Overview

- Reading
  B. Razavi Chapter 11.

- Introduction
  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. In this lecture, we deal with the design of reference generators in CMOS technology, focusing on well-established “bandgap” techniques.

  In most applications, the required temperature dependence assumes one of three forms:
  
  (1) proportional to absolute temperature (PTAT);

  (2) constant-$G_m$ behavior;

  (3) temperature independent.

  In addition, several parameters of reference generators, such as output impedance, output noise, and power dissipation, may be critical as well.
Supply-independent biasing

- Current-mirror biasing using (a) an ideal current source, (b) a resistor.

  \[
  \Delta I_{out} = \frac{\Delta V_{DD}}{R_1 + 1/g_{m1}} \frac{(W/L)_2}{(W/L)_1}
  \]

  How do we generate \( I_{REF} \) independent of the supply voltage?

- Simple circuit to establish supply-independent currents.

  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 \( I_{out} \).

  If \( M_1-M_4 \) operate in saturation and \( \lambda = 0 \), then \( I_{out} = K I_{REF} \), and hence can support any current level.

Supply-independent biasing

- Addition of \( R_S \) to define the currents

  Assuming \( \lambda = 0 \), then \( I_{out} = I_{REF} \) and \( V_{GS1} = V_{GS2} + I_{D2} R_S \)

  Neglecting body effect, we have

  \[
  I_{out} = \frac{2L_{out}}{\mu_n C_{ox}(W/L)_n} \left( 1 - \frac{1}{\sqrt{K}} \right) \sqrt{V_{TH2} + I_{out} R_S}
  \]

  That is, \( I_{out} = \frac{2}{\mu_n C_{ox}(W/L)_n} \cdot \frac{1}{R_S} \left( 1 - \frac{1}{\sqrt{K}} \right)^2 \)

  The current is independent of the supply voltage (but still a function of process and temperature).
Supply-independent biasing (cont’d)

- Addition of $R_S$ to define the currents (assuming $\lambda \neq 0$). Determine $\Delta I_{out} / \Delta V_{DD}$.

$$R_1 = r_o \parallel (1/g_{m1}), \quad R_3 = r_o \parallel (1/g_{m3})$$

$$\frac{V_{DD} - V_X}{r_o} + I_{out} R_3 g_{m4} = \frac{V_X}{R_1}$$

The equivalent transconductance of $M_2$ and $R_S$ is $G_{m2} = \frac{I_{out}}{V_X} = \frac{g_{m2} r_o}{R_3 + r_o + (g_{m2} + g_{m3}) R_o r_o}$

Thus, $\frac{I_{out}}{V_{DD}} = \frac{1}{r_o} \left[ \frac{1}{G_{m2} (r_o + R_1)} - g_{m4} R_3 \right]^{-1} \rightarrow 0$, if $r_o = \infty$.

---

Supply-independent biasing (cont’d)

- Addition of $R_S$ to define the currents

An important issue in supply-independent biasing is the existence of “degenerate” bias points. For example, if all 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.

In other words, the circuit can settle in one of two different operating conditions.

- Addition of start-up device

  - The diode-connected device $M_5$ provides a current path from $V_{DD}$ through $M_3$ and $M_1$ to ground upon start-up.

  This technique is practical on if $V_{TH1} + V_{THS} + |V_{THS}| < V_{DD}$ and $V_{GS1} + V_{THS} + |V_{GS1}| > V_{DD}$, the latter to ensure $M_5$ remains off after start-up.
Temperature-independent reference

- How to generate a quantity that remains constant with temperature?
  - If two quantities having opposite temperature coefficients (TCs) are added with proper weighting, the result displays a zero TC.
  - Ex. $V_{\text{REF}} = \alpha_1 V_1 + \alpha_2 V_2$, with zero TC, i.e., $\alpha_1 \partial V_1/\partial T + \alpha_2 \partial V_2/\partial T = 0$.

Negative-TC voltage

- For a bipolar device, where $I_C = I_S \exp(V_{BE} / V_T)$, where $V_T = kT / q$.
  The saturation current $I_S$ is proportional to $\mu kT n_i^2$, where $\mu$ denotes the mobility of minority carries and $n_i$ is the intrinsic minority carrier concentration of silicon.

- Temperature dependence: $\mu \propto \mu_0 T_m$, and $n_i^2 \propto T^3 \exp[-E_g/(kT)]$, where $m \approx -3/2$ and $E_g \approx 1.12\text{eV}$ is the bandgap energy of silicon.

- Find TC of $V_{BE}$: since $V_{BE} = V_T \ln(I_C / I_S)$ and $I_s = b T^{4+\mu} \exp(-E_s / kT)$, then

$$\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \ln \frac{I_C}{I_S} \cdot \frac{\partial I_S}{\partial T} + \frac{V_{BE} - (4 + m) V_T - E_g / q}{T}$$

$$\Rightarrow \text{With } V_{BE} \approx 750\text{mV and } T = 300^\circ\text{K, } \partial V_{BE} / \partial T \approx -1.5\text{mV/^\circ\text{K}}.$$

- 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.
Positive-TC voltage

- Generation of PTAT voltage

\[ \Delta V_{BE} = V_{BE1} - V_{BE2} = V_T \ln \frac{nI_0}{I_{S1}} - V_T \ln \frac{I_0}{I_{S2}} = V_T \ln n \]

Thus, \( \Delta V_{BE} \) exhibits a positive temperature coefficient:

\[ \frac{\partial \Delta V_{BE}}{\partial T} = \frac{k}{q} \ln n \]

- Another type

\[ \frac{\partial \Delta V_{BE}}{\partial T} = \frac{k}{q} \ln (mn) \]

The TC is independent of temperature or behavior of collector currents.

Bandgap reference

- Develop a reference having a nominally zero TC:

Let \( V_{REF} = \alpha_1 V_{BE1} + \alpha_2 (V_T \ln n) \), where \( V_T \ln n \) is the difference between the base-emitter voltage of the two bipolar transistors operating at different current densities.

\[ \alpha_1 = 1, \alpha_2 \ln n \approx 17.2, \text{indicating that for zero TC: } V_{REF} \approx V_{BE1} + 1.72 V_T = 1.25V. \]

Conceptual generation of temperature-independent voltage

Two modifications:

- A mechanism must be added to guarantee \( V_{O1} = V_{O2} \).
- Since \( \ln n = 1.72 \) translates to a prohibitively large \( n \), the term \( RI = V_T \ln n \) must be scaled up by a reasonable factor.
Actual implementation of bandgap reference

\[
V_{out} = V_{BE2} + \frac{\Delta V_{BE}}{R_3} (R_3 + R_2) = V_{BE2} + \left( V_T \ln n \right) \left( 1 + \frac{R_2}{R_3} \right)
\]

where \( \Delta V_{BE} = V_{BE1} - V_{BE2} = V_T \ln n \).

- For a zero TC, we 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 these results do not depend on the TC of the resistors.

Collector current variation

- What happens to the temperature coefficient of \( V_{BE} \) if the collector currents are PTAT?

\[
\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \ln \left( \frac{I_c}{I_S} \right) + V_T \left( \frac{1}{I_c} \frac{\partial I_c}{\partial T} + \frac{1}{I_S} \frac{\partial I_S}{\partial T} \right)
\]

Since \( I_{C1} = I_{C2} = (V_T \ln n) / R_3 \Rightarrow \partial I_c / \partial T = (V_T \ln n) / (R_3 T) = I_c / T \),

we have

\[
\frac{\partial V_{BE}}{\partial T} = \frac{\partial V_T}{\partial T} \ln \left( \frac{I_c}{I_S} \right) + \frac{V_T}{T} \frac{\partial I_c}{\partial T} - \frac{V_T}{T} \frac{\partial I_S}{\partial T}
\]

\[
= \frac{V_{BE} - (3 + m) V_T - E_S / q}{T}
\]

indicating that the TC is slightly less negative than \(-1.5 \text{mV/oK} \) at \( T = 300^\circ \text{K} \).
Compatibility with CMOS technology

- Realization of a $pnp$ bipolar transistor in CMOS technology.

![Image of pnp transistor]

The $p$-type substrate acts as the collector and it is inevitably connected to the most negative supply (usually ground).

- Circuit implemented with $pnp$ transistors

![Image of circuit diagram]

OP amp offset and output impedance

- Effect of op amp offset on the reference voltage

- If $A_1$ is large, $V_{BE1} - V_{OS} \approx V_{BE2} + R_3 I_{C2}$ and

  $$V_{out} = V_{BE2} + (R_3 + R_2)I_{C2}.$$ Thus,

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

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

- The key point is that $V_{OS}$ is amplified by $1 + R_2/R_3$, introducing error in $V_{out}$. More importantly, $V_{OS}$ itself varies with temperature, raising the temperature coefficient of the output voltage.
Reduction of the effect of op amp offset

- $R_1$ and $R_2$ are ratioed by a factor of $m$, producing $I_1 \approx mI_2$.

- Neglecting base currents and assuming $A_1$ is large, we have $V_{BE1} + V_{BE2} - V_{OS} = V_{BE3} + V_{BE4} + R_3I_2$ and $V_{out} = V_{BE3} + V_{BE4} + (R_3 + R_2)I_2$. It follows that

$$V_{out} = V_{BE3} + V_{BE4} + (R_3 + R_2) \frac{2V_t \ln(mn) - V_{OS}}{R_3} = 2V_t + \left(1 + \frac{R_2}{R_1}\right) \left[2V_t \ln(mn) - V_{OS}\right]$$

- The effect of the offset voltage is reduced by increasing the first term in the square brackets.

- The implementation is not feasible in a standard CMOS technology because the collectors of $Q_2$ and $Q_4$ are not grounded. We modify the series combination of the diodes as illustrated in Fig. (a), converting to one of the diodes to an emitter follower.

Reduction of the effect of op amp offset (cont’d)

- Reference generator incorporating two series base-emitter voltage

  Discussion:

  - **Advantage:**
    - The op amp experiences no resistive loading.

  - **Disadvantage:**
    - The mismatch and channel-length modulation of the PMOS devices introduce error at the output.
    - Since $Q_2$ and $Q_4$ have a finite current gain $\beta$, they generate an error in the emitter currents of $Q_1$ and $Q_3$ and introduce error at the output.
Feedback polarity

- The negative feedback factor is given by $\beta_N = \frac{1}{g_{m2} + R_3 + R_2}$
- The positive feedback factor is given by $\beta_P = \frac{1}{g_{m1} + R_1}$
- To ensure an overall negative feedback, $\beta_P$ must be less than $\beta_N$, preferably by roughly a factor of two so that the transient response remains well-behaved with large capacitive loads.

Bandgap reference

- Bandgap reference $V_{REF} = V_{BE} + V_T \ln n$, then $\frac{\partial V_{REF}}{\partial T} = \frac{\partial V_{BE}}{\partial T} + \frac{V_T}{T} \ln n$

Setting $\frac{\partial V_{REF}}{\partial T} = 0$ and $\frac{\partial V_{BE}}{\partial T} = \frac{V_{BE} - (4 + m)V_T - E_g/q}{T}$, we have $\frac{V_{BE} - (4 + m)V_T - E_g/q}{T} = -\frac{V_T}{T} \ln n$

Thus, we obtain $V_{REF} = \frac{E_g}{q} + (4 + m)V_T$

- 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 $T \to 0$, $V_{REF} \to E_g/q$. 
Supply dependence and start-up

- 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.
- The supply rejection of the circuit typically degrades at high frequencies owing to the op amp’s rejection properties, often mandating “supply regulation.”

Curvature correction

- Curvature in temperature dependence of a bandgap voltage
  - Bandgap voltages exhibit a finite “curvature,” i.e., their TC is typically zero at one temperature and positive or negative at other temperatures.
  - The curvature arises from temperature variation of base-emitter voltages, collector currents, and offset voltages.
- Variation of the zero-TC temperature for differences samples
  - Many curvature correction techniques have been devised to suppress the variation of $V_{REF}$ 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 (right Figure), making it difficult to correct the curvature reliably.
PTAT current generation

- Generation of a PTAT current
- Generation of a PTAT current using a simple amplifier

- $M_1 = M_2, M_3 = M_4 \rightarrow I_{D1} = I_{D2} \rightarrow V_X = V_Y$
- $I_{D1} = I_{D2} = \frac{V_T \ln n}{R_1}$
- In practice, due to mismatches between the transistors and the TC of $R_1$, the variation of $I_{DS}$ deviates from the ideal equation.

PTAT current generation (cont’d)

- Generation of a temperature-independent voltage

- $M1 = M2, M3 = M4 = M5$, the output equals

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

  If $\frac{\partial}{\partial T} V_{REF} = 0$, we can find the required reference value.

- In reality, mismatches of the PMOS devices introduce error in $V_{out}$. 
Constant-$G_m$ biasing

- Supply-independent bias

- It is often desirable to bias the transistors such that their transconductance does not depend on the temperature, process, or supply voltage.

- Supply-independent bias circuit: $I_{in} = \frac{2 \mu_n C_{ox} (W / L)_N}{R_s} \left(1 - \frac{1}{\sqrt{K}}\right)^2$

  The transconductance of $M_1$ equals $g_{m1} = 2 \mu_n C_{ox} \left(\frac{W}{L}\right)_N \frac{I_{D1}}{R_s} \left(1 - \frac{1}{\sqrt{K}}\right)$ independent of the supply voltage and MOS device parameters.

- In reality, the value of $R_s$ does vary with temperature and process.

Constant-$G_m$ biasing (cont’d)

- Constant-$G_m$ biasing by means of a switched-capacitor resistor.

  - Switched-capacitor resistor $R_s = \frac{1}{C_{ox} f_{CK}}$

  - 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.

  - Voltage-to-current conversion by means of a switched-capacitor resistor.
Speed issue

- Effect of circuit transients on reference voltages and currents
  - 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.
  - The critical node $P$ 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:

- The stability of op amp must not degrade with the addition of $C_B$, requiring the op amp to be of one-stage nature.
- Since $C_B$ generally slow down the transient response of the op amp, its value must be much greater than the capacitance that couples the disturbance to node $P$.

Speed issue (cont’d)

- Effect of increasing bypass capacitor on the response of a reference generator

- Setup for testing the transient response of a reference generator
Noise issue

- A/D converter using a reference generator
  - If a high-precision A/D converter employs a bandgap voltage as the reference with which the analog input signal is compared, then the noise in the reference is directly added to the input.

- Circuit for calculation of noise in a reference generator
  \[ i_{d1} = i_{d2} = -g_{mP}V_P = V_{n, out}/(R_1 + g_{mN}^{-1}) \rightarrow V_P = -g_{mN}^{-1}V_{n, out}/(R_1 + g_{mN}^{-1}) \]

  and \[ V_P = A_0V_{in, op} \]

  Node A:
  \[ \frac{V_{n, out}}{R_1 + g_{mN}^{-1}} + \frac{1}{g_{mN}} = \frac{V_{n, op}}{V_{n, out}} + V_{n, out} \]

  Since typically \( g_{mP}A_0 \gg g_{mN} \gg R_1^{-1} \), we have \( |V_{n, out}| \approx V_{n, op} \).

  - The noise of the op amp directly appears at the output.
  - 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 application.

---

Simplified core of a bandgap circuit

(a) Addition of cascode devices to improve supply rejection
(b) Use of self-biased cascode to eliminate \( V_{b1} \) and \( V_{b2} \)
Generation of a floating reference voltage

\[ V_{\text{out}} = \frac{R_2}{R_6} V_{BE4} + 2 \frac{R_3}{R_1} V_T \ln n \]

Regulation of supply voltage to improve supply rejection