Behzad Razavi

DESIGN OF Analog CMOS Integrated Circuits

Mc Graw Hill Education

SECOND EDITION

CHAPTER

12

Bandgap References

Analog circuits incorporate voltage and current references extensively. Such references are dc quantities that exhibit little dependence on supply and process parameters and a *well-defined* dependence on the temperature. For example, the bias current of a differential pair must be generated according to a reference, for it affects the voltage gain and noise of the circuit. We have also seen the need for precise voltages to define common-mode levels in op amps. Moreover, in systems such as A/D and D/A converters, a reference is required to define the input or output full-scale range.

In this chapter, we deal with the design of reference generators in CMOS technology, focusing on well-established "bandgap" techniques. First, we study supply-independent biasing and the problem of start-up. Next, we describe temperature-independent references and examine issues such as the effect of offset voltages. Finally, we present constant-G_m biasing and study an example of state-of-the-art bandgap references.

12.1 ■ General Considerations

As mentioned above, the objective of reference generation is to establish a dc voltage or current that is independent of the supply and process and has a well-defined behavior with temperature. In most applications, the required temperature dependence assumes one of three forms: (1) proportional to absolute temperature (PTAT); (2) constant-G_m behavior, i.e., such that the transconductance of certain transistors remains constant; (3) temperature independent. We can therefore divide the task into two design problems: supply-independent biasing and definition of the temperature variation.

In addition to supply, process, and temperature variability, several other parameters of reference generators may be critical as well. These include output impedance, output noise, and power dissipation. We return to these issues later in this chapter.

12.2 ■ Supply-Independent Biasing

Our use of bias currents and current mirrors in previous chapters has implicitly assumed that a "golden" reference current is available. As shown in Fig. 12.1(a), if I_{REF} does not vary with V_{DD}, and channel-length modulation of M_2 and M_3 is neglected, then I_{D2} and I_{D3} remain independent of the supply voltage. The question then is—How do we generate I_{REF}?

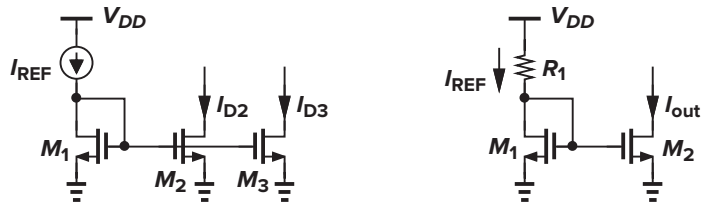


Figure 12.1 Current mirror biasing using (a) an ideal current source and (b) a resistor.

As an approximation of a current source, we tie a resistor from V_{DD} to the gate of M_1 [Fig. 12.1(b)]. However, the output current of this circuit is quite sensitive to V_{DD}:

$$\Delta I_{out} = \frac{\Delta V_{DD}}{R_1 + 1/g_{m1}} \cdot \frac{(W/L)_2}{(W/L)_1} \tag{12.1}$$

In order to arrive at a less sensitive solution, we postulate that the circuit must bias *itself*, i.e., I_{REF} must be somehow derived from I_{out}. The idea is that if I_{out} is to be ultimately independent of V_{DD}, then I_{REF} can be a replica of I_{out}. Figure 12.2 illustrates an implementation where M_3 and M_4 copy I_{out}, thereby defining I_{REF}. In essence, I_{REF} is "bootstrapped" to I_{out}. With the sizes chosen here, we have $I_{out} = K I_{REF}$ if channel-length modulation is neglected. Note that, since each diode-connected device feeds from a current source, I_{out} and I_{REF} are relatively independent of V_{DD}.

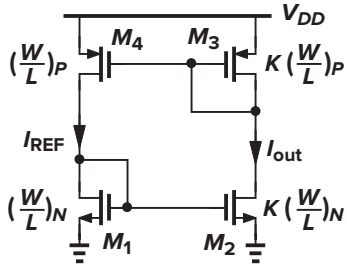


Figure 12.2 Simple circuit to establish supply-independent currents.

Since I_{out} and I_{REF} in Fig. 12.2 display little dependence on V_{DD}, their magnitude is set by other parameters. How do we calculate these currents? Interestingly, if M_1–M_4 operate in saturation and $\lambda \approx 0$, then the circuit is governed by only one equation, $I_{out} = K I_{REF}$, and hence can support *any* current level! For example, if we initially force I_{REF} to be 10 μA, the resulting I_{out} of $K \times 10\ \mu$A "circulates" around the loop, sustaining these current levels in the left and right branches indefinitely.

To uniquely define the currents, we add another constraint to the circuit, e.g., as shown in Fig. 12.3(a). Here, resistor R_S decreases the current of M_2 while the PMOS devices require that $I_{out} = I_{REF}$ because they have identical dimensions and thresholds. We can write $V_{GS1} = V_{GS2} + I_{D2} R_S$, or

$$\sqrt{\frac{2I_{out}}{\mu_n C_{ox}(W/L)_N}} + V_{TH1} = \sqrt{\frac{2I_{out}}{\mu_n C_{ox} K (W/L)_N}} + V_{TH2} + I_{out} R_S \tag{12.2}$$

Neglecting body effect, we have

$$\sqrt{\frac{2I_{out}}{\mu_n C_{ox}(W/L)_N}} \left(1 - \frac{1}{\sqrt{K}}\right) = I_{out} R_S \tag{12.3}$$

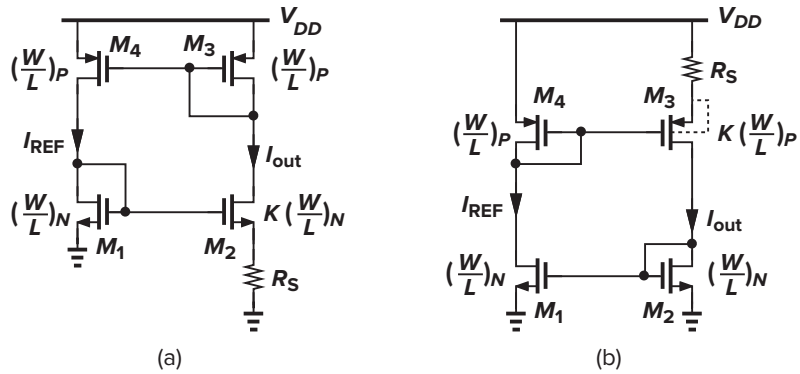


Figure 12.3 (a) Addition of R_S to define the currents; (b) alternative implementation eliminating body effect.

and hence

$$I_{out} = \frac{2}{\mu_n C_{ox}(W/L)_N} \cdot \frac{1}{R_S^2}\left(1 - \frac{1}{\sqrt{K}}\right)^2 \tag{12.4}$$

As expected, the current is independent of the supply voltage (but still a function of process and temperature).

The assumption $V_{TH1} = V_{TH2}$ introduces some error in the foregoing calculations because the sources of M_1 and M_2 are at different voltages. Shown in Fig. 12.3(b) is to place the resistor in the source of M_3 while eliminating body effect by tying the source and bulk of each PMOS transistor. Another solution is described in Problem 12.1.

The circuits of Figs. 12.3(a) and (b) exhibit little supply dependence if channel-length modulation is negligible. For this reason, relatively long channels are used for all of the transistors in the circuit. This also helps reduce their flicker noise.

▶ Example 12.1

Assuming $\lambda \neq 0$ in Fig. 12.3(a), estimate the change in I_{out} for a small change ΔV_{DD} in the supply voltage.

Solution

Simplifying the circuit as depicted in Fig. 12.4, where $R_1 = r_{O1} \| (1/g_{m1})$ and $R_3 = r_{O3} \| (1/g_{m3})$, we calculate the "gain" from V_{DD} to I_{out}. The small-signal gate-source voltage of M_4 equals $-I_{out} R_3$, and the current through r_{O4} is $(V_{DD} - V_X)/r_{O4}$. Thus,

$$\frac{V_{DD} - V_X}{r_{O4}} + I_{out} R_3 g_{m4} = \frac{V_X}{R_1} \tag{12.5}$$

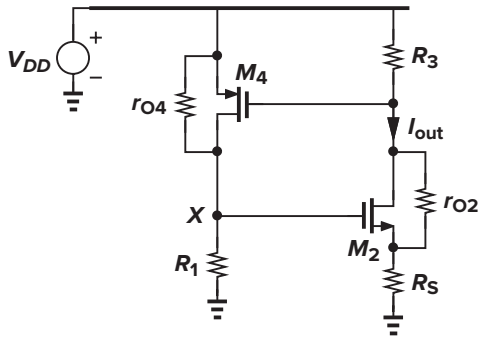


Figure 12.4

If we denote the equivalent transconductance of M_2 and R_S by $G_{m2} = I_{out}/V_X$, then

$$\frac{I_{out}}{V_{DD}} = \frac{1}{r_{O4}} \left[\frac{1}{G_{m2}(r_{O4} \| R_1)} - g_{m4} R_3 \right]^{-1} \tag{12.6}$$

Note from Chapter 3 that

$$G_{m2} = \frac{g_{m2} r_{O2}}{R_S + r_{O2} + (g_{m2} + g_{mb2}) R_S r_{O2}} \tag{12.7}$$

Interestingly, the sensitivity vanishes if $r_{O4} = \infty$. ◀

In some applications, the sensitivity given by (12.6) is prohibitively large. Also, owing to various capacitive paths, the supply sensitivity of the circuit rises at high frequencies. For these reasons, the supply voltage of the core is often derived from a locally-generated, less sensitive voltage. We return to this point in Sec. 12.8.

An important issue in supply-independent biasing is the existence of "degenerate" bias points. In the circuit of Fig. 12.3(a), for example, if all of the transistors carry zero current when the supply is turned on, they may remain off indefinitely because the loop can support a zero current in both branches. This condition is not predicted by (12.4) because in manipulating (12.3), we divided both sides by $\sqrt{I_{out}}$, tacitly assuming that $I_{out} \neq 0$. In other words, the circuit can settle in one of *two* different operating conditions.

Called the "start-up" problem, the above issue is resolved by adding a mechanism that drives the circuit out of the degenerate bias point when the supply is turned on. Shown in Fig. 12.5(a) is a simple example, where the diode-connected device M_5 provides a current path from V_{DD} through M_3 and M_1 to ground upon start-up. Thus, M_3 and M_1, and hence M_2 and M_4, cannot remain off. Of course, this technique is practical only if $V_{TH1} + V_{TH5} + |V_{TH3}| < V_{DD}$ and $V_{GS1} + V_{TH5} + |V_{GS3}| > V_{DD}$, the latter to ensure that M_5 remains off after start-up. Another start-up circuit is analyzed in Problem 12.2.

The problem of start-up generally requires careful analysis and simulation. The supply voltage must be ramped from zero in a dc sweep simulation (such that parasitic capacitances do not cause false start-up) as well as in a transient simulation and the behavior of the circuit examined for each supply voltage. Figure 12.5(b) depicts an example of the observed behavior in the presence of the start-up circuit. In complex implementations, more than one degenerate point may exist.

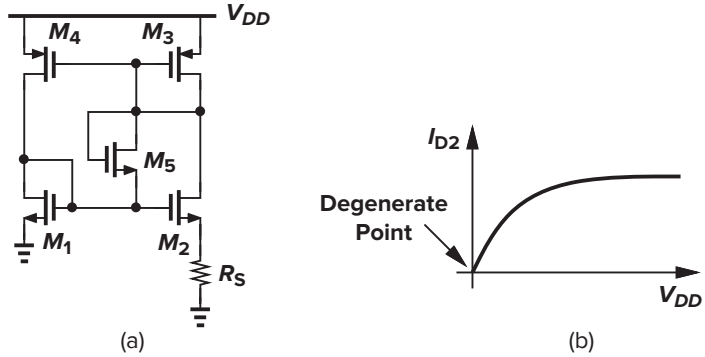


Figure 12.5 (a) Addition of start-up device to the circuit of Fig. 12.3(a), and (b) illustration of degenerate point.

12.3 ■ Temperature-Independent References

Reference voltages or currents that exhibit little dependence on temperature prove essential in many analog circuits. It is interesting to note that, since most process parameters vary with temperature, if a reference is temperature-independent, then it is usually process-independent as well.

How do we generate a quantity that remains constant with temperature? We postulate that if two quantities having opposite temperature coefficients (TCs) are added with proper weighting, the result displays a zero TC. For example, for two voltages V_1 and V_2 that vary in opposite directions with temperature, we choose α_1 and α_2 such that $\alpha_1 \partial V_1/\partial T + \alpha_2 \partial V_2/\partial T = 0$, obtaining a reference voltage, $V_{REF} = \alpha_1 V_1 + \alpha_2 V_2$, with zero TC.

We must now identify two voltages that have positive and negative TCs. Among various device parameters in semiconductor technologies, the characteristics of bipolar transistors have proven the most reproducible and well-defined quantities that can provide positive and negative TCs. Even though many parameters of MOS devices have been considered for the task of reference generation [1, 2], bipolar operation still forms the core of such circuits.

12.3.1 Negative-TC Voltage

The base-emitter voltage of bipolar transistors or, more generally, the forward voltage of a *pn*-junction diode exhibits a negative TC. We first obtain the expression for the TC in terms of readily-available quantities.

For a bipolar device, we can write $I_C = I_S \exp(V_{BE}/V_T)$, where $V_T = kT/q$. The saturation current I_S is proportional to $\mu k T n_i^2$, where μ denotes the mobility of minority carriers and n_i is the intrinsic carrier concentration of silicon. The temperature dependence of these quantities is represented as $\mu \propto \mu_0 T^m$, where $m \approx -3/2$, and $n_i^2 \propto T^3 \exp[-E_g/(kT)]$, where $E_g \approx 1.12$ eV is the bandgap energy of silicon. Thus,

$$I_S = bT^{4+m} \exp \frac{-E_g}{kT} \tag{12.8}$$

where b is a proportionality factor. Writing $V_{BE} = V_T \ln(I_C/I_S)$, we can now compute the TC of the base-emitter voltage. In taking the derivative of V_{BE} with respect to T, we must know the behavior of I_C as a function of the temperature. To simplify the analysis, we assume for now that I_C is held *constant*. Thus,

$$\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \ln \frac{I_C}{I_S} - \frac{V_T}{I_S} \frac{\partial I_S}{\partial T} \tag{12.9}$$

From (12.8), we have

$$\frac{\partial I_S}{\partial T} = b(4+m)T^{3+m} \exp \frac{-E_g}{kT} + bT^{4+m} \left(\exp \frac{-E_g}{kT} \right) \left(\frac{E_g}{kT^2} \right) \tag{12.10}$$

Therefore,

$$\frac{V_T}{I_S} \frac{\partial I_S}{\partial T} = (4+m)\frac{V_T}{T} + \frac{E_g}{kT^2} V_T \tag{12.11}$$

With the aid of (12.9) and (12.11), we can write

$$\frac{\partial V_{BE}}{\partial T} = \frac{V_T}{T} \ln \frac{I_C}{I_S} - (4+m)\frac{V_T}{T} - \frac{E_g}{kT^2} V_T \tag{12.12}$$

$$= \frac{V_{BE} - (4+m)V_T - E_g/q}{T} \tag{12.13}$$

Equation (12.13) gives the temperature coefficient of the base-emitter voltage at a given temperature T, revealing dependence on the magnitude of V_{BE} itself. With $V_{BE} \approx 750$ mV and $T = 300$ K, we have $\partial V_{BE}/\partial T \approx -1.5$ mV/K.

In old bipolar technologies, where I_C/I_S was relatively small (because the transistors were large), $V_{BE} \approx 700$ mV and $\partial V_{BE}/\partial T \approx -1.9$ mV/K at room temperature. Modern bipolar transistors typically operate at much higher current densities, exhibiting $V_{BE} \approx 800$ mV and hence $\partial V_{BE}/\partial T \approx -1.5$ mV/K at $T = 300$ K.

From (12.13), we note that the temperature coefficient of V_{BE} itself depends on the temperature, creating error in constant reference generation if the positive-TC quantity exhibits a *constant* temperature coefficient.

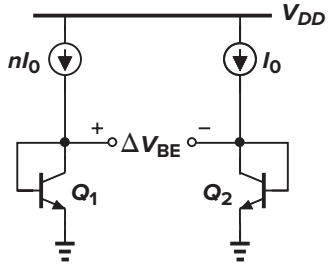


Figure 12.6 Generation of PTAT voltage.

12.3.2 Positive-TC Voltage

It was recognized in 1964 [3] that if two bipolar transistors operate at unequal current densities,[1] then the *difference* between their base-emitter voltages is directly proportional to the absolute temperature. For example, as shown in Fig. 12.6, if two identical transistors ($I_{S1} = I_{S2}$) are biased at collector currents of nI_0 and I_0 and their base currents are negligible, then

$$\Delta V_{BE} = V_{BE1} - V_{BE2} \tag{12.14}$$

$$= V_T \ln \frac{nI_0}{I_{S1}} - V_T \ln \frac{I_0}{I_{S2}} \tag{12.15}$$

$$= V_T \ln n \tag{12.16}$$

Thus, the V_{BE} difference exhibits a positive temperature coefficient:

$$\frac{\partial \Delta V_{BE}}{\partial T} = \frac{k}{q} \ln n \tag{12.17}$$

Interestingly, this TC is independent of the temperature or behavior of the collector currents.[2]

▶ **Example 12.2**

How must n be chosen to yield a TC of +1.5 mV/K so as to cancel the TC of the base-emitter voltage at $T = 300$ K?

Solution

We choose n so that $(k/q) \ln n = 1.5$ mV/K. Since $k/q = V_T/T = 0.087$ mV/K, we have $\ln n \approx 17.2$ and hence $n = 2.95 \times 10^7$!! We must therefore modify the circuit to avoid such a large disparity between the two currents. ◀

[1] Current density is defined as the ratio of the collector current, I_C, and the saturation current, I_S.

[2] Nonidealities in the characteristics of bipolar transistors introduce a small temperature dependence in this TC.

▶ **Example 12.3**

Calculate ΔV_{BE} in the circuit of Fig. 12.7, where Q_2 is formed as the parallel combination of m units, each identical to Q_1.

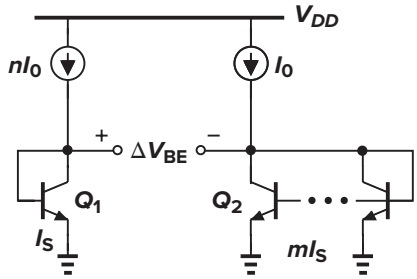


Figure 12.7

Solution

Neglecting base currents, we can write

$$\Delta V_{BE} = V_T \ln \frac{nI_0}{I_S} - V_T \ln \frac{I_0}{mI_S} \tag{12.18}$$

$$= V_T \ln(nm) \tag{12.19}$$

The temperature coefficient is therefore equal to $(k/q)\ln(nm)$. In this circuit, the two transistors' current densities differ by a factor of nm.

◀

12.3.3 Bandgap Reference

With the negative- and positive-TC voltages obtained above, we can now develop a reference that has a nominally zero temperature coefficient. We write $V_{REF} = \alpha_1 V_{BE} + \alpha_2(V_T \ln n)$, where $V_T \ln n$ is the difference between the base-emitter voltages of the two bipolar transistors operating at different current densities. How do we choose α_1 and α_2? Since at room temperature, $\partial V_{BE}/\partial T \approx -1.5$ mV/K whereas $\partial V_T/\partial T \approx +0.087$ mV/K, we may set $\alpha_1 = 1$ and choose $\alpha_2 \ln n$ such that $(\alpha_2 \ln n)(0.087\ \text{mV/K}) = 1.5$ mV/K. That is, $\alpha_2 \ln n \approx 17.2$, indicating that for zero TC

$$V_{REF} \approx V_{BE} + 17.2V_T \tag{12.20}$$

$$\approx 1.25\ \text{V} \tag{12.21}$$

Let us now devise a circuit that adds V_{BE} to $17.2V_T$. First, consider the circuit shown in Fig. 12.8, where the base currents are assumed to be negligible, transistor Q_2 consists of n unit transistors in parallel, and Q_1 is a unit transistor. Suppose we somehow force V_{O1} and V_{O2} to be equal. Then, $V_{BE1} = RI + V_{BE2}$ and $RI = V_{BE1} - V_{BE2} = V_T \ln n$. Thus, $V_{O2} = V_{BE2} + V_T \ln n$, suggesting that V_{O2} can serve as a temperature-independent reference if $\ln n \approx 17.2$ (while V_{O1} and V_{O2} remain equal).

The circuit of Fig. 12.8 requires three modifications to become practical. First, a mechanism must be added to guarantee that $V_{O1} = V_{O2}$. Second, since $\ln n = 17.2$ translates to a prohibitively large n, the term $RI = V_T \ln n$ must be scaled up by a reasonable factor. Third, V_{O2}, which is somehow forced to be equal to V_{O1}, *cannot* become temperature-independent because $V_{O2} \approx V_{BE1} \approx 800$ mV whereas, for temperature independence, we must have $V_{O2} = V_{BE2} + 17.2V_T \approx 1.25$ V. Shown in Fig. 12.9 is an implementation accomplishing all tasks [4]. Here, amplifier A_1 senses V_X and V_Y, driving the top terminals of R_1 and R_2 ($R_1 = R_2$) such that X and Y settle to approximately equal voltages. The

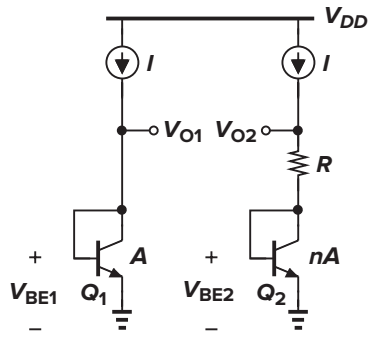


Figure 12.8 Conceptual generation of temperature-independent voltage.

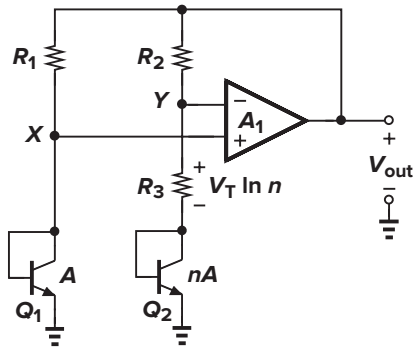


Figure 12.9 Actual implementation of the concept shown in Fig. 12.8.

reference voltage is obtained at the output of the amplifier (rather than at node Y). Following the analysis of Fig. 12.8, we have $V_{BE1} - V_{BE2} = V_T \ln n$, arriving at a current equal to $V_T \ln n / R_3$ through the right branch and hence an output voltage of

$$V_{out} = V_{BE2} + \frac{V_T \ln n}{R_3}(R_3 + R_2) \tag{12.22}$$

$$= V_{BE2} + (V_T \ln n)\left(1 + \frac{R_2}{R_3}\right) \tag{12.23}$$

For a zero TC, we must have $(1 + R_2/R_3)\ln n \approx 17.2$. For example, we may choose $n = 31$ and $R_2/R_3 = 4$. Note that these results do not depend on the TC of the resistors.

It is interesting to understand how the third issue mentioned above is resolved in the topology of Fig. 12.9: we do not attempt to make V_Y ($\approx V_{BE1}$) temperature-independent; rather, we amplify the PTAT voltage drop across R_3 by a factor of $1 + R_2/R_3$ and then add the result to V_{BE2}.

▶ Example 12.4

In Fig. 12.9, R_1 and R_2 are equal and sustain equal voltages, each carrying a current of $(V_T \ln n)/R_3$. We therefore have

$$V_{out} = V_{BE1} + (V_T \ln n)\frac{R_1}{R_3} \tag{12.24}$$

But the second term is *not* equal to $17.2V_T$ if we have already chosen $(V_T \ln n)(1 + R_2/R_3) = 17.2V_T$. Explain this discrepancy.

Solution

The first terms in (12.23) and (12.24) are different. We substitute $V_{BE1} = V_{BE2} + V_T \ln n$ in Eq. (12.13):

$$\frac{\partial V_{BE1}}{\partial T} = \frac{V_{BE2} + V_T \ln n - (4+m)V_T - E_g/q}{T} \tag{12.25}$$

$$= \frac{\partial V_{BE2}}{\partial T} + \frac{k}{q} \ln n \tag{12.26}$$

Thus,

$$\frac{\partial V_{out}}{\partial T} = \frac{\partial V_{BE1}}{\partial T} + \left(\frac{k}{q} \ln n\right) \frac{R_1}{R_3} \tag{12.27}$$

$$= \frac{\partial V_{BE2}}{\partial T} + \left(\frac{k}{q} \ln n\right) \left(1 + \frac{R_1}{R_3}\right) \tag{12.28}$$

which is consistent with (12.23). ◀

The circuit of Fig. 12.9 entails a number of design issues. We consider each one below.

Collector Current Variation The circuit of Fig. 12.9 violates one of our earlier assumptions: the collector currents of Q_1 and Q_2, given by $(V_T \ln n)/R_3$, are proportional to T, whereas $\partial V_{BE}/\partial T \approx -1.5$ mV/K was derived for a *constant* current. What happens to the temperature coefficient of V_{BE} if the collector currents are PTAT? As a first-order iterative solution, let us assume that $I_{C1} = I_{C2} \approx (V_T \ln n)/R_3$. Returning to Eq. (12.9) and including $\partial I_C/\partial T$, we have

$$\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \ln \frac{I_C}{I_S} + V_T \left(\frac{1}{I_C}\frac{\partial I_C}{\partial T} - \frac{1}{I_S}\frac{\partial I_S}{\partial T}\right) \tag{12.29}$$

Since $\partial I_C/\partial T \approx (V_T \ln n)/(R_3 T) = I_C/T$, we can write

$$\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \ln \frac{I_C}{I_S} + \frac{V_T}{T} - \frac{V_T}{I_S}\frac{\partial I_S}{\partial T} \tag{12.30}$$

Equation (12.13) is therefore modified as

$$\frac{\partial V_{BE}}{\partial T} = \frac{V_{BE} - (3+m)V_T - E_g/q}{T} \tag{12.31}$$

indicating that the TC is slightly less negative than -1.5 mV/K. In practice, accurate simulations are necessary to predict the temperature coefficient.

Compatibility with CMOS Technology Our derivation of a temperature-independent voltage relies on the exponential characteristics of bipolar devices for both negative- and positive-TC quantities. We must therefore seek structures in a standard CMOS technology that exhibit such characteristics.

In n-well processes, a *pnp* transistor can be formed as depicted in Fig. 12.10. A p^+ region (the same as the S/D region of PFETs) inside an n-well serves as the emitter and the n-well itself as the base. The p-type substrate acts as the collector and it is inevitably connected to the most negative supply (usually ground). The circuit of Fig. 12.9 can therefore be redrawn as shown in Fig. 12.11.

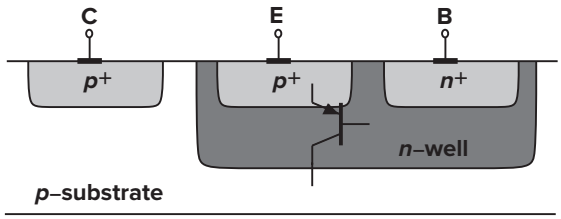


Figure 12.10 Realization of a *pnp* bipolar transistor in CMOS technology.

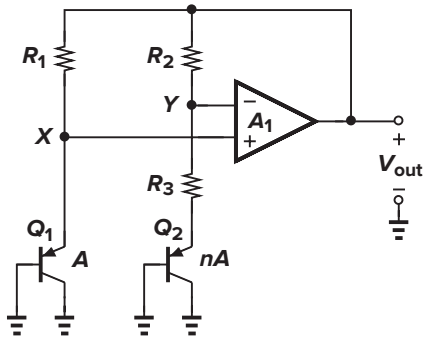


Figure 12.11 Circuit of Fig. 12.9 implemented with *pnp* transistors.

Op Amp Offset and Output Impedance As explained in Chapter 14, owing to asymmetries, op amps suffer from input "offsets," i.e., the output voltage of the op amp is not zero if the input is set to zero. The input offset voltage of the op amp in Fig. 12.9 introduces error in the output voltage. Included in Fig. 12.12, the effect is quantified as $V_{BE1} - V_{OS} \approx V_{BE2} + R_3 I_{C2}$ (if A_1 is large) and $V_{out} = V_{BE2} + (R_3 + R_2) I_{C2}$. Thus,

$$V_{out} = V_{BE2} + (R_3 + R_2)\frac{V_{BE1} - V_{BE2} - V_{OS}}{R_3} \tag{12.32}$$

$$= V_{BE2} + \left(1 + \frac{R_2}{R_3}\right)(V_T \ln n - V_{OS}) \tag{12.33}$$

where we have assumed that $I_{C2} \approx I_{C1}$ despite the offset voltage. The key point here is that V_{OS} is amplified by $1 + R_2/R_3$, introducing error in V_{out}. More important, as explained in Chapter 14, V_{OS} itself varies with temperature, raising the temperature coefficient of the output voltage.

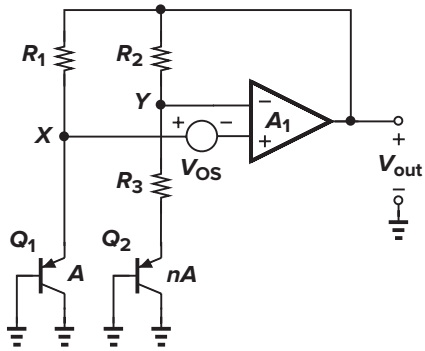


Figure 12.12 Effect of op amp offset on the reference voltage.

▶ Example 12.5

Assuming an ideal op amp, determine the small-signal gain from V_{OS} to V_{out} in Fig. 12.12.

Solution

In the absence of the op amp offset, the two diode-connected bipolar transistors carry equal bias currents, exhibiting a transconductance of g_m. Replacing Q_1 and Q_2 with a small-signal resistance equal to $1/g_m$ and noting that $V_X - V_{OS} \approx V_Y$, we write the following small-signal equation:

$$\frac{1/g_m}{1/g_m + R_1} V_{out} - V_{OS} = \frac{1/g_m + R_3}{1/g_m + R_3 + R_2} V_{out} \tag{12.34}$$

Since $R_1 = R_2$,

$$\frac{V_{out}}{V_{OS}} = -\left[1 + \frac{1}{g_m R_2} + \frac{(1/g_m + R_2)^2}{R_2 R_3}\right] \tag{12.35}$$

If $g_m R_2 \gg 1$, then $V_{out}/V_{OS} \approx -(1+R_2/R_3)$, agreeing with the results obtained previously. (After all, if $1/g_m \approx 0$, V_{OS} simply sees a noninverting amplifier with a gain of $1 + R_2/R_3$.)

Why does (12.35) not completely agree with the $-V_{OS}(1 + R_2/R_3)$ component in (12.33)? Recall that (12.33) was derived with the assumption that $I_{C1} \approx I_{C2}$ despite the offset voltage. Since $V_X - V_{OS} = V_Y$, we have $I_{C1}R_1 - V_{OS} = I_{C2}R_2$, and hence $I_{C1} = I_{C2} + V_{OS}/R_2$. Let us return to (12.32) and write

$$V_{BE1} - V_{BE2} - V_{OS} = V_T \ln \frac{I_{C1}}{I_{S1}} - V_T \ln \frac{I_{C2}}{I_{S2}} - V_{OS} \tag{12.36}$$

$$= V_T \ln n - V_T \ln \frac{I_{C1}}{I_{C2}} - V_{OS} \tag{12.37}$$

$$= V_T \ln n - V_T \ln \left(1 + \frac{V_{OS}}{R_2 I_{C2}}\right) - V_{OS} \tag{12.38}$$

$$\approx V_T \ln n - V_T \frac{V_{OS}}{R_2 I_{C2}} - V_{OS} \tag{12.39}$$

$$\approx V_T \ln n - \left(1 + \frac{1}{g_m R_2}\right) V_{OS} \tag{12.40}$$

The output offset contribution therefore amounts to $-[1 + 1/(g_m R_2)](1 + R_2/R_3)V_{OS}$, which is approximately the same as (12.35). ◀

Several methods are employed to lower the effect of V_{OS}. First, the op amp incorporates large devices in a carefully chosen topology so as to minimize the offset (Chapter 19). Second, as illustrated in Fig. 12.7, the collector currents of Q_1 and Q_2 can be ratioed by a factor of m such that $\Delta V_{BE} = V_T \ln(mn)$. Third, each branch may use two pn junctions in series to double ΔV_{BE}. Figure 12.13 depicts a realization using the last two techniques. Here, R_1 and R_2 are ratioed by a factor of m, producing $I_1 \approx mI_2$. Neglecting base currents and assuming that A_1 is large, we can now write $V_{BE1} + V_{BE2} - V_{OS} = V_{BE3} + V_{BE4} + R_3 I_2$ and $V_{out} = V_{BE3} + V_{BE4} + (R_3 + R_2)I_2$. It follows that

$$V_{out} = V_{BE3} + V_{BE4} + (R_3 + R_2)\frac{2V_T \ln(mn) - V_{OS}}{R_3} \tag{12.41}$$

$$= 2V_{BE} + \left(1 + \frac{R_2}{R_3}\right)[2V_T \ln(mn) - V_{OS}] \tag{12.42}$$

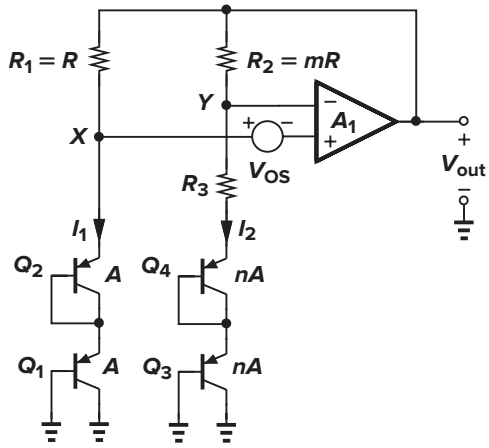


Figure 12.13 Reduction of the effect of op amp offset.

Thus, the effect of the offset voltage is reduced by increasing the first term in the square brackets. The issue, however, is that $V_{out} \approx 2 \times 1.25\ \text{V} = 2.5\ \text{V}$, a value difficult to generate by the op amp at low supply voltages.

In the circuits studied above, the op amp drives two resistive branches and must therefore provide a low output impedance. Fortunately, it is possible to avoid this issue by a simple modification described below.

The implementation of Fig. 12.13 is not feasible in a standard CMOS technology because the collectors of Q_2 and Q_4 are not grounded. In order to utilize the bipolar structure shown in Fig. 12.10, we modify the series combination of the diodes as illustrated in Fig. 12.14(a), converting one of the diodes to an emitter follower. However, we must ensure that the bias currents of both transistors have the same behavior with temperature. Thus, we bias each transistor by a PMOS current source rather than a resistor [Fig. 12.14(b)]. The overall circuit then assumes the topology shown in Fig. 12.15, where the op amp adjusts the gate voltage of the PMOS devices so as to equalize V_X and V_Y. Interestingly, in this circuit, the op amp experiences no resistive loading, but the mismatch and channel-length modulation of the PMOS devices introduce error at the output (Problem 12.3).

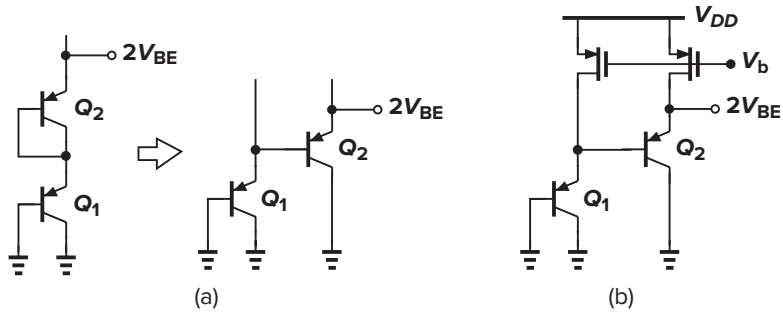


Figure 12.14 (a) Conversion of series diodes to a topology with grounded collectors; (b) circuit of part (a) biased by PMOS current sources.

An important concern in the circuit of Fig. 12.15 is the relatively low current gain of the "native" *pnp* transistors. Since the base currents of Q_2 and Q_4 generate an error in the emitter currents of Q_1 and Q_3, a means of base current cancellation may be necessary (Problem 12.5).

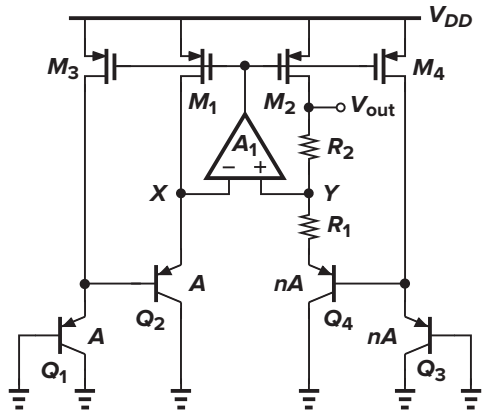


Figure 12.15 Reference generator incorporating two series base-emitter voltages.

Feedback Polarity In the circuit of Fig. 12.9, the feedback signal produced by the op amp returns to both of its inputs. The negative-feedback factor is given by

$$\beta_N = \frac{1/g_{m2} + R_3}{1/g_{m2} + R_3 + R_2} \tag{12.43}$$

and the positive-feedback factor by

$$\beta_P = \frac{1/g_{m1}}{1/g_{m1} + R_1} \tag{12.44}$$

To ensure an overall negative feedback, β_P must be less than β_N, preferably by roughly a factor of two so that the circuit's transient response remains well behaved with large capacitive loads.

Bandgap Reference The voltage generated according to (12.20) is called a "bandgap reference." To understand the origin of this terminology, let us write the output voltage as

$$V_{REF} = V_{BE} + V_T \ln n \tag{12.45}$$

and hence:

$$\frac{\partial V_{REF}}{\partial T} = \frac{\partial V_{BE}}{\partial T} + \frac{V_T}{T} \ln n \tag{12.46}$$

Setting this to zero and substituting for $\partial V_{BE}/\partial T$ from (12.13), we have

$$\frac{V_{BE} - (4+m)V_T - E_g/q}{T} = -\frac{V_T}{T} \ln n \tag{12.47}$$

If $V_T \ln n$ is found from this equation and inserted in (12.45), we obtain

$$V_{REF} = \frac{E_g}{q} + (4+m)V_T \tag{12.48}$$

Thus, the reference voltage exhibiting a nominally-zero TC is given by a few *fundamental* numbers: the bandgap voltage of silicon, E_g/q, the temperature exponent of mobility, m, and the thermal voltage, V_T. The term "bandgap" is used here because as $T \to 0$, $V_{REF} \to E_g/q$.

▶ **Example 12.6**

Prove directly that, as $T \to 0$, $V_{BE} \to E_g/q$, and hence $V_{REF} = V_{BE} + V_T \ln n \to E_g/q$.

Solution

From Eq. (12.8), we have

$$V_{BE} = V_T \ln \frac{I_C}{I_S} \tag{12.49}$$

$$= V_T \left[\ln I_C - \ln b - (4+m) \ln T + \frac{E_g}{kT} \right] \tag{12.50}$$

Thus, $V_{BE} \to E_g/q$ if $T \to 0$ and I_C is constant. ◀

Supply Dependence and Start-Up In the circuit of Fig. 12.9, the output voltage is relatively independent of the supply voltage so long as the open-loop gain of the op amp is sufficiently high. The circuit may require a start-up mechanism because if V_X and V_Y are equal to zero, the input differential pair of the op amp may turn off. Start-up techniques similar to those of Fig. 12.5 can be added to ensure that the op amp turns on when the supply is applied.

The supply rejection of the circuit typically degrades at high frequencies owing to the op amp's rejection properties, often mandating "supply regulation." An example is described in Sec. 12.8.

Curvature Correction If plotted as a function of temperature, bandgap voltages exhibit a finite "curvature," i.e., their TC is typically zero at one temperature and positive or negative at other temperatures (Fig. 12.16). The curvature arises from temperature variation of base-emitter voltages, collector currents, and offset voltages.

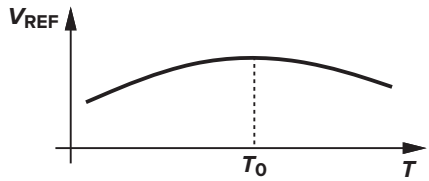


Figure 12.16 Curvature in temperature dependence of a bandgap voltage.

Many curvature correction techniques have been devised to suppress the variation of V_{REF} [5, 6] in bipolar bandgap circuits, but they are seldom used in CMOS counterparts. This is because, due to large offsets and process variations, samples of a bandgap reference display substantially different zero-TC temperatures (Fig. 12.17), making it difficult to correct the curvature reliably.

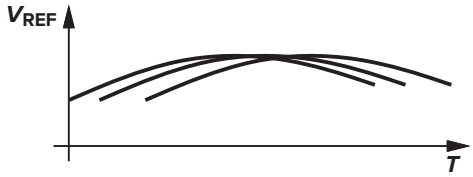


Figure 12.17 Variation of the zero-TC temperature for different samples.

12.4 ■ PTAT Current Generation

In the analysis of bandgap circuits, we noted that the bias currents of the bipolar transistors are in fact proportional to absolute temperature. Useful in many applications, PTAT currents can be generated by a topology such as that shown in Fig. 12.18. Alternatively, we can combine the supply-independent biasing scheme of Fig. 12.2 with a bipolar core, arriving at Fig. 12.19.[3] Assuming for simplicity that M_1-M_2 and M_3-M_4 are identical pairs, we note that for $I_{D1} = I_{D2}$, the circuit must ensure that $V_X = V_Y$. Thus, $I_{D1} = I_{D2} = (V_T \ln n)/R_1$, yielding the same behavior for I_{D5}. In practice, due to mismatches between the transistors and, more important, the temperature coefficient of R_1, the variation of I_{D5} deviates from the ideal equation. For low-voltage operation, the topology of Fig. 12.18 is preferred.

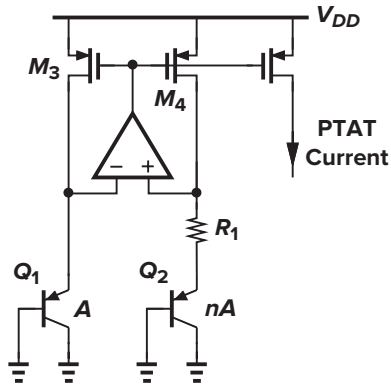


Figure 12.18 Generation of a PTAT current.

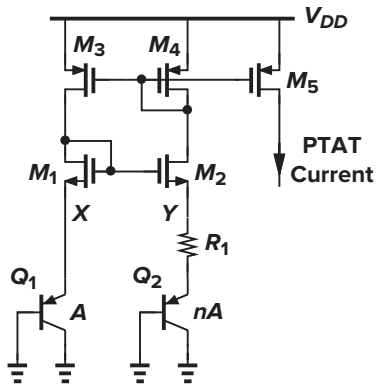


Figure 12.19 Alternative method of generating a PTAT current.

The circuit of Fig. 12.18 can be readily modified to provide a bandgap reference voltage as well. Illustrated in Fig. 12.20, the idea is to add a PTAT voltage $I_{D5}R_2$ to a base-emitter voltage. The output therefore equals

$$V_{REF} = |V_{BE3}| + \frac{R_2}{R_1} V_T \ln n \tag{12.51}$$

[3]The two circuits in Figs. 12.18 and 12.19 exhibit different supply rejections. With a carefully-designed op amp, the former achieves a higher rejection.

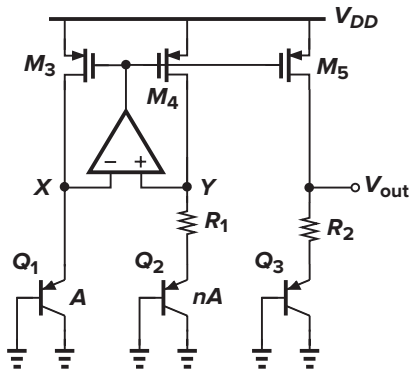


Figure 12.20 Generation of a temperature-independent voltage.

where all of the PMOS transistors are assumed identical. Note that the value of V_{BE3} and hence the size of Q_3 are somewhat arbitrary so long as the sum of the two terms in (12.51) gives a zero TC. In reality, mismatches of the PMOS devices introduce error in V_{out}.

12.5 ■ Constant-G_m Biasing

The transconductance of MOSFETs plays a critical role in analog circuits, determining such performance parameters as noise, small-signal gain, and speed. For this reason, it is often desirable to bias the transistors such that their transconductance does not depend on the temperature, process, or supply voltage.

A simple circuit used to define the transconductance is the supply-independent bias topology of Fig. 12.3. Recall that the bias current is given by

$$I_{out} = \frac{2}{\mu_n C_{ox}(W/L)_N} \frac{1}{R_S^2} \left(1 - \frac{1}{\sqrt{K}}\right)^2 \tag{12.52}$$

Thus, the transconductance of M_1 equals

$$g_{m1} = \sqrt{2\mu_n C_{ox} \left(\frac{W}{L}\right)_N I_{D1}} \tag{12.53}$$

$$= \frac{2}{R_S}\left(1 - \frac{1}{\sqrt{K}}\right) \tag{12.54}$$

a value independent of the supply voltage and MOS device parameters.

In reality, the value of R_S in (12.54) does vary with temperature and process. If the temperature coefficient of the resistor is known, bandgap and PTAT reference generation techniques can be utilized to cancel the temperature dependence. *Process* variations, however, limit the accuracy with which g_{m1} is defined.

In systems where a precise clock frequency is available, the resistor R_S in Fig. 12.3 can be replaced by a switched-capacitor equivalent (Chapter 13) to achieve a somewhat higher accuracy. Depicted in Fig. 12.21, the idea is to establish an average resistance equal to $(C_S f_{CK})^{-1}$ between the source of M_2 and ground, where f_{CK} denotes the clock frequency. Capacitor C_B is added to shunt the high-frequency components resulting from switching to ground. Since the absolute value of capacitors is typically more tightly controlled and since the TC of capacitors is much smaller than that of resistors, this technique provides a higher reproducibility in the bias current and transconductance.

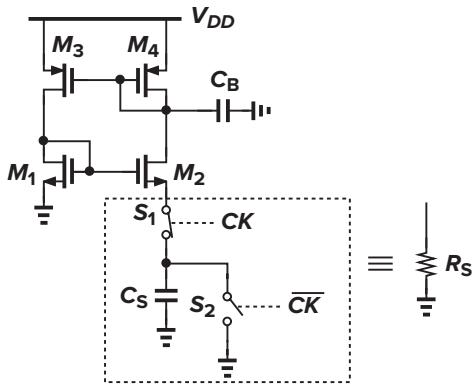


Figure 12.21 Constant-G_m biasing by means of a switched-capacitor "resistor."

The switched-capacitor approach of Fig. 12.21 can be applied to other circuits as well. For example, as shown in Fig. 12.22, a voltage-to-current converter with a relatively high accuracy can be constructed.

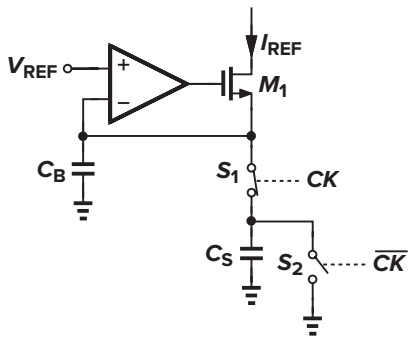


Figure 12.22 Voltage-to-current conversion by means of a switched-capacitor resistor.

12.6 ■ Speed and Noise Issues

Even though reference generators are low-frequency circuits, they may affect the speed of the circuits that they feed. Furthermore, various building blocks may experience "crosstalk" through reference lines. These difficulties arise because of the finite output impedance of reference voltage generators, especially if they incorporate op amps. As an example, let us consider the configuration shown in Fig. 12.23, assuming that the voltage at node N is heavily disturbed by the circuit fed by M_5. For fast changes in V_N, the op amp cannot maintain V_P constant, and the bias currents of M_5 and M_6 experience large transient changes. Also, the duration of the transient at node P may be quite long if the op amp suffers from a slow response. For this reason, many applications may require a high-speed op amp in the reference generator.

In systems where the power consumed by the reference circuit must be small, the use of a high-speed op amp may not be feasible. Alternatively, the critical node, e.g., node P in Fig. 12.23, can be bypassed to ground by means of a large capacitor (C_B) so as to suppress the effect of external disturbances. This approach involves two issues. First, the stability of the op amp must not degrade with the addition of the capacitor, requiring the op amp to be of a one-stage nature (Chapter 10). Second, since C_B generally slows down the transient response of the op amp, its value must be much greater than the capacitance that couples the disturbance to node P. As illustrated in Fig. 12.24, if C_B is not sufficiently large, then V_P

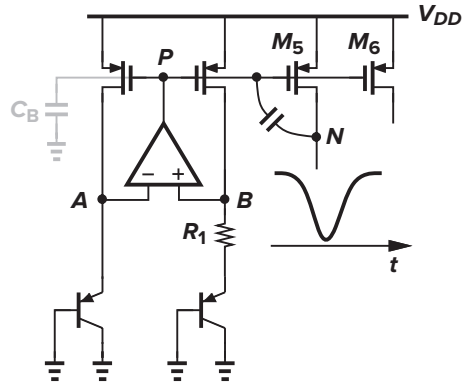


Figure 12.23 Effect of circuit transients on reference voltages and currents.

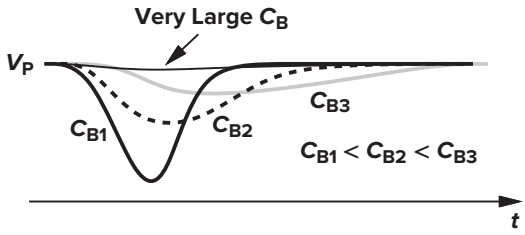


Figure 12.24 Effect of increasing bypass capacitor on the response of a reference generator.

experiences a change and takes a long time to return to its original value, possibly degrading the settling speed of the circuits biased by the reference generator. In other words, depending on the environment, it may be preferable to leave node P agile so that it can quickly recover from transients. In general, as depicted in Fig. 12.25, the response of the circuit must be analyzed by applying a disturbance at the output and observing the settling behavior.

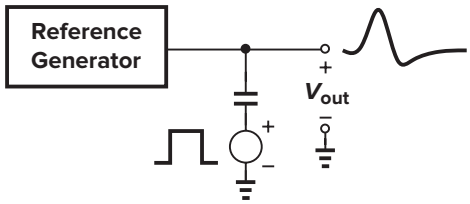


Figure 12.25 Setup for testing the transient response of a reference generator.

▶ Example 12.7

Determine the small-signal output impedance of the bandgap reference shown in Fig. 12.23 and examine its behavior with frequency.

Solution

Figure 12.26 depicts the equivalent circuit, modeling the open-loop op amp by a one-pole transfer function $A(s) = A_0/(1 + s/\omega_0)$ and an output resistance R_{out} and each bipolar transistor by a resistance $1/g_{mN}$. If M_1 and M_2 are identical, each having a transconductance of g_{mP}, then their drain currents are equal to $g_{mP}V_X$, producing a differential voltage at the input of the op amp equal to

$$V_{AB} = -g_{mP}V_X\frac{1}{g_{mN}} + g_{mP}V_X\left(\frac{1}{g_{mN}} + R_1\right) \tag{12.55}$$

$$= g_{mP}V_X R_1 \tag{12.56}$$

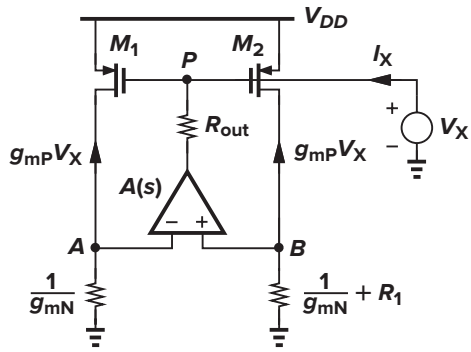


Figure 12.26 Circuit for calculation of the output impedance of a reference generator.

The current flowing through R_{out} is therefore given by

$$I_X = \frac{V_X + g_{mP} V_X R_1 A(s)}{R_{out}} \tag{12.57}$$

yielding

$$\frac{V_X}{I_X} = \frac{R_{out}}{1 + g_{mP} R_1 A(s)} \tag{12.58}$$

$$= \frac{R_{out}}{1 + g_{mP} R_1 \dfrac{A_0}{1 + s/\omega_0}} \tag{12.59}$$

$$= \frac{R_{out}}{1 + g_{mP} R_1 A_0} \frac{1 + \dfrac{s}{\omega_0}}{1 + \dfrac{s}{(1 + g_{mP} R_1 A_0)\omega_0}} \tag{12.60}$$

Thus, the output impedance exhibits a zero at ω_0 and a pole at $(1 + g_{mP} R_1 A_0)\omega_0$, with the magnitude behavior plotted in Fig. 12.27. Note that $|Z_{out}|$ is small for $\omega < \omega_0$, but it rises to a high value as the frequency approaches the pole. In fact, setting $\omega = (1 + g_{mP} R_1 A_0)\omega_0$ and assuming $g_{mP} R_1 A_0 \gg 1$, we have

$$|Z_{out}| = \frac{R_{out}}{1 + g_{mP} R_1 A_0} \left| \frac{1 + j(1 + g_{mP} R_1 A_0)}{1 + j} \right| \tag{12.61}$$

$$= \frac{R_{out}}{\sqrt{2}} \tag{12.62}$$

which is only 30% lower than the open-loop value.

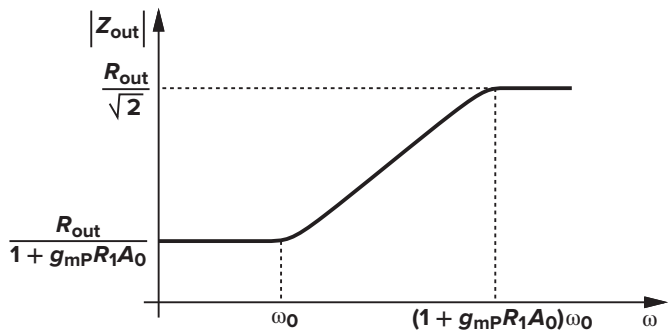


Figure 12.27 Variation of the reference generator output impedance with frequency.

The output noise of reference generators may affect the performance of low-noise circuits considerably. Figure 12.28 illustrates an example: the load current source of a common-source stage is driven by a bandgap circuit with a current multiplication factor of N. Thus, the noise current of M_1 (or M_2) is amplified by the same factor as it appears in M_3. Note that M_1–M_3 carry noise due to the op amp A_1 as well.

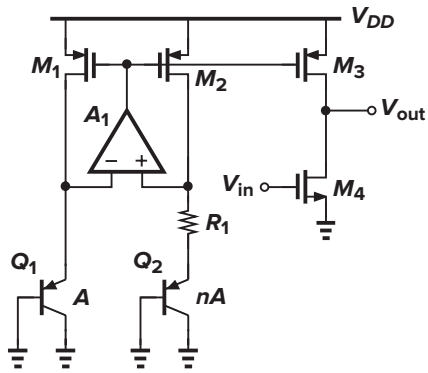


Figure 12.28 Effect of bandgap circuit noise on a CS stage.

As another example, if a high-precision A/D converter employs a bandgap voltage as the reference with which the analog input signal is compared (Fig. 12.29), then the noise in the reference is directly added to the input.

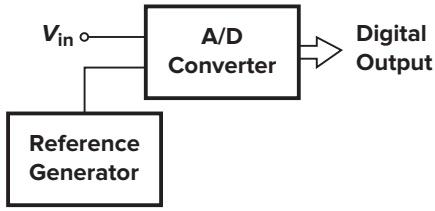


Figure 12.29 A/D converter using a reference generator.

As a simple example, let us calculate the output noise voltage of the circuit shown in Fig. 12.30, taking into account only the input-referred noise voltage of the op amp, $V_{n,op}$. Since the small-signal drain currents of M_1 and M_2 are equal to $V_{n,out}/(R_1 + g_{mN}^{-1})$, we have $V_P = -g_{mP}^{-1} V_{n,out}/(R_1 + g_{mN}^{-1})$, obtaining the differential voltage at the input of the op amp as $-g_{mP}^{-1} A_0^{-1} V_{n,out}/(R_1 + g_{mN}^{-1})$. Beginning

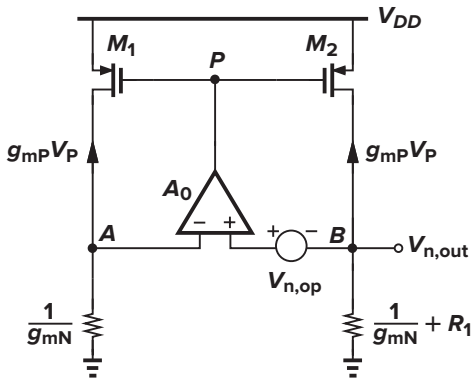


Figure 12.30 Circuit for calculation of noise in a reference generator.

from node A, we can then write

$$\frac{V_{n,out}}{R_1 + g_{mN}^{-1}} \cdot \frac{1}{g_{mN}} - \frac{V_{n,out}}{g_{mP}A_0\left(R_1 + g_{mN}^{-1}\right)} = V_{n,op} + V_{n,out} \tag{12.63}$$

and hence

$$V_{n,out}\left[\frac{1}{R_1 + g_{mN}^{-1}}\left(\frac{1}{g_{mN}} - \frac{1}{g_{mP}A_0}\right) - 1\right] = V_{n,op} \tag{12.64}$$

Since typically $g_{mP}A_0 \gg g_{mN} \gg R_1^{-1}$,

$$|V_{n,out}| \approx V_{n,op} \tag{12.65}$$

suggesting that the noise of the op amp directly appears at the output. Note that even the addition of a large capacitor from the output to ground may not suppress low-frequency $1/f$ noise components, a serious difficulty in low-noise applications. The noise contributed by other devices in the circuit is studied in Problem 12.6.

12.7 ■ Low-Voltage Bandgap References

The bandgap voltage expressed by Eq. (12.20) is around 1.25 V, eluding implementation with today's low supplies. The fundamental limitation is that we must add about $17.2V_T$ to one V_{BE} so as to achieve a net zero temperature coefficient.

Is it possible to add two *currents* with positive and negative TCs and then convert the result to an arbitrary voltage that has a zero TC (Fig. 12.31)? Recall from Fig. 12.18 that we can readily generate a PTAT current given by $V_T \ln n/R$. We also envision another current of the form V_{BE}/R serving as that with a negative TC, but how can we generate such a current with minimal complexity?

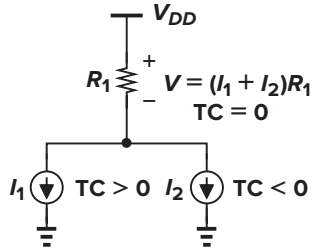


Figure 12.31 Summation of two currents with opposite TCs to obtain a result with zero TC.

Let us return to the circuit of Fig. 12.18, assume that M_3 and M_4 are identical, and note that $|I_{D4}| = V_T \ln n/R_1$ is a PTAT current. We place a resistor in parallel with Q_2 as shown in Fig. 12.32(a). We recognize that R_1 now carries an additional current equal to $|V_{BE2}|/R_2$, i.e., a current with a negative TC. Unfortunately, however, the PTAT behavior is now disturbed because $I_{C1} \neq I_{C2}$. Fortunately, a simple modification resolves this issue: as shown in Fig. 12.32(b), we tie R_2 from Y to ground and place another resistor in parallel with Q_1. Proposed by Banba et al. [8], this topology lends itself to low-voltage implementation, requiring a minimum V_{DD} of $V_{BE1} + |V_{DS3}|$.

To analyze the circuit, we observe that $V_X \approx V_Y \approx |V_{BE1}|$ and $I_{D3} = I_{D4}$. Thus,

$$I_{C1} + \frac{|V_{BE1}|}{R_3} = I_{C2} + \frac{|V_{BE1}|}{R_2} \tag{12.66}$$

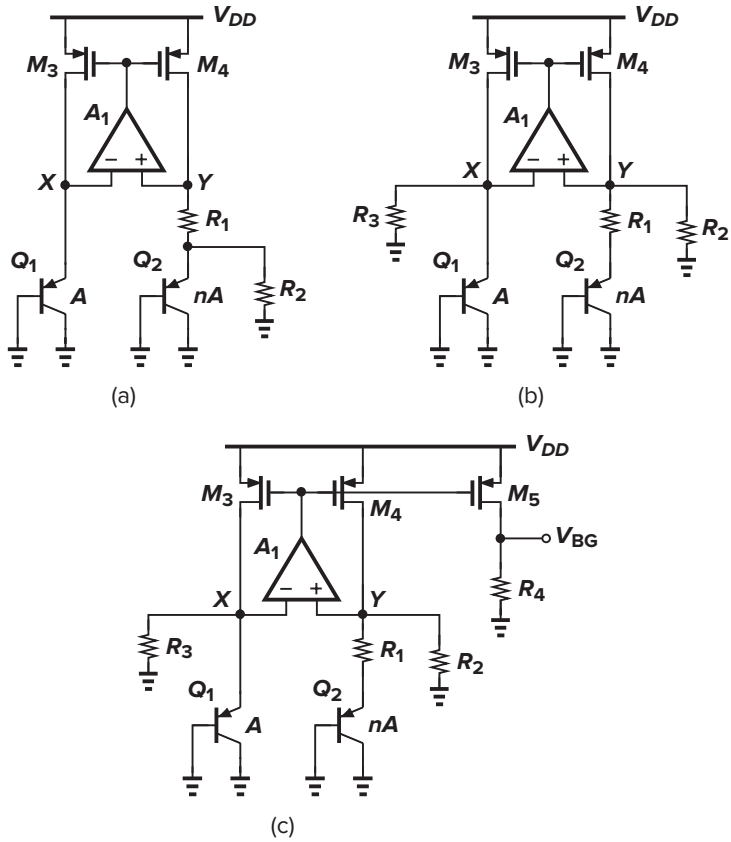


Figure 12.32 (a) Attempt to make drain current of M_4 temperature-independent, (b) circuit modification resulting in a zero-TC current, and (c) generation of arbitrarily small voltage with zero TC.

which yields $I_{C1} = I_{C2}$ if $R_2 = R_3$. We still have $|V_{BE1}| = |V_{BE2}| + I_{C2}R_1$ and hence $I_{C2} = V_T \ln n/R_1$. This current and the current flowing through R_2, $|V_{BE1}|/R_2$, constitute $|I_{D4}|$:

$$|I_{D4}| = \frac{V_T \ln n}{R_1} + \frac{|V_{BE1}|}{R_2} \tag{12.67}$$

$$= \frac{1}{R_2}\left(|V_{BE1}| + \frac{R_2}{R_1} V_T \ln n\right) \tag{12.68}$$

Selecting $(R_2/R_1)V_T \ln n$ approximately equal to $17.2V_T$ renders a zero TC for I_{D4}. This current is then copied and passed through a resistor to generate a zero-TC voltage [Fig. 12.32(c)] [8]:

$$V_{BG} = \frac{R_4}{R_2}\left(|V_{BE1}| + \frac{R_2}{R_1} V_T \ln n\right) \tag{12.69}$$

(if M_5 is identical to M_4). We choose $(R_2/R_1)\ln n \approx 17.2$, observing that V_{BG} has a zero TC and its value can be lower than the conventional limit of 1.25 V.

▶ Example 12.8

If the op amp in Fig. 12.32(c) has an input-referred offset voltage, V_{OS}, determine V_{BG}.

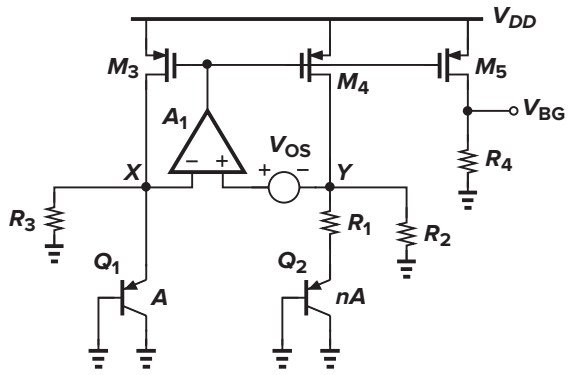


Figure 12.33

Solution

As shown in Fig. 12.33, we now have $V_X \approx V_Y + V_{OS} \approx |V_{BE1}|$ and

$$I_{C1} + \frac{|V_{BE1}|}{R_3} = I_{C2} + \frac{|V_{BE1}| - V_{OS}}{R_2} \tag{12.70}$$

which implies that $I_{C1} = I_{C2} - V_{OS}/R_2$ if $R_2 = R_3$. Since $|V_{BE1}| = |V_{BE2}| + R_1 I_{C2} + V_{OS}$, we have $I_{C2} = (V_T \ln n - V_{OS})/R_1$. This current and the current flowing through R_2, $(|V_{BE1}| - V_{OS})/R_2$, add up to $|I_{D4}|$:

$$|I_{D4}| = \frac{V_T \ln n - V_{OS}}{R_1} + \frac{|V_{BE1}| - V_{OS}}{R_2} \tag{12.71}$$

It follows that

$$V_{BG} = \frac{R_4}{R_2}\left(|V_{BE1}| + \frac{R_2}{R_1} V_T \ln n\right) - \frac{R_4}{R_1 || R_2} V_{OS} \tag{12.72}$$

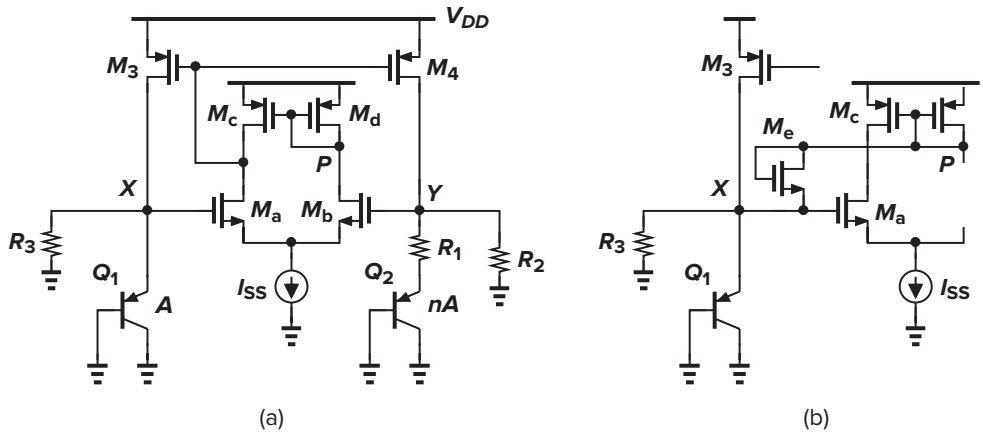
revealing that the op amp offset is amplified by a factor of $R_4/(R_1||R_2)$. Alternatively, we can write

$$V_{BG} = \frac{R_4}{R_2}\left[|V_{BE1}| + \frac{R_2}{R_1} V_T \ln n - \left(1 + \frac{R_2}{R_1}\right) V_{OS}\right] \tag{12.73}$$

concluding that the effect of V_{OS} can be minimized only by maximizing n. ◀

It is instructive to estimate the lowest supply voltage with which the circuit of Fig. 12.32(c) can operate properly. With large bipolar transistors and a small bias current, e.g., 10 μA, the base-emitter voltage can be as low as 0.7 V. Similarly, wide PMOS devices allow a $|V_{DS}|$ of about 50 mV. The circuit can thus operate with a minimum V_{DD} of around 0.75 V. In this case, R_4 tends to be a large resistor, e.g., 50 kΩ, producing significant noise and requiring a bypass capacitor at the output. Also, if the PMOS drain currents are copied to generate a larger current, say, 0.5 mA, then their noise is amplified by the same factor. This noise contains thermal and flicker components due to the PMOS devices and the noise of the op amp. In Problem 12.24, we analyze the noise behavior of this circuit, but from Example 12.8, we observe that the op amp input noise is amplified by a factor of $R_4/(R_1||R_2)$.

The op amp in Fig. 12.32(c) can be realized as a five-transistor OTA. Depicted in Fig. 12.34(a) is an example. The OTA design proceeds according to the following guidelines. (1) Large transistor dimensions are chosen so as to minimize their flicker noise and offset. (2) The gate-source voltage of M_a and M_b



(a) (b)

Figure 12.34 (a) Implementation of low-voltage BG circuit using a five-transistor OTA, and (b) addition of start-up device.

plus the headroom required by I_{SS} must not exceed $|V_{BE1}|$. (3) The transistors are chosen long enough to yield a reasonable loop gain, e.g., 5 to 10.

The foregoing topology must incorporate a start-up mechanism. Otherwise, the circuit begins with $V_X = V_Y = 0$, M_a and M_b remain off, and so do M_3 and M_4. Since, with $V_{DD} < 1$ V, the voltage difference between node P and node X is initially positive but finally negative (why?), we can tie a diode-connected NMOS transistor between these two nodes to ensure start-up [Fig. 12.34(b)]. Alternatively, the NMOS device can be connected between X and V_{DD}.

Another low-voltage bandgap circuit can be derived from the topology of Fig. 12.20 by simply tying a resistor from the output node to ground [9]. Shown in Fig. 12.35, the circuit now allows some of I_{D5} to flow through R_3:

$$|I_{D5}| = \frac{V_{out}}{R_3} + \frac{V_{out} - |V_{BE3}|}{R_2} \tag{12.74}$$

If the PMOS devices are identical, $|I_{D5}| = V_T \ln n / R_1$, yielding

$$V_{out} = \frac{R_3}{R_2 + R_3}\left(|V_{BE3}| + \frac{R_2}{R_1} V_T \ln n\right) \tag{12.75}$$

The standard bandgap voltage is thus scaled down by a factor of $R_3/(R_2 + R_3)$. The reader is encouraged to compute the effect of the op amp offset at the output and compare the result with (12.72).

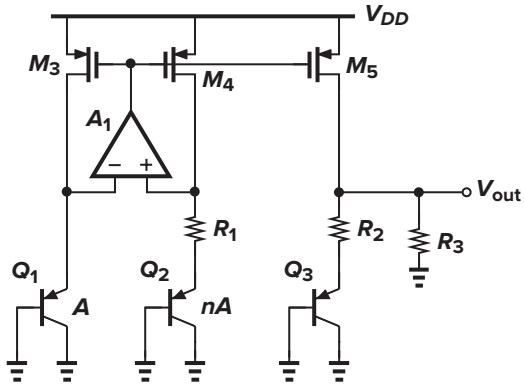


Figure 12.35 Alternative low-voltage BG circuit.

It is possible to add other bias branches to the foregoing circuits so as to provide curvature correction, but such schemes typically rely on trimming because the various mismatches within the circuit tend to shift the zero-TC temperature randomly. Other low-voltage bandgaps are described in [10].

12.8 ■ Case Study

In this section, we study a bandgap reference circuit designed for high-precision analog systems [7]. The reference generator incorporates the topology of Fig. 12.19, but with two series base-emitter voltages in each branch so as to reduce the effect of MOSFET mismatches. A simplified version of the core is depicted in Fig. 12.36, where the PMOS current mirror arrangement ensures equal collector currents for Q_1–Q_4. While requiring a high supply voltage, this design exemplifies issues that prove important in practice.

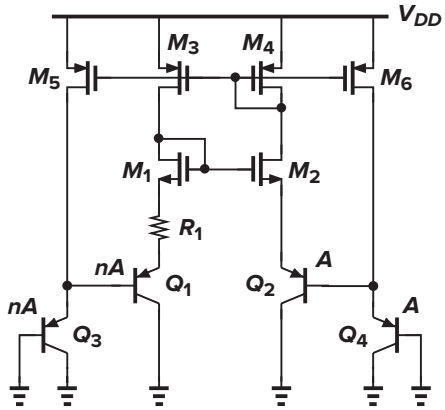


Figure 12.36 Simplified core of the bandgap circuit reported in [7].

Channel-length modulation of the MOS devices in Fig. 12.36 still results in significant supply dependence. To resolve this issue, each branch can employ both NMOS and PMOS cascode topologies. Figure 12.37(a) shows an example in which the low-voltage cascode current mirror described in Chapter 5

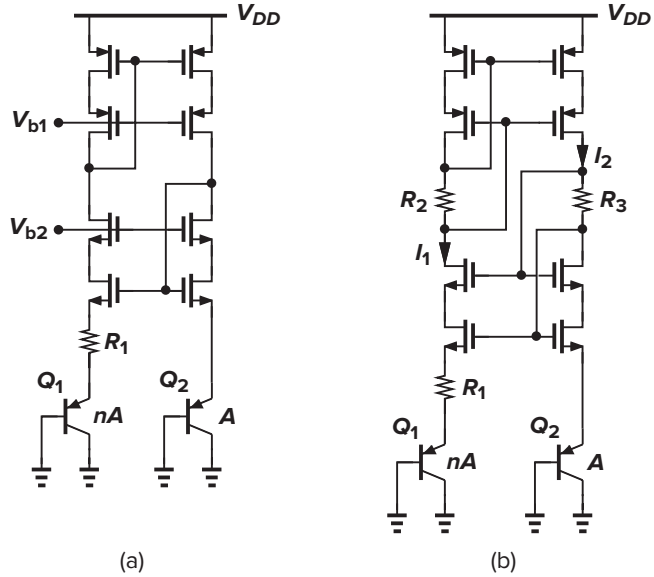


Figure 12.37 (a) Addition of cascode devices to improve supply rejection; (b) use of self-biased cascode to eliminate V_{b1} and V_{b2}.

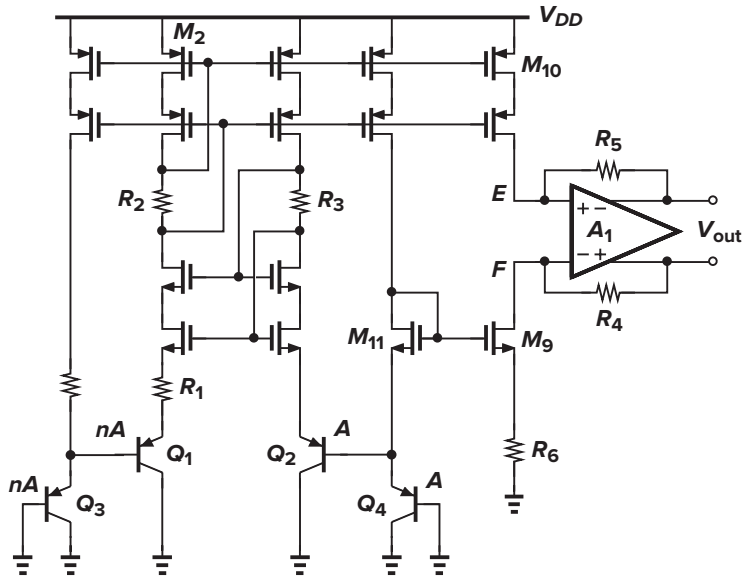


Figure 12.38 Generation of a floating reference voltage.

is utilized. To obviate the need for V_{b1} and V_{b2}, this design actually introduces a "self-biased" cascode, shown in Fig. 12.37(b), where R_2 and R_3 sustain proper voltages to allow all MOSFETs to remain in saturation. This cascode topology is analyzed in Problem 12.7.

The bandgap circuit reported in [7] is designed to generate a *floating* reference. This is accomplished by the modification shown in Fig. 12.38, where the drain currents of M_9 and M_{10} flow through R_4 and R_5, respectively. Note that M_{11} sets the gate voltage of M_9 at $V_{BE4} + V_{GS11}$, establishing a voltage equal to V_{BE4} across R_6 if M_9 and M_{11} are identical. Thus, $I_{D9} = V_{BE4}/R_6$, yielding $V_{R4} = V_{BE4}(R_4/R_6)$. Also, if M_{10} is identical to M_2, then $|I_{D10}| = 2(V_T \ln n)/R_1$, and hence $V_{R5} = 2(V_T \ln n)(R_5/R_1)$. Since the op amp ensures that $V_E \approx V_F$, we have

$$V_{out} = \frac{R_4}{R_6} V_{BE4} + 2\frac{R_5}{R_1} V_T \ln n \tag{12.76}$$

Proper choice of the resistor ratios and n therefore provides a zero temperature coefficient.

In order to further enhance the supply rejection, this design regulates the supply voltage of the core and the op amp. Illustrated in Fig. 12.39, the idea is to generate a local supply, V_{DDL}, that is defined by a reference V_{R1} and the ratio of R_{r1} and R_{r2} and hence remains relatively independent of the global supply voltage. But how is V_{R1} itself generated? To minimize the dependence of V_{R1} upon the supply,

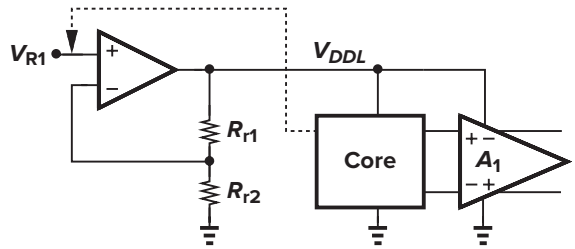


Figure 12.39 Regulation of the supply voltage of the core and op amp to improve supply rejection.

this voltage is established *inside* the core, as depicted in Fig. 12.40. In fact, R_M is chosen such that V_{R1} is a bandgap reference.

Figure 12.41 shows the overall implementation, omitting a few details for simplicity. A start-up circuit is also used. Operating from a 5-V supply, the reference generator produces a 2.00-V output while consuming 2.2 mW. The supply rejection is 94 dB at low frequencies, dropping to 58 dB at 100 kHz [7].

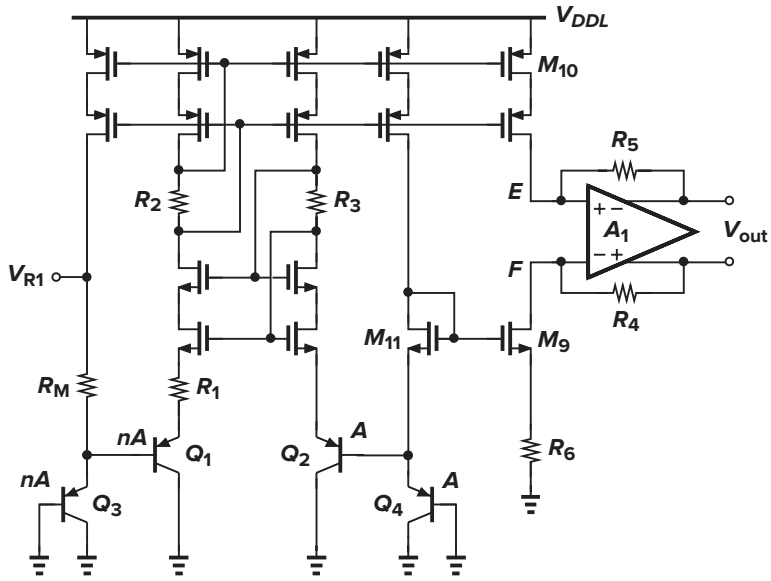


Figure 12.40 Generation of V_{R1}, used in Fig. 12.39.

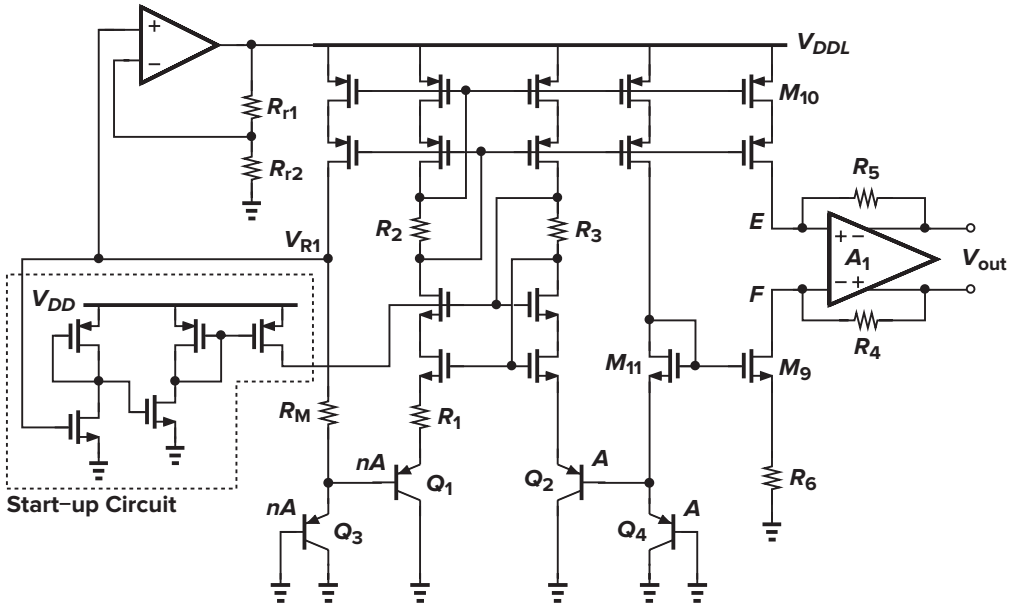


Figure 12.41 Overall circuit of the bandgap generator reported in [7].

References

[1] R. A. Blauschild et al., "A New NMOS Temperature-Stable Voltage Reference," *IEEE J. of Solid-State Circuits,* vol. 13, pp. 767–774, December 1978.

[2] Y. P. Tsividis and R. W. Ulmer, "A CMOS Voltage Reference," *IEEE J. of Solid-State Circuits*, vol. 13, pp. 774–778, December 1978.

[3] D. Hilbiber, "A New Semiconductor Voltage Standard," *ISSCC Dig. of Tech. Papers*, pp. 32–33, February 1964.

[4] K. E. Kujik, "A Precision Reference Voltage Source," *IEEE J. of Solid-State Circuits*, vol. 8, pp. 222–226, June 1973.

[5] G. C. M. Meijer, P. C. Schmall, and K. van Zalinge, "A New Curvature-Corrected Bandgap Reference," *IEEE J. of Solid-State Circuits*, vol. 17, pp. 1139–1143, December 1982.

[6] M. Gunawan et al., "A Curvature-Corrected Low-Voltage Bandgap Reference," *IEEE J. of Solid-State Circuits*, vol. 28, pp. 667–670, June 1993.

[7] T. Brooks and A. L. Westwisk, "A Low-Power Differential CMOS Bandgap Reference," *ISSCC Dig. of Tech. Papers,* pp. 248–249, February 1994.

[8] H. Banba et al., "A CMOS Bandgap Reference Circuit with Sub-1-V Operation," *IEEE J. of Solid-State Circuits*, vol. 34, pp. 670–674, May 1999.

[9] H. Neuteboom et al., "A DSP-Based Hearing Instrument IC," *IEEE J. of Solid-State Circuits*, vol. 32, pp. 1790–1806, November 1997.

[10] C. J. B. Fayomi et al., "Sub-1-V CMOS Bandgap Reference Design Techniques: A Survey," *Analog Integrated Circuits and Signal Processing*, vol. 62, pp. 141–157, February 2010.

[11] B. Gilbert, "Monolithic Voltage and Current References: Themes and Variations," pp. 269–352 in *Analog Circuit Design*, J. H. Huijsing, R. J. van de Plassche, and W. M. C. Sansen, eds. (Boston: Kluwer Academic Publishers, 1996).

Problems

Unless otherwise stated, in the following problems, use the device data shown in Table 2.1 and assume that $V_{DD} = 3\,\text{V}$ where necessary.

12.1. Derive an expression for I_{out} in Fig. 12.42.

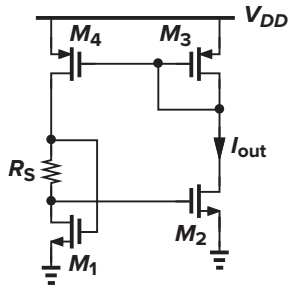


Figure 12.42

12.2. Explain how the start-up circuit shown in Fig. 12.43 operates. Derive a relationship that guarantees that $V_X < V_{TH}$ after the circuit turns on.

12.3. Consider the circuit of Fig. 12.15.

(a) If M_1 and M_2 suffer from channel-length modulation, what is the error in the output voltage?

(b) Repeat part (a) for M_3 and M_4.

(c) If M_1 and M_2 have a threshold mismatch of ΔV, i.e., $V_{TH1} = V_{TH}$ and $V_{TH2} = V_{TH} + \Delta V$, what is the error in the output voltage?

(d) Repeat part (c) for M_3 and M_4.

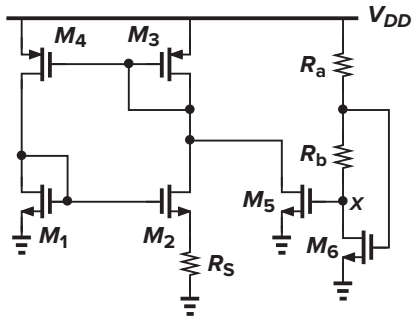


Figure 12.43

12.4. In Fig. 12.15, if the open-loop gain of the op amp A_1 is not sufficiently large, then $|V_X - V_Y|$ exceeds V_e, where V_e is the maximum tolerable error. Calculate the minimum value of A_1 in terms of V_e such that the condition $|V_X - V_Y| < V_e$ is satisfied.

12.5. In the circuit of Fig. 12.15, assume that Q_2 and Q_4 have a finite current gain β. Calculate the error in the output voltage.

12.6. Calculate the output noise voltage of the circuit shown in Fig. 12.30 due to the thermal and flicker noise of M_1 and M_2.

12.7. Consider the self-biased cascode shown in Fig. 12.44. Determine the minimum and maximum values of RI_{REF} such that both M_1 and M_2 remain in saturation.

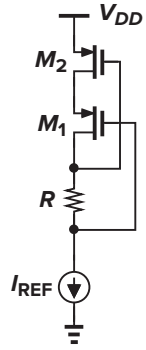


Figure 12.44

12.8. The circuit of Fig. 12.3(a) sometimes turns on even with no explicit start-up mechanism. Identify the capacitive path(s) that couple the transition on V_{DD} to the internal nodes and hence provide the start-up current.

12.9. Sketch the temperature coefficient of V_{BE} [Eq. (12.13)] versus temperature. Some iteration may be necessary.

12.10. Determine the derivative of Eq. (12.13) with respect to temperature and sketch the result versus T. This quantity reveals the curvature of the voltage.

12.11. Suppose that in Fig. 12.9, the amplifier has an output resistance R_{out}. Calculate the error in V_{out}.

12.12. The circuit of Fig. 12.9 is designed with $R_3 = 1$ kΩ and a current of 50 μA through it. Calculate $R_1 = R_2$ and n for a zero TC.

12.13. In the circuit of Fig. 12.15, Q_1 and Q_2 are biased at 100 μA and Q_3 and Q_4 at 50 μA. If $R_1 = 1$ kΩ, calculate R_2 and $(W/L)_{1-4}$ such that the circuit operates with $V_{DD} = 3$ V. Which op amp topology can be used here?

12.14. Since the bandgap of silicon exhibits a small temperature coefficient, Eq. (12.48) suggests that $\partial V_{REF}/\partial T \propto (4 + m)k/q$, a relatively large value, whereas we derived V_{REF} such that it has a zero TC. Explain the flaw in this argument.

12.15. A differential pair with resistive loads is designed such that its voltage gain, $g_m R_D$, has a zero TC at room temperature. If only the temperature dependence of the mobility is considered, determine the required temperature behavior of the tail current. Design a circuit that roughly approximates this behavior.

12.16. In Problem 12.15, assume that the tail current is constant, but the load resistors exhibit a finite TC. What resistor temperature coefficient cancels the variation of the mobility at room temperature?

12.17. In the circuit of Fig. 12.32(b), how should R_1–R_3 be chosen so that the negative-feedback loop is stronger than the positive-feedback loop?

12.18. Does the five-transistor OTA in Fig. 12.34(a) impose additional supply voltage constraints?

12.19. Figure 12.45 illustrates a "single-junction" bandgap design [11]. Here, switches S_1 and S_2 are driven by complementary clocks.

(a) What is V_{out} when S_1 is on and S_2 is off?

(b) What is the change in V_{out} when S_1 turns off and S_2 turns on?

(c) How are I_1, I_2, C_1, and C_2 chosen to produce a zero-TC output when S_1 is off?

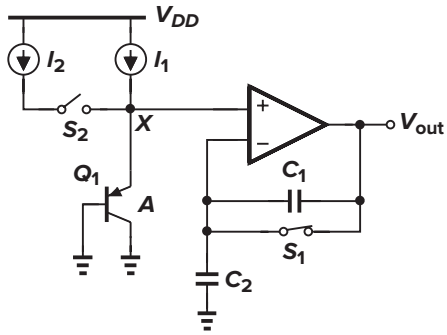


Figure 12.45

12.20. Suppose that in Fig. 12.45, I_2/I_1 deviates from its nominal value by a small error ϵ. Calculate V_{out} when S_1 is off.

12.21. The circuit of Fig. 12.20 is designed with $(W/L)_{1-4} = 50/0.5$, $I_{D1} = I_{D2} = 50\ \mu$A, $R_1 = 1$ kΩ, and $R_2 = 2$ kΩ. Assume that $\lambda = \gamma = 0$ and Q_3 is identical to Q_1.

(a) Determine n and $(W/L)_5$ such that V_{out} has a zero TC at room temperature.

(b) Neglecting the noise contribution of Q_1–Q_3, calculate the output thermal noise.

12.22. Consider the circuit of Fig. 12.21. Assume $K = 4$, $f_{CK} = 50$ MHz, and a power budget of 1 mW. Determine the aspect ratio of M_1–M_4 and the value of C_S such that $g_{m1} = 1/(500\ \Omega)$.

12.23. Suppose $(W/L)_3 = K(W/L)_4$ in Fig. 12.32(c). How should R_2 and R_3 be chosen?

12.24. Determine the output noise voltage of the circuit in Fig. 12.32(c).

12.25. Analyze the circuit of Fig. 12.3(a) if R_S is placed in series with the source of M_1.