hence, the transistor speed of operation. The data in Table G.2 can be used to show that the unity-gain frequency of the lateral *pnp* is 2 orders of magnitude lower than that of the *nnp* transistor fabricated in the same process. Another important difference between the lateral *pnp* and the corresponding *nnp* transistor is the value of collector current at which their β values reach their maximums: For the high-voltage process, for example, this current is in the tens of microamperes range for the *pnp* and in the milliampere range for the *nnp*. On the positive side, the problem of the lack of high-quality *pnp* transistors has spurred analog circuit designers to come up with highly innovative circuit topologies that either minimize the use of *pnp* transistors or minimize the dependence of circuit performance on that of the *pnp*. We encounter some of these ingenious circuits at various locations in this book.

The dramatic reduction in device size achieved in the advanced low-voltage process should be evident from Table G.2. As a result, the scale current I_S also has been reduced by about three orders of magnitude. Here we should note that the base width, W_B , achieved in the advanced process is on the order of 0.1 μm , as compared to a few microns in the standard high-voltage process. Note also the dramatic increase in speed; for the low-voltage *nnp* transistor, $\tau_F = 10$ ps as opposed to 0.35 ns in the high-voltage process. As a result, f_T for the modern *nnp* transistor is 10 GHz to 25 GHz, as compared to the 400 MHz to 600 MHz achieved in the high-voltage process. Although the Early voltage, V_A , for the modern process is lower than its value in the

Table G.2 Typical Parameter Values for BJTs*

Parameter	Standard High-Voltage Process		Advanced Low-Voltage Process	
	<i>nnp</i>	Lateral <i>pnp</i>	<i>nnp</i>	Lateral <i>pnp</i>
A_E (μm^2)	500	900	2	2
I_S (A)	5×10^{-15}	2×10^{-15}	6×10^{-18}	6×10^{-18}
β_0 (A/A)	200	50	100	50
V_A (V)	130	50	35	30
V_{CEO} (V)	50	60	8	18
τ_F	0.35 ns	30 ns	10 ps	650 ps
C_{je0}	1 pF	0.3 pF	5 fF	14 fF
$C_{\mu0}$	0.3 pF	1 pF	5 fF	15 fF
r_x (Ω)	200	300	400	200

*Adapted from Gray et al. (2001); see Appendix I.

old high-voltage process, it is still reasonably high at 35 V. Another feature of the advanced process—and one that is not obvious from Table G.2—is that β for the *npn* peaks at a collector current of 50 μA or so. Finally, note that as the name implies, *npn* transistors fabricated in the low-voltage process break down at collector–emitter voltages of 8 V, versus 50 V or so for the high-voltage process. Thus, while circuits designed with the standard high-voltage process utilize power supplies of $\pm 15\text{ V}$ (e.g., in commercially available op amps of the 741 type), the total power-supply voltage utilized with modern bipolar devices is 5 V (or even 2.5 V to achieve compatibility with some of the submicron CMOS processes).

G.3 Comparison of Important Characteristics

Table G.3 provides a compilation of the important characteristics of the NMOS and the *npn* transistors. The material is presented in a manner that facilitates comparison. In the following, we comment on the various items in Table G.3. As well, a number of numerical examples and exercises are provided to illustrate how the wealth of information in Table G.3 can be put to use. Before proceeding, note that the PMOS and the *pnp* transistors can be compared in a similar way.

Table G.3 Comparison of MOSFET and the BJT		
	NMOS	<i>npn</i>
Circuit Symbol		
To Operate in the Active Mode, Two Conditions Have to Be Satisfied	<p>(1) <i>Induce a channel:</i></p> $v_{GS} \geq V_t, \quad V_t = 0.3\text{--}0.5\text{ V}$ <p>Let $v_{GS} = V_t + v_{OV}$</p> <p>(2) <i>Pinch-off channel at drain:</i></p> $v_{GD} < V_t$ <p>or equivalently,</p> $v_{DS} \geq V_{OV}, \quad V_{OV} = 0.1\text{--}0.3\text{ V}$	<p>(1) <i>Forward-bias EBJ:</i></p> $v_{BE} \geq V_{BEon}, \quad V_{BEon} \simeq 0.5\text{ V}$ <p>(2) <i>Reverse-bias CBJ:</i></p> $v_{BC} < V_{BCon}, \quad V_{BCon} \simeq 0.4\text{ V}$ <p>or equivalently,</p> $v_{CE} \geq 0.3\text{ V}$
Current–Voltage Characteristics in the Active Region	$i_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} (v_{GS} - V_t)^2 \left(1 + \frac{v_{DS}}{V_A} \right)$ $= \frac{1}{2} \mu_n C_{ox} \frac{W}{L} v_{OV}^2 \left(1 + \frac{v_{DS}}{V_A} \right)$ $i_G = 0$	$i_C = I_S e^{v_{BE}/V_T} \left(1 + \frac{v_{CE}}{V_A} \right)$ $i_B = i_C / \beta$

Table G.3		
	NMOS	npn
Low-Frequency, Hybrid- π Model		
Low-Frequency T Model		
Transconductance g_m	$g_m = I_D / (V_{OV}/2)$ $g_m = (\mu_n C_{ox}) \left(\frac{W}{L} \right) V_{OV}$ $g_m = \sqrt{2(\mu_n C_{ox}) \left(\frac{W}{L} \right) I_D}$	$g_m = I_C / V_T$
Output Resistance r_o	$r_o = V_A' / I_D = \frac{V_A' L}{I_D}$	$r_o = V_A' / I_C$
Intrinsic Gain $A_0 \equiv g_m r_o$	$A_0 = V_A' / (V_{OV}/2)$ $A_0 = \frac{2V_A' L}{V_{OV}}$ $A_0 = \frac{V_A' \sqrt{2\mu_n C_{ox} W L}}{\sqrt{I_D}}$	$A_0 = V_A' / V_T$
Input Resistance with Source (Emitter) Grounded	∞	$r_\pi = \beta / g_m$

(continued)

Table G.3 continued		
	NMOS	npn
High-Frequency Model		
Capacitances	$C_{gs} = \frac{2}{3} W L C_{ox} + W L_{ov} C_{ox}$ $C_{gd} = W L_{ov} C_{ox}$	$C_{\pi} = C_{de} + C_{je}$ $C_{de} = \tau_F g_m$ $C_{je} \approx 2 C_{je0}$ $C_{\mu} = C_{\mu0} \left[1 + \frac{V_{CB}}{V_{C0}} \right]^m$
Transition Frequency f_T	$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$ <p>For $C_{gs} \gg C_{gd}$ and $C_{gs} \approx \frac{2}{3} W L C_{ox}$,</p> $f_T \approx \frac{1.5 \mu_n V_{OV}}{2\pi L^2}$	$f_T = \frac{g_m}{2\pi(C_{\pi} + C_{\mu})}$ <p>For $C_{\pi} \gg C_{\mu}$ and $C_{\pi} \approx C_{de}$,</p> $f_T \approx \frac{2\mu_n V_T}{2\pi W_B^2}$
Design Parameters	$I_D, V_{OV}, L, \frac{W}{L}$	I_C, V_{BE}, A_E (or I_S)
Good Analog Switch?	Yes, because the device is symmetrical and thus the $i_D - v_{DS}$ characteristics pass directly through the origin.	No, because the device is asymmetrical with an offset voltage V_{CEoff} .

G.3.1 Operating Conditions

At the outset, note that we shall use **active mode** or **active region** to denote both the active mode of operation of the BJT and the saturation mode of operation of the MOSFET.

The conditions for operating in the active mode are very similar for the two devices: The explicit threshold V_t of the MOSFET has V_{BEon} as its implicit counterpart in the BJT. Furthermore, for modern processes, V_{BEon} and V_t are almost equal.

Also, pinching off the channel of the MOSFET at the drain end is very similar to reverse biasing the CBJ of the BJT; the first makes i_D nearly independent of v_D , and the second makes I_C nearly independent of v_C . Note, however, that the asymmetry of the BJT results in V_{BCon} and V_{BEon} being unequal, while in the symmetrical MOSFET the operative threshold voltages at the source and the drain ends of the channel are identical (V_t). Finally, for both the MOSFET and the BJT to operate in the active mode, the voltage across the device (v_{DS}, v_{CE}) must be at least 0.1 V to 0.3 V.

G.3.2 Current–Voltage Characteristics

The square-law control characteristic, i_D-v_{GS} , in the MOSFET should be contrasted with the exponential control characteristic, i_C-v_{BE} , of the BJT. Obviously, the latter is a much more sensitive relationship, with the result that i_C can vary over a very wide range (five decades or more) within the same BJT. In the MOSFET, the range of i_D achieved in the same device is much more limited. To appreciate this point further, consider the parabolic relationship between i_D and v_{OV} , and recall from our discussion above that v_{OV} is usually kept in a narrow range (0.1 V to 0.3 V).

Next we consider the effect of the device dimensions on its current. For the bipolar transistor, the control parameter is the area of the emitter–base junction (EBJ), A_E , which determines the scale current I_S . It can be varied over a relatively narrow range, such as 10 to 1. Thus, while the emitter area can be used to achieve current scaling in an I_C (as we can see in Section 8.2 in connection with the design of current mirrors), its narrow range of variation reduces its significance as a design parameter. This is particularly so if we compare A_E with its counterpart in the MOSFET, the aspect ratio W/L . MOSFET devices can be designed with W/L ratios in a wide range, such as 1.0 to 500. As a result, W/L is a very significant MOS design parameter. Like A_E , it is also used in current scaling, as we can see in Section 8.2. Combining the possible range of variation of v_{OV} and W/L , one can design MOS transistors to operate over an i_D range of four decades or so.

The channel-length modulation in the MOSFET and the base-width modulation in the BJT are similarly modeled and give rise to the dependence of i_D (i_C) on v_{DS} (v_{CE}) and, hence, to the finite output resistance r_o in the active region. Two important differences, however, exist. In the BJT, V_A is solely a process-technology parameter and does not depend on the dimensions of the BJT. In the MOSFET, the situation is quite different: $V_A = V'_A L$, where V'_A is a process-technology parameter and L is the channel length used. Also, in modern deep submicron processes, V'_A is very low, resulting in V_A values that are lower than the corresponding values for the BJT.

The last, and perhaps most important, difference between the current–voltage characteristics of the two devices concerns the input current into the control terminal: While at low frequencies the gate current of the MOSFET is practically zero and the input resistance looking into the gate is practically infinite, the BJT draws base current i_B that is proportional to the collector current; that is, $i_B = i_C/\beta$. The finite base current and the corresponding finite input resistance looking into the base comprise a definite disadvantage of the BJT in comparison to the MOSFET. Indeed, it is the infinite input resistance of the MOSFET that has made possible analog and digital circuit applications that are not feasible with the BJT. Examples include dynamic digital memory (Chapter 16) and switched-capacitor filters (Chapter 17).

Example G.1

- For an NMOS transistor with $W/L = 10$ fabricated in the 0.18- μm process whose data are given in Table G.1, find the values of V_{OV} and V_{GS} required to operate the device at $I_D = 100 \mu\text{A}$. Ignore channel-length modulation.
- Find V_{BE} for an *n*pn transistor fabricated in the low-voltage process specified in Table G.2 and operated at $I_C = 100 \mu\text{A}$. Ignore base-width modulation.

Example G.1 *continued*

Solution

(a)

$$I_D = \frac{1}{2}(\mu_n C_{ox}) \left(\frac{W}{L}\right) V_{OV}^2$$

Substituting $I_D = 100 \mu\text{A}$, $W/L = 10$, and, from Table G.1, $\mu_n C_{ox} = 387 \mu\text{A}/\text{V}^2$ results in

$$100 = \frac{1}{2} \times 387 \times 10 \times V_{OV}^2$$

$$V_{OV} = 0.23 \text{ V}$$

Thus,

$$V_{GS} = V_m + V_{OV} = 0.5 + 0.23 = 0.73 \text{ V}$$

(b)

$$I_C = I_S e^{V_{BE}/V_T}$$

Substituting $I_C = 100 \mu\text{A}$ and, from Table G.2, $I_S = 6 \times 10^{-18} \text{ A}$ gives,

$$V_{BE} = 0.025 \ln \frac{100 \times 10^{-6}}{6 \times 10^{-18}} = 0.76 \text{ V}$$

EXERCISE

G.1 (a) For NMOS transistors fabricated in the 0.18- μm technology specified in Table G.1, find the range of I_D obtained for active-mode operation with V_{OV} ranging from 0.2 V to 0.4 V and $W/L = 0.1$ to 100. Neglect channel-length modulation.

(b) If a similar range of current is required in an *npn* transistor fabricated in the low-voltage process specified in Table G.2, find the corresponding change in its V_{BE} .

Ans. (a) $I_{D\min} = 0.8 \mu\text{A}$ and $I_{D\max} = 3.1 \text{ mA}$ for a range of about 4000:1; (b) for I_C varying over a 4000:1 range, $\Delta V_{BE} = 207 \text{ mV}$

G.3.3 Low-Frequency Small-Signal Models

The low-frequency models for the two devices are very similar except, of course, for the finite base current (finite β) of the BJT, which gives rise to r_π in the hybrid- π model and to the unequal emitter and collector currents in the T models $\alpha < 1$. Here it is interesting to note that the low-frequency, small-signal models become identical if one thinks of the MOSFET as a BJT with $\beta = \infty$ ($\alpha = 1$)

For both devices, the hybrid- π model indicates that the **open-circuit voltage gain** obtained from gate to drain (base to collector) with the source (emitter) grounded is $-g_m r_o$. It follows that $g_m r_o$ is the *maximum gain available from a single transistor* of either type. This important transistor parameter is given the name **intrinsic gain** and is denoted A_0 . We have more to say about the intrinsic gain in Section 8.3.2.

Although not included in the MOSFET low-frequency model shown in Table G.3, the body effect can have some implications for the operation of the MOSFET as an amplifier. In simple terms, if the body (substrate) is not connected to the source, it can act as a second gate for the MOSFET. The voltage signal that develops between the body and the source, v_{bs} , gives rise to a drain current component $g_{mb} v_{bs}$, where the body transconductance g_{mb} is proportional to g_m ; that is, $g_{mb} = \chi g_m$, where the factor χ is in the range of 0.1 to 0.2. The body effect has no counterpart in the BJT.

G.3.4 The Transconductance

For the BJT, the transconductance g_m depends *only* on the dc collector current I_C . (Recall that V_T is a physical constant $\simeq 0.025$ V at room temperature.) It is interesting to observe that g_m does not depend on the geometry of the BJT, and its dependence on the EBJ area is only through the effect of the area on the total collector current I_C . Similarly, the dependence of g_m on V_{BE} is only through the fact that V_{BE} determines the total current in the collector. By contrast, g_m of the MOSFET depends on I_D , V_{OV} , and W/L . Therefore, we use three different (but equivalent) formulas to express g_m of the MOSFET.

The first formula given in Table G.3 for the MOSFET's g_m is the most directly comparable with the formula for the BJT. It indicates that for the same operating current, g_m of the MOSFET is smaller than that of the BJT. This is because $V_{OV}/2$ is the range of 0.05 V to 0.15 V, which is two to six times the corresponding term in the BJT's formula, namely V_T .

The second formula for the MOSFET's g_m indicates that for a given device (i.e., given W/L), g_m is proportional to V_{OV} . Thus a higher g_m is obtained by operating the MOSFET at a higher overdrive voltage. However, we should recall the limitations imposed on the magnitude of V_{OV} by the limited value of V_{DD} . Put differently, the need to obtain a reasonably high g_m constrains the designer's interest in reducing V_{OV} .

The third g_m formula shows that for a given transistor (i.e., given W/L), g_m is proportional to $\sqrt{I_D}$. This should be contrasted with the bipolar case, where g_m is directly proportional to I_C .

G.3.5 Output Resistance

The output resistance for both devices is determined by similar formulas, with r_o being the ratio of V_A to the bias current (I_D or I_C). Thus, for both transistors, r_o is inversely proportional to the bias current. The difference in nature and magnitude of V_A between the two devices has already been discussed.

G.3.6 Intrinsic Gain

The intrinsic gain A_0 of the BJT is the ratio of V_A , which is solely a process parameter (5 V to 100 V), and V_T , which is a physical parameter (0.025 V at room temperature). Thus A_0 of a BJT is independent of the device junction area and of the operating current, and its value ranges from 200 V/V to 5000 V/V. The situation in the MOSFET is very different: Table G.3 provides three different (but equivalent) formulas for expressing the MOSFET's intrinsic

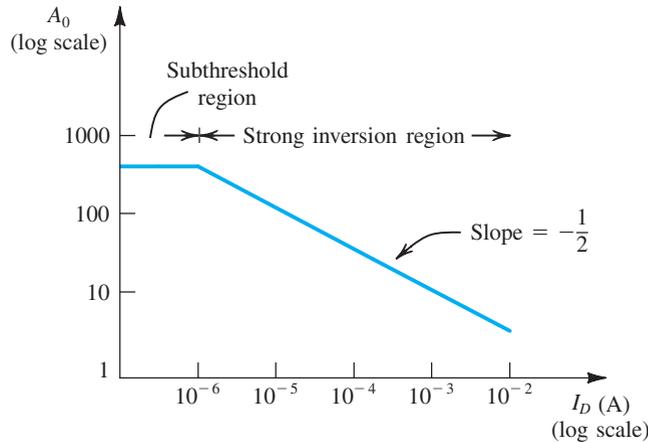


Figure G.1 The intrinsic gain of the MOSFET versus bias current I_D . Outside the subthreshold region, this is a plot of $A_0 = V'_A \sqrt{2\mu_n C_{ox} WL/I_D}$, for the case: $\mu_n C_{ox} = 20 \mu\text{A/V}^2$, $V'_A = 20 \text{V}/\mu\text{m}$, $L = 2 \mu\text{m}$, and $W = 20 \mu\text{m}$.

gain. The first formula is the one most directly comparable to that of the BJT. Here, however, we note the following:

1. The quantity in the denominator is $V_{OV}/2$, which is a design parameter, and although it is becoming smaller in designs using short-channel technologies, it is still at least two to four times larger than V_T . Furthermore, as we have seen, there are reasons for selecting larger values for V_{OV} .
2. The numerator quantity V_A is both process- and device-dependent, and its value has been steadily decreasing.

As a result, the intrinsic gain realized in a single MOSFET amplifier stage fabricated in a modern short-channel technology is only 20 V/V to 40 V/V, at least an order of magnitude lower than that for a BJT.

The third formula given for A_0 in Table G.3 points out a very interesting fact: For a given process technology (V'_A and $\mu_n C_{ox}$) and a given device (L and W), the intrinsic gain is inversely proportional to $\sqrt{I_D}$. This is illustrated in Fig. G.1, which shows a typical plot of A_0 versus the bias current I_D . The plot confirms that the gain increases as the bias current is lowered. The gain, however, levels off at very low currents. This is because the MOSFET enters the subthreshold region of operation (Section 5.1.9), where it becomes very much like a BJT with an exponential current–voltage characteristic. The intrinsic gain then becomes constant, just like that of a BJT. Note, however, that although a higher gain is achieved at lower bias currents, the price paid is a lower g_m and less ability to drive capacitive loads and thus a decrease in bandwidth. This point will be further illustrated shortly.

Example G.2

We wish to compare the values of g_m , input resistance at the gate (base), r_o , and A_0 for an NMOS transistor fabricated in the 0.25- μm technology specified in Table G.1 and an *npn* transistor fabricated in the

Example G.2 *continued*

low-voltage technology specified in Table G.2. Assume both devices are operating at a drain (collector) current of 100 μA . For the MOSFET, let $L = 0.4 \mu\text{m}$ and $W = 4 \mu\text{m}$, and specify the required V_{OV} .

Solution

For the NMOS transistor,

$$I_D = \frac{1}{2}(\mu_n C_{ox})\left(\frac{W}{L}\right)V_{OV}^2$$

$$100 = \frac{1}{2} \times 267 \times \frac{4}{0.4} \times V_{OV}^2$$

Thus,

$$V_{OV} = 0.27 \text{ V}$$

$$g_m = \sqrt{2(\mu_n C_{ox})\left(\frac{W}{L}\right)I_D}$$

$$= \sqrt{2 \times 267 \times 10 \times 100} = 0.73 \text{ mA/V}$$

$$R_{in} = \infty$$

$$r_o = \frac{V_A' L}{I_D} = \frac{5 \times 0.4}{0.1} = 20 \text{ k}\Omega$$

$$A_0 = g_m r_o = 0.73 \times 20 = 14.6 \text{ V/V}$$

For the *npn* transistor,

$$g_m = \frac{I_C}{V_T} = \frac{0.1 \text{ mA}}{0.025 \text{ V}} = 4 \text{ mA/V}$$

$$R_{in} = r_\pi = \beta_0 / g_m = \frac{100}{4 \text{ mA/V}} = 25 \text{ k}\Omega$$

$$r_o = \frac{V_A}{I_C} = \frac{35}{0.1 \text{ mA}} = 350 \text{ k}\Omega$$

$$A_0 = g_m r_o = 4 \times 350 = 1400 \text{ V/V}$$

EXERCISE

G.2 For an NMOS transistor fabricated in the 0.5- μm process specified in Table G.1 with $W = 5 \mu\text{m}$ and $L = 0.5 \mu\text{m}$, find the transconductance and the intrinsic gain obtained at $I_D = 10 \mu\text{A}$, $100 \mu\text{A}$, and 1 mA .

Ans. 0.2 mA/V, 200 V/V; 0.6 mA/V, 62 V/V; 2 mA/V, 20 V/V

G.3.7 High-Frequency Operation

The simplified high-frequency equivalent circuits for the MOSFET and the BJT are very similar, and so are the formulas for determining their unity-gain frequency (also called **transition frequency**) f_T . As we demonstrate in Chapter 10, f_T is a measure of the *intrinsic* bandwidth of the transistor itself and does *not* take into account the effects of capacitive loads. We address the issue of capacitive loads shortly. For the time being, note the striking similarity between the approximate formulas given in Table G.3 for the value of f_T of the two devices. In both cases f_T is inversely proportional to the square of the critical dimension of the device: the channel length for the MOSFET and the base width for the BJT. These formulas also clearly indicate that shorter-channel MOSFETs⁴ and narrower-base BJTs are inherently capable of a wider bandwidth of operation. It is also important to note that while for the BJT the approximate expression for f_T indicates that it is entirely process determined, the corresponding expression for the MOSFET shows that f_T is proportional to the overdrive voltage V_{OV} . Thus we have conflicting requirements on V_{OV} : While a higher low-frequency gain is achieved by operating at a low V_{OV} , wider bandwidth requires an increase in V_{OV} . Therefore the selection of a value for V_{OV} involves, among other considerations, a trade-off between gain and bandwidth.

For *npn* transistors fabricated in the modern low-voltage process, f_T is in the range of 10 GHz to 20 GHz as compared to the 400 MHz to 600 MHz obtained with the standard high-voltage process. In the MOS case, NMOS transistors fabricated in a modern submicron technology, such as the 0.18- μm process, achieve f_T values in the range of 5 GHz to 15 GHz.

Before leaving the subject of high-frequency operation, let's look into the effect of a capacitive load on the bandwidth of the common-source (common-emitter) amplifier. For this purpose we shall assume that the frequencies of interest are much lower than f_T of the transistor. Hence we shall not take the transistor capacitances into account. Figure G.2(a) shows a common-source amplifier with a capacitive load C_L . The voltage gain from gate to drain can be found as follows:

$$\begin{aligned} V_o &= -g_m V_{gs} (r_o \parallel C_L) \\ &= -g_m V_{gs} \frac{r_o \frac{1}{sC_L}}{r_o + \frac{1}{sC_L}} \\ A_v &= \frac{V_o}{V_{gs}} = -\frac{g_m r_o}{1 + sC_L r_o} \end{aligned} \quad (\text{G.1})$$

Thus the gain has, as expected, a low-frequency value of $g_m r_o = A_0$ and a frequency response of the single-time-constant (STC) low-pass type with a break (pole) frequency at

$$\omega_p = \frac{1}{C_L r_o} \quad (\text{G.2})$$

Obviously this pole is formed by r_o and C_L . A sketch of the magnitude of gain versus frequency is shown in Fig. G.2(b). We observe that the gain crosses the 0-dB line at

⁴Although the reason is beyond our capabilities at this stage, f_T of MOSFETs that have very short channels varies inversely with L rather than with L^2 .

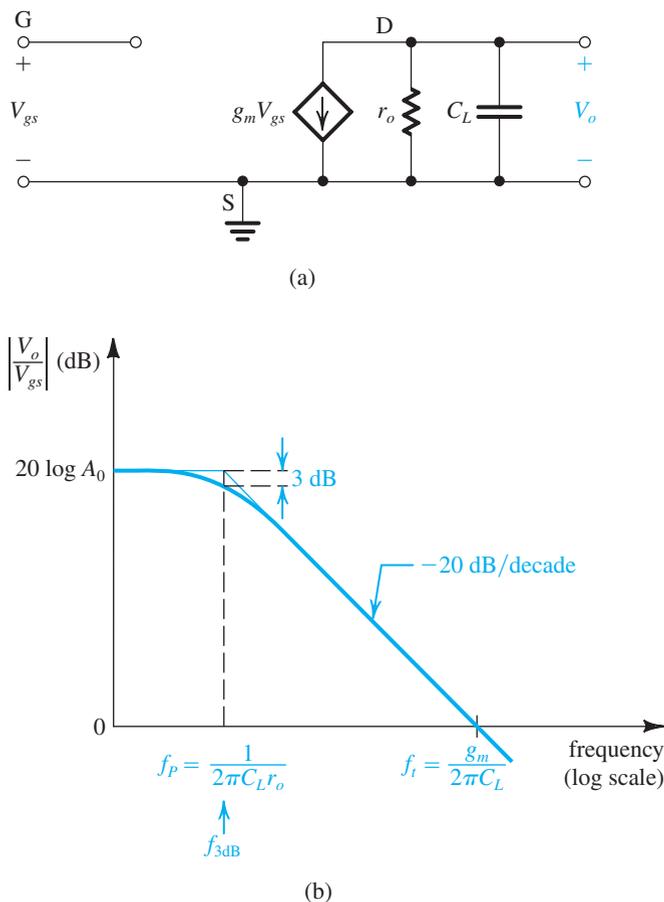


Figure G.2 Frequency response of a CS amplifier loaded with a capacitance C_L and fed with an ideal voltage source. It is assumed that the transistor is operating at frequencies much lower than f_T , and thus the internal capacitances are not taken into account.

frequency ω_t ,

$$\omega_t = A_0 \omega_p = (g_m r_o) \frac{1}{C_L r_o}$$

Thus,

$$\omega_t = \frac{g_m}{C_L} \quad (\text{G.3})$$

That is, the **unity-gain frequency** or, equivalently, the **gain–bandwidth product**⁵ ω_t is the ratio of g_m and C_L . We thus clearly see that for a given capacitive load C_L , a larger gain–bandwidth product is achieved by operating the MOSFET at a higher g_m . Identical analysis and conclusions apply to the case of the BJT. In each case, bandwidth increases as bias current is increased.

⁵The unity-gain frequency and the gain–bandwidth product of an amplifier are the same when the frequency response is of the single-pole type; otherwise the two parameters may differ.

G.3.8 Design Parameters

For the BJT there are three design parameters— I_C , V_{BE} , and I_S (or, equivalently, the area of the emitter–base junction)—and the designer can select any two. However, since I_C is exponentially related to V_{BE} and is very sensitive to the value of V_{BE} (V_{BE} changes by only 60 mV for a factor of 10 change in I_C), I_C is much more useful than V_{BE} as a design parameter. As mentioned earlier, the utility of the EBJ area as a design parameter is rather limited because of the narrow range over which A_E can vary. It follows that for the BJT there is only one effective design parameter: the collector current I_C . Finally, note that we have not considered V_{CE} to be a design parameter, since its effect on I_C is only secondary. Of course, as we learned in Chapter 7, V_{CE} affects the output-signal swing.

For the MOSFET there are four design parameters— I_D , V_{OV} , L , and W —and the designer can select any three. For analog circuit applications the trade-off in selecting a value for L is between the higher speed of operation (wider amplifier bandwidth) obtained at lower values of L and the higher intrinsic gain obtained at larger values of L . Usually one selects an L of about 25% to 50% greater than L_{\min} .

The second design parameter is V_{OV} . We have already made numerous remarks about the effect of the value of V_{OV} on performance. Usually, for submicron technologies, V_{OV} is selected in the range of 0.1 V to 0.3 V.

Once values for L and V_{OV} have been selected, the designer is left with the selection of the value of I_D or W (or, equivalently, W/L). For a given process and for the selected values of L and V_{OV} , I_D is proportional to W/L . It is important to note that the choice of I_D or, equivalently, of W/L has no bearing on the value of intrinsic gain A_0 and the transition frequency f_T . However, it affects the value of g_m and hence the gain–bandwidth product. Figure G.3 illustrates this point by showing how the gain of a common-source amplifier operated at a constant V_{OV} varies with I_D (or, equivalently, W/L). Note that while the dc gain remains unchanged, increasing W/L and, correspondingly, I_D , increases the bandwidth proportionally. This, however, assumes that the load capacitance C_L is not affected by the device size, an assumption that may not be entirely justified in some cases.

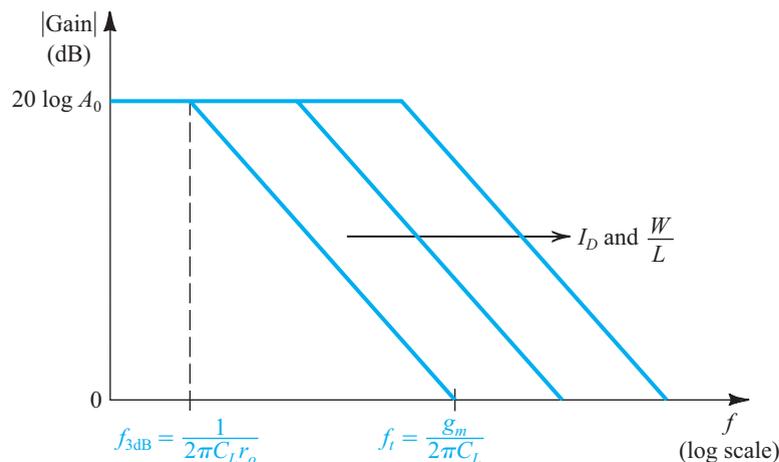


Figure G.3 Increasing I_D or W/L increases the bandwidth of a MOSFET amplifier operated at a constant V_{OV} and loaded by a constant capacitance C_L .

Example G.3

In this example we investigate the gain and the high-frequency response of an *npn* transistor and an NMOS transistor. For the *npn* transistor, assume that it is fabricated in the low-voltage process specified in Table G.2, and assume that $C_\mu \simeq C_{\mu 0}$. For $I_C = 10 \mu\text{A}$, $100 \mu\text{A}$, and 1 mA , find g_m , r_o , A_0 , C_{de} , C_{je} , C_π , C_μ , and f_T . Also, for each value of I_C , find the gain–bandwidth product f_i of a common-emitter amplifier loaded by a 1-pF capacitance, neglecting the internal capacitances of the transistor. For the NMOS transistor, assume that it is fabricated in the 0.25- μm CMOS process with $L = 0.4 \mu\text{m}$. Let the transistor be operated at $V_{OV} = 0.25 \text{ V}$. Find W/L that is required to obtain $I_D = 10 \mu\text{A}$, $100 \mu\text{A}$, and 1 mA . At each value of I_D , find g_m , r_o , A_0 , C_{gs} , C_{gd} , and f_T . Also, for each value of I_D , determine the gain–bandwidth product f_i of a common-source amplifier loaded by a 1-pF capacitance, neglecting the internal capacitances of the transistor.

Solution

For the *npn* transistor,

$$g_m = \frac{I_C}{V_T} = \frac{I_C}{0.025} = 40I_C \text{ A/V}$$

$$r_o = \frac{V_A}{I_C} = \frac{35}{I_C} \Omega$$

$$A_0 = \frac{V_A}{V_T} = \frac{35}{0.025} = 1400 \text{ V/V}$$

$$C_{de} = \tau_F g_m = 10 \times 10^{-12} \times 40I_C = 0.4 \times 10^{-9} I_C \text{ F}$$

$$C_{je} \simeq 2C_{je0} = 10 \text{ fF}$$

$$C_\pi = C_{de} + C_{je}$$

$$C_\mu \simeq C_{\mu 0} = 5 \text{ fF}$$

$$f_T = \frac{g_m}{2\pi(C_\pi + C_\mu)}$$

$$f_i = \frac{g_m}{2\pi C_L} = \frac{g_m}{2\pi \times 1 \times 10^{-12}}$$

We thus obtain the following results:

I_C	g_m (mA/V)	r_o (k Ω)	A_0 (V/V)	C_{de} (fF)	C_{je} (fF)	C_π (fF)	C_μ (fF)	f_T (GHz)	f_i (MHz)
10 μA	0.4	3500	1400	4	10	14	5	3.4	64
100 μA	4	350	1400	40	10	50	5	11.6	640
1 mA	40	35	1400	400	10	410	5	15.3	6400

Example G.3 *continued*

For the NMOS transistor,

$$I_D = \frac{1}{2} \mu_n C_{ox} \frac{W}{L} V_{ov}^2$$

$$= \frac{1}{2} \times 267 \times \frac{W}{L} \times \frac{1}{16}$$

Thus,

$$\frac{W}{L} = 0.12 I_D$$

$$g_m = \frac{I_D}{V_{ov}/2} = \frac{I_D}{0.25/2} = 8 I_D \text{ A/V}$$

$$r_o = \frac{V_A' L}{I_D} = \frac{5 \times 0.4}{I_D} = \frac{2}{I_D} \Omega$$

$$A_0 = g_m r_o = 16 \text{ V/V}$$

$$C_{gs} = \frac{2}{3} W L C_{ox} + C_{ov} = \frac{2}{3} W \times 0.4 \times 5.8 + 0.6 W$$

$$C_{gd} = C_{ov} = 0.6 W$$

$$f_T = \frac{g_m}{2\pi(C_{gs} + C_{gd})}$$

$$f_i = \frac{g_m}{2\pi C_L}$$

We thus obtain the following results:

I_D	W/L	g_m (mA/V)	r_o (k Ω)	A_0 (V/V)	C_{gs} (fF)	C_{gd} (fF)	f_T (GHz)	f_i (MHz)
10 μ A	1.2	0.08	200	16	1.03	0.29	9.7	12.7
100 μ A	12	0.8	20	16	10.3	2.9	9.7	127
1 mA	120	8	2	16	103	29	9.7	1270

EXERCISE

G.3 Find I_D , g_m , r_o , A_0 , C_{gs} , C_{gd} , and f_T for an NMOS transistor fabricated in the 0.5- μ m CMOS technology specified in Table G.1. Let $L = 0.5 \mu\text{m}$, $W = 5 \mu\text{m}$, and $V_{ov} = 0.3 \text{ V}$.

Ans. 85.5 μ A; 0.57 mA/V; 117 k Ω ; 66.7 V/V; 8.3 fF; 2 fF; 8.8 GHz

G.4 Combining MOS and Bipolar Transistors—BiCMOS Circuits

From the discussion above it should be evident that the BJT has the advantage over the MOSFET of a much higher transconductance (g_m) at the same value of dc bias current. Thus, in addition to realizing higher voltage gains per amplifier stage, bipolar transistor amplifiers have superior high-frequency performance compared to their MOS counterparts.

On the other hand, the practically infinite input resistance at the gate of a MOSFET makes it possible to design amplifiers with extremely high input resistances and an almost zero input bias current. Also, as mentioned earlier, the MOSFET provides an excellent implementation of a switch, a fact that has made CMOS technology capable of realizing a host of analog circuit functions that are not possible with bipolar transistors.

It can thus be seen that each of the two transistor types has its own distinct and unique advantages: Bipolar technology has been extremely useful in the design of very-high-quality general-purpose circuit building blocks, such as op amps. On the other hand, CMOS, with its very high packing density and its suitability for both digital and analog circuits, has become the technology of choice for the implementation of very-large-scale integrated circuits. Nevertheless, the performance of CMOS circuits can be improved if the designer has available (on the same chip) bipolar transistors that can be employed in functions that require their high g_m and excellent current-driving capability. A technology that allows the fabrication of high-quality bipolar transistors on the same chip as CMOS circuits is aptly called **BiCMOS**. At appropriate locations throughout this book we present interesting and useful BiCMOS circuit blocks.

G.5 Validity of the Square-Law MOSFET Model

We conclude this appendix with a comment on the validity of the simple square-law model we have been using to describe the operation of the MOS transistor. While this simple model works well for devices with relatively long channels ($>1 \mu\text{m}$), it does *not* provide an accurate representation of the operation of short-channel devices. This is because a number of physical phenomena come into play in these submicron devices, resulting in what are called **short-channel effects**. Although a detailed study of short-channel effects is beyond the scope of this book, it should be mentioned that MOSFET models have been developed that take these effects into account. However, they are understandably quite complex and do not lend themselves to hand analysis of the type needed to develop insight into circuit operation. Rather, these models are suitable for computer simulation and are indeed used in SPICE (Appendix B). For quick, manual analysis, however, we will continue to use the square-law model, which is the basis for the comparison of Table G.3.

G.1 Find the range of I_D obtained in a particular NMOS transistor as its overdrive voltage is increased from 0.15 V to 0.4 V. If the same range is required in I_C of a BJT, what is the corresponding change in V_{BE} ?

G.2 What range of I_C is obtained in an *npn* transistor as a result of changing the area of the emitter–base junction by a factor of 10 while keeping V_{BE} constant? If I_C is to be kept constant, by what amount must V_{BE} change?

G.3 For each of the CMOS technologies specified in Table G.1, find the $|V_{OV}|$ and hence the $|V_{GS}|$ required to operate a device with a W/L of 10 at a drain current $I_D = 100 \mu\text{A}$. Ignore channel-length modulation.

G.4 Consider NMOS and PMOS devices fabricated in the 0.25- μm process specified in Table G.1. If both devices are to operate at $|V_{OV}| = 0.25 \text{ V}$ and $I_D = 100 \mu\text{A}$, what must their W/L ratios be?

G.5 Consider NMOS and PMOS transistors fabricated in the 0.25- μm process specified in Table G.1. If the two devices are to be operated at equal drain currents, what must the ratio of $(W/L)_p$ to $(W/L)_n$ be to achieve equal values of g_m ?

G.6 An NMOS transistor fabricated in the 0.18- μm CMOS process specified in Table G.1 is operated at $V_{OV} = 0.2 \text{ V}$. Find the required W/L and I_D to obtain a g_m of 10 mA/V. At what value of I_C must an *nnpn* transistor be operated to achieve this value of g_m ?

G.7 For each of the CMOS process technologies specified in Table G.1, find the g_m of an NMOS and a PMOS transistor with $W/L = 10$ operated at $I_D = 100 \mu\text{A}$.

G.8 An NMOS transistor operated with an overdrive voltage of 0.25 V is required to have a g_m equal to that of an *nnpn* transistor operated at $I_C = 0.1 \text{ mA}$. What must I_D be? What value of g_m is realized?

G.9 It is required to find the incremental (i.e., small-signal) resistance of each of the diode-connected transistors shown in Fig. PG.9. Assume that the dc bias current $I = 0.1 \text{ mA}$. For the MOSFET, let $\mu_n C_{ox} = 200 \mu\text{A}/\text{V}^2$ and $W/L = 10$.

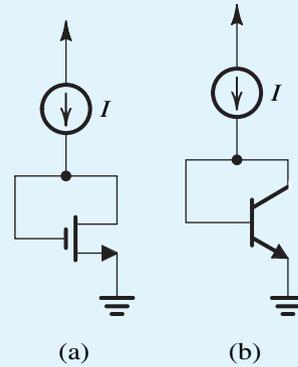


Figure PG.9

G.10 For an NMOS transistor with $L = 1 \mu\text{m}$ fabricated in the 0.8- μm process specified in Table G.1, find g_m , r_o , and A_0 if the device is operated with $V_{OV} = 0.5 \text{ V}$ and $I_D = 100 \mu\text{A}$. Also, find the required device width W .

G.11 For an NMOS transistor with $L = 0.3 \mu\text{m}$ fabricated in the 0.18- μm process specified in Table G.1, find g_m , r_o , and A_0 obtained when the device is operated at $I_D = 100 \mu\text{A}$ with $V_{OV} = 0.2 \text{ V}$. Also, find W .

G.12 Fill in the table below. For the BJT, let $\beta = 100$ and $V_A = 100 \text{ V}$. For the MOSFET, let $\mu_n C_{ox} = 200 \mu\text{A}/\text{V}^2$, $W/L = 40$, and $V_A = 10 \text{ V}$. Note that R_{in} refers to the input resistance at the control input terminal (gate, base) with the (source, emitter) grounded.

Bias Current	BJT		MOSFET	
	$I_C = 0.1 \text{ mA}$	$I_C = 1 \text{ mA}$	$I_D = 0.1 \text{ mA}$	$I_D = 1 \text{ mA}$
g_m (mA/V)				
r_o (k Ω)				
A_0 (V/V)				
R_{in} (k Ω)				

G.13 For an NMOS transistor fabricated in the 0.18- μm process specified in Table G.1 with $L = 0.3 \mu\text{m}$ and $W = 6 \mu\text{m}$, find the value of f_T obtained when the transistor is operated at $V_{OV} = 0.2 \text{ V}$. Use both the formula in terms of C_{gs} and C_{gd} and the approximate formula. Why does the approximate formula overestimate f_T ?

G.14 An NMOS transistor fabricated in the 0.18- μm process specified in Table G.1 and having $L = 0.3 \mu\text{m}$ and $W = 6 \mu\text{m}$ is operated at $V_{OV} = 0.2 \text{ V}$ and used to drive a

capacitive load of 100 fF. Find A_0 , f_p (or $f_{3\text{ dB}}$), and f_T . At what I_D value is the transistor operating? If it is required to double f_T , what must I_D become? What happens to A_0 and f_p in this case?

G.15 For an *nnp* transistor fabricated in the high-voltage process specified in Table G.2, evaluate f_T at $I_C = 10\ \mu\text{A}$, $100\ \mu\text{A}$, and $1\ \text{mA}$. Assume $C_\mu \simeq C_{\mu 0}$. Repeat for the low-voltage process.

G.16 Consider an NMOS transistor fabricated in the $0.8\text{-}\mu\text{m}$ process specified in Table G.1. Let the transistor have $L = 1\ \mu\text{m}$, and assume it is operated at $I_D = 100\ \mu\text{A}$.

- For $V_{OV} = 0.25\ \text{V}$, find W , g_m , r_o , A_0 , C_{gs} , C_{gd} , and f_T .
- To what must V_{OV} be changed to double f_T ? Find the new values of W , g_m , r_o , A_0 , C_{gs} , and C_{gd} .

G.17 For a lateral *ppn* transistor fabricated in the high-voltage process specified in Table G.2, find f_T if the

device is operated at a collector bias current of $1\ \text{mA}$. Compare to the value obtained for a vertical *nnp*.

G.18 Show that for a MOSFET the selection of L and V_{OV} determines A_0 and f_T . In other words, show that A_0 and f_T will not depend on I_D and W .

G.19 Consider an NMOS transistor fabricated in the $0.18\text{-}\mu\text{m}$ technology specified in Table G.1. Let the transistor be operated at $V_{OV} = 0.2\ \text{V}$. Find A_0 and f_T for $L = 0.2\ \mu\text{m}$, $0.3\ \mu\text{m}$, and $0.4\ \mu\text{m}$.

D G.20 Consider an NMOS transistor fabricated in the $0.5\text{-}\mu\text{m}$ process specified in Table G.1. Let $L = 0.5\ \mu\text{m}$ and $V_{OV} = 0.3\ \text{V}$. If the MOSFET is connected as a common-source amplifier with a load capacitance $C_L = 1\ \text{pF}$ (as in Fig. G.2a), find the required transistor width W and bias current I_D to obtain a unity-gain bandwidth of $100\ \text{MHz}$. Also, find A_0 and $f_{3\text{ dB}}$.