

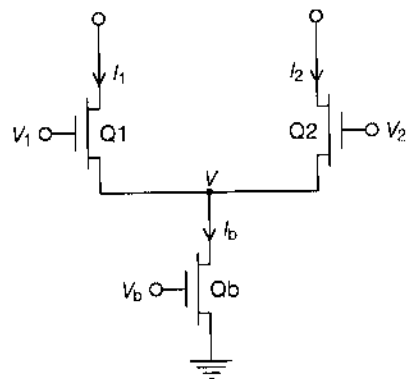
## TRANSCONDUCTANCE AMPLIFIER

In this chapter, we will discuss our first circuit. It is the most important circuit that we will treat, because we will use it for almost everything we do. The circuit is called the transconductance amplifier. We will explain how it works and examine some of its characteristics.

This amplifier is a device that generates as its output a current that is a function of the difference between two input voltages,  $V_1$  and  $V_2$ ; that difference is called the **differential input** voltage. The circuit is called a **differential transconductance amplifier**. An ordinary *conductance* turns a voltage difference across two terminals into a current through the *same* two terminals. A *transconductance* turns a voltage difference *somewhere* into a current *somewhere else*. In the transconductance amplifier, a voltage difference between two inputs creates a current as the output.

### DIFFERENTIAL PAIR

Many circuits take an input signal represented as a *difference between two voltages*. These circuits all use some variant of the **differential pair** shown in Figure 5.1 as an input stage. Because the differential pair is so universally useful, we will analyze its characteristics here; then we will show how it is used in the transconductance amplifier.



**FIGURE 5.1** Schematic of the differential pair. The bias current  $I_b$  is set by the bias voltage  $V_b$ , and is divided between  $I_1$  and  $I_2$  depending on the difference between  $V_1$  and  $V_2$ .

The bottom transistor Qb is used as a current source; under normal circumstances, its drain voltage  $V$  is large enough that the drain current  $I_b$  is saturated at a value set by the gate voltage  $V_b$ . The manner in which  $I_b$  is divided between Q1 and Q2 is a sensitive function of the difference between  $V_1$  and  $V_2$ , and is the essence of the operation of the stage.

From Equation 3.15 (p. 39), we know that the saturated drain current  $I_{\text{sat}}$  is exponential in the gate and source voltages

$$I_{\text{sat}} = I_0 e^{\kappa V_g - V_s}$$

Applying this expression to Q1 and Q2, we obtain

$$I_1 = I_0 e^{\kappa V_1 - V} \quad \text{and} \quad I_2 = I_0 e^{\kappa V_2 - V} \quad (5.1)$$

The sum of the two drain currents must be equal to  $I_b$ :

$$I_b = I_1 + I_2 = I_0 e^{-V} (e^{\kappa V_1} + e^{\kappa V_2})$$

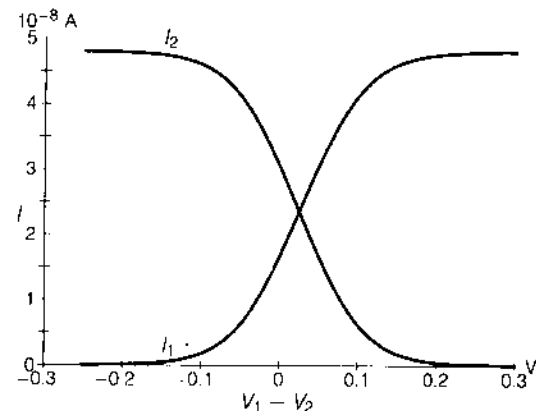
We can solve this equation for the voltage  $V$ :

$$e^{-V} = \frac{I_b}{I_0} \frac{1}{e^{\kappa V_1} + e^{\kappa V_2}} \quad (5.2)$$

Substituting Equation 5.2 into Equation 5.1, we obtain expressions for the two drain currents:

$$I_1 = I_b \frac{e^{\kappa V_1}}{e^{\kappa V_1} + e^{\kappa V_2}} \quad \text{and} \quad I_2 = I_b \frac{e^{\kappa V_2}}{e^{\kappa V_1} + e^{\kappa V_2}} \quad (5.3)$$

If  $V_1$  is more positive than  $V_2$  by many  $kT/(q\kappa)$ , transistor Q2 gets turned off, so essentially all the current goes through Q1,  $I_1$  is approximately equal to  $I_b$ , and  $I_2$  is approximately equal to 0. Conversely, if  $V_2$  is more positive than  $V_1$  by many  $kT/(q\kappa)$ , Q1 gets turned off,  $I_2$  is approximately equal to  $I_b$ , and  $I_1$  is approximately equal to 0. The two currents out of a differential pair are shown as a function of  $V_1 - V_2$  in Figure 5.2.



**FIGURE 5.2** Output currents of the differential pair as a function of differential input voltage. The sum of the two currents is equal to the bias current  $I_b$ . The curves do not cross at zero input-voltage difference due to mismatch between Q1 and Q2.

The differential transconductance amplifier, like many other circuits that we will study, uses various kinds of current mirrors to generate an output current that is proportional to the difference between the two drain currents. This difference is

$$I_1 - I_2 = I_b \frac{e^{\kappa V_1} - e^{\kappa V_2}}{e^{\kappa V_1} + e^{\kappa V_2}} \quad (5.4)$$

Multiplying both the numerator and denominator of Equation 5.4 by  $e^{-(V_1+V_2)/2}$ , we can express every exponent in terms of voltage differences. The result is

$$\begin{aligned} I_1 - I_2 &= I_b \frac{e^{\kappa(V_1-V_2)/2} - e^{-\kappa(V_1-V_2)/2}}{e^{\kappa(V_1-V_2)/2} + e^{-\kappa(V_1-V_2)/2}} \\ &= I_b \tanh \frac{\kappa(V_1 - V_2)}{2} \end{aligned} \quad (5.5)$$

The tanh is one of the few functions in the world that displays truly civilized behavior. It goes through the origin with unity slope, becomes +1 for large positive arguments, and becomes -1 for large negative arguments. How about the factor of two in the denominator of Equation 5.5—where did it come from? Let us examine what happens when only small changes are made in  $V_1$  and  $V_2$ . We increase  $V_1$  and decrease  $V_2$  such that  $V$  is kept constant. The current through Q2 goes down exponentially and the current through Q1 goes up exponentially. The difference in voltages, however, is large as  $V_2$  relative to  $V$ , or as  $V_1$  relative to  $V$ . That is why the curves of Figure 5.2 take twice as much voltage to saturate as do the single transistor curves in Chapter 3; this is the origin of the factor of two in Equation 5.5.

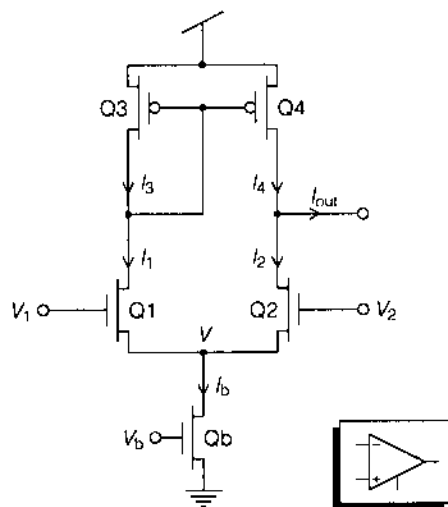
## SIMPLE TRANSCONDUCTANCE AMPLIFIER

The schematic for the transconductance amplifier is shown in Figure 5.3. The circuit consists of a differential pair and a single current mirror, like the one shown in Figure 3.9 (p. 40), which is used to subtract the drain currents  $I_1$  and  $I_2$ . The current  $I_1$  drawn out of Q3 is reflected as an equal current out of Q4; the output current is thus equal to  $I_1 - I_2$ , and is therefore given by Equation 5.5.

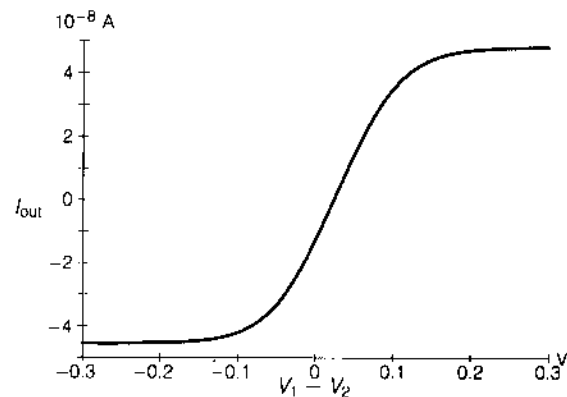
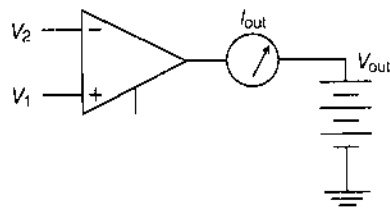
We can measure the current out of the amplifier as a function of the input voltages using the setup shown in Figure 5.4. We are using a current meter with its primary input connected to the amplifier output and its reference input connected to a voltage source  $V_{out}$ . A perfect current meter has zero resistance; real current meters have sophisticated feedback arrangements to make their input resistances very small. For that reason, the voltage on the output node of the amplifier will be  $V_{out}$ . For now, we will simply set  $V_{out}$  in the midrange between  $V_{DD}$  and ground. In later sections, we will investigate the effect of  $V_{out}$  on the performance of the circuit.

The current out of the simple amplifier is plotted as a function of  $V_1 - V_2$  in Figure 5.5. The curve is very close to a tanh, as expected. We can determine the effective value of  $kT/(q\kappa)$  by extrapolating the slope of the curve at

**FIGURE 5.3** Schematic diagram of the simple transconductance amplifier. The current mirror formed by Q3 and Q4 is used to form the output current, which is equal to  $I_1 - I_2$ . The symbol used for the circuit is shown in the inset.



**FIGURE 5.4** Arrangement for measuring the output current of the transconductance amplifier. The circuit of Figure 5.3 is represented symbolically.



**FIGURE 5.5** Output current of the transconductance amplifier as a function of differential input voltage. The mismatch between transistor characteristics can be seen in two ways. For this particular amplifier, the input offset voltage is approximately 25 millivolts, typical for a digital CMOS process. The limiting current for positive inputs is approximately 6 percent larger than that for negative inputs; a more typical variation would be 20 percent.

the origin to the two asymptotes. The difference between the positive and negative intercepts should be  $4kT/(q\kappa)$ . Using this procedure on Figure 5.5, we obtain the result that  $kT/(q\kappa)$  is approximately 43 millivolts, giving  $\kappa$  as approximately 0.58. This value is in good agreement with that obtained from the voltage dependence of the saturation current (Figure 3.7 (p. 38)). The **transconductance**  $G_m$  of the amplifier is just the slope of the tanh in Equation 5.5 at the origin. For Figure 5.5, the output current changes  $5.6 \times 10^{-8}$  amp for a 100-millivolt change in  $V_{in} = V_1 - V_2$ .  $G_m$  is therefore  $5.6 \times 10^{-7}$  mho. In terms of the circuit variables,

$$G_m = \frac{\partial I_{out}}{\partial V_{in}} = \frac{I_b}{2kT/(q\kappa)} \quad (5.6)$$

Notice that the transconductance is proportional to the bias current  $I_b$ , a fact that will become important when the differential circuit is used to produce a voltage-type output, or as part of a multiplier.

The layout of a typical implementation of the simple amplifier is shown in Plate 7(a).

## CIRCUIT LIMITATIONS

Now that we know how an ideal transconductance amplifier works, we can investigate the limitations and imperfect behavior of such circuits in the real world. Deviations from ideal behavior are of two basic sorts:

1. Mismatch between transistors

2. Deviation of a transistor from perfect current-source behavior; this second class of nonideality manifests itself in two ways:
  - a. Voltage limitations due to transistors coming out of saturation
  - b. Finite slope of the drain curves in saturation

### Transistor Mismatch

Unlike people in the United States, not all transistors are created equal. Some are created with higher values of  $I_0$  than are others. The effects of differences of  $I_0$  between transistors can be seen in Figure 5.5. The circuit from which these data were taken is typical: The tanh curve is shifted by about 25 millivolts. In addition, the saturated current coming out of Q4 is not the same as the current coming out of Q2. In other words, the negative asymptote is not the same as the positive asymptote. In Figure 5.5, the difference is about 6 percent.

The Q3–Q4 current mirror does not have 100-percent reflectivity. What we take out of Q3 does not necessarily come out of Q4, because Q4 may have a slightly larger or smaller value of  $I_0$  than does Q3. Differences of a factor of two between  $I_0$  values of nominally identical transistors are observed in circuits such as this. A more typical number for transistors that are physically adjacent is  $\pm 20$  percent, corresponding to a difference in gate voltage of  $\pm 10$  millivolts.

Note that the difference across a whole chip is not much bigger than a difference between two reasonably closely spaced transistors. The  $I_0$  variation occurs on a small distance scale. For this reason, putting transistors physically close to one another will not eliminate the problem.

One of the things we will emphasize in this book is that we can design circuits in a manner such that these variations *can* be tolerated. The voltage difference matters in some applications; it does not matter in others. It is a good habit not to trust the transistors to be closer than a factor of two in current (approximately  $\pm 30$  millivolts in gate voltage). We will try to build circuits that will work in any application given this limitation. The best way to ensure that a circuit will tolerate such variations is to have it **self-compensate** for the voltage offsets.

Self-compensation has another advantage: As circuits age and change and shift with time, the system tunes itself up. More important, in the kinds of applications we will be examining (such as vision and hearing) the systems *must* be self-adjusting. Because we have no idea what the input signals will be, the circuit had better be self-adjusting for whatever operating conditions it encounters.

### Output-Voltage Limitation

One small detail we have not mentioned is where the current  $I_{out}$  goes. To understand the circuit, we have to be more specific about the  $I_{out}$  destination:  $I_{out}$  goes into a current meter so that we can measure it. That current meter may be fictitious (invented so that we can reason about the next circuit that we are driving), or it may be a real. In either case, the other end of the meter is

at some voltage, which we call  $V_{out}$ . This voltage is imposed by the circuit into which this output is fed.

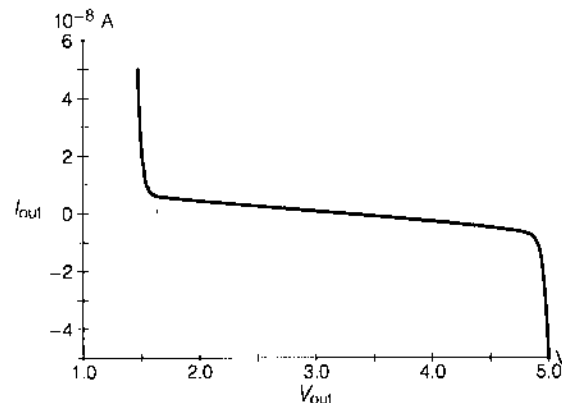
When we construct large integrated systems, it is important that we realize how few *primary inputs* there are compared with the enormous number of interconnections in the computing machinery itself. A 1988 chip may contain approximately  $10^6$  interconnected transistors in a 40-pin package. The entire auditory system in the brain is stimulated with only two audio input channels. The human visual system, which occupies 40 percent of the brain, is driven by the light impinging on the retina through two small lenses. For this reason, the internal rules of processing, whether of brain or of chip, can (in fact *must*) be made to facilitate the computation process. Interfaces to the outside world will be handled as exceptions.

Thus, from a practical point of view, we should assume that all inputs come from some other circuit, and all outputs go to some other circuit. For the system to work, the preceding circuit had to have an opinion, and the next circuit will certainly have an opinion. The preceding circuit generates the input, and the next circuit receives the output. We have to do something with  $I_{out}$ : dump it into a capacitor, for example, or hook it into a current mirror. What we do with it depends on what we are building.

Our first question is: How far we can trust this current  $I_{out}$  to be independent of  $V_{out}$ ? For unlimited variation in  $V_{out}$ , the answer is: not at all. Nothing has unlimited range. There is, however, a range of output voltages over which the output current will be nearly constant. For the circuit of Figure 5.3, a plot of the variation of  $I_{out}$  as a function of  $V_{out}$  for fixed values of  $V_1$  and  $V_2$  is shown in Figure 5.6. There is an upper limit near  $V_{DD}$ , where the output current increases rapidly. There also is a lower limit, below which the output current decreases rapidly. In between, the current is not independent of output voltage, but it has a near-constant slope corresponding to a **finite output conductance**. We will discuss the two limits first; then we will examine the effect of the slope in the middle.

**Upper limit** We should not expect (or desire) this device to be able to put out a constant current at voltages larger than  $V_{DD}$  or smaller than zero. That limitation is important; it means that the circuit cannot generate for the next stage a voltage that is outside those limits. If we raise the output voltage above  $V_{DD}$ , the drain of Q4 becomes the source, and we start draining current out of the ammeter down through Q4 to  $V_{DD}$ . Q4 will be turned on, because the voltage on its gate is less than  $V_{DD}$ —because that voltage is generated by Q3. Even if Q1 is turned off, the worst that can happen is that the voltage on the gate of Q4 will approach  $V_{DD}$ . So, if  $V_{out}$  gets a few tenths of a volt above  $V_{DD}$ , we will start to get an exponential negative increase in  $I_{out}$ .

**Lower limit** The lower limit to  $V_{out}$  is a little more tricky. First, let us consider the case where  $V_1$  is greater than  $V_2$  by several  $kT/(q\kappa)$ . Under these conditions,  $V$  is approximately equal to  $\kappa(V_1 - V_2)$ ,  $I_2$  is approximately equal to 0, and  $I_{out}$



**FIGURE 5.6** Dependence of output current of the simple transconductance amplifier on output voltage. The exponential dependencies at the ends are due to the high ( $V_{DD}$ ) and low ( $V_{min}$ ) limits discussed in the text. The slope in the midrange, due to the nonzero drain conductance of the output transistors Q2 and Q4, is the output conductance of the amplifier. The curve shown is for  $V_1$  approximately equal to  $V_2$ ; input-voltage differences cause the entire curve to shift up or down, according to the relationship shown in Figure 5.5.

is positive, approximately equal to  $I_4$ . As we lower the output voltage, all is well until  $V_{out}$  decreases to less than  $V$ , after which the output node becomes the *source* of Q2, and the  $V$  node becomes the *drain*. The interchange of source and drain of Q2 results in a reversal of current through Q2— $I_2$  becomes negative instead of positive. The reversal occurs when  $V_{out}$  is equal to  $\kappa(V_1 - V_b)$ , but is not noticeable in the output current until  $V_{out}$  is approximately equal to  $\kappa(V_2 - V_b)$ , where  $I_2$  becomes comparable with  $I_1$ . A further decrease in output voltage results in an exponential increase in  $I_{out}$ , because the gate-source voltage of Q2 is increasing. This negative  $I_2$  is supplied by an increase in  $I_1$ , which results in an equal increase in output current through Q4. The output current thus increases from two equal contributions of the same sign.

If  $V_2$  is greater than  $V_1$  by several  $kT/(q\kappa)$ , the same effect can be observed. The output current is negative, and  $V$  is equal to  $\kappa(V_2 - V_b)$ . As we decrease the output voltage, we make the voltage between the source and the drain of Q2 smaller and smaller, Q2 comes out of saturation, and  $V$  begins to decrease. As both  $V_{out}$  and  $V$  decrease, the gate-source voltage of Q2 increases, causing Q2 to conduct more current. The voltage  $V$  follows  $V_{out}$  more and more closely. There is no noticeable change in output current, however, until  $V$  approaches  $\kappa(V_1 - V_b)$ , at which point the current through Q1 becomes comparable to  $I_b$ . As we decrease the voltage at the output node further,  $I_1$  exceeds  $I_b$ , and  $V$  does not decrease as fast as does  $V_{out}$ . Once  $V$  is greater than  $V_{out}$ , the drain and source of Q2 are interchanged, and the situation is exactly as it was for  $V_2$  greater than  $V_1$ . Transistor Q2 starts siphoning charge away from the  $V$  node, and the output current increases exponentially.

We call the limitation on the operation of the simple transconductance amplifier imposed by this behavior the “ $V_{min}$  problem.” We can express the minimum output voltage as

$$V_{min} = \kappa(\min(V_1, V_2) - V_b) \quad (5.7)$$

In other words, the amplifier will work with its output voltage up to nearly  $V_{DD}$ , and down to  $V_b$  below the lowest input signal that we have applied to it, but not lower than that.

We run into two walls, one on the top and one on the bottom. The wall on the top side is not serious; all it does is to prevent us from going right up to  $V_{DD}$ . When  $V_1$  is greater than  $V_2$ , the current comes out of Q4—so, if we make the output node equal to  $V_{DD}$ , we will not get any current out. We cannot quite work up against the rail, but we can get very close. The upper limit on the output voltage is set by the saturation properties of Q4. As long as we stay a few  $kT/(q\kappa)$  below  $V_{DD}$ , we are fine.

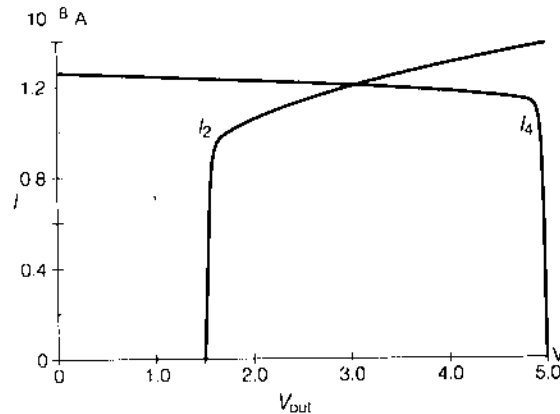
The bottom  $V_{min}$  limit is much more serious. It is the biggest problem with this circuit. It forms a *hard limit* below which the circuit does not work, and that limit depends on the input voltage.

## VOLTAGE OUTPUT

We call these circuits “transconductance amplifiers” because that is the way in which they are usually used. They also can be used, however, to take a difference in voltage at the input and turn it into a voltage at the output. Instead of measuring  $I_{out}$  with an ammeter, we measure  $V_{out}$  with a voltmeter. The drain conductances of Q2 and Q4 are used to convert the output current into an output voltage.

The drain current of a transistor is not completely independent of its drain voltage, even in saturation. There is a finite slope of  $I_d$  versus  $V_d$  given by the **Early effect**, discussed in Appendix B. This effect is responsible for the dependence of output current on output voltage seen between the two limits in Figure 5.6.

The finite output conductance of the circuit can be used to convert the current-type output signal into a voltage-type signal. We take away the  $V_{out}$  voltage source completely, and use a voltmeter instead of an ammeter. An ideal voltmeter draws no current from the circuit it is measuring. It is an ideal open-circuit, having zero input conductance (infinite input resistance). The input conductances of real voltmeters vary over many orders of magnitude. A good electrometer has an input resistance greater than  $10^{12}$  ohm; ordinary voltmeters can have resistances many orders of magnitude less. A typical oscilloscope has an input resistance of  $10^6$  ohm. We must always check that the current drawn by our measuring equipment is small compared with the current flowing in the circuit. Using a good voltmeter, we can observe the voltage at which the output current is equal to zero. That is the **open-circuit output voltage** of the device.



**FIGURE 5.7** Current out of the  $n$ -channel (Q2) and  $p$ -channel (Q4) output transistors of a simple transconductance amplifier, as a function of output voltage, for  $V_1$  approximately equal to  $V_2$ . The output current is the difference between the two curves; therefore, the open-circuit output voltage is given by the intersection of these curves.

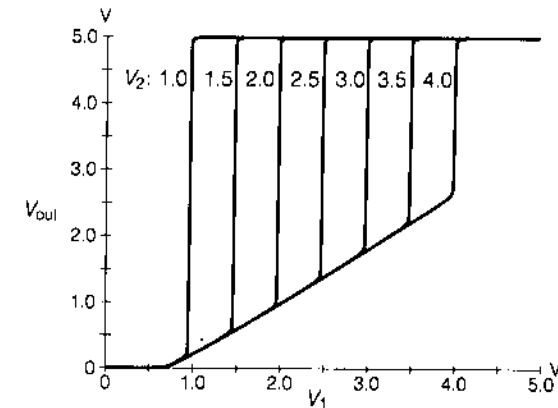
Consider the combination of Q4 and Q2. In Figure 5.7, we plot the magnitude of  $I_4$  and  $I_2$  versus the output voltage. Current is flowing down from  $V_{DD}$  to ground— $I_4$  is flowing *into* the output node, and  $I_2$  is flowing *out of* the output node. For any given output voltage, the output current is the difference between the two curves. For any particular input voltages, the value of  $V_{min}$  will be somewhat below the lowest input— for the operating conditions shown in Figure 5.7,  $I_2$  goes to zero when  $V_{out}$  is approximately equal to  $V_{min}$ . Because Q4 is a  $p$ -channel device, its drain voltage is plotted downward from  $V_{DD}$ . The open-circuit voltage is the value at which the two currents are equal; *open-circuit* means *no current*.

We can easily see what the output voltage would do as a function of a difference in the input voltages, if the drain curves for Q2 and Q4 were absolutely flat. When  $V_1$  is a little less than  $V_2$ , the output voltage will decrease to nearly  $V_{min}$ ; when  $V_1$  is a little bigger than  $V_2$ , the voltage will increase to almost  $V_{DD}$ . The experimental dependence of  $V_{out}$  on  $V_1$ , for several values of  $V_2$ , is shown in Figure 5.8. The output voltage stays at 5 volts until  $V_1$  gets very close to  $V_2$ , then it drops rapidly to  $V_{min}$ . The sloping line where all curves merge is thus  $V_{min}$ .

Instead of an infinite slope in the transition region, which would correspond to infinite gain, which corresponds to output transistors that have zero drain conductance, the actual circuit has a finite slope as a result of its real transistors, which have finite drain conductance.

### Voltage Gain

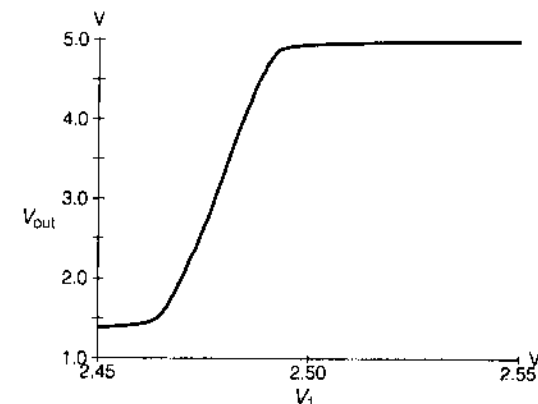
The **voltage gain**  $A$  is defined as  $\partial V_{out}/\partial V_{in}$ , where  $V_{in}$  is equal to  $V_1 - V_2$ . An enlargement of the steep part of the  $V_2$  at 2.5 volt curve in Figure 5.8 is



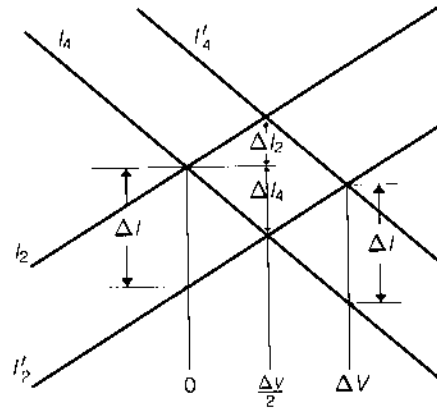
**FIGURE 5.8** Open-circuit output voltage of the simple transconductance amplifier as a function of  $V_1$ , for several values of  $V_2$ . The steep part of the curves occur at  $V_1$  approximately equal to  $V_2$ . The sloping lower limit of the output is due to the  $V_{min}$  problem.

shown in Figure 5.9. The maximum gain is approximately 143. We can easily compute what the gain of this circuit should be by considering the properties of the output transistors.

An enlargement of the intersection of the Q2 and Q4 drain curves in Figure 5.7 is shown in Figure 5.10. For a certain input-voltage difference, the curves are marked  $I_2$  and  $I_4$ . When the input-voltage difference is increased by  $\Delta v$ , both curves change:  $I_2$  decreases to  $I'_2$ , and  $I_4$  increases to  $I'_4$ . Because the bias current  $I_b$  is constant, an increase  $\Delta I$  in  $I_4$  due to a change in input voltage will result in an equal decrease  $\Delta I$  in  $I_2$ , as shown. The total change in output current per unit change in input-voltage difference was defined in Equation 5.6 as the transconductance  $G_m$  of the circuit.



**FIGURE 5.9** Expanded view of the center curve ( $V_2$  equal to 2.5 volts) of Figure 5.8.



**FIGURE 5.10** Expanded view of the intersection of the two curves of Figure 5.7 for two slightly different input voltages. The dependence of output voltage on transistor currents and drain conductances is obtained from the construction shown.

From Figure 5.6, we have seen that the output current *decreases* as the output voltage *increases*. The decrease in the output current per unit change in the output voltage is called the **output conductance** of the circuit.

$$G_{\text{out}} = -\frac{\partial I_{\text{out}}}{\partial V_{\text{out}}} = \frac{\partial I_2}{\partial V_{\text{out}}} - \frac{\partial I_A}{\partial V_{\text{out}}} \quad (5.8)$$

This quantity is the negative of the slope of the central region of Figure 5.6. The output conductance is just the sum of the contributions due to the two output transistors. In Figure 5.6, the output current changes  $3.6 \times 10^{-9}$  amp for a 1 volt change in output voltage. Thus,  $G_{\text{out}}$  is approximately  $3.6 \times 10^{-9}$  mho.

When the output is open-circuited, the total *increase*  $2\Delta I$  in output current due to an increase  $\Delta v$  in input voltage difference is compensated by an equal *decrease* in output current due to the increase  $\Delta V$  in the output voltage.

$$2\Delta I = \frac{\partial I_{\text{out}}}{\partial V_{\text{in}}} \Delta v = -\frac{\partial I_{\text{out}}}{\partial V_{\text{out}}} \Delta V \quad (5.9)$$

Substituting Equation 5.6 and Equation 5.8 into Equation 5.9, we obtain the open-circuit voltage gain  $A = \Delta V / \Delta v$ .

$$A = \frac{G_m}{G_{\text{out}}}$$

From the values of  $G_m$  and  $G_{\text{out}}$  measured on the simple transconductance circuit, the open-circuit voltage gain should be

$$\frac{5.6 \times 10^{-7}}{3.6 \times 10^{-9}} = 148$$

The result is in good agreement with the measured gain of 143.

In Appendix B, we derive the slope of the transistor drain curve in saturation. The slope of the saturated part of the drain curves of a given transistor has a slope proportional to the current level. In other words, the intercept on the  $V_{\text{ds}}$

axis occurs at a voltage  $V_0$  that is approximately independent of the absolute current level:

$$\frac{\partial I_{\text{sat}}}{\partial V_{\text{ds}}} \approx \frac{I_{\text{sat}}}{V_0}$$

$V_0$  is called the **Early voltage** for the transistor. For open-circuit operation, the steep part of the curves of Figure 5.8 will occur when the current through the two output transistors is equal, and therefore is half of the bias current  $I_b$ . Thus, *the output conductance of the amplifier is proportional to the bias current  $I_b$* . The transconductance also is proportional to the bias current (Equation 5.6).

Because both  $G_m$  and  $G_{\text{out}}$  are proportional to the bias current, the explicit dependence on current level cancels out and *the voltage gain is independent of bias current*. We can express the  $\partial I / \partial V$  terms in Equation 5.8 in terms of the  $V_0$  values ( $V_N$  for Q2 and  $V_P$  for Q4):

$$\frac{1}{A} = \left( \frac{1}{V_N} + \frac{1}{V_P} \right) \frac{2}{\kappa} \quad (5.10)$$

where the  $V$ s are expressed in  $kT/q$  units.

Equation 5.10 allows us to compute the gain of any output stage composed of complementary  $p$ - and  $n$ -channel transistors. In Appendix B, we note that the  $V_0$  of a given transistor is proportional to its length. Hence, we can make the gain arbitrarily high by committing a large silicon area to long output transistors. For typical 1988 processes, output transistors 20 microns long will give a voltage gain of approximately 2000.

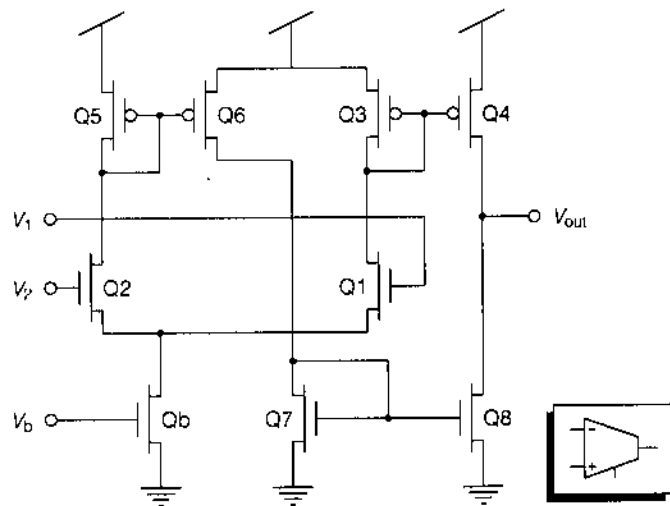
## WIDE-RANGE AMPLIFIERS

A simple transconductance amplifier will not generate output voltages below  $V_{\text{min}}$ , which, in turn, is dependent on the input voltages. This limitation often is a source of problems at the system level, because it is not always possible to restrict the range of input voltages. We can remove this restriction, however, by a simple addition to the circuit, as shown in Figure 5.11.

Instead of feeding the output directly, the drain of Q2 is connected to the current mirror formed by Q5 and Q6. The currents coming out of Q4 and Q6 now are just the two halves of the current in the differential pair. We then reflect the Q6 current one more time, through Q7 and Q8, and subtract it from  $I_4$  to form the output. As in the simple circuit, the output current is just the difference between  $I_1$  and  $I_2$ .

The major advantage of the wide-range amplifier over the simple circuit is that both input and output voltages can run almost up to  $V_{\text{DD}}$  and almost down to ground, without affecting the operation of the circuit. In other words, we have eliminated the  $V_{\text{min}}$  problem.

The other nice thing about this circuit is that the current mirrors, such as Q3 and Q5, hold the drain voltages of Q1 and Q2 very nearly constant. In



**FIGURE 5.11** Schematic of the wide-range transconductance amplifier. This circuit has many advantages over the simple transconductance amplifier of Figure 5.3. The output-voltage range is not affected by the input voltages. The symbol for the circuit is shown in the inset.

diode-connected transistors, the current increases exponentially with the gate voltage, so the drain voltages never get very far below  $V_{DD}$ . For that reason, Q2 no longer has a problem associated with its drain conductance; its source-drain voltage is nearly equal to that of Q1. So the drain conductances of Q1 and Q2 are not critical in this circuit. The same thing is true of Q6: Q7 is a diode-connected transistor; it holds the drain voltage of Q6 very nearly constant. The only transistors that work over a large voltage range are Qb, Q4, and Q8, and we can make their channels long to get a low drain conductance (output current that is *nearly independent* of output voltage). Because of their low output conductance, long Q4 and Q8 transistors give the circuit a high voltage gain.

A layout of the improved circuit is shown in Plate 7(b). This new circuit is about twice the size of the simple transconductance amplifier of Plate 7(a). Such wide-range amplifiers have about 10 times the gain of the simple amplifier, and they work all the way down to ground and all the way up to  $V_{DD}$ . When we are willing to tolerate an increase of a factor of two in area, we can build a much better amplifier.

## ABSTRACTION

The improved wide-range circuit can be used just like our original transconductance amplifier. When we design complex analog systems, we use an amplifier as an elementary component. We will not always distinguish between the simple

and wide-range transconductance amplifiers, until we work through a complete implementation. We use the symbol shown in Figure 5.3 as an *abstraction* of the detailed circuit diagram. By convention, the minus input is shown at the top and the plus input at the bottom.

If we wish to distinguish the two circuits, we use a symbol with a wider flattened end to indicate a wide-range amplifier (see Figure 5.11). When we are not sure which circuit to use, we can think about the application and work out its operating range. Most of the time, we need to deal with only the abstraction, in which the output current is a simple function of the difference between the input voltages.

The open-circuit voltage gain of either kind of transconductance amplifier is large. The voltage gain of a simple amplifier like that shown in Plate 7(a) can be 100 to 300; that of a wide-range amplifier like that shown in Plate 7(b) is 1000 to 2000. For this reason, we will often use these amplifiers as “operational amplifiers,” as the term is used in classical linear-circuit design. The gain of a classical operational amplifier usually is considerably larger than that we can achieve with the designs described in this chapter. We have chosen to use a symbol for the transconductance amplifier that is similar to that commonly used for an operational amplifier, with the addition of the transconductance control input. This convention is not as confusing as it might appear to be, at first sight, to people familiar with the conventional lore. All amplifiers have a well-defined limit to the current they can supply. Hence, the conventional operational amplifier has all the limitations described for the transconductance amplifier when the former is used in the open-circuit output mode. The classical operational amplifier, however, does not allow its user to control its output-current level. This additional degree of freedom provided by the transconductance amplifier is, as we will see in the following chapters, essential to the full range of techniques necessary for large-scale analog computation. The contrast between common usage and the convention adopted in this book can thus be viewed as follows: The usual treatment of an operational amplifier emphasizes the latter’s open-circuit output properties, and treats output-current limitation as a nuisance— as a deviation from ideality. Instead, we choose to view this limitation as a virtue, and to give the designer control over it. The open-circuit behavior then can be viewed as an idealization, achievable only as a limiting case. These issues will become much more clear when we consider the response of systems to time-varying inputs.

## SUMMARY

We mentioned that analog circuits can do computations that are difficult or time consuming (or both) when implemented in the conventional digital paradigm. We already have seen that the transconductance amplifier computes a  $\tanh$ , which is an extremely useful function. It has a very smooth transition from linear behavior for small inputs to behavior that is saturated and does not blow up if we push the input out of limits.



The circuit also is a multiplier, as we can see from Equation 5.5. It multiplies a current by a difference in voltages. Of course, the current  $I_b$  is the exponential of the voltage  $V_b$ . This circuit can multiply the exponential of one signal by the tanh of some other signal.

We might not, a priori, have chosen this particular function as the most desirable primitive for general analog computation. The same can be said of any synapse in the brain. The point is that we have no reason to expect our preconceptions concerning elementary functions to be particularly reliable. We can be sure, however, that we can learn to build systems out of *any* reasonable set of primitives. We thus follow the time-proven example of evolution, and use primitives that are efficient in the implementation medium. The transconductance amplifier is versatile, and makes efficient use of silicon real estate. Excellent collections of circuits and techniques useful in the design of primitives have been compiled [Gregorian et al., 1986; Vittoz et al., 1977; Vittoz, 1985a; Vittoz, 1985b].

## REFERENCES

- Gregorian, R. and Temes, G.C. *Analog MOS Integrated Circuits for Signal Processing*. New York: Wiley, 1986.
- Vittoz, E.A and Fellrath, J. CMOS analog integrated circuits based on weak inversion operation. *IEEE Journal of Solid-State Circuits*, SC-12:224, 1977.
- Vittoz, E.A. Dynamic analog techniques. In Tsividis, Y. and Antognetti, P. (eds), *Design of MOS VLSI Circuits for Telecommunications*. Englewood Cliffs, NJ: Prentice-Hall, 1985a, p. 145.
- Vittoz, E.A. Micropower techniques. In Tsividis, Y. and Antognetti, P. (eds), *Design of MOS VLSI Circuits for Telecommunications*. Englewood Cliffs, NJ: Prentice-Hall, 1985b, p. 104.

# ELEMENTARY ARITHMETIC FUNCTIONS

We have introduced the idea that computation can be done with analog circuits. A computation often is described as a **function** that maps **arguments** into a **result**. In any practical computation system, the arguments have **types**. In digital computers, those types are integer, floating point, Boolean, character, and so on.

In the analog world (physical computation), there are two signal types: *voltages* and *currents*. It would be wonderful if there were elegant ways to do every computation just the way we would like with every signal type, but there are not. The analog art form consists of figuring out how to use particular signal types that lend themselves to certain computations, and thus to end up with the computation that we want.

Introductory computer-science courses often start with the basic algorithms for certain elementary functions that you may want to compute. In this chapter, we present such an introduction for analog circuits.

## IDENTITY FUNCTIONS

In any formal system, the first operation anyone discusses is the **identity** operation. In analog systems, we need an identity function to make *copies* of signal values, so they can be used as inputs to several different computations.