A Floating CMOS Bandgap Voltage Reference for Differential Applications

MARCO FERRO, FRANCO SALERNO, AND RINALDO CASTELLO, MEMBER, IEEE

Abstract—The first floating CMOS bandgap voltage reference is reported. The circuit is very suitable for precision mixed-mode systems, using the differential approach, especially for the case of single supply operation. An experimental prototype, realized in a 2-μm CMOS technology, generates a continuous-time low-impedance voltage of 2.48 V ± 24 mV before trimming. The temperature coefficient measured on 30 samples ranges from −20 to +32 ppm/°C in the temperature range from 0 to 100°C. Thanks to the differential approach, a high-frequency power supply rejection of −50 dB at 100 kHz was achieved, which is the best ever reported. The active area of the chip is 1800 mil² and the circuit dissipates 6 mW when operated from a single 5-V supply.

I. INTRODUCTION

In recent years fully differential techniques for the realization of high-precision analog circuits like filters, A/D, D/A converters, etc. have enjoyed an increasing popularity. This has been particularly true in mixed-mode systems realized in MOS technology [1]-[3] where precision analog functions and a large number of digital circuits must be realized on the same substrate.

The reason for this tendency is the superior immunity of differential circuits to any unwanted signal that can be coupled into the signal path via the power supply lines, the substrate, or any other parasitic coupling, particularly switching noise associated with digital circuits. A second reason is the effective doubling of the dynamic range (assuming the same noise level) thanks to the availability of negative and positive signal polarities.

The advantages of using differential circuits become especially evident when only a single low-voltage supply is available, as often occurs in circuits realized in scaled technologies [4]. In fact, in this case, the reduced signal swing severely limits the achievable dynamic range. Furthermore, the lack of a solid ground renders it very difficult to isolate the analog and digital portions of the circuit from each other.

A precision voltage reference is a very commonly used analog block in mixed-mode systems, however, to date, only single-ended realizations have been reported [5]-[7].

Fig. 1. Basic bandgap configurations: (a) with positive output respect to ground; and (b) with negative output respect to the positive supply.

Nonetheless, a differential version of such a device could be advantageous, especially when the reference must directly interface itself with other differential blocks [1], [3], [8]. This should result in an improved noise immunity, particularly at high frequency, for the overall system.

Fig. 2. Basic bandgap configurations: (a) with positive output respect to ground; and (b) with negative output respect to the positive supply.

This paper reports on the first integrated floating MOS bandgap voltage reference. The circuit generates a low-impedance voltage equal to twice the bandgap of silicon and can directly and continuously drive a differential load as in an A/D or D/A converter. The temperature coefficient measured on 30 samples ranges from −20 to +32 ppm/°C in the temperature range from 0 to 100°C. Due to the differential configurations and the careful operational-amplifier design and layout, the supply rejection curve stays remarkably flat over frequency and is still better than −50 dB at 100 kHz, which is the best value reported to date.

II. PRINCIPLE OF OPERATION OF THE FLOATING BANDGAP REFERENCE

The basic CMOS bandgap configuration for the case of an n-well process is shown in Fig. 1(a) [13]. The circuit produces an output voltage referenced to the negative rail which is equal to the bandgap voltage of silicon. An alternative version of the same circuit that produces the same voltage but referred to the positive supply is shown in Fig. 1(b) [9].

Manuscript received September 20, 1988; revised January 24, 1989.
M. Ferro and F. Salerno are with the Centro Studi e Laboratori Telecomunicazioni (CSELT), 10148 Torino, Italy.
R. Castello is with the Dipartimento di Elettronica, Università di Pavia, 27100 Pavia, Italy.
IEEE Log Number 8927272.
For differential applications, however, it is often necessary to supply a precise reference voltage which is floating between the two rails and is centered around some suitable common-mode value. In single supply systems this common-mode voltage must be generated on-chip and therefore is unavoidably contaminated with spurious signals. In this case it is particularly important for the reference voltage generator to be able to effectively reject both the noise coupled on the supply lines and also the noise present on the common-mode voltage itself. Then, instead of obtaining the differential voltage starting from a generator referenced to either supply [1], it is more effective to directly generate a floating reference voltage. This can be obtained conceptually with the circuit of Fig. 2.

The configuration of Fig. 2 can be thought of as the superposition of the two complementary single-ended versions shown in Fig. 1(a) and (b). The floating reference is obtained by substituting the single-ended amplifier used in both the circuits of Fig. 1(a) and (b) with a differential amplifier whose positive output drives resistors $R_2$ and $R_3$ and whose negative output drives the base terminals of the bipolar transistors. The bandgap voltage is now available as the difference between the two outputs of the amplifier, therefore its common-mode voltage is controlled by the common-mode feedback circuit of the amplifier itself.

The main limitation of all these circuits comes from the offset of the amplifier that in CMOS technology is typically in the range $\pm 15$ mV [10]. This offset voltage is amplified by about a factor of 10 and leads to large variation in the output reference voltage; furthermore it degrades its temperature stability.

One way to solve this problem is to use a sampled voltage reference [6]. In this implementation the output voltage is available during only one phase of the clock, with the other phase used to sample and null-off the amplifier offset. This solution, however, is only viable if the entire system is synchronous with the reference clock. On the other hand, if the reference voltage has to be available continuously to the system or if the system clock rate is very high (i.e., several megahertz), this solution cannot be used. In this case the problem can be alleviated by designing a reference whose output is equal to a multiple of the bandgap voltage [5], thereby reducing the relative importance of the offset error relative to the reference value.

Another obvious way to improve the situation is to use a low offset amplifier. It has been shown, in fact, that it is possible to produce, in CMOS technology amplifiers, with an offset voltage of only a few millivolts or less [7], [11], [12]. This can be obtained in two different ways. One approach [11] is to use a very complicated layout technique to match the two transistors of the amplifier differential input pair. This, however, results in extra area and/or power consumption. The other approach is to use compatible lateral bipolar transistors as input devices in the operational amplifier [7], [12]. This is a very interesting solution, however, for the architecture used in our design, it has some potential drawbacks. These are the possible high base current, and the need to connect the controlling gates of the lateral bipolar devices to some appropriate stable voltage reference which in some cases can be lower than the negative supply [14]. Because of these considerations and since no data on the lateral bipolar was available at the moment of the design, such a solution was not adopted in the voltage reference reported here. This remains, in any case, a potentially interesting alternative for future revisions of the circuit. In our design, a $2V_{BE}$ bandgap architecture combined with a relatively low offset amplifier ($\pm 3$ mV standard deviation) has been adopted. This low offset voltage is the result of a careful amplifier design which guarantees that the random offset is totally dominated by the input devices and also of a very symmetrical and closely packed layout. Furthermore, the use of a perfectly balanced structure automatically gives zero systematic offset.

The conceptual schematic of the $2V_{BE}$ floating bandgap reference is shown in Fig. 3. For proper operation of the circuit the current in the two internal branches should match that in the two external ones. This is obtained by deriving the current in the internal branches, via current mirrors $M_1$–$M_3$, from a current generator proportional to absolute temperature (PTAT) which is a replica of the bandgap structure. Notice that current mirrors $M_1$–$M_3$...
are referenced to the positive output of the bandgap generator instead of to the positive supply. This is in order to improve power supply rejection.

Besides being used to bias the internal branches of the bandgap generator, the PTAT current source is used to bias all the active circuits in the chip. This is done to compensate for the mobility degradation associated with temperature increase, thereby maintaining a reasonable constant speed of response over the temperature range of operation.

III. Actual Implementation

The complete differential bandgap schematic is shown in Fig. 4. The structure on the left-hand side of the figure is the PTAT current generator which is a replica of the input structure of the bandgap generator. In both cases there are four branches, each corresponding to one of the bipolar transistors used to generate the $2V_{BE}$ voltage drop. However, while for the bandgap these branches are powered by the outputs of the amplifier, for the PTAT they are powered by the two supplies. To improve power supply rejection the p-type current mirrors used in the PTAT block are cascoded. The n-type current mirrors $M10$ and $M11$, on the other hand, are not cascoded. This is due to the large value of the threshold voltages of the MOS transistors available in the digital technology used (up to 1.3-V worst case) which would prevent proper operation if both types of current mirrors were cascoded.

The use of n-type simple current mirrors limits the power supply rejection achievable by the bandgap reference. An improved version of the reference to be realized in a lower threshold technology using cascoded n and p current mirrors should give a better PSRR.

As with many self-biased circuits, this PTAT generator has two possible operating points: its normal one and a spurious one where all currents are equal to zero. To insure that the circuit is always biasing itself in the normal mode of operation, start-up transistors $MS7$ and $MS8$ are added to the circuit as shown in Fig. 4. The start-up sequence is as follows: at power-on the current level in all the MOS devices making up the PTAT circuit is zero with the exception of transistors $MS5$ and $MS6$. As a consequence, node $A$ rises to the positive supply forcing $MS7$ and $MS8$ into heavy conduction therefore starting up the circuit. $MS3$ is designed in such a way that as the current in all the branches reaches its nominal value the voltage at node $A$ will approach the negative supply therefore shutting off $MS7$ and $MS8$. Capacitor $CS1$ is used to compensate the loop and ensure stable operation.

The start-up problem is more complicated for the bandgap circuit itself due to the fact that both the differential and common-mode signal path must be started in the op amp. Transistor $MS1$ in the bandgap loop plays the same role as transistors $MS7$ and $MS8$ in the PTAT, however, its turning off is delayed by inverter $I1$ and capacitors $CS$. This is done to ensure that the current in
the PTAT generator reaches its steady-state value and therefore the operational amplifier reaches its nominal operating point (for both the differential and common-mode portion), before the start-up transistor has turned itself off. In fact, if MS1 turns off simultaneously to the turning on of the bias current generators in the op amp, the bandgap reference may fall back into its spurious bias point with its output voltage equal to zero.

The purpose of transistors M4, M5 and resistor RD is to balance the load seen by the two op-amp outputs. In fact, without the insertion of the dummy load, the positive op-amp output will have to supply a current which is approximately 2.5β times that supplied by the negative output, where β is the current gain of the parasitic p-n-p transistors used in the bandgap circuit. The current flowing in transistors M4 and M5 balances out the current flowing in M1–M3 while RD, whose value is equal to the parallel combination of R2 and R3, compensates for the loading of these two resistors. Notice that although RD goes from the amplifier negative output to the common-mode reference voltage while R2 and R3 are connected between the amplifier positive output and its inputs, the voltage drop across these resistors is almost exactly the same. This is because the amplifier common-mode input and output voltages practically coincide since the two components making up the bandgap voltage \( V_{BE} \) and \( n \Delta V_{BE} \) are very much equal to each other. The purpose of inverter I2 and transistor MS2 is to turn on the dummy load only after the bandgap has been started. This is necessary to insure proper start-up.

The effect of the balancing circuit is that of reducing the requirements on the amplifier common-mode feedback (CMFB) circuit and improving the matching between the two outputs of the voltage reference, therefore improving its supply rejection. The presence of the dummy load also improves the symmetry of the structure and therefore the symmetry of the layout. This has a key beneficial effect on the high-frequency PSRR of the circuit.

Another source of thermal drift of the bandgap reference is the variation of the β of the bipolar transistors with the current density. A plot of β versus current is shown in Fig. 5 for the p-n-p bipolar transistor available in the used technology. From this figure it can be seen that, although the value of β is fairly small, it stays quite flat over several decades of current variation. By properly choosing the size of the bipolar transistor unit module, it is possible to guarantee excellent beta matching between the small device (one unit module) and the large device (ten unit modules) for the desired current level. This ensures a negligible contribution to the bandgap thermal drift due to beta mismatch compared with the contribution due to amplifier offset.

The value of the resistor ratio \( R1/R2 \) can be adjusted with a trimming circuit. Trimming is accomplished using fusible polysilicon links which turn on one of the eight MOS transistors connecting the negative input of the amplifier to one of eight possible tapping points of a voltage divider placed between \( R1 \) and \( R2 \). By using this approach the series resistance of the MOS transistors used in the trimming circuit does not influence the value of the closed-loop gain and therefore small trimming transistors can be used. An identical eight-tap voltage divider is placed in series with \( R3 \) together with a dummy resistor simulating the effect of \( R1 \) for symmetry reasons. The positive input of the amplifier, however, is always connected to the same tap and the dummy resistor is shorted with a metal strip.

IV. OPERATIONAL AMPLIFIER

The bandgap reference uses a fully differential amplifier whose block diagram is shown in Fig. 6. The input stage is a differential folded cascode with a dc gain of approximately 60 dB. The second stage is a class A/B differential circuit capable of providing the large current necessary to drive the resistive load. It has a gain of more than 20 dB when driving a resistor of 2 kΩ.

As shown in Fig. 6, two CMFB circuits are used: one at the output of the first stage and one at the output of the amplifier. Two CMFB circuits are necessary because the signal runs differentially from input to output and two high-impedance points are present in the amplifier.

The first CMFB circuit uses the class A structure shown in Fig. 7(a). It is a simple one-stage amplifier with a gain of approximately 0.5, preceded by two PMOS source fol-
Fig. 7. CMFB circuits: (a) input-stage CMFB; and (b) output-stage CMFB.

lowers driving the resistor pair $R_F1$ and $R_F2$ that are used to sense the common-mode signal. This circuit has a limited linear range, however, this is not a problem because of the small differential swing present at the output of the first stage, thanks to the 20-dB minimum gain provided by the second stage.

The output CMFB circuit is shown in Fig. 7(b). This circuit must be linear over the whole differential output voltage swing of approximately 5 V peak to peak. For this reason the common-mode sensing resistors $R_F3$ and $R_F4$ are inserted directly between the two output nodes. The rest of this common-mode circuit is a two-stage class A circuit connected in parallel to the output nodes with a dc open-loop gain of approximately 70 dB.

The complete schematic of the differential portion of the operational amplifier is shown in Fig. 8, together with the two CMFB circuits represented as two black boxes. Circuit stability is ensured by placing two Miller capacitances of 2 pF from the outputs of the first stage to the outputs of the amplifier. This value of the compensation capacitance ensures ac stability for a closed-loop gain of 10 and results in a gain-bandwidth product of 40 MHz with a power dissipation of 5 mW. The offset voltage has a standard deviation of approximately 3 mV, thanks to the careful design and layout as mentioned above. The circuit is capable of driving a differential capacitive load of up to 300 pF and/or a resistive load of 2 kΩ or more. The amplifier area is approximately 500 mil².

V. EXPERIMENTAL RESULTS

A prototype version of the floating bandgap reference voltage was realized using an n-well CMOS technology featuring a minimum gate channel of 2 μm for digital circuits and 2.5 μm for analog circuits. A microphotograph of the test circuit is shown in Fig. 9. The active area is 1800 mil². The nominal output voltage is 2.48 V (about two times the bandgap voltage of silicon) with a standard deviation of 24 mV without trimming.

The bit configuration of the trimming circuit that gives the best average temperature stability of the output voltage was obtained experimentally by characterizing a large number of samples from different wafers. Fig. 10 shows the temperature stability of the reference output voltage for various samples, all trimmed at the optimum bit configuration obtained, as explained above. The various curves correspond to circuits taken from three different wafers of the same lot. The various wafers, however, received different threshold adjustment implants in order to simulate processing variations. The same bit configuration in the
trimming circuit for all the samples was used to estimate the temperature stability that can be expected from the bandgap reference if no trimming adjustment is performed.

The prototype bandgap reference exhibits an average temperature variation of 21 ppm/°C and a standard deviation of 10 ppm/°C over the temperature range from 0 to 100°C. This is an extremely low value and cannot be fully explained with theoretical calculations. The number of samples measured (about 30) and the fact that they are all coming from the same run (although with different threshold implants) cannot guarantee that the above measured performance will be maintained over the entire production spread. Further data are necessary to fully characterize the behavior of this voltage reference in a production environment.

It is possible that the differential nature of the bandgap could have some beneficial effects on the temperature stability, however this was not studied in detail.

The dominant contribution to the thermal drift is the amplifier offset. Beta-mismatch-induced drift was negligible as explained before. The power supply rejection (de-
fined as the ratio between the signal on the reference voltage and the noise injected on the supply) of the bandgap reference, as a function of frequency, for both supplies, is shown in Fig. 11. Notice the remarkable flatness of the curves and their similarity. Although the low-frequency value is comparable, and in some cases inferior, to other single-ended reference voltage [5]-[7], the value at 100 kHz is the best ever reported. This is the main benefit of using a differential approach and is also due to the very symmetrical design and layout.

The bandgap reference reported here was designed to be used as part of a U interface chip for ISDN applications. In this case the energy of the signal processed by the chip is concentrated around 80 kHz with a relatively narrow spectrum. It was, therefore, essential for the circuit to be able to reject noise in this frequency range to avoid crosstalk and other degradations. On the other hand, the power supply rejection of the circuit at very low frequencies was not so critical.

It is believed that the power supply rejection value is limited primarily by the PTAT current generator. By using cascoded NMOS mirrors, which is feasible for a lower threshold process, a sizable improvement of the low-frequency supply rejection is expected. Furthermore, a more symmetrical loading condition for the amplifier should further improve the flatness of the supply rejection curve over frequency. This can be achieved by forcing all four output transistors to carry approximately the same current whereas in the present case two carry a large current and two carry a much smaller one. Ultimately the finite output resistance of the bipolar transistors should limit the achievable negative supply rejection.

The noise performance of the reference is depicted in Fig. 12. The total noise integrated over 250 kHz is 310-μV rms. The overall performance of the differential reference is summarized in Table I.

### VI. CONCLUSIONS

A precision floating CMOS bandgap reference has been reported. A monolithic prototype realized in an n-well CMOS technology exhibits a temperature coefficient in the temperature range from 0 to 100°C ranging from −20 to +32 ppm/°C. The main feature of the differential bandgap reference is its high power supply rejection which remains practically flat into the 100-kHz range.

The floating implementation of the reference makes it suitable for high-precision mixed-mode systems, using the differential approach, especially for the case of single supply operation. In particular, the superior high-frequency power supply rejection is especially important when the reference is used in a system where the signal to be processed has a spectrum centered around one or more relatively high frequencies, as in the case of modems.
ACKNOWLEDGMENT

The authors wish to thank S. Mariani for the careful layout and P. O'Donnell for editing the manuscript.

REFERENCES


Marco Ferro was born in Turin, Italy, on September 11, 1955. He graduated in electronics at the Istituto Tecnico "G. B. Pininfarina," Turin, Italy, in 1974. In late 1974 he joined the Electronic Circuits Techniques Department of Centro Studi e Laboratori Telecomunicazioni (CSELT), Turin, Italy, where he has been involved in the circuit design of transmission systems over both cables and optical fiber. Having made a specialty of analog circuit design and system development, he currently is a consultant in analog systems and circuits for integration of components in the field of ISDN transmission systems.

Franco Salerno was born in Santa Caterina Albanese, Cosenza, Italy, on May 18, 1936. He received the degree in electronic engineering from the Politecnico di Torino, Turin, Italy, in 1982. In late 1982 he joined the Integrated Circuits Department of the Centro Studi e Laboratori Telecomunicazioni (CSELT), Turin, Italy, where he has been involved in the design of analog/digital interfaces in NMOS and CMOS technologies and device modeling for circuit simulation. In 1984 his interests turned to analog signal processing, switched-capacitor filters, and analog front-end circuits for the ISDN system. His present interest is in the design of high-frequency bipolar and MOS analog circuits for telecommunication.

Rinaldo Castello (S'78–M'83) was born in Genova, Italy, in 1953. He received the degree of ingegneri (summa cum laude) from the Universita' di Genova, Genova, Italy, in 1977. In 1979 he began his graduate study at the University of California, Berkeley, where he received the M.S.E.E. degree in 1981 and the Ph.D. degree in 1984. While at Berkeley he was a Teaching Assistant and a Research Assistant. Both in 1983 and 1984 he was a Visiting Professor during part of the academic year at the Universita' di Genova. From 1984 to 1986 he was a Visiting Assistant Professor at the University of California, Berkeley. Since 1987 he has been an Associate Professor with the Department of Electrical Engineering at the University of Pavia, Pavia, Italy. His main interest is in MOS integrated circuit design, particularly in the area of telecommunications and analog/digital interfaces. Dr. Castello has been a member of the Program Committee of the European Solid State Circuit Conference (ESSCIRC) since 1987.