Chapter 6

Bipolar Current Mirrors

Barrie Gilbert

6.1 Introduction - The Ideal Current Mirror

Few readers of this book will need to be introduced to the concept of the current mirror; it has become a familiar icon of modern analog design. Accordingly, we will deal only briefly with the well-established foundations of the subject, already adequately presented in many excellent texts\cite{13}, and concentrate instead on developments of the basic current-mirror forms, having properties suited to special, though not uncommon, applications. Some of these developments have been included only to illustrate the wide variety of possibilities, and may not have any immediate practical value; this is true of the high-ratio forms presented in Section 6.5.2.

In principle, the basic function could have been realized before the advent of the bipolar junction transistor (BJT), but the unique properties of this device opened the door to efficient and eminently practical forms of mirrors, now to be found in a large proportion of analog ICs (and in many digital ones, too). The design approach used here is based strongly on the translinear view of the BJT. In many cases, this will mean the invocation of strict-TL forms of which the classical current mirror is the simplest possible example. Just as probable, however, will be a strong dependence on the basic translinearity of the BJT, that is, the predictable behavior of many special-purpose current mirrors will frequently hinge on the unique relationship between collector current, $I_C$, and base-emitter voltage, $V_{BE}$; they are “TN” circuits, although not always “TL” (see Chapter 2).

Current mirrors find endless uses not only in biasing applications of low to moderate accuracy, where their high output impedance makes them valuable as good approximations to ideal current sources. More complex mirrors provide special capabilities, such as high accuracy over many decades of current, exceptionally high output resistance, very low or high transfer ratios, and so on. Depending on how broad is one’s definition of a “current mirror”, the term can also embrace various types of useful nonlinear behavior and temperature shaping; we will certainly want to examine these important aspects of current-mirror design. Mirrors are also employed as broadband signal conveyors (as, for example, in current-mode amplifiers, described in other chapters of this book). We shall therefore spend some time examining the noise performance of common mirrors. A thorough discussion of dynamic behavior (AC gain and phase, and large-signal
few hundred millivolts above the common node to many volts above it; that is, the incremental output resistance, $r_o$ (more generally, impedance, $Z_o$) should be relatively high, ideally infinite.

2) Both the large-signal mirror ratio, $M = I_2/I_1$, and the small-signal ratio $G = DI_2/DI_1$, should be essentially independent of the magnitude of the currents over many decades; that is, the ideal mirror is a linear element (so $GM$). In signal-path applications, the gain magnitude and phase response should be weak functions of frequency, ideally, completely independent of it.

3) The DC voltage, $V_1$, at the input node N1 should be small (say, within a few hundred millivolts of the common node) and it, and any AC voltage generated at this node, should be essentially independent of the input current, $I_1$; that is, the input resistance, $r_i$ (more generally, impedance, $Z_i$) should be relatively low. In a formal context, it might be convenient to define an ideal mirror as one having an input impedance which approaches zero, but in practice this is rarely as important as achieving a near-zero output admittance (Item 1).

We shall see that it is possible to design mirrors in which one or more of these characteristics can be highly refined. Frequently, such optimization is at the expense of other properties. For example, in directing attention to improving output resistance, the input resistance often has to increase. The particular way in which a mirror is optimized depends, of course, on the application. In many bias applications, for example, the emphasis will be on maintaining the highest possible output impedance (both the conductance and capacitance may need to be minimized): the input characteristics are less important here. On the other hand, in signal-path applications, it will often be found that quite high output conductances can be tolerated, while the input impedance (the reactive part often containing both capacitive and inductive components) must be minimized.

6.1.1 Mirrors, Reflectors, Conveyors, Sources

Before proceeding, it will be useful to compare the general properties of the current mirror with other cells having similar properties. The term "current reflector", for example, is sometimes applied to a three terminal network (Figure 6.1b) which satisfies all of the basic criteria for a mirror, but in which the direction of the output current is reversed. However, this

---

1Lower-case variables are used here to denote small-signal parameters.

2The terms "mirror" and "reflector" invite comparison with their optical counterparts. In conventional ray tracing, the direction of the ray is shown as reversing at the reflection surface.
terminology is not standardized. As signal-path cells, the mirror is an inverting stage; while the reflector is non-inverting; a folded cascode can be viewed as a reflector. Current conveyors, dealt with at length in Chapter 3, can be viewed as a special kind of “double-mirror” capable of both sinking and sourcing currents at input and output. Figure 6.1c depicts an ideal current conveyor. Nodes N1 and N2 can now accept and deliver currents of either polarity. The voltage at N1 closely follows that at N0; little (ideally zero) current flows in N0. In practice, two extra terminals, NP and NN are required to provide biasing and a source of positive and negative current to the output. Many of the techniques for improving the performance of current mirrors to be described in this chapter can be applied to current conveyors.

Finally, Figure 6.1d shows a floating current source, which requires only two terminals. Since there is no external control of the current at the output, sources of this kind are generally only used in DC biasing applications. While it is perfectly possible to design two-terminal current-source circuits, it is customary to utilize three-terminal controlled sources because of the simplicity (particularly, as is often the case, when multiple sources are required) and flexibility they afford. Current mirrors are widely used in such applications, although where the emphasis is on current generation rather than replication even simpler circuits often suffice. Some transducers generate current-mode signals directly (such as photodiodes), but most signals, and all accurate fixed references, are in the form of voltages, requiring voltage-to-current (V/I) conversion. This is a large topic, and there have been some interesting developments, particularly in the field of wideband V/I conversion, in recent years. However, it is beyond the scope of the present discussion, which for the most part assumes that we are dealing with variables (bias levels and signals) which are already in the form of currents. Section 6.5.3 will show some simple techniques for V/I conversion.

6.2 One-transistor Mirrors

In the simplest possible scenario, a single BJT can be used as a mirror: node N1 (of Figure 6.1a) is the base, N2 the collector and N0 the emitter. Of course, the practical objection to this proposal is that the mirror ratio, M, is much higher than generally needed and poorly-controlled, being just the common-emitter current-gain, β, and not very linear. While these objections are all true, it is nevertheless useful to begin here, because we will discover that many of the properties of more familiar mirrors can be predicted from the behavior of the single transistor circuit.

Figure 6.2a shows the biasing for an NPN device. We wish to determine the input and output resistances. It will be assumed that β is independent of current over the range of interest. If β were also independent of V2, the output I2 would always be simply βI1, and the output resistance r0 would be infinite. In fact, β increases with V2 due to base-width modulation, and in the customary BJT model would be exactly doubled\(^4\) when the collector bias voltage, VCB (roughly, V2), is equal to the forward Early voltage, VAE.

It is important to note that for an ideal BJT operating at a fixed base current (I1 in this case) VBE does not vary at all with the collector voltage V2, surprising as that may seem. It follows that if we were to replace the fixed current drive I1 with a fixed voltage drive precisely equal to VBE=I1 the output would vary in exactly the same way with V2, and a constant current I2=I1 would flow in the base. To model this, the collector current can be written

\[
I_C = I_2 = (1+V_{CB}/V_{AP})I_1 V_{CB} = 0
\] (6.1)

Thus, the output resistance, near V_{CB} = 0, is unaffected by whether the base is current- or voltage-driven, and is simply

\[
r_0 = V_{AP}/I_2
\] (6.1a)

For a transistor having some fixed but general value of V_{BE}, and having any beta (even infinite), we will find that Eq. 6.1 still applies, when the

\(^3\)It is worth mentioning, however, that two-terminal current-source circuits can be devised in which the current is a function of some environmental factor, such as temperature.

\(^4\)Indeed, one can usefully define VAF as that value of VCB which exactly halves the base width, thus doubling the collector current for a given base charge.
expression

\[ I_C = I_E = (1 + \frac{V_{CB}}{V_{AP}}) I_E(0) \exp(\frac{V_{BE} - V_T}{V_T}) \]  
\[ (6.1b) \]

is more appropriate. Here, \( I_E(T) \) is the "saturation current" for the device, typically \( 10^{-10}\)A, and \( V_T = kT/q \) is the thermal voltage, approximately 26mV at \( T = 300K \).

To trace the improvement in the design of current sources as we progress, it will be useful to assign bench-mark values to some of the parameters of a hypothetical "standard BJT" and the operating conditions. Thus, let \( V_{AP} \) be 100V, \( B_0 \) (the value of \( B \) at \( V_{CB} = 0 \) and also at zero frequency) be 100, the mirror's output current, \( I_2 \), be 1mA and the temperature, \( T \), be 300K. Then, this primitive mirror has a calculated output resistance of 100k\( \Omega \). At low values of \( V_2 \), \( r_o \) may fall due to quasi-saturation effects. When \( V_2 \) is less than \( V_{BE} \) (that is, the transistor enters saturation) the output resistance falls suddenly; thus, the lower limit on \( V_2 \) is about 200mV above the common node. At high values of \( V_2 \), \( r_o \) may again diminish, due to the onset of avalanche multiplication in the collector. The absolute upper limit on \( V_2 \) is constrained by the \( B \) of the transistor. These effects and limitations are found to a greater or lesser extent, in fully-developed current mirrors.

As well as being poorly defined, the mirror ratio \( M \) is a function of \( V_2 \). This raises an aspect of current mirror behavior which may be overlooked: not only does the output node exhibit a finite resistance, but the small-signal gain, \( G \), varies with \( V_2 \), being in this case \( B_0 \) at \( V_2 = V_{BE} \) and, for example, \( B_0 \) at \( V_2 = V_{BE} + V_{AP} \). Beta is also a function of current and temperature.

The small-signal input resistance, \( r_i = \Delta I_1/\Delta V_1 \), which in this case is the incremental emitter resistance \( r_o = \Delta V_{BE}/\Delta I_2 \) multiplied by \( B_0 \), that is, \( \beta_0 \Delta V_{BE} \). Stated in terms of the input current, \( r_i = V_T/\beta I_1 \), independent of beta. For \( I_1 = 0.1 \)mA, \( r_i \) is about 2.6k\( \Omega \). The incremental input resistance is often of less significance than the absolute variation in input voltage over the current range: following classical junction behavior, this varies by roughly 18mV for a 2:1 change, or 60mV for a decade change, in \( I_2 \).

Uncontrolled mirrors of this sort are rarely used, although there are special applications in monolithic design where the beta-dependence of output current can be of value in compensating for the base current of other transistors in the circuit. An indirect example is the input bias-current cancellation scheme for an op-amp, described by Laude(2).

In some cases, an inverting transistor can be used as a single-transistor mirror (Figure 6.2b). This approaches the behavior of a more typical current mirror, having a much lower, and more predictable, value of \( M \), which is now the inverse beta, typically ranging from less than 1 to barely more than 10, depending on the fabrication process and device geometry.

Within any given process, inverse beta has a much narrower range of values, typically varying by less than \( \pm 20\% \) for all production lots under given operating conditions, and often as close as \( \pm 3\% \) for all devices within a circuit, thus allowing fairly predictable design. \( V_2 \) must not exceed the open-base emitter-collector breakdown voltage, \( B \), typically 6V. Many modern circuits operate from a single 5V supply, so this need not be a serious shortcoming. An example of the use of an inverting transistor as a near-unity-gain mirror is found (as an incidental feature of the output-stage design) in a voltage-frequency converter described by Gilbert [3].

The output resistance, \( r_o \), is now \( V_{BE}/I_2 \), where \( V_{BE} \) is the inverse Early voltage, typical values being 2 to 10V. Since this is very much lower than \( V_{BE} \), the output resistance is correspondingly lower than for normal-mode operation. However, the output capacitance of this rudimentary mirror is very low, being just \( C_{BE} \), whereas the normally-operated device has an output capacitance of \( C_{BC} + C_{CS} \). It is usually possible to realize a factor of at least ten in the reduction of the output capacitance. This may commend the use of inverting transistors as current sources in biasing applications of high-frequency circuits. For example, in biasing a differential pair, the capacitance at the common emitter node may be troublesome in creating a nuisance pole in the open-loop transfer function, sometimes called the "tail-pole" [4]. In such cases, the inexact value of the bias current may be more tolerable than the high capacitance of a standard current source. One further advantage of the inverted-mode operation is that it is an easy matter to create multiple current outputs from a single multi-emitter transistor (Figure 6.2c). This consumes very little chip area, because only a single collector region is needed and each additional emitter may require only 100 or 200 square microns including its surround. When carefully designed, the matching between these currents can be surprisingly good - typically within \( \pm 1\% \). The effective value of the mirror ratio depends on the number of emitters and the particular geometry; the interested reader should consult the classical literature on dual designs, which uses NPN transistors in the inverse mode [5, 50].

6.2.1 Single-Transistor Mirrors Using Lateral PNPs

Many junction-isolated IC processes are limited by the unavailability of a high-performance isolated-collector PNP transistor, and recourse to the lateral PNP is necessary for mirrors which source current from the positive supply. Later, we will discuss accurate mirrors using these devices, but at this point we will mention briefly some special aspects of lateral PNPs that affects their use as current sources.

The lateral PNP is often treated as a very poor relative in monolithic design texts. True, it has poor frequency response, begins to suffer from high-level injection effects at much lower currents than an NPN transistor, and exhibits other non-ideal aspects. Nevertheless, when used creatively, this device often provides very adequate performance in biasing applications, and

\(^2\)For simplicity, we will uniformly assume that \( \exp(V_{BE}/VT - 1) \) and \( \exp(V_{BE}/VT) \) are essentially equal for most practical values of \( V_{BE} \) (see Chapter 2).
can even have advantages over fully-isolated vertical PNP transistors in some cases. Its structure can be found in most standard texts on monolithic design, but certain important peculiarities of its behavior are rarely discussed. For example, it is often said that the beta of a lateral PNP is quite low. In fact, in a well-controlled fabrication process, the low-current gain can be as high as that of the NPN transistor; values of 200-300 at \( I_C = 1 \mu A \) are not uncommon. However, due to the low doping concentration in the epi-layer, which forms the base of a lateral PNP, its beta is a very strong function of current, since the injected carriers modulate the effective base doping and thereby reduce emitter efficiency. This is simply the nature of high-level injection, the effects of which begin to be felt at about 1mA for a typical minimum-geometry NPN, but at current of perhaps one-hundredth of this for a lateral PNP. It can be shown that in high-level injection its large-signal beta, \( \beta(I_C) = I_C/I_B \), is more appropriately specified, not as a constant, but as a function of the collector current and a parameter which will be called the “beta scaling current”, \( I_{BS} \), such that

\[
\beta(I_C) = I_{BS}/I_C
\]  
(6.2)

For a typical structure, \( I_{BS} \) is about 10mA, and is surprisingly stable over production lots, being traceable to the doping profile, device geometry, and other well-controlled parameters. \( I_{BS} \) is only a weak function of temperature. Thus, the beta would have a reliable value of one at \( I_C = 10mA \), ten at \( I_C = 1mA \) and so on. Collector currents of devices having their emitter-base junctions in parallel, and driven by a single source of \( I_B \), or driven independently, will exhibit good matching, even though the \( V_{BE} \) is not “classical”, increasing by much more than 60mV/decade in the high-injection region, and also having a considerable excess component due to ohmic resistance. In low current operation, the beta matching of lateral PNPs is little different than for an NPN (possibly a little better).

Many interesting consequences of operation at high-level injection follow. First, the base current is

\[
I_B = I_C^2/I_{BS}
\]  
(6.3)

that is, a parabolic function of \( I_C \). Alternatively, the collector current is seen to be a square-root function of \( I_B \), that is,

\[
I_C = \sqrt{I_B I_{BS}}.
\]  
(6.4)

At lower injection levels, the base current becomes asymptotic to its maximum (low-current) value, \( \beta_{max} \), and at very low currents it shows the customary decline again, as space-charge recombination begins to dominate the base current. We can empirically formulate the base current as

\[
I_B = I_C \sqrt{(I_C/I_{BS})^2 + (1/\beta_{max})^2}
\]  
(6.5)

Figure 6.3 Measured base current of lateral PNPs compared to the empirical fit given in Eq. 6.5.

Figure 6.3 shows the good agreement between this suggested formulation and measurements on a lateral PNP from a production monolithic process. In this case, \( I_{BS} \) was 8.1mA and \( \beta_{max} \) was 90. Similar close conformance has been found for a variety of different processes.

While it would be unwise to depend too heavily on this generally-reliable high-level behavior to implement some desired nonlinear function, it does suggest that a single lateral PNP may be valuable as a single-transistor mirror in biasing applications. An example is shown in Figure 6.4, where the “slewing” node of a high-speed op-amp must be driven by a fairly large current, say, 2mA. Using a conventional current mirror, using high-beta transistors and a mirror ratio of, say, 2, would require a driving current of 1mA. By relying on the nonlinear beta of a lateral PNP, with \( I_{BS} = 10mA \), we can calculate (using Eq. 6.4) that \( I_B \) needs to be only 0.4mA.

There are other advantages to using single lateral PNPs in this way. Obviously, the chip area is less than that required by a two-transistor mirror; note in this connection that in multiple-output mirrors the common-base node corresponds to a single isolated epi region, another factor leading to
reduced chip area. The output capacitance is only \( C_{JC} \) (the epi-substrate junction now only affects the input capacitance) and since the mirror gain is usually low, Miller multiplication of this capacitance is only moderate. A more subtle advantage is that the output resistance is higher than might be expected from low-current measurements of \( V_{AF} \). This is because of another consequence of high-level injection, namely, the progressively-increasing charge concentration in the base at higher collector currents reduces the rate at which the depletion layer moves into the base, that is, the Early voltage of a lateral PNP increases at high values of \( I_C \) as the reduction of base width with collector bias decreases. In fact, it can be shown that the output conductance, \( g_o \), is proportional to the square-root of collector current; for a typical device, its value can be pragmatically stated as \( 5\sqrt{I_C} \) microsiemens, when \( I_C \) is in milliamps. Figure 6.5 shows the close agreement with experimental results for \( g_o \) versus \( I_C \) measured on seven samples.

Where a very low output capacitance is of dominant importance, a lateral PNP can also be operated in an inverse mode, with the output being taken from the emitter - usually a small circular diffusion of 5 to 20\( \mu \)m in diameter - with the collector ring diffusion treated as the emitter common terminal. The inverse \( \beta \) is lower, partly because more than half of the current injected from this ring flows outward to the surrounding isolation wall. However, the base scaling current is higher, due to the larger area of the effective emitter, so useful amounts of output can be generated before the lowered \( \beta \) renders the device inefficient. Since the breakdown voltage of the emitter and collector are the same (that is, the \( B V_{CEO} \) of the corresponding NPN) the bias restrictions which arose with the inverted-mode operation of the NPN do not apply here. Note, however, that lateral PNP transistors can exhibit breakdown due to punch-through if the emitter-collector spacing is insufficient for the bias voltage; this occurs when the depletion layers penetrate into the base region far enough to reduce the effective base width to zero.

Current sources and mirrors based on lateral PNP's can take advantage of another possibility afforded by the structure, namely, the use of a split collector diffusion. Figure 6.6a shows a typical geometry, in which the two segments are of equal size; unequal segments can be used, and of course more
than two segments are possible. The unique nature of this device makes it valuable in numerous biasing applications. The current matching between halves of a two-collector lateral PNP is very good, routinely better than \( \pm 1\% \), and using special layout precautions, employing cross-connected pairs of split-collector transistors, sources matched to better than \( \pm 0.1\% \) (at equal \( V_{CE} \)) have been maintained in production. This is valuable when using the device to provide a pair of active loads for an amplifier input stage, for example. This matching does not depend on the \( V_{BE} \) matching of two separate transistors, as is the case when using normally-biased PNP transistors as current generators. One way of visualizing this is to imagine that the single 'dot' of the emitter is like a light bulb positioned at the centre of the collector ring - emitting holes rather than photons - which illuminates all points on the ring with equal 'flux'. Since the emitter and collector diffusions are defined on the same mask, the radial symmetry is excellent. Even more intriguing is the almost-zero sensitivity of this current match to thermal- or stress-gradientes of the chip; these can severely degrade the matching of collector currents in a pair of NPNs, via their effect on \( V_{BE} \), which results in a current ratio error of 8\% for a roughly 2mV \( \Delta V_{BE} \) induced by a 1°C temperature difference between devices.

A split-collector lateral PNP transistor (two 180° collector segments) is occasionally used to realize a near-unity-gain mirror, by connecting one of the segments back to the base (Figure 6.6(c)). This is particularly effective at low currents, where the beta is high. Obviously, this idea can be extended to provide a ratio of 2 or 0.5 (using three 120° segments), or 3 or 0.33 (using four 90° segments). It is difficult to use more segments, limited by the minimum feature dimension and required spacing between these segments. Non-integral splits can be used, with reduced ratio accuracy.

6.3 Two-transistor Mirrors

The basic NPN current mirror (Figure 6.7(a)) is the simple theme from which endless variations and transformations can be spun, many of fundamental and pervasive value, others of which are mere curiosities awaiting exploitation. Although much has been written about this circuit in the many fine standard texts, there remain certain points of view that may not have been expressed or sufficiently stressed. It happens, for example, to be the simplest translinear circuit of practical interest, and we will begin here with the translinear view of this mirror.

Recall from Chapter 2 that in any loop containing an equal number of voltages across clockwise- and counterclockwise-biased PN junctions, the product of the current-densities in the two directions must be equal. Here, the
loop contains only one junction voltage in each direction, and these are one and the same voltage in this trivial case, namely, the $V_{BB}$ of both transistors. Since the current densities are equal, the ratio of the collector currents must be identical to the emitter area ratio, $A$. In practice, this would only be exactly true if this ratio were an integer and realized by exact replication of a basic geometry, for example, by making Q2 from four basic units of Q1, and if both transistors were operated at identical values of $V_{CE}$, which demands in this case that $V_2 = V_{BB}$. Note that the finite ohmic resistances do not affect the accuracy of the currents, since they can be assumed to scale in proportion to $A$ and therefore the excess $V_{BB}$ which they contribute is identical for each device. Nor is the finite, temperature-dependent beta, including its nonlinearity with current, a source of error in the ratio of the collector currents.

Errors in the actual mirror ratio, $M$, arise because

1. $V_2$ is, in general, not equal to $V_{BB}$, and the finite output resistance will cause $I_2$ to increase with $V_2$.

2. Some of $I_1$ is lost to the bases of both Q1 and Q2, so that $I_{C1} < I_1$.

3. Unless strict integer replication of geometry is used, there will be errors in the effective value of $A$ and the ratio of junction resistances.

4. Even then, small random variations in emitter area and other sources of $V_{BB}$ mismatch (including thermal gradients and strain) will occur.

### 6.3.1 Effect of Practical Device Imperfections

Errors in delineating the emitter area ratio $A$ will result in essentially identical errors in the actual mirror ratio $M$. This is sometimes explained in terms of "$V_{BB}$ mismatch", but this expression is helpful only as a reminder that a very small voltage difference between emitter nodes can have a large effect on accuracy. Indeed, there really can be no $V_{BB}$ mismatch in this circuit, since in the current mirror both devices operate at identical $V_{BB}$, having their emitter-base junctions in parallel. However, we can translate an uncertainty in area $\Delta A$ into an equivalent $V_{BB}$ error for the case (not true in the current mirror) where the transistors are operated at identical collector currents. This is well-known to be

$$\Delta V_{BB} = V_T \ln (1 + \Delta A)$$  \hspace{1cm} (6.6)

which simplifies to

$$\Delta V_{BB} = V_T \Delta A$$  \hspace{1cm} (6.6a)

For $\Delta A \ll 1$, amounting to $\pm 260 \mu V$ at $T = 300 K$ for a $\pm 1\%$ area uncertainty. $V_{BB}$ is also a function of emitter depth and other local variations in the device; these, and any further variations in $V_{BB}$, can be converted to a modified area ratio $A'$, which may have a value greater or less than $A$, by rearranging Eq. 6.6 to first calculate $\Delta A$ then

$$A' = A \exp(\Delta V_{BB}/V_T)$$  \hspace{1cm} (6.7)

The error due to finite beta can be quickly determined by noting that

$$I_1 = I_{C1} + I_{B1} + I_{B2} = I_{C1} + I_{C2}/\beta_1 + I_{C2}/\beta_2$$  \hspace{1cm} (6.8)

and, if $V_2 = V_{BB}$ and good scaling disciplines have been observed, we can assume that $\beta_1 = \beta_2 = \beta_o$, so that

$$I_1 = I_{C1} + I_{C2}/\beta_o + AI_{C2}/\beta_o$$

while

$$I_2 = AI_{C1}$$

The actual mirror ratio, $I_2/I_1$ is therefore

$$M = \frac{A}{1 + (1 + A)/\beta_o}$$  \hspace{1cm} (6.9)

For a classical mirror having $A = 1$ and our "standard" condition of $\beta_o = 100$, the error in $M$ is $-2\%$. Note that for very large values of $A$ the above expression correctly shows the current ratio is asymptotic to $\beta_o$.

Finite Early voltage causes the output current to increase with the collector bias as given in Eq. 6.1. Combined with the loss due to finite beta, the complete expression is

$$I_2 = \frac{A(1 + V_2/V_{AR})}{1 + (1 + A)/\beta_o} I_1$$  \hspace{1cm} (6.10)

Note that this expression includes the voltage-dependence of beta, since $I_2$ increases at the same rate as $\beta_2$, so the base current of Q2 is constant with $V_2$. The output resistance remains $V_{AR}/I_2$, independent of $\beta$ or $A$. The small-signal input resistance $r_1$ is now essentially just the $r_o$ of Q1, that is, $V_1/I_1$, and is also independent of $\beta$, $A$, and $V_{AR}$. Note, however, that this is exactly the same as for the rudimentary one-transistor mirror.

### 6.3.2 Use of Emitter Resistors to Improve Accuracy

We will return to further developments of the translinear view of the current mirror later, but first examine the effect of adding resistors in the emitter branches (Figure 6.7b). This is usually referred to as emitter degeneration,
because the naturally-high gm of the devices is lowered in this way.

To understand how this works, first note that the $V_{BE}$ of $Q_2$ is a function of both its collector current $I_C$ and collector voltage $V_{CB}$. Starting with Eq. 6.1b

$$V_{BE} = V_{BE0} - V_T \ln (1 + V_{CB}/V_{AF})$$

(6.1c)

Here, $V_{BE0}$ the value of $V_{BE}$ for some $I_C$ and temperature, and at $V_{CB} = 0$, given by

$$V_{BE0} = V_T \ln (I_C/2I(T))$$

(6.1d)

So, if the mirror is to have a high output resistance, it must somehow be desensitized to the variation in the $V_{BE}$ of $Q2$ caused by output voltage variations. This is the function of emitter degeneration.

With reference to Figure 6.7b, it can be seen that the larger the "degeneration voltage", $V_{BE}$, that is allowed to develop across resistors $R_1$ and $R_2$, the less any variation of $V_{BE}$ affects the output current $I_2$. When $V_{BE}$ is very large, $Q2$ might be considered to be operating essentially as a common-base stage, current-driven in its emitter, but there is a subtle and important difference. We can readily calculate the output resistance for this limiting case by considering just the variation of $\beta_2$ with $V_{BE}$. For now, assume that the resistor ratio $R_2/R_1$ is made equal to the emitter area ratio $A$, as indicated in the figure. This would normally be the case; the consequences of altering it are discussed later. Then, we can immediately write

$$V_{BE} = I_2(1+\beta_2)R = (I_1/L_2/\beta_2)AR$$

from which it follows that

$$M = \frac{I_2}{I_1} = \frac{\beta_2}{1 + (1 + A)/\beta_2}$$

(6.11)

Now, if $\beta_2$ were fixed, this is essentially the same result as given already in Eq. 6.9, but in fact

$$\beta_2 = \beta_0(1+V_{CB}/V_{AF})$$

(6.12)

Inserting this into Eq. 6.11 and finding the derivative $dM/dV_{CB}$ yields the output conductance, from which the general output resistance can be shown to be

$$r_o' = \frac{V_{AF}}{A(1+V_{CB}/V_{AF})^2 + 1 + A}$$

(6.13a)

This is not very insightful, so let us consider some special cases. First, for the classical mirror with $A = 1$ and $\beta_0 >> 1$, operating at with $V_2$ much less than

Figure 6.8 Output resistance of a highly-degenerated mirror operating at $I_C = 1mA$ varies from about 5M$\Omega$ at $V_{CB} = 0$ to 20M$\Omega$ at $V_{CB} = V_{AF}$.

$V_{AF}$, Eq. 6.13a simplifies to approximately

$$r_o' = \frac{V_{AF}\beta_0}{2I_2}$$

(6.13b)

with $I_2$ replaced by $I_2$, appropriate for the conditions stated.

For example, using our benchmark device having $\beta_0 = 100$ and $V_{AF} = 100\Omega$, $r_o'$ at 1mA is 5M$\Omega$, 50 times the value for the basic mirror. The reason it is not beta times as high, as might be expected, is that a reduction in the base current of Q2 due to an increase in $V_2$ simultaneously reduces two sources of error: the first is in the loss of current conveyed from the emitter branch of Q2 to its collector (that is, alpha increases with $V_2$); the second is in the reduction of the base current lost from the input current $I_1$. If the base node were grounded and the emitter of Q2 driven by an ideal current source of value $I_2$, then $r_o'$ would be simply $V_{AF}\beta_0/I_2$, the value sometimes given for the asymptotic output resistance of the highly-degenerated mirror. Errors in analysis can usually be traced to the use of small-signal modelling procedures.
Figure 6.9 Improvement in output resistance versus degeneration voltage $V_E$, normalized to case where $V_E = 0$, for $\beta = 50, 100$ and 200.

which do not accurately reflect alteration in the operating point of all devices.

Eq. 6.13a shows that the output resistance is also a function of the collector bias voltage, which is not often recognized; for the case where $A = 1$ and $\beta_0 >> 1$, it simplifies to

$$r'_o = \frac{V_{AF} \beta_0}{2I_2} (1+V_{CB}/V_{AF})^2$$  \hspace{1cm} (6.13c)

Thus, at $V_{CB} = V_{AF}$, we would expect $r'_o$ to be four times higher than at $V_{CB} = 0$, and that a curve of $I_2$ versus $V_2$ should be nonlinear. This is exactly what happens for our semi-ideal transistor, having $\beta = 100$, $V_{AF} = 100$, operating at $I_2 = 1$mA and using a very large value for $R$. Figure 6.8 shows a simulation result; the incremental $r'_o$ which varies from about 5MΩ to 20MΩ between $V_{CB} = 0$ and $V_{CB} = V_{AF}$, as predicted by Eq. 6.13c. Nonlinearity in $r'_o$ is rarely of concern; at high mirror ratios, and for lower values of $V_B$, the nonlinearity is reduced.

When the emitter resistors are of more practical value, the output resistance does not reach these asymptotic values. Figure 6.9 plots the ratio $r'_o/r_o$, that is, the improvement in output resistance compared to the basic mirror (Figure 6.7a), versus degeneration voltages of zero to 10V, with $\beta_0$ set to 50, 100 and 200; note that the asymptotic value at $\beta_0 = 100$ is 50, as previously found. These curves were calculated for $V_{AF} = 100V$.

It is sometimes desirable to achieve a mirror ratio other than unity entirely by the use of unequal emitter resistors, that is, keeping $A = 1$ but making $R_2 \neq R_1$. This situation could arise, for example, when a current gain of, say, five is needed, while the collector capacitance of the output device has to be minimal. The use of equal transistor sizes operating at unequal currents will inevitably result in a $\Delta V_{BE}$ between the emitter voltages. In the special case where $I_1$ is PTAT, this will track the PTAT voltages across the resistors, and the ratio will remain stable over temperature, for all combinations of emitter areas and resistor ratios. More commonly, $I_1$ will be temperature-stable and this $\Delta V_{BE}$ will now add (when $M < 1$) or subtract (when $M > 1$) from the stable voltage $V_B = I_1 R_1$ across emitter resistor $R_1$ (Figure 6.7b), resulting in a temperature dependence in the voltage across $R_2$, and hence in the actual mirror ratio. If $M$ is not greatly in error (because it is close to one, or because $V_B$ is large, or because the resistor ratio has been adjusted to first-order correct the error), it can be shown that the residual drift has a fractional magnitude $(V_{AF}/V_B \ln M)/300$ at $T = 300K$. For example, if the adjusted $M$ were 5 and a $V_B$ of 500mV were used, the drift would be $(26mV/500mV) \ln 5)/300 = 0.0028$ or 280 ppm/°C.

The input resistance of a mirror using emitter degeneration is approximately $V_T/I_1 + R_1$. This may be troublesome; if nothing else, the voltage at the common base node $N_1$ erodes the output range, which now cannot go below $I_1 R_1 + V_{BE}$. In signal-path applications, the pole formed by $R_1$ and the $C_{22}$ of Q1 plus other parasitic capacitive components, will affect the HF response. Note that emitter resistors do not reduce the output deficit caused by finite beta. Beta compensation can be included by adding a resistor $R_B$ of the correct value in the base of Q1, as shown in Figure 6.10. It can be shown that for $V_{CB}$ close to zero (that is, neglecting base-width modulation errors) the necessary value is

$$R_B = \frac{\beta + 1}{\beta - A} \frac{(1 + A) (r_o + R_1)}{}$$  \hspace{1cm} (6.14)

For example, using $A = 2$, $I_1 = 1mA$ (so $r_o = 26Ω$) and $R_1 = 500Ω$ (so introducing about 500mV of degeneration) $R_B$ should be 1.57kΩ if the factor involving beta is ignored, or 1.62kΩ if it is included and beta is assumed to be 100. Clearly, the method becomes unpredictable when $A$ is comparable with $\beta$, and can never be precise since $\beta_1$ and $\beta_2$ are not in general equal; nevertheless, the technique is of considerable practical utility.
6.3.3 Noise and Drift Considerations

Current mirrors generate noise. This arises from two sources: shot noise in the junctions and Johnson noise in the resistances (extrinsic and diffusion). At low frequencies, flicker noise will also be present.

Consider first the ideal non-degenerated mirror, that is, one using semi ideal transistors (having high beta and Early voltage, and no ohmic resistances), with a general emitter ratio $A$. It is easy to show that the flatband noise spectral density (NSD) in the output due to shot noise is

$$s_t = \sqrt{2qI_1(A+1)}$$

(6.15a)

where $q$ is the electronic charge, $1.6 \times 10^{-19}$ C. This can be pragmatically stated as

$$s_t = \sqrt{I_1 \cdot (A(A+1)) \cdot 17.9 \text{pA/Hz}}$$

(6.15b)

when $I_1$ is expressed in mA. Thus, an ideal unity-gain current mirror operating at 1mA generates a NSD of 25.3pA/Hz. As a practical aside, it is important in verifying these expressions by simulation experiments using SPICE to note that collector currents must first be converted to voltages in order to use the NOISE analysis mode, and that this should be performed using the "H" element (current-controlled voltage-source, or ideal transresistance element) controlled by the bias voltage source for the collector, rather than by the use of a load resistor, which will also be modeled as having Johnson noise. Even a 1Ω current-monitoring resistor will generate about 128pV/Hz, far more than the voltage generated in this same resistor by the 25pA/Hz of the mirror in our example.

The inclusion of Johnson noise arising in the base resistance, $r_{bb}$, in Q1 and Q2 is most simply carried out by invoking superposition, that is, by first calculating its contribution to the noise spectral density and later calculating the total noise. Base resistance noise can be shown to be

$$s_f = 2qI_1 \sqrt{(A(A+1)r_{bb}/kT)}$$

(6.16a)

where $r_{bb}$ is the base resistance of Q1, which evaluates to

$$s_f = I_1 \sqrt{(A(A+1)r_{bb}) \times 4.976 \text{pA/Hz}}$$

(6.16b)

when $I_1$ is expressed in mA and $r_{bb}$ is in ohms. Thus, the contribution to the total NSD due to 100Ω of $r_{bb}$ in the otherwise-ideal unity-gain mirror at 1mA is about 70.4pA/Hz, considerably more than the shot noise under these conditions. These expressions assume that $r_{bb}$ scales with emitter area. Emitter contact resistance $r_{ec}$ does not increase the overall noise in a properly-scaled mirror, because it reduces the transconductance of Q2 faster.
than it increases the noise contribution of \( Q1 \).

The total NSD of the uncorrelated shot- and Johnson-noise sources sum to

\[
s_T = \sqrt{(s_q^2 + s_j^2)}
\]

(6.17)

Thus, the total NSD at the output of \( Q2 \) calculates to 74.8pA/√Hz for our example. To provide a better feel for the likely impact of this noise level, we can express 75pA/√Hz on a bias of 1mA as a dynamic range of 82.5dB in a 1MHz bandwidth. Emitter resisters reduce the effect of both the shot noise and Johnson noise. The full analysis is beyond the scope of this Chapter, but Figure 6.11 shows the reduction possible as the equal emitter resistors \( R \) are increased from 1Ω to 10kΩ in the "standard mirror" (\( A = 1, I_1 = 1mA \)) with an \( r_{ep} \) of zero and 100Ω.

Emitter resistors are also used to reduce errors due to strain- and thermally-induced shifts in \( V_{BE} \). Such desensitizing is important in maintaining accurate balance in many monolithic circuits, where a power-dissipating transistor may be located close to a critical mirror\(^6\), and the resulting thermal gradient induces a \( \Delta V_{BE} \), or where inadequate die-attach uniformity leads to chip stresses. The improvement arises for the same reasons that degeneration reduces random noise, and can be readily calculated; it is roughly proportional to \((R+r_p)/r_p \) or \((V_T+r_p)/V_T \), where \( V_T \) is the voltage across the emitter resistors. Thus, by including just 240mV (about 9\( V_T \)) of degeneration the sensitivity is reduced by an order of magnitude.

Degeneration can also reduce the absolute error in the mirror ratio, by transferring the matching requirement to resistors, rather than emitter areas. Well-designed resistors can match to ±0.1% or better, while the emitter-area match of small, single NPNs can be as poor as ±1% (corresponding to a \( \Delta V_{BE} \) of 260μV at 300K). However, there is an interesting trap, here, which is not generally recognized. Imagine a unity-gain current mirror with nominally-identical emitter areas, but mismatched by \( \Delta A \). In the absence of emitter resistors, there would be an essentially identical error, \( \Delta M \), in the mirror ratio, but this would remain fixed over temperature; that is, there would be no drift in \( M \). Likewise, if we inserted emitter resistors of sufficiently large value, the error due to \( \Delta A \) would be swamped, and any error in the mirror ratio would be caused solely by \( AR \). But for moderate values of \( R \) the \( \Delta V_{BE} \) arising across the emitter nodes due to the \( \Delta A \) will induce a drift in the output current, and at some critical value of \( R \) this drift will be maximized. It's easy to show that this will occur when the voltage across the resistors is \( kT/q \), at which point the drift rate for small values of \( \Delta A \) will be given approximately by

\[
\Delta I_c/\Delta T = \Delta A \times 750\text{ppm/°C}
\]

(6.18)

\(^6\)As well as fixed VBE errors due to a stationary thermal gradient, distortion often arises in a poorly designed IC when the gradient fluctuates with load variations.

For example, when \( \Delta A=0.1 \) (a rather severe 2.5mV \( V_{BE} \) mismatch at \( T=300K \)) the peak drift in \( I_c \) is 75ppm/°C; over the full military temperature range this results in a change of 1.4% in \( I_c \).

### 6.3.4 Nonlinear Two-Transistor Mirrors

Numerous ways exist to alter the two-transistor mirror to effect various types of useful nonlinear behavior. Such circuits can be used to modify the shape of the transfer characteristic and the shape of the output over temperature. One of the earliest examples of this is the Widlar mirror \(^{6,6a}\) in which a resistor is added only to the emitter of \( Q2 \). This was intended to be useful in producing the small bias currents often needed in monolithic circuits without requiring the use of a large resistor \( R_p \) in the supply branch. Figure 6.12a shows the circuit, with \( Q1 \) shown as optionally having a larger area than \( Q2 \), since it is expected to be used in a current-lowering mode, aided by \( R_2 \). The governing equation is

\[
I_2R_2 = V_{BB1} - V_{BE2} = V_T \ln(I_1/AI_2)
\]

(6.19)

the solution of which is transcendental for \( I_2 \) but which can easily be solved
for \( R_2 \) given the values of \( I_2 \) and \( I_1 \), when

\[
R_2 = \frac{(V_T/\beta)}{ln(\beta/\alpha)}
\]  
(6.20a)

or, for the needed input \( I_1 \) given the values of \( I_2 \) and \( R_2 \), when

\[
I_1 = \frac{A_I}{\beta} \exp\left(\frac{I_2 R_2}{V_T}\right)
\]  
(6.20b)

It is sometimes said of this mirror that \( I_2 \) is proportional to absolute temperature (PTAT) but clearly this is only approximately true under certain conditions, when the ratio \( I_1/\alpha \) is very large. Figure 6.13 shows how \( I_2 \) varies with \( I_1 \) for various values of \( R_2 \), using the general operating conditions maintained throughout the earlier examples; in this case \( A = 1 \).

Note that the ratio \( I_2/\beta \) is invariant if \( I_1 \) is PTAT; this "persistence of PTAT" is a common aspect of the current-mode transfer function of mixed junction/resistor circuits.

By altering the topology, placing \( R_1 \) in the collector of \( Q_1 \), as shown in Figure 6.12b, a dramatic change in behavior occurs. This is sometimes called the "gm compensated" mirror, because when \( R_1 \) is made equal to \( 1/\beta \), that is, the \( r_\beta \) of \( Q_1 \) operating at some specified \( I_1 \), the increase in \( V_{BE1} \) caused by an increase in \( I_1 \) is exactly compensated by the increase in the voltage drop across \( R_1 \). Consequently, the voltage delivered to the base-emitter junction of \( Q_2 \) is rendered insensitive to \( I_1 \) and therefore so is the output \( I_2 \).

In fact, the behavior is much more interesting and general than this, and the circuit has many uses. In this case, an analytic expression for \( I_2 \) can be found:

\[
V_{BE1} - \beta R_1 = V_{BE2}
\]

thus

\[
I_2 = \frac{A_I}{\beta} \exp\left(-\frac{I_1 R_1}{V_T}\right)
\]  
(6.21)

It is readily shown that \( I_2 \) peaks when \( I_1 R_1 = V_T \) which is the condition previously noted as resulting in zero sensitivity to \( I_1 \), at which point the...
Figure 6.15 A fusion of the Widlar and gm-compensated mirrors provides relatively good regulation of a variable input current.

The exponential factor has a value of $e^{-t}$ or 0.368. Note that the area factor, $A$, has been placed in the emitter of $Q_2$, so that we can optionally restore the mirror ratio, $M$, to unity by making $A = e = 2.72$ (roughly $11/4$ in integer terms). Of course, in these recent analyses, we have neglected to account for the effects of beta and Early voltage; their inclusion is sometimes tractable, sometimes not, and always obscures the basic concepts. If high accuracy is essential, additional circuit elements invariably are required.

Figure 6.16 shows $I_2$ versus $I_1$ from 0 to 5mA, using $R_1 = 26\Omega$ and our "standard" transistors, with $A = 1$. It will be apparent that in addition to its utility in reducing the effect of input current variations when operated near the peak point, the circuit is useful in generating small currents, operating with voltage drops of much more than $V_T$ across $R_1$. In this way, much smaller resistor values are needed to produce a given output current than for the circuit of Figure 6.12a. For example, suppose we are using a Widlar mirror as part of a bias scheme to operate from a 5V supply and generate $I_2 = 1\mu A$. In order to keep the total resistance within reasonable bounds, we might set $I_1$ at 100\mu A, requiring an $R_o$ of about 43k\Omega, and choose to use two minimum-geometry transistors, so $A = 1$. Then, $R_1$ must be 59.5k\Omega (from Eq. 6.20b). On the other hand, using the circuit of Figure 6.12b we find $R_1$ can be 100 times smaller - 595\Omega - since it has 100\mu A, rather than 1\mu A, flowing in it.

Now, in the Widlar mirror $I_2$ increases slightly with the supply voltage, while using the "gm-compensated" form it decreases. It is therefore likely that a lower sensitivity to the supply can be obtained by combining the two

Figure 6.16 Output of the circuit of Figure 6.15; see text for details.

circuits, as shown in Figure 6.15. Under certain conditions, this can prove useful. Figure 6.16 shows the result for one such optimization, in which $I_2$ is designed to be 5\mu A for a 5V supply; here, $A = 1$, $R_0 = 18k\Omega$, $R_1 = 100\Omega$ and $R_2 = 15k\Omega$; the output is within 4\% of its nominal value for supply voltages from 3V to 10V.

6.4 Three-transistor Mirrors

As is apparent, several variations can be found on the two-transistor theme: it is not surprising, therefore, to find that the addition of a third transistor considerably broadens the palette of the designer. Only a few of the numerous possibilities can be included in this brief survey. Most will have as their aim the improvement of accuracy, which most often means the reduction of beta-related errors.

6.4.1 The EF-Augmented Mirror

Figure 6.17a shows a commonly-used current-mirror, in which $Q_3$ operates as an emitter-follower (EF) and provides the base currents of $Q_1$ and $Q_2$.
augmenting their current gain. This mirror seems to have escaped being named, but, because we will refer to it several times in the rest of this Chapter, it will here be called the EF-augmented, or EFA, mirror. The use of emitter degeneration resistors and a non-unity area ratio is optional. It is readily shown that, assuming $R_i = AR_3$ the mirror ratio is now

$$M = \frac{A}{1 + \left( \frac{1}{\beta_1} + \frac{A}{\beta_2} \right) \frac{1}{\beta_3}} \quad (6.22)$$

so that even for quite large values of $A$, the beta errors are usually negligible, certainly now much less than other practical errors. For $A = 1$ and $\beta >> 1$, this simplifies to

$$M = 1 - \frac{2}{\beta_1 \beta_3} \quad (6.22a)$$

where it has been assumed that the betas of Q1 and Q2 are equal. When this is not true (for example, when the output collector is very positively biased) the more complete expression must be used.

The output resistance, $r_0$, using no emitter degeneration is the same as for the basic mirror - $V_{AF}$ divided by the output current - but the maximum $r_0$ is now $V_{AF} \beta_2 / I_2$, twice as high, for large amounts of degeneration.

The high internal current gain introduces the possibility of HF instability, since there are two "beta poles" in the basic circuit: from the base of Q3 to the base to Q1 is the beta pole of Q3; from here back to the base of Q3 is the beta pole of Q1. In fact, a third pole may be generated by the emitter resistance of Q3 and the total capacitance on the base node, which includes the $C_{e}$ of Q1 and Q2 (or more devices); this may become complex since the driving-point impedance is somewhat reactive. A complete discussion of the general HF response will not be included here. Suffice it to say that this system is often stable only because the much lower current in Q3 lowers its effective $r_t$ to the point where it creates a dominant pole. Also, there is a small amount of feedforward around this slow device via the $C_{e}$ of Q1. By adding a further small capacitance across the base-emitter nodes of Q3 its beta is effectively removed from the HF loop. This is shown in Figure 6.17b, which also shows a more general scheme, including emitter degeneration and multiple outputs, now possible because of the higher effective beta.

### 6.4.2 Use of Cascodes to Raise Output Resistance

Some applications of current mirrors and sources require the highest possible output resistance. In these cases it is sometimes possible to place less emphasis on accuracy of the mirror ratio. A cascode transistor - a common-base stage, inserted between the output node and the mirror proper - can sometimes help. The maximum output resistance of a simple cascode, that is, when driven by a perfect current-source at its emitter, is $V_{AF} / I$, and since the last mirror discussed can already attain that value using large amounts of emitter degeneration, we may need to look for further improvement.

However, the use of a cascode allows the input resistance of the main mirror cell to be kept lower than possible using degeneration, and it can operate with a minimum output bias voltage which is lower, for the same output resistance.

Working on the assumption that it is the variation of beta with $V_{CB}$ that limits $r_0$, we might try a Darlington configuration, to lower the base current variation, or an MOS or JFET device, which has essentially zero gate current. Such combinations are becoming very popular with the advent of combined Bipolar-MOS processes. In some cases, however, beta is not the limiting factor: to understand why, we need to look again at Eq. 6.1c, which shows the dependence of $V_{BE}$ on $V_{CB}$:

$$V_{BE} = V_{BE0} - V_T \ln (1 + V_{CB}/V_{AF})$$

For small variations $\Delta V_{CB}$ near $V_{CB} = 0$, we can calculate the reverse transfer coefficient of the cascode to be

$$DV_{BE}/DV_{CB} = -V_T/V_{AF}$$

that is, neglecting the beta-modulation effects of the output resistance, any

---

7The term "Darlington" configuration here refers to the use of a pair of transistors with common collectors and the emitter of the first connected to the base of the second so as to result in a very high effective beta, equal to the product of the individual betas.
variation in the voltage at the final output will cause the output node of the basic mirror to move and thus modulate the current. In short, it is still important to maximize the output resistance of the main mirror cell if the full benefits of a cascode are to be realized. The variation of $V_{GS}$ with $V_{DS}$ in a FET differs markedly, but the general idea of a reverse transfer ratio in the cascode still applies; in fact, the magnitude of the error is even greater.

### 6.4.3 The Wilson Mirror and Improvements

Sometime in 1967, George Wilson and the author set each other a mutual challenge as an overnight assignment: find a way of making a better current mirror using only three transistors. George won, and the Wilson mirror has become the first choice in many applications where good beta-innunity and high output resistance are needed.

The standard form (which, like the basic mirror, hardly needs repeating) is shown in Figure 6.18a. The circuit is not a panacea; its low beta-sensitivity, as we shall show, depends on beta-matching (unlike the EFA mirror just discussed) and the trick only works at unity-gain; significant errors can arise due to inequalities in the collector voltages of Q1 and Q2; the voltage at the output cannot approach the common node as any of the other mirrors presented so far; its noise performance is worse than the basic mirror; its AC response is strongly peaked. Nevertheless, it has proven to be a valuable cell, and we will later show some ways in which its utility can be extended.

The circuit is interesting in several ways. First, note that it contains a sub-cell, Q1/Q2, which is just a basic mirror; the input of this sub-cell is essentially the output of the complete mirror, with Q3 viewed as a cascode, and the output of the sub-cell becomes the input of the complete mirror. Thus, the Wilson mirror is a current-feedback circuit, the only mirror so far discussed that relies on feedback to establish its mirror ratio. We can think of its operation in this way: following the application of the input current, $I_1$, Q3 would like to respond by generating a current beta times as large, but its emitter current is sensed by Q2, and mirrored back to the input, where a current-balance between $I_1$ and $I_{C3}$ is established, requiring only the base current of Q3 to maintain, this current being in the nature of the error signal in a feedback loop. We might at this point be inclined to suggest that Q3 should therefore be made a Darlington, to reduce this "error" still further. If the sub-cell had no beta errors of its own (essentially true of the EFA mirror) that might be the correct thing to do. In fact, the finite base current of Q3 is needed to ensure that the mirror ratio is close to unity, as we will now see.

Figure 6.18b shows an approximate way of analyzing this circuit. We begin by assuming that the emitter currents of Q1 and Q2 are equal and of value I. This follows from the fact that they are assumed to be identical transistors and are operating with the same $V_{BE}$. Then, using the convention $\delta = I_P/IB$, we find the collector current of Q1 to be $(1-\delta)I$, while the combined collector current of Q2 and the base currents of Q1 and Q2 sum to $(1+\delta)I$, which becomes the emitter current of Q3. To a good approximation, the base current of Q3 is $\delta I$, so the mirror ratio is

$$M = \frac{I_2}{I_1} = \frac{(1+\delta)I - \delta I}{(1+\delta)I + \delta I} = 1$$

(6.23)

whatever the value of $\delta$ (which is roughly $1/\beta$). A more careful analysis, but still assuming all betas are equal, yields

$$M = 1 - \frac{2}{\beta^2 + 2\beta + 2} \approx 1 - \frac{2}{\beta^2}$$

(6.24)

This result, given in most texts on the subject, suggests that the error in M due to finite beta is essentially the same as that for the EFA mirror (see Eq. 6.22a). In fact, it obscures an important detail, namely, that when the betas of Q1 and Q2 are different, there is a significant reduction in the accuracy promised by Eq. 6.24 and the Wilson mirror can very possibly be an order of magnitude worse that the EFA mirror in this regard. The complete expression is

$$M = 1 - \frac{2(\beta_1 - \beta_2)}{\beta_1 \beta_3 + 2 \beta_1 + 2}$$

(6.24a)

Thus, if we assume $\beta_1 = 100$ and $\beta_3 = 95$ (a plausible 5% beta mismatch) the numerator evaluates to 12 rather than 2, so the error in the mirror ratio is
six times worse than for the EFA mirror. The actual error in this case is still a fairly harmless 0.124%, but if this were a mirror built from lateral PNP transistors, where the betas might not just be low and mismatched, but strongly different because of high-level injection effects, the errors can be quite large.

It can be shown that the output resistance for the Wilson mirror is the same as for the heavily-degenerated basic mirror, that is

\[ r_e' = \frac{V_{AE}R_o}{2I_2} (1 + V_{CB}/V_{AE})^2 \]  

(6.25)

Emitter degeneration does not raise the output resistance of the Wilson mirror, because the output bias \( V_2 \) does not materially alter the \( V_{CB} \) of Q1; in fact, it is easily shown that the "attenuation factor" is \( V_{AE}/V_T \), typically 4000:1.

However, degeneration is often used for other reasons: it can transfer the burden of ratio-control from the emitter areas to a resistor ratio; it may obviate an error due mismatch in Q1 and Q2 (see below); it lowers the output noise.

Before discussing various improvements which can be made, one other useful property is worth mentioning. The impedance at the emitter of Q3, call it \( Z_3 \), is quite low. Imagine a signal current of \( i \) applied to this node, and a bias current of \( I_1 \) applied to the normal mirror input. Q1 is forced to conduct \( I_1 \), which requires that the \( V_{BE} \) of Q1 must remain at the constant value required by \( I_1 \), whatever the value of \( i \), which is therefore absorbed by Q3, and thus appears at the output. It follows that \( Z_3 \) ought to be nearly zero. Of course, a component \( \delta i \) flows in the base of Q3, thus modifying the collector current of Q1 and perturbing its \( V_{BE} \). We can guess (correctly!) that at low frequencies \( Z_3 \) approximates to a resistance of \( V_T/\beta I_1 \), or about
Figure 6.21 Comparison of AC responses of the three most popular
mirrors: the Wilson is the most peaky; the EFA mirror has about half the
bandwidth of the basic mirror, but the same 1.8dB of peaking.

0.26Ω for I₁ = 1mA. At high frequencies, the circuit becomes resonant (at
about fₗ/3) but Zₕ remains under 10Ω up to about fₗ/10. Of course, device
capacitances play a more important role at lower currents. Figures 6.19 and
6.20 show typical results for a circuit using equally-sized transistors with an
fₗ of about 3GHz. The flatness of both the impedance and the AC transfer
function, with respect to this non-standard input, can be improved by the
addition of a small capacitor across the collector-base junction of Q₁.

While on the subject of AC response, we show in Figure 6.21 a comparison
of the three most popular current mirrors: the basic mirror, the EFA mirror
and the Wilson mirror. In all cases, equal transistor geometries were used,
typical of a 3GHz NPN process, operating at a bias current of 1mA and
without the use of emitter degeneration. The severe peaking of the Wilson
mirror is apparent. In designing mirrors for current-path applications, much
can be done to control the gain and phase response. In fact, the first
improvement to the Wilson mirror to be described has a flatter HF response
as one of its benefits.

Figure 6.22a shows the circuit, in which an extra transistor Q₃b has been
added to Q₃ to supply the base currents of Q₁ and Q₂; in practice, this would
simply be an additional emitter, as shown in Figure 6.22b. Figure 6.23 shows
a useful improvement in the AC response. This is due to the radical
alteration of the incremental resistance driving the common-base node. In the
basic Wilson mirror, this resistance is the $r_e$ of Q₃; in the modified circuit it
is increased by the factor β/2, thus forming a pole with the $C_e$ of Q₁ and Q₂
at a much lower frequency, resulting in a more stable feedback path.

This modification has a second advantage, only of interest in low-voltage
applications, which is that it allows the output node to drop closer to the
common node before finally crashing (Figure 6.24). A more important
improvement is the addition of an extra diode-connected transistor, Q₄, in
series with the collector of Q₁ (Figure 6.22b). Its purpose is to equalize the
$V_{CB}$ of Q₁ and Q₂ and thereby remove a significant source of error in the
mirror ratio. The magnitude of this error is quickly estimated: a $AV_{CB}$ of one
$V_{BE}$ can be shown to alter the ratio of the sub-cell Q₁/Q₂ from unity to

$$M' = \exp(V_{BE}/V_{AP})$$  (6.26)