CHAPTER 1

THE BASICS

Analog and digital circuits ultimately need a reference, be it voltage, current, or time. The reference establishes a stable point used by other subcircuits to generate predictable and repeatable results. This reference point should not fluctuate significantly under various operating conditions such as moving power supply voltages, temperature variations, and transient loading events. A few examples of circuit applications where references are intrinsically required are digital-to-analog converters, analog-to-digital converters, DC-DC converters, AC-DC converters, operational amplifiers, and linear regulators. These subsystems, of course, are the fundamental elements that make up cellular phones, pagers, laptops, and many other popular electronic products.

Though often neglected in discussions, references play a pivotal role in the design of integrated circuits. Their accuracy requirements, in fact, are more stringent in today's marketplace than ever before. The need for increased overall performance is the driving force behind this trend. In the future, as more complex and more compact systems emerge, this tendency toward high performance is expected only to increase. The prerequisite for high performance is partially exemplified by the growing demand for battery-operated circuits, which calls for high accuracy as well as low current overhead and low voltage operation. In addition to the increasing market demand for

precision references, however, there is still, and will be, a necessity for crude low-precision voltage references. Among other functions, these low-precision references are indispensable for properly biasing circuit blocks.

The loading characteristics imposed on the reference by other circuits in the system determine the design constraints of the reference. For instance, a time-varying current load will require the reference to respond quickly to rapid transitions. In fact, depending on the speed of the transitions, a load capacitor may also be necessary to prevent the reference voltage from drooping excessively. As a result, not only will the circuit have to react quickly but it will also have to be able to drive relatively large capacitor values. These characteristics inevitably allude to a stable circuit with high closed-loop bandwidth response. For reference circuits, usually, the load-current transitions are not extensive (i.e., peak-to-peak currents of roughly less than 100 μA). If the load demands more than 100 μA, a voltage regulator is commonly used, which, by definition, regulates the output voltage against various loading conditions, including, most importantly, load-current changes. A voltage regulator is essentially a buffered reference. It is, in other words, a voltage reference cascaded with a unity-gain amplifier capable of driving higher currents.

Designing a voltage reference merits the scrutiny of several factors, which are mostly governed by the overall system. Temperature-drift performance is one of the most important issues with which to contend. For references where accuracy is paramount, a temperature-compensated reference is normally warranted. The general design approach for such a reference is to sum predictable, well-characterized, temperature-dependent components, voltages or currents, to yield a well-adjusted temperature-compensated response. For instance, a voltage that increases with temperature is summed with another that decreases with temperature to produce a temperature-independent voltage. The summing ratio, of course, is balanced such that their collective effect is low-voltage variations across the whole operating temperature range, which may span from -40 to 125 °C (commercial range). Ultimately, for precision references, summing parabolic temperature-dependent terms, like quadratic temperature-dependent components, are used to approximately cancel the undesired second-order effects exhibited by the diode voltage. The first step in the process, though, is to identify voltages and/or currents that are well characterized and that do not vary significantly with process. Typically, a *p-n* junction diode voltage is chosen for this

purpose. The diode voltage is predictable, ±2 to 5%, and is well characterized with temperature. Metal oxide semiconductors' (MOS) threshold voltages, although theoretically still viable, are less conducive for high accuracy since the initial accuracy is not as good as that of the *p-n* junction diode, typically ±15 to 20%.

Integrated circuits, in general, are fabricated in a variety of process technologies ranging from standard bipolar and vanilla complementary metal oxide semiconductor (CMOS) processes to state-of-the-art silicon-on-insulator (SOI) and biCMOS technologies. As such, references may take one of several forms depending on the particular process for which they were designed. As the accuracy requirements of the applications increase, the complexity of the circuit also tends to increase to compensate for first-order, second-order, and even third-order parasitic effects. Similarly, references that must operate under low quiescent current flow and/or low supply voltage conditions are inherently more complex than those that need not meet such strict requirements. Whatever the level of complexity, though, all reference circuits stem from the same basic principles and components. Most process technologies, in fact, have a *p-n* junction diode available, which is the basis for most precision references. The techniques used to ultimately design the reference in different process technologies may differ slightly in practical terms but not in the conceptual sense. A bipolar reference, for instance, may use the base-emitter junction diode of an NPN transistor while a CMOS design will, more than likely, use the source-bulk junction diode of a *p*-type MOS transistor as the basic building block, which is none other than a *p-n* junction diode. Now, if another, more stable voltage or current were to become available, the techniques and the design approach would still remain essentially the same but under the guise of a different building block.

The performance of a reference is gauged by its variation and is described by its allowable operating conditions. The specifications of the reference include line regulation, temperature drift, quiescent current flow, input (power supply) voltage range, and loading conditions. Line regulation and temperature drift refer to the variations in reference voltage resulting from steady-state changes in power supply voltage and temperature. The typical metric used for variations across temperature is temperature coefficient (TC) and it is normally expressed in parts-per-million per degree Celsius (ppm/°C),

$$\mathrm{TC}_{\mathrm{ref}} = \frac{1}{\mathrm{Reference}} \cdot \frac{\partial \mathrm{Reference}}{\partial \mathrm{Temperature}}, \tag{1.1}$$

where the reference is either in volts, amps, or seconds. Overall accuracy is determined, primarily, by the inherent or initial accuracy of the reference and, secondarily, by line regulation and temperature-drift performance of the same, which is ultimately described by

$$\text{Accuracy} = \frac{\Delta\text{Reference}_{\text{IC}} + \Delta\text{Reference}_{\text{TC}} + \Delta\text{Reference}_{\text{LNR}}}{\text{Reference}}, \tag{1.2}$$

where the subscripts IC, TC, and LNR refer to initial accuracy, temperature coefficient, and line regulation performance, respectively. Accuracy is specified in parts-per-million (ppm), percent, or bits [1]. For example, a 100-ppm, 2.5 V reference varies ± 0.25 mV ($\Delta V = 2.5V^{*}100 \cdot 10^{-6}$), which is equivalent to a $\pm 1\%$ reference ($\% = 0.25$ mV $\div$ 2.5 V). Similarly, the same reference is said to have 13 bits of accuracy while varying by the same amount ($\Delta V =$ 0.25 mV $\leq$ 2.5 V $\div 2^{\text{Bits}}$ or Bits $\leq \log(2.5\text{ V} \div 0.25\text{ mV}) \div \log(2)$, where Bits is the maximum number of bits that still satisfies the relation). Load regulation is also often used to describe a reference and it refers to the effects of load on the reference, such as load current for the case of voltage references. Load regulation may be included in the metric for accuracy but it is most appropriately specified for regulator structures capable of handling a wider range of load currents, as discussed earlier.

This chapter, in particular, aside from introducing the topic, deals with the basics of designing a reference. The general topics discussed include diodes and current mirrors. They are the necessary building blocks used in the synthesis of any reference design, current or voltage reference. A design example is also presented to supplement the theory with a practical problem, within the context of a real working environment. The following chapter will then combine these circuit blocks to generate all sorts of current references. In the end, the current generators are then used to design voltage references.

1.1 THE DIODE

Figure 1.1 illustrates the current-voltage (I-V) relationship of a junction diode. The diode can be in the forward-biased, reverse-biased, or reverse-breakdown region. When the voltage across the terminals of

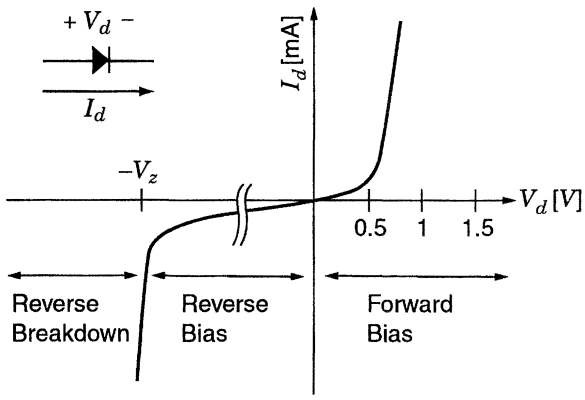

Figure 1.1 Typical *I-V* curve for a junction diode.

the diode (from anode to cathode) is between the breakdown voltage (denoted as $-V_z$) and approximately 0.5 V, the current flowing through the diode is significantly low. Once the current reaches the forward-biased region, it increases exponentially with the diode voltage,

$$I_D = I_S\left[\exp\left(\frac{V_D}{nV_T}\right) - 1\right], \tag{1.3}$$

where I_D is the current flowing through the diode, I_S is the saturation current, V_D is the voltage across the diode, n is the ideality factor (a process-dependent constant), and V_T is the thermal voltage. The ideality factor is typically around 1. The thermal voltage is directly proportional to temperature,

$$V_T = \frac{kT}{q}, \tag{1.4}$$

where k is Boltzmann's constant (1.38×10^{-23} Joules/Kelvin), T is the absolute temperature in Kelvin degrees (273.15 + °C), and q is the magnitude of the electronic charge (1.602×10^{-19} Coulomb). At a temperature of 25 °C, the thermal voltage equals roughly 25.7 mV.

The reverse-biased current is approximately constant and nearly negligible. In other words, the device exhibits a large output resistance when reverse-biased. However, significant conduction into the cathode begins to occur when the voltage across the junction approaches

the reverse-breakdown voltage $-V_z$. This conduction marks the onset of reverse breakdown. Operation in this region is not necessarily destructive but it may change some of the characteristics of the device. However, it is necessary to limit the reverse current flowing through the diode to prevent excessive power dissipation; otherwise, the device could experience irreversible damage. There are two different mechanisms responsible for diode breakdown: *avalanche multiplication* and *Zener tunneling*. Avalanche multiplication predominates at reverse-biased voltages of 7 V or greater. Zener tunneling is responsible for breakdown at reverse-biased voltages of 5 V or less. A combination of the avalanche and the Zener phenomena occurs for breakdown voltages between 5 and 7 V [2]. Irrespective of the mechanism, breakdown diodes are commonly referred as *Zener diodes*.

Zener tunneling is characterized by a negative temperature coefficient (TC). Avalanche multiplication, on the other hand, exhibits a positive TC [3]. As a result, breakdown diodes below 5 V have negative TCs, while those above 7 V have positive TCs. For example, the breakdown voltage of an emitter-base diode is typically 6 to 8 V with a TC of approximately $+2$ to 4 mV/°C. Generally, the temperature coefficient tends to become more positive (less negative) as the breakdown voltage increases, i.e., the 40 V base-collector Zener diode has a TC of roughly 35–40 mV/°C [3]. Zener diodes with breakdown voltages between 5 and 7 V, which have low TC values, are therefore appropriately used in voltage reference designs. This breakdown voltage range corresponds to the transition from Zener tunneling to avalanche multiplication.

1.1.1 Breakdown Region

The behavior of the diode in its reverse and breakdown regions is exploited in many reference applications. Since the diode is virtually an open circuit in its reverse region and a short circuit in its reverse-breakdown region, the device exhibits a voltage of $-V_z$ when current is flowing into the cathode terminal. The output impedance of the reference is simply the reciprocal of the slope of the *I-V* curve. As long as the diode remains in reverse breakdown, large changes in current produce small changes in voltage. In other words, the reference has a low output resistance. Diodes fabricated specifically for use in reverse breakdown are called *Zener diodes*. Their symbolic representation is illustrated in Figure 1.2, where V_z denotes the breakdown voltage (also called the *Zener voltage*) and current flows from the

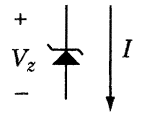

Figure 1.2 Symbolic representation of a Zener diode.

cathode to the anode terminal of the diode. Typical output resistance values range from 10 to 300 Ω. This value includes the parasitic Ohmic resistances associated with both terminals of the actual *p-n* junction, diffusion resistance. It is the voltage span of most Zener diode references; 5 to 7 V Zener diodes have low TCs, which restricts them to relatively high voltage applications, power supply voltages greater than the respective breakdown voltage of the device. The temperature dependence of the familiar Zener voltage exhibits a positive temperature coefficient ranging from approximately +1.5 to 5 mV/°C.

1.1.2 Forward-Biased Region

The current-voltage (*I-V*) relationship in the forward-biased region is well represented by equation (1.3). For most of the practical operating range (currents ranging from a few to several hundred micro-amps), the diode voltage is approximately 0.6 V. This constancy results because of the exponential nature of the diode. Consequently, a forward-biased diode is useful in the generation of a repeatable and predictable voltage by forcing current into its anode. When operated in this mode, rearranging and differentiating equation (1.3) with respect to the diode current yields its output resistance (R_{out}), which is simply the reciprocal of the slope of the *I-V* curve,

$$R_{\text{out}} = \frac{\partial V_D}{\partial I_D} = \frac{nV_T}{I_S}\exp\left(-\frac{V_D}{nV_T}\right) \approx \frac{nV_T}{I_D}. \tag{1.5}$$

It is important to note that the typical temperature dependence of this voltage is approximately −2.2 mV/°C!

The temperature dependence of a forward-biased diode is not linear, however. Its dependence on temperature is described by

$$V_D \approx V_{go} - \frac{T}{T_r}\left[V_{go} - V_D(T_r)\right] - (\eta - x)V_T \ln\left(\frac{T}{T_r}\right), \tag{1.6}$$

where V_{go} is the diode voltage at 0 °K, T is the absolute temperature in degrees Kelvin (°K), $V_D(T_r)$ is the voltage across the diode at temperature T_r, η is a temperature-independent and process-dependent constant ranging from 3.6 to 4, and x refers to the temperature dependence of the current forced through the diode ($I_D = DT^x$, where D is a temperature-independent constant and x equals 1 for a proportional-to-absolute temperature current). The form of this relationship is a variant of the one presented by [4]. A current having a positive temperature coefficient (TC) is intuitively a good choice for a diode current. The positive TC can partially compensate for the negative TC of the diode voltage. The wisdom behind this choice is also obvious from the relation above since the logarithmic coefficient decreases as the order of the positive temperature-dependent current (x) increases. The diode relation is intrinsic for understanding and designing accurate references where the second-order components contained in the logarithmic term of equation (1.6) are compensated. As a result, a full characterization of the linear and quadratic dependence of the diode voltage is sometimes useful in the design process. Obtaining the Taylor series expansion of the logarithmic component and collecting the appropriate terms aptly describe its dependence to temperature, which is

$$V_{BE} \approx \left[V_{go} + (\eta - x)V_{T_r}\right] - \left[\frac{V_{go} - V_{BE}(T_r) + (\eta - x)V_{T_r}}{T_r}\right]T$$

$$- \left\{\frac{(\eta - x)V_{T_r}}{T_r}\left[T \ln\left(\frac{T}{T_r}\right) - T + T_r\right]\right\}. \tag{1.7}$$

It is noted that the logarithm changes the coefficient of the linear component as well as that of the temperature-independent component. The derivation for equations (1.6) and (1.7) is presented in Appendix A.1 at the end of the chapter.

1.2 CURRENT MIRRORS

Current mirrors are widely used in basically all reference circuits. Consequently, a brief look at their operation and design considerations is justified. Figures 1.3 through 1.6 illustrate various bipolar and CMOS implementations of the current mirror, ranging from "simple" to regulated cascode mirrors. The basic function of a mirror is to take

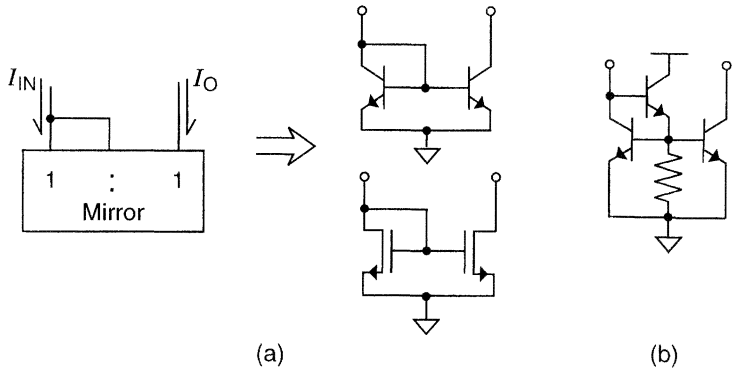

Figure 1.3 Simple current mirrors.

an input current and duplicate a multiplied ratio of the same to another terminal. Intuitively, by looking at Figure 1.3(a), the input current charges the base node of the bipolar mirror until the sum of the collector current and the base current is equal to the input current, which occurs when the circuit reaches equilibrium. Well, given that the transistor sizes are the same and the output current is mostly a function of the base voltage, as described in equation (1.3) where the diode voltage is the base-emitter voltage, the output current is equal to the input current minus the base currents. Fortunately, the base currents are much smaller than the collector current, a beta factor smaller (β), which is between 70 to 100 times smaller. Consequently, the output current is roughly equal to the input current. Had the output device been twice the size of the input transistor, the output current would have been double the input current.

Figure 1.3(b) shows a version of the same current mirror where the approximation error introduced by the base current is reduced by another β factor since another NPN transistor is added. The resistor is not always necessary but it is added to ensure that current flows through the extra NPN transistor, thereby exhibiting a more stable and predictable transient response. The CMOS version of the mirror, also shown in Figure 1.3(a), does not require this so-called β helper since its input resistance is virtually infinite; its gate current is on the order of pico-amps. Although neglected in the intuitive analysis, the output current dependence to output voltage, collector or drain voltage, is a key parameter of the mirror. Inherently, transistors exhibit an Early voltage effect, also known as channel-length modulation for MOS devices. This variation in output current, more commonly described by its output resistance, is considered to be negligible in many

integrated circuits but not so in the design of references, where small changes in current can be significant. Cascode mirrors, consequently, are often used since they produce more accurate results.

1.2.1 The Simple Mirror

The predominant source of error in the CMOS current mirror is channel-length modulation (λ),

$$I_D = I_{D\text{-sat}}(1 + \lambda V_{DS}) \tag{1.8}$$

where I_D is the drain-source current and $I_{D\text{-sat}}$ is the current once the device is in saturation. This basic MOS relation is used to derive the output current of the simple CMOS mirror (I_O), as depicted in Table 1.1. It is seen that the output current is not exactly equal to the input current; in fact, it is a function of both drain-source voltages. The bipolar circuit, as mentioned in the previous discussion, has an additional source of error: base current. Since

$$I_C = \beta I_B, \tag{1.9}$$

$$I_{\text{IN}} = I_{C1} + \frac{I_{C1} + I_{C2}}{\beta}, \tag{1.10}$$

and

$$I_C = I_S \exp\left(\frac{V_{BE}}{V_T}\right)\left(1 + \frac{V_{CE}}{V_A}\right), \tag{1.11}$$

where I_B is base current and V_A is Early voltage, the relationship of the output current to β and Early voltage effects is simply determined

TABLE 1.1. Output Current-Mirror Comparison Between Simple and Cascode Mirrors

	Simple Mirror—I_O	Cascode Mirror—I_O
CMOS	$I_{\text{IN}}\left(\frac{1 + \lambda V_{GS}}{1 + \lambda V_{DS}}\right)$	I_{IN}
bipolar	$I_{\text{IN}} \div \left[\frac{1}{\beta} + \left(\frac{1 + \lambda V_{BE}}{1 + \lambda V_{CE}}\right)\left(1 + \frac{1}{\beta}\right)\right]$	$I_{\text{IN}} \div \left[\frac{2}{\beta} + 1\right]$

by combining the above relations and collecting terms. The resulting equation is shown in Table 1.1. Connecting another NPN transistor, as shown in Figure 1.3(b), from the collector to the base of the input transistor minimizes the aforementioned β error at the cost, of course, of headroom limitations, two base-emitter voltages instead of just one. Overall, the output resistance of the simple mirror is comparable in both CMOS and bipolar technologies. The bandwidth is greater, the noise is lower, and the matching capability is better for the bipolar version, though. However, these benefits come at the expense of more current error resulting from β effects.

1.2.2 Cascode Mirrors

The evolution from the simple mirror to the cascode counterpart is readily apparent, shown in Figure 1.4. A couple of cascode devices are added to ensure that the drain-source or collector-emitter voltages of both input and output transistors are the same, thereby practically eliminating channel-length modulation and Early voltage effects. The voltage at the base or gate of these devices is simply a low-precision steady-state bias voltage, which may be generated by running some current through a resistor and/or diode-connected transistor. The resulting output resistance is therefore increased. In particular, it is increased by the product of the transconductance and the output resistance of the cascading device for the CMOS case and by a β factor for the bipolar circuit. Consequently, the much-improved input-to-output current relations shown in Table 1.1 result where the output current is essentially independent of drain-source and collector-emitter voltages.

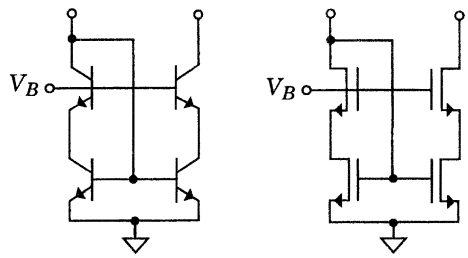

Figure 1.4 Cascode current mirrors.

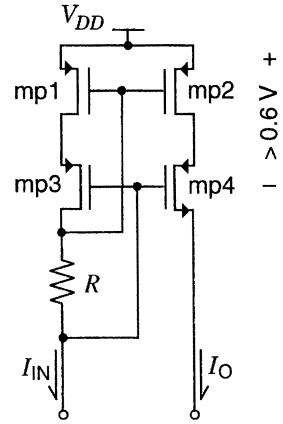

Figure 1.5 *P*-type cascode current-mirror example.

Design Example 1.1: Design a low-voltage, one-to-one, cascode current mirror (input current equals 10 μA, DC) whose output voltage is greater than 0.6 V below the positive supply voltage. The input and the output currents flow away from the mirror. Assume that a biCMOS process, with a minimum channel length for MOS devices of 1 μm, is to be used.

Since MOS devices are available and accuracy is best achieved with these transistors, *p*-type MOS transistors are chosen for the task. Figure 1.5 illustrates the circuit. Transistors mp1 and mp2 constitute the basic mirror while transistors mp3 and mp4 are the cascoding devices. Resistor R is simply used to generate a bias voltage for the gates of mp3 and mp4. Since the output voltage can swing up to within 0.6 V of positive supply V_{DD}, the sum of the saturation voltages across mp2 and mp4 must be less than 0.6 V. Since the mirror ratio is one-to-one, mp1 and mp2 must also be equal in size. For best results, mp3 and mp4 should also be the same size to ensure equal voltages at the drains of mp1 and mp2. The low overhead voltage of 0.6 V is arbitrarily chosen to be distributed equally between mp2 and mp4. Thus, choosing a saturation voltage (V_{sat}) of 0.25 V for all devices yields

$$V_{sat1} = V_{sat2} = V_{sat3} = V_{sat4} = \sqrt{\frac{2I_D}{K'(W/L)}}$$

$$= \sqrt{\frac{2I_{IN}}{K'(W/L)}} < 0.25\ V,$$

where K' is assumed to be 15 $\mu A/V^2$. The aspect ratio of all four devices is therefore chosen to be 25 $\mu m/\mu m$,

$$(W/L) > \frac{2I_{IN}}{K'(0.25^2)} = \frac{2(10\mu)}{(15\mu)(0.25^2)} = 21.3\ \mu m/\mu m.$$

Critical device mp1 and device mp2 must match well for good overall performance of the mirror; hence, their channel length is chosen to be six times larger than the minimum allowed by the process to minimize channel-length modulation (i.e., 6μm). The channel-length modulation effects of the cascoding devices on the output current are not as significant. As a result, their channel length is chosen to minimize area (i.e., 2 μm). Consequently, the dimensions (W/L) chosen for mp1 through mp4 are 150/6 $\mu m/\mu m$, 150/6 $\mu m/\mu m$, 50/2 $\mu m/\mu m$, and 50/2 $\mu m/\mu m$, respectively.

Finally, to ensure proper operation when the output voltage is within 0.6 V of the positive supply, the source-drain voltage of mp1 and mp2 is designed to be roughly 0.3 V, which would allow the output voltage to swing within 0.55 V of the positive supply (0.3 V + V_{sat}). The following loop equation is consequently used to ascertain the value of the resistor:

$$V_{SG1} + V_R - V_{SG3} - V_{SD1} = 0 = V_{SG1} + V_R - V_{SG3} - 0.3\ V.$$

Since V_{SG1} is equal to V_{SG3} (same aspect ratio and equal current densities), they cancel and the voltage across the resistor is simply designed to be 0.3 V, which means R is 30 kΩ,

$$R = \frac{0.3\ V}{I_{IN}} = \frac{0.3\ V}{10\mu} = 30\ k\Omega.$$

1.2.3 Regulated Cascode Mirrors

A current mirror is actually a single-stage amplifier with negative feedback. The input NPN transistor in Figure 1.3(a) is a common-emitter amplifier whose output, the collector terminal, is connected to its input, the base terminal, for negative feedback. The output current is therefore regulated against the input current. Similarly, a cascode device uses a common-gate or common-base amplifier stage to increase the output resistance of the overall circuit. However, its regulation performance may be further increased if higher open-loop gain

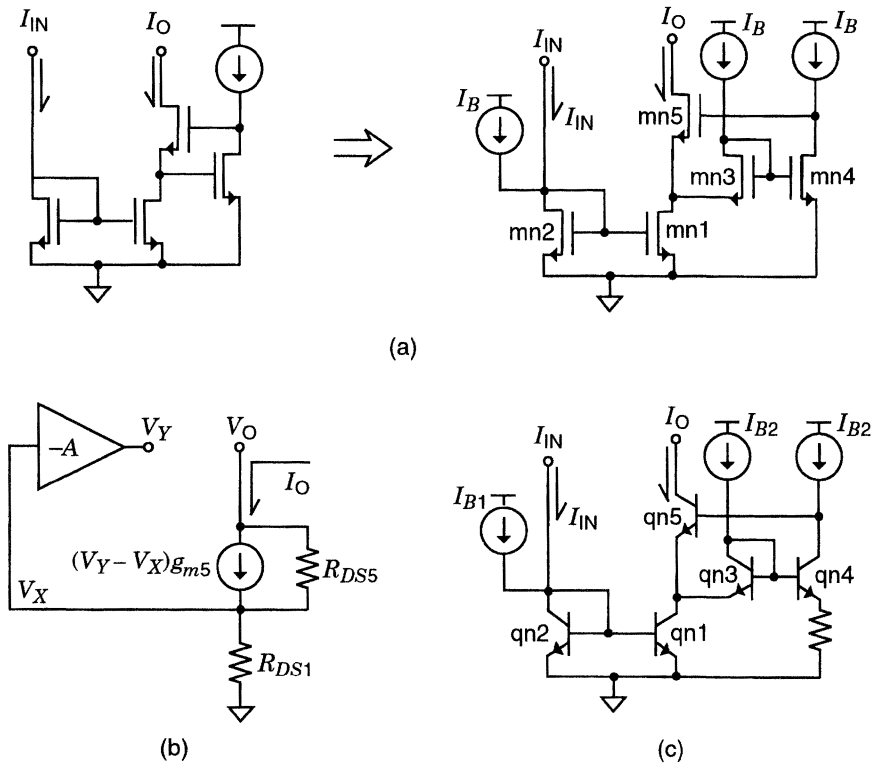

Figure 1.6 Regulated cascode mirrors.

were to be introduced in the cascode gain stage. Regulated cascode mirrors do just that. The output resistance of the current mirror is consequently boosted by a factor equal to that additional gain. Figure 1.6 shows some CMOS as well as bipolar implementations of regulated cascode mirrors [5, 6].

Figure 1.6(a), in particular, is a two-step evolution of the CMOS-regulated cascode circuit, from a high-voltage to a low-voltage circuit. In the final version, there is local feedback in addition to level shifting to assure a low voltage across the drain-source voltage of mn1 (V_{DS1}). Low voltage is desired to extend the working range of the mirror (i.e., the output voltage range for which the current is still regulated). The additional current source that is added to device mn2 (I_B) is used to match the current flowing through mn1, which is $I_B + I_O$, to avoid output current errors. Transistors mn3 and mn4 are designed to have a gate-source voltage difference equal to V_{DS1}, which is designed to be close to the saturation voltage of mn1 to maximize output voltage range.

Figure 1.6(b) shows the equivalent circuit model used to derive the output resistance of the mirror at hand. It is apparent from the circuit that the following four relations apply:

$$I_O = \frac{V_X}{R_{DS1}}, \tag{1.12}$$

$$V_Y = -AV_X, \tag{1.13}$$

$$V_O - V_X = [I_O - (V_Y - V_X)g_{m5}]R_{DS5}, \tag{1.14}$$

and

$$R_O = \frac{V_O}{I_O} = \frac{I_O R_{DSI} + (V_O - V_X)}{I_O}, \tag{1.15}$$

where g_m denotes transconductance and A is the additional gain associated with the cascode circuit ($A \approx g_{m4}R_{DS4}$). Recombining the aforementioned equations and collecting terms yields

$$R_O \approx R_{DS1}(Ag_{m5}R_{DS5}) = R_{DS1}(g_{m4}R_{DS4}g_{m5}R_{DS5}). \tag{1.16}$$

By equating A to 1, the relation for the regular cascode circuit is ascertained.

In the bipolar version of the same circuit, shown in Figure 1.6(c), a resistor is used to define the voltage across mirroring device qn1. More current error has now been introduced to the circuit, though. In particular, the culprits are qn3's and qn5's base currents (β errors). The following relation highlights the nature of these errors:

$$I_O\left(1 + \frac{1}{\beta}\right) + I_{B2}\left(1 - \frac{1}{\beta}\right) = (I_{IN} + I_{B1})\left(1 - \frac{2}{\beta}\right). \tag{1.17}$$

The error, of course, is minimized by appropriately sizing I_{B1} and I_{B2}. In the end, however, the complexity and the degraded frequency response of the regulated cascode mirror (loop-bandwidth response) limit its use. This trait encourages the designer to use, whenever possible, the simple and cascode mirrors in most analog design applications.

1.3 SUMMARY

Junction diodes are key elements in the design of references. They are not necessarily a requirement but they are certainly useful when seeking high performance at a reasonable cost. This trait arises because the *p-n* junction voltage has a relatively high initial accuracy. Additionally, the electrical characteristics are repeatable, predictable, and well characterized over a wide range of currents and temperatures. Consequently, most current and voltage references ultimately use these *p-n* junction diodes as the intrinsic building blocks in their respective designs.

Generally, references are an essential part of most, if not all, electrical systems. Although highly accurate precision references are usually desired, they are not always necessary. Rudimentary references are often sufficiently accurate to satisfy the demands of many applications. As a result, the complexity of the circuits varies from one design to the next. A simple reference may be a naturally existing voltage that does not change significantly with operating conditions, like the *p-n* junction diode. Precision references, however, improve accuracy by attempting to cancel the linear and nonlinear components of a given voltage. At this point, though, the basic tools with which references are designed have been explored (i.e., diodes and current mirrors). These are used in a variety of combinations to produce temperature-dependent currents and voltages, which eventually become key elements in the design of precision reference circuits. The next chapter deals with the generation of these temperature-dependent current references.

APPENDIX A.1 TEMPERATURE DEPENDENCE OF THE DIODE VOLTAGE

The collector current of a bipolar transistor exhibits an exponential relationship to the base-emitter voltage and is nominally expressed as

$$I_C = I_S \exp\left(\frac{V_{BE}}{V_T}\right), \tag{A.1.1}$$

where I_C is the collector current, I_S is the saturation current in the forward-active region, V_{BE} is the base-emitter (diode) voltage, and V_T is the thermal voltage. The effects of Early voltage are neglected for

this derivation. Consequently, V_{BE} is derived to be

$$V_{BE} = V_T \ln\left(\frac{I_C}{I_S}\right), \tag{A.1.2}$$

The saturation current I_S is defined by the electron charge q (1.6×10^{-19} Coulomb), the intrinsic carrier concentration n_i (approximately 1.5×10^{10} cm^{-3} at 300 °K for silicon), the diffusion constant for electrons D_n, the emitter cross-sectional area A_e, the effective width of the base W_B, and the base-doping density N_A (assumed to be constant) [7],

$$I_S = \frac{qn_i^2 D_n A}{W_B N_A} = \frac{qn_i^2 \overline{D_n} A_e}{Q_B}, \tag{A.1.3}$$

where Q_B is the number of doping atoms in the base per unit area of emitter and $\overline{D_n}$ is the average effective value of the electron diffusion constant in the base. The notation is consistent with NPN transistors but the theory also applies to PNP devices. The temperature dependence of the intrinsic carrier concentration and the average electron diffusion constant can be described by [8]

$$n_i^2 = AT^3 \exp\left(\frac{-V_{go}}{V_T}\right), \tag{A.1.4}$$

and

$$\overline{D_n} = V_T \overline{\mu_n}, \tag{A.1.5}$$

where A is a temperature-independent constant, T is absolute temperature, V_{go} is the extrapolated diode voltage at 0 °K, and $\overline{\mu_n}$ is the average mobility for minority carriers in the base,

$$\overline{\mu_n} = BT^{-n}, \tag{A.1.6}$$

where B and n are temperature-independent constants. The relationship for the saturation current I_S is more explicitly expressed by

substituting equations (A.1.4) – (A.1.6) in (A.1.3), resulting in

$$I_S = \frac{q\left[AT^3\exp\left(\frac{-V_{go}}{V_T}\right)\right](V_T BT^{-n})A_e}{Q_B} \tag{A.1.7}$$

or

$$I_S = CT^{(4-n)}\exp\left(\frac{-V_{go}}{V_T}\right), \tag{A.1.8}$$

where C is a temperature-independent constant defined by all the constants in equation (A.1.7), such as q, A, B, A_e, Q_B, and k/q from the thermal voltage term ($V_T = kT/q$, where k is Boltzmann's constant: 8.62×10^{-5} eV/° K). Finally, the collector current can be assumed to have a temperature dependence whose behavior can be described by

$$I_C = DT^X, \tag{A.1.9}$$

where D is a constant and x is an arbitrary number defined by the temperature dependence of the current forced through the collector; i.e., x is 1 for a proportional-to-absolute temperature (PTAT) current ($I_C \propto T^1$). Consequently, the temperature dependence of the base-emitter voltage can be reexpressed by substituting equations (A.1.8)–(A.1.9) in (A.1.2), resulting in

$$\begin{aligned} V_{BE} &= V_T \ln\left[\frac{D}{C}T^{(x-(4-n))}\exp\left(\frac{V_{go}}{V_T}\right)\right] \\ &= V_{go} + V_T \ln\left(\frac{D}{C}\right) - [(4-n)-x]V_T \ln T. \end{aligned} \tag{A.1.10}$$

However, a more appropriate form of the base-emitter relationship, for the purpose of design, is its temperature dependence as a function of a reference temperature (T_r). This form can be derived by obtaining the relation of the base-emitter voltage (V_{BE}) at a reference temperature (T_r), solving for a constant, and substituting it back in equation

(A.1.10). The relation of V_{BE} at T_r is

$$V_{BE}(T_r) = V_{go} + V_{T_r} \ln\left(\frac{D}{C}\right) - [(4-n)-x]V_{T_r} \ln T_r, \quad \text{(A.1.11)}$$

where V_{T_r} is the thermal voltage evaluated at the reference temperature T_r. At this point, the constant is derived to be

$$\ln\left(\frac{D}{C}\right) = \frac{V_{BE}(T_r) - V_{go} + [(4-n)-x]V_{T_r} \ln T_r}{V_{T_r}}. \quad \text{(A.1.12)}$$

Now, equation (A.1.12) is substituted back in (A.1.10) to yield the well-known temperature dependence relationship of the base-emitter voltage,

$$V_{BE} = V_{go} - \frac{T}{T_r}\left[V_{go} - V_{BE}(T_r)\right] - [(4-n)-x]V_T \ln\left(\frac{T}{T_r}\right). \quad \text{(A.1.13)}$$

It is often useful to develop the Taylor series expansion of the logarithmic term and substitute it back in equation (A.1.13). The purpose for the expansion is to more accurately design the cancellation of the linear as well as the curvature-correcting component of the bandgap reference. The process-dependent constant $(4-n)$ is sometimes expressed as η with an approximate value between 3.6 and 4 [4]. The V_{BE} relationship can be rewritten as

$$V_{BE} = A + BT + Cf(T), \quad \text{(A.1.14)}$$

where A, B, and C are constants and $f(T)$ represents all the terms whose order is greater than 1 (i.e., $C_2T^2 + C_3T^3 + \cdots + C_nT^n$. Now the Taylor series expansion of the logarithmic term about the reference temperature T_r can be developed,

$$\begin{aligned} &-(\eta - x)V_{T_r}\frac{T}{T_r}\ln\left(\frac{T}{T_r}\right) \\ &\quad = \frac{-(\eta - x)V_{T_r}}{T_r}\left[a_0 + \frac{a_1(T-T_r)^1}{1!}\right. \\ &\qquad\qquad \left. + \frac{a_2(T-T_r)^2}{2!} + \cdots + \frac{a_n(T-T_r)^n}{n!}\right], \end{aligned} \quad \text{(A.1.15)}$$

where the coefficients $a_0, a_1, \ldots, a_n$ are described by

$$a_k = \left.\frac{\partial^k \left[T \ln\left(\frac{T}{T_r}\right)\right]}{\partial^k T}\right|_{T=T_r}. \tag{A.1.16}$$

Equations (A.1.13), (A.1.15), and (A.1.16) are then used to derive the coefficients of equation (A.1.14) explicitly,

$$\begin{aligned} V_{BE} \approx & \left[V_{go} - \frac{(\eta - x)V_{T_r}}{T_r}(a_0 - a_1 T_r)\right] \\ & - \left[\frac{V_{go} - V_{BE}(T_r)}{T_r} + \frac{a_1(\eta - x)V_{T_r}}{T_r}\right]T \\ & - \left\{\frac{(\eta - x)V_{T_r}}{T_r}\left[T \ln\left(\frac{T}{T_r}\right) - a_0 - a_1(T - T_r)\right]\right\} \end{aligned} \tag{A.1.17}$$

or

$$\begin{aligned} V_{BE} = & \left[V_{go} + (\eta - x)V_{T_r}\right] - \left[\frac{V_{go} - V_{BE}(T_r) + (\eta - x)V_{T_r}}{T_r}\right]T \\ & - \left\{\frac{(\eta - x)V_{T_r}}{T_r}\left[T \ln\left(\frac{T}{T_r}\right) - T + T_r\right]\right\}, \end{aligned} \tag{A.1.18}$$

which are variations of the relations offered by [8] and [9]. Given equations (A.1.14) through (A.1.18), higher-order terms like T^2, T^3, etc. can also be derived. It is noteworthy to mention that these higher-order terms will affect the lower terms since the expansion is done about temperature T_r, e.g., $(T - T_r)^2 = (T_r^2) - (2T_r)T + T^2$.

BIBLIOGRAPHY

[1] R. Kenyon, "A Quick Guide to Voltage References," *EDN*, no. 8, pp. 161–167, April 13, 2000.

[2] A.S. Sedra and K.C. Smith, *Microelectronic Circuits*. New York: Holt, Rinehart and Winston, 1987.

[3] A. Hastings, *The Art of Analog Layout*. New Jersey: Prentice-Hall, Inc., 2001.

[4] M. Gunawan et. al., "A Curvature-Corrected Low-Voltage Bandgap Reference," *IEEE Journal of Solid-State Circuits*, vol. 28, no. 6, pp. 667–670, June 1993.

[5] A.L. Coban and P.E. Allen, "A 1.75 V Rail-to-Rail CMOS Op Amp," *Proceedings IEEE International Symposium on Circuits and Systems*, vol. 5, pp. 497–500, 1994.

[6] M. Helfenstein et. al., "90 dB, 90 MHz, 30 mW OTA with the Gain-Enhancement Implemented by One- and Two-Stage Amplifiers," *Proceedings IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 1732–1735, 1995.

[7] P.R. Gray and R.G. Meyer, *Analysis and Design of Analog Integrated Circuits*. New York: Wiley, 1993.

[8] Y.P. Tsividis, "Accurate Analysis of Temperature Effects in $I_c - V_{be}$ Characteristics with Application to Bandgap Reference Sources," *IEEE Journal of Solid-State Circuits*, vol. SC-15, no. 6, pp. 1076–1084, December 1980.

[9] G.M. Meijer et. al., "A New Curvature-Corrected Bandgap Reference," *IEEE Journal of Solid-State Circuits*, vol. SC-17, no. 6, pp. 1139–1143, December 1982.