

# A Simple Three-Terminal IC Bandgap Reference

A. PAUL BROKAW, MEMBER, IEEE

**Abstract**—A new configuration for realization of a stabilized bandgap voltage is described. The new two-transistor circuit uses collector current sensing to eliminate errors due to base current. Because the stabilized voltage appears at a high impedance point, the application to circuits with higher output voltage is simplified. Incorporation of the new two-transistor cell in a three-terminal 2.5-V monolithic reference is described. The complete circuit is outlined in functional detail together with analytical methods used in the design. The analytical results include sensitivity coefficients, gain and frequency response parameters, and biasing for optimum temperature performance. The performance of the monolithic circuit, which includes temperature coefficients of 5 ppm/°C over the military temperature range, is reported.

## I. INTRODUCTION

THE REQUIREMENT for a stable reference voltage is almost universal in electronic design. The temperature-compensated avalanche breakdown diode fills many of the needs, but cannot be used with low voltage supplies and often suffers from long-term stability problems. Use of a transistor base emitter diode temperature compensated to the bandgap voltage of silicon is a technique which overcomes some of the avalanche diode limitations. Bandgap circuits can be operated from low voltage sources and depend mainly upon subsurface effects which tend to be more stable than the surface breakdowns generally obtained with avalanche diodes.

The conventional three-transistor bandgap cell works well for very low voltage two-terminal or "synthetic Zener diode" requirements. The three-transistor cell is less flexible in three-terminal applications and in circuits where the desired output is not an integral multiple of the bandgap voltage.

The two-transistor cell presented here is simpler, more flexible in three-terminal applications, and eliminates sources of error inherent in the three-transistor cell. The two-transistor cell offers separate control over output voltage and temperature coefficient in a circuit using only a single control loop.

The new bandgap circuit has been used as the basis of a monolithic three-terminal reference circuit supplying a stable 2.5-V output and operating down to 4-V input.

## II. CONVENTIONAL CIRCUIT

Conventional bandgap circuits are based on the concept illustrated in Fig. 1. The transistors  $Q_1$  and  $Q_2$  are

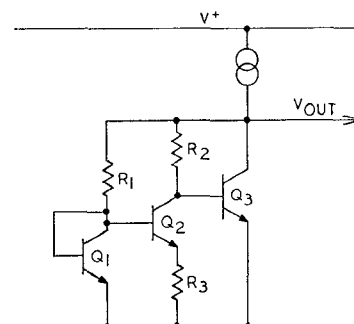


Fig. 1. Conventional bandgap circuit.

operated at different current densities to produce temperature proportional voltages across  $R_3$  and  $R_2$ . A third transistor  $Q_3$  is used to sense the output voltage through  $R_2$ . As a result,  $Q_3$  drives the output to a voltage which is the sum of its  $V_{BE}$  and the temperature-dependent voltage across  $R_2$ . When the output voltage is set to approximate the bandgap voltage of silicon, the voltage across  $R_2$  will compensate the temperature coefficient of  $V_{BE}$ , and the output voltage will be temperature invariant [1]. To minimize the output voltage temperature coefficient the collector current of  $Q_3$  must be made proportional to temperature, as are the currents in  $Q_1$  and  $Q_2$ . This large temperature-dependent current at the point where the stabilized voltage appears makes it inconvenient to produce an output greater than the bandgap voltage. Higher voltages can be generated by stacking several junctions to produce, in effect, several circuits like Fig. 1 in series [2].

The theory used to predict the temperature behavior of circuits like Fig. 1 neglects the effect of base current flowing in  $R_1$  and  $R_2$ . The variability in this current due to processing and temperature effects on  $h_{FE}$  gives rise to an output voltage error and drift. This effect is particularly severe when the current in  $Q_2$  is made much smaller than currents in  $Q_1$  and  $Q_3$  to produce the required current density difference. Use of "Super Beta" processing to reduce this problem results in low-voltage transistors not suitable for a three-terminal reference.

An additional temperature stability problem arises out of the nonlinearity and nonuniformity of the temperature characteristics of diffused resistors. The nonlinearity cannot easily be compensated, and the nonuniformity cannot be accommodated in the design.

The idealized circuit shown in Fig. 2 minimizes the difficulties of obtaining outputs above the bandgap voltage, reduces the problem of  $h_{FE}$  variability to one of  $\alpha$  match, and can be implemented with thin-film resistors on the monolithic chip to virtually eliminate nonlinear

Manuscript received May 6, 1974; revised July 25, 1974. This paper was presented at the International Solid-State Circuits Conference, Philadelphia, Pa., February 1974.

The author is with the Semiconductor Division, Analog Devices, Inc., Wilmington, Mass.

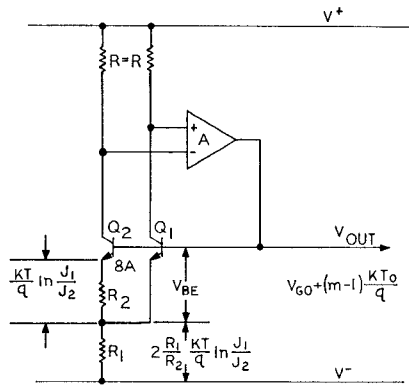


Fig. 2. Idealized circuit illustrating two-transistor bandgap cell.

temperature coefficients of resistance (TCR) as an error factor. The circuit uses two transistors and collector-current sensing to establish the bandgap voltage. The voltage appears at active transistor (as opposed to diode-connected) bases, so that it is a straight forward and simple matter to obtain overall output voltages above the bandgap voltage.

### III. BASIS OF THE NEW CONFIGURATION

#### A. Generating the Bandgap Voltage

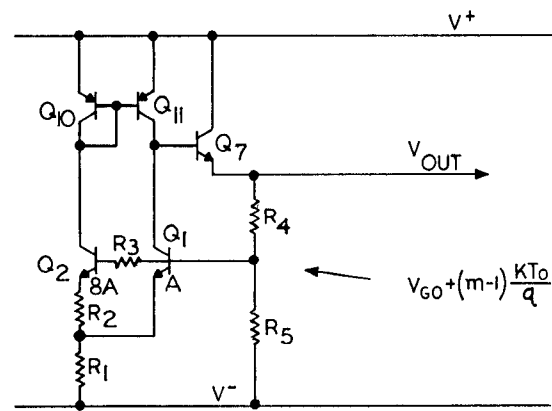
In the circuit of Fig. 2 the emitter area of  $Q_2$  is made larger than that of  $Q_1$  (by a ratio of 8-to-1 in the example given). When the voltage at their common base is small, so that the voltage drop across  $R_2$  is small, the larger area of  $Q_2$  causes it to conduct more of the total current available through  $R_1$ . The resulting imbalance in collector voltages drives the op amp so as to raise the base voltage. Alternatively, if the base voltage is high, forcing a large current through  $R_1$ , the voltage developed across  $R_2$  will limit the current through  $Q_2$  so that it will be less than the current in  $Q_1$ . The sense of the collector voltage imbalance will now be reversed, causing the op amp to reduce the base voltage. Between these two extreme conditions is a base voltage at which the two collector currents match, toward which the op amp drives from any other condition. Assuming equal  $\alpha$  or common-base current transfer ratio for  $Q_1$  and  $Q_2$ , this will occur when the emitter current densities are in the ratio 8-to-1, the emitter area ratio.

When this difference in current density has been produced by the op amp, there will be a difference in  $V_{BE}$ , between  $Q_1$  and  $Q_2$ , which will appear across  $R_2$ . This difference will be given by the expression

$$\Delta V_{BE} = \frac{kT}{q} \ln \frac{J_1}{J_2}. \quad (1)$$

Since the current in  $Q_1$  is equal to the current in  $Q_2$ , the current in  $R_1$  is twice that in  $R_2$  and the voltage across  $R_1$  is given by

$$V_1 = 2 \frac{R_1}{R_2} \frac{kT}{q} \ln \frac{J_1}{J_2}. \quad (2)$$



Eliminate error due to base current in  $R_4$  by setting:

$$R_3 = \frac{R_2}{R_1} \frac{R_4 R_5}{R_4 + R_5}$$

Fig. 3. Simplified circuit for developing higher output voltages.

Assuming that the resistor ratio and current density ratio are invariant, this voltage varies directly with  $T$ , the absolute temperature. This is the voltage which is used to compensate the negative temperature coefficient of  $V_{BE}$ .

The voltage at the base of  $Q_1$  is the sum of the  $V_{BE}$  of  $Q_1$  and the temperature-dependent voltage across  $R_1$ . This is analogous to the output voltage of the conventional bandgap circuit and can be set, by adjustment of  $R_1/R_2$ , to a temperature stable value, as described in the Appendix.

#### B. Increasing the Stabilized Output Voltage

Assuming that the amplifier of Fig. 2 has sufficient gain, it will balance the collector currents of  $Q_1$  and  $Q_2$  despite an additional voltage drop added between its output and the common-base connection. This additional drop will not affect the base voltage which results in collector current balance. If the voltage is introduced by means of a resistive voltage divider, the op amp output voltage will be proportional to the common-base voltage.

The circuit of Fig. 3 uses an active load to sense the collector current of  $Q_1$  and  $Q_2$  more directly. The function of the op amp is replaced by  $Q_{10}$ ,  $Q_{11}$ , and  $Q_7$ . The p-n-p transistors form a simple current mirror which takes the difference of the collector currents of  $Q_1$  and  $Q_2$ . This difference current drives the base of  $Q_7$  which supplies the circuit output voltage. This voltage is divided by  $R_4$  and  $R_5$  and applied to the base of  $Q_1$ . The sense of the signal to  $Q_7$  drives  $Q_1$  and  $Q_2$  to minimize the collector current difference. By designing the circuit to stabilize the base voltage at the bandgap voltage the output will be stabilized at a higher voltage. Since the output voltage depends upon  $R_4$  and  $R_5$  it can be set to any convenient value and need not be an integral multiple of the bandgap voltage.



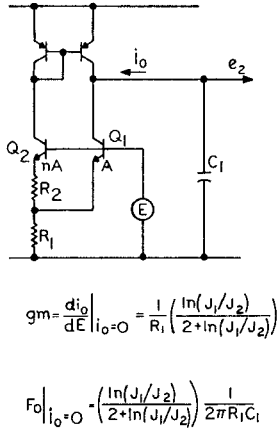


Fig. 5. Transconductance and frequency compensation model.

### V. FREQUENCY COMPENSATION

The amplifier in this circuit operates in a closed loop to regulate the output voltage. A composite junction-MOS capacitor  $C_1$  is used to control the open-loop cross-over frequency and stabilize the closed-loop response. The analytical basis for this compensation is illustrated by Fig. 5. When the two-transistor bandgap cell is operated into a current mirror an output current is produced whenever the common-base voltage departs from the nominal voltage determined by the current density ratio and by  $R_1$  and  $R_2$ . The change in this current as a function of the departure of base voltage from its nominal value has the dimensions of transconductance and can be used as such in design. The following incremental approximation gives a simple result which is more than adequate for most design procedures.

Incremental changes in the base voltage of  $Q_1$  give rise to changes in collector current which can be approximated by the ratio of the voltage change to  $R_e$ , the incremental emitter resistance. This same voltage increment also drives  $R_2$  and the incremental impedance of  $Q_2$ . If the transistors are operating at equal currents, the two  $R_e$  terms will be equal, making the total effective resistance in the  $Q_2$  branch higher. This will result in a lower incremental current in  $Q_2$ . Equating the incremental base voltage changes gives

$$\Delta i_{e1} R_{e1} = \Delta i_{e2} (R_{e2} + R_2). \quad (5)$$

Substituting for  $R_e$  and for  $R_2$  in terms of the voltage across it and current through it converts (5) to

$$\Delta i_{e1} \frac{kT}{q i_{e1}} = \Delta i_{e2} \left( \frac{kT}{q i_{e2}} + \frac{kT}{q i_{e2}} \ln \frac{J_1}{J_2} \right). \quad (6)$$

A second approximation made is that the total incremental current is due to the voltage change across  $R_1$  resulting from a voltage change  $\Delta E$  at the common bases. That is

$$\Delta i_{e1} + \Delta i_{e2} = \frac{\Delta E}{R_1}, \quad (7)$$

which neglects  $R_e$  as compared with the value of  $R_1$ . After eliminating common factors and assuming  $i_{e1} = i_{e2}$ , the combination of (6) and (7) can be manipulated to form

$$\frac{\Delta i_{e1} - \Delta i_{e2}}{\Delta E} = \frac{1}{R_1} \left( \frac{\ln(J_1/J_2)}{2 + \ln(J_1/J_2)} \right). \quad (8)$$

The difference in the  $Q_1$ ,  $Q_2$  collector currents is the output current  $i_o$  and is approximately given by the difference of the emitter current increments. Taking (8) to the limit at  $i_o = 0$  yields

$$\left. \frac{d i_o}{d E} \right|_{i_o=0} = \frac{1}{R_1} \left( \frac{\ln(J_1/J_2)}{2 + \ln(J_1/J_2)} \right). \quad (9)$$

With a current density ratio of 8:1, the term  $\ln J_1/J_2$  is approximately 2, so that the "transconductance" of the entire circuit is approximately  $1/(2R_1)$ .

A capacitive load on the current output of the circuit in Fig. 5 will give a 6 dB/octave rolloff of voltage transfer ratio. The frequency at which the capacitive reactance equals the transconductance will be the unity-gain frequency of the simple circuit. This is given by the expression for  $F_0$  in the figure. In the circuit of Fig. 4, the loop attenuation due to  $R_4$  and  $R_5$  reduces the overall unity-gain frequency by the ratio of the bandgap voltage to 2.5 V, which is approximately two.

The transconductance can also be used to estimate a low frequency "gain." In the simple circuit of Fig. 3, the gain is expressed as the ratio of the voltage at the base of  $Q_7$  to a small-signal input applied to the base of  $Q_1$  at balance. Using a value of 3 k $\Omega$  for  $R_1$  and estimating the output impedance of  $Q_{11}$  at about 300 k $\Omega$  gives a gain of about 50. In the monolithic circuit, the effective open-loop gain is increased several orders of magnitude by the bootstrap connection to the current mirror.

### VI. MONOLITHIC CIRCUIT PERFORMANCE

The circuit of Fig. 4 is shown in Fig. 6 as it appears in monolithic form. Several diffusion lots have been made and measurements of these units indicate the typical properties given by the following table.

Typical Reference Circuit Parameters (-55 to +125°C)	
Output voltage	2.5 V $\pm$ 2 percent
Minimum input voltage	4 V
Load regulation, 0 to 10 mA	3 mV
Supply rejection, 4.5 to 7 V	0.25 mV
Supply rejection, 7 to 30 V	0.25 mV
Standby current	1 mA
Output voltage temperature coefficient ( $\gamma$ )	5 to 60 ppm/°C
$\left( \gamma = \frac{(V_{\max} - V_{\min})}{(V_{\text{nominal}})(\Delta T)} \right)$	

The observed variation in temperature coefficients

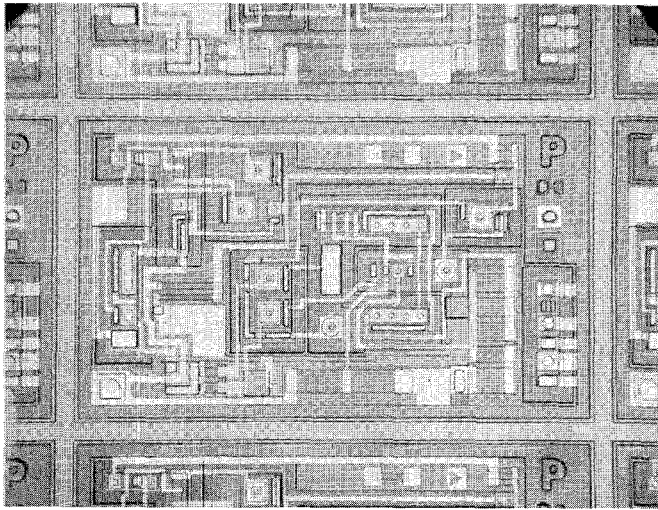


Fig. 6. Illustration of  $37 \times 62$  mil monolithic die showing thin-film resistors and balanced thermal layout.

arises mainly from variability in absolute voltage at the base of  $Q_1$  and variation in the coefficient  $m$  (see Appendix). Units showing a very constant temperature coefficient generally show a good correlation between absolute output voltage and drift.

Several units showing large but constant temperature coefficients have been adjusted to very small final temperature coefficients. This has been done by measuring the temperature coefficient calculating the ideal zero temperature coefficient voltage correction and laser-trimming  $R_1$  or  $R_2$  as required. Other units which are closer to the nominal output voltage exhibit very stable temperature characteristics initially. In extremely low-drift units, the performance of the basic cell appears to be masked by other drifts arising in the remainder of the circuit. These residual drifts are on the order of 2 to 4 ppm/°C over a temperature range of  $-55$  to  $+125^\circ\text{C}$ .

Some diffusion lots have shown a greater curvature in the temperature characteristic. These units exhibit the roughly parabolic temperature characteristic, which is implied by (14) of the Appendix. They show good temperature performance around a peak which occurs at a temperature related to initial output voltage. At temperature extremes the temperature coefficient may increase to 60 ppm/°C or more, if the peak is not centered in the  $-55$  to  $+125^\circ\text{C}$  range.

The elapsed time since obtaining the first completed units has not been sufficient to accumulate long-term drift results. Accelerated life tests have been made at high temperatures to uncover any gross drift problems. The temperature stability and monitoring equipment have not been adequate to determine the ultimate stability of the device. Examination of data taken over the course of 1000 hours at  $+125^\circ\text{C}$  does not reveal any trends or systematic drifts at the 100 ppm level, which approximates the repeatability of the measurements.

## APPENDIX

### THEORETICAL TEMPERATURE BEHAVIOR

#### A. Optimum Cell Voltage

The optimum voltage at the base of  $Q_1$  is approximately the bandgap voltage. A general analysis of the two-transistor bandgap cell yields a more-exact result and some insight into the residual temperature drift in an optimally adjusted circuit.

As a matter of convenience, the circuit description has involved the assumption that the required current density ratio is established by the use of emitter area ratio alone. The circuit can also be based on different collector currents and equal areas or a combination of unequal areas and currents.

The equation describing the voltage across  $R_1$  can be generalized by including a parameter  $P_1$  which is the ratio of emitter currents  $i_{e1}/i_{e2}$ , as follows:

$$V_1 = (P_1 + 1) \frac{R_1}{R_2} \frac{kT}{q} \ln \frac{J_1}{J_2}. \quad (10)$$

An expression given by Brugler [3] has been modified by the addition of a current-dependent term to give the  $V_{BE}$  of  $Q_1$  as

$$V_{BE1} = V_{so} \left( 1 - \frac{T}{T_o} \right) + V_{BEo} \frac{T}{T_o} + \frac{mkT}{q} \ln \frac{T_o}{T} + \frac{kT}{q} \ln \frac{J}{J_o} \quad (11)$$

where  $V_{so}$  is the bandgap voltage of silicon and  $V_{BEo}$  is the value of  $V_{BE1}$  at a reference temperature  $T_o$ .

Since the voltages across both  $R_1$  and  $R_2$  are proportional to temperature, it follows that the current and the current density in  $Q_1$  are also proportional to temperature. Therefore, the current density ratio in the last term of (11) can be equated to a temperature ratio as  $J/J_o = T/T_o$ . This relation can be used to reduce the sum of  $V_1$  and  $V_{BE1}$  to the form

$$E = V_{so} + \frac{T}{T_o} (V_{BEo} - V_{so}) + (m - 1) \frac{kT}{q} \ln \frac{T_o}{T} + (P_1 + 1) \frac{R_1}{R_2} \frac{kT}{q} \ln \frac{J_1}{J_2}. \quad (12)$$

This represents the stable voltage established at the base of  $Q_1$  in the circuit of Fig. 1. Differentiating this result twice with respect to temperature yields

$$\frac{dE}{dT} = \frac{1}{T_o} (V_{BEo} - V_{so}) + (P_1 + 1) \frac{R_1}{R_2} \frac{k}{q} \ln \frac{J_1}{J_2} + (m - 1) \frac{k}{q} \ln \frac{T_o}{T} - 1 \quad (13)$$

and

$$\frac{d^2 E}{dT^2} = -(m-1) \frac{k}{q} \frac{1}{T}. \quad (14)$$

Equating the first derivative to zero results in the equation

$$V_{BE_o} + (P_1 + 1) \frac{R_1}{R_2} \frac{kT_o}{q} \ln \frac{J_1}{J_2} = V_{oo} + (m-1) \frac{kT_o}{q}. \quad (15)$$

The left side of (15) is the value of  $E$  at  $T = T_o$ . This means that, in principle, if the base voltage of  $Q_1$  is set to  $V_{oo} + (m-1) kT_o/q$  at temperature  $T_o$ , the temperature coefficient of the output voltage will be zero. Assuming values of  $m$  greater than one in (14) implies, however, a nonzero temperature coefficient at temperatures other than  $T_o$ . Experimental data indicate that values of  $m$  as low as 1.2 have been achieved.

Examination of (13) and (15) shows that departures from the optimum output voltage will result in an approximately constant temperature coefficient. The magnitude of this coefficient will be the absolute error voltage divided by the absolute temperature. For example, a 3 percent absolute voltage error at 300 K will result in a 3 percent/300 K = 100 ppm/°C temperature coefficient.

#### B. Effects of Resistor Temperature Coefficients

Both the common and differential temperature coefficients of resistance of  $R_1$  and  $R_2$  enter into the overall output voltage stability. The differential TCR may be treated by assuming that the total change in resistor ratio is due to change in  $R_1$ . Since the remainder of the circuit forces a predetermined current in  $R_1$ , its temperature coefficient of resistance is translated to an error voltage in equal proportion. The effect on the overall output voltage is reduced by the fact that the voltage across  $R_1$  accounts for only a fraction of the total output. This fraction is variable, but is about 1/2 at room temperature. Therefore, the differential TCR of  $R_1$  and  $R_2$  appears as about half the equivalent proportional change in the output voltage.

The common TCR affecting both resistors equally does not affect the voltage across  $R_1$ , which depends only upon their ratio. It does affect the total current in the transistors, however, and therefore affects the  $V_{BE}$  of  $Q_1$ .

This effect can be evaluated by expressing  $V_{BE_1}$  as

$$V_{BE_1} = \frac{kT}{q} \ln \left( \frac{i_{e_1}}{I_s} \right). \quad (16)$$

Differentiating

$$\frac{dV_{BE_1}}{dR_2} = \frac{kT}{q} \frac{I_s}{i_{e_1}} \left( \frac{-i_{e_1}}{R_2 I_s} \right) = \frac{-kT}{qR_2} \quad (17)$$

and

$$\frac{dV_{BE_1}}{dR_2} R_2 = \frac{-kT}{q} \quad (18)$$

which at room temperature is about -26 mV or -260  $\mu$ V/percent. Dividing this coefficient by the cell's output voltage of about 1.22 V gives a relative coefficient of about  $-2.13 \times 10^{-4}$ /percent of output voltage change as a function of relative resistance change. This is a reduction in sensitivity of about 47 times.

Despite this reduction, the TCR of diffused resistors is so large, nonlinear, and nonuniform that it still represents a serious temperature drift. The monolithic circuit described here uses thin-film-on-silicon resistors with a total TCR of about -60 ppm so that the 47 times reduction makes the uncorrectable component of TCR sensitivity less than 1 ppm/°C.

The effects of the TCR of other resistors, or of ratios such as  $R_4$  and  $R_5$ , are readily evaluated by standard analytical techniques. The low differential TCR inherent in thin-film resistor pairs keeps these effects on the output voltage at a minimum.

#### REFERENCES

- [1] R. J. Widlar, "New developments in IC voltage regulators," *IEEE J. Solid-State Circuits*, vol. SC-6, pp. 2-7, Feb. 1971.
- [2] K. E. Kuijk, "A precision reference voltage source," *IEEE J. Solid-State Circuits*, vol. SC-8, pp. 222-226, June 1973.
- [3] J. S. Brugler, "Silicon transistor biasing for linear collector current temperature dependence," *IEEE J. Solid-State Circuits* (Corresp.), vol. SC-2, pp. 57-58, June 1967.



**A. Paul Brokaw** (M'73) was born in Minneapolis, Minn., on January 18, 1935. He received the B.S. degree in physics from Oklahoma State University, Stillwater, in 1962.

From 1962 to 1963, he was a Design Engineer at Labko Scientific, Inc. In 1963, he joined the technical staff of Arthur D. Little, Inc. as an Electronic Circuit Designer. After six years at ADL, he joined the design staff of Communication Technology, Inc. In 1971, he started at Nova Devices, Inc., now the Semiconductor Division of Analog Devices, Inc., where he is the Director of Advanced Product Development.