Stellingen behorende bij het proefschrift

HIGH-ACCURACY CMOS SMART TEMPERATURE SENSORS

door

Anton Bakker

Delft, 17 april 2000

1. Door verhoging van de intelligentie van sensoren kan ineens een enorme nieuwe markt ontstaan.

(Dit proefschrift, hoofdstuk 1)

2. De markt voor goedkope intelligente temperatuur sensoren heeft gigantische groeimogelijkheden door toepassingen in thermisch management, uitlezing van thermische sensoren en correctie van ongewenste temperatuurafhankelijkheid.

(Dit proefschrift, hoofdstuk 1)

3. De nested-chopper techniek kan de offset van een versterker verlagen tot onder de 100nV terwijl het ruisniveau gelijk blijft aan de thermische ruisvloer.

(Dit proefschrift, hoofdstuk 2)

4. Hoewel de werking van een intelligente temperatuur sensor gebaseerd is op een bipolaire transistor kan in CMOS technologie een hogere nauwkeurigheid worden verkregen dan in bipolaire technologie.

(Dit proefschrift, hoofdstuk 3)

5. De Sigma-Delta A-D converter is voor intelligente temperatuur sensoren en veel andere data-acquisitie systemen het beste compromis tussen vermogensverbruik en nauwkeurigheid.

(Dit proefschrift, hoofdstuk 4)

6. Als de offset in een intelligente temperatuur sensor dynamisch is weggeregeld, wordt de onnauwkeurigheid voornamelijk bepaald door de spreiding in de basis-emitter spanning van de gebruikte bipolaire transistor.

(Dit proefschrift, hoofdstuk 5)

7. Wanneer meer dan de helft van de automobilisten op een bepaalde plek te hard rijdt, zou volgens de regels van de democratie de maximumsnelheid aangepast moeten worden.

8. De situatie in het Oost-blok laat zien dat het milieu gebaat is bij een gezonde economische groei.

9. Gedoogbeleid draagt ten onrechte het woord "beleid" in zich.

10. Het opvoeden van een kind is voornamelijk het heropvoeden van jezelf.

11. Het spanningsveld tussen geloof en wetenschap ontstaat door het gebrek aan kennis van één van beiden.

12. Als we echt van de apen zouden afstammen, zouden we daar niet eens over nadenken.

HIGH-ACCURACY CMOS SMART TEMPERATURE SENSORS

HIGH-ACCURACY CMOS SMART TEMPERATURE SENSORS

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K.F. Wakker,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College voor Promoties aangewezen,
op maandag 17 april 2000 te 10:30 uur

door

Anthonius BAKKER

elektrotechnisch ingenieur
geboren te Amsterdam

Dit proefschrift is goedgekeurd door de promotor:

Prof.dr.ir. J.H. Huijsing

Leden van de promotiecommissie:

De Rector Magnificus	Technische Universiteit Delft
Prof.dr.ir. J.H. Huijsing	Technische Universiteit Delft
Prof.dr. P.J. French	Technische Universiteit Delft
Prof.dr.ir. A.H.M. van Roermund	Technische Universiteit Delft
Prof.ir. A.J.M. van Tuijl	Universiteit Twente
Prof.dr.ir. P.P.L. Regtien	Universiteit Twente
Dr.ir. G.C.M. Meijer	Technische Universiteit Delft
Dr.ir. K.-J. de Langen	Philips Semiconductors, USA

Dr.ir. G.C.M. Meijer heeft als begeleider in belangrijke mate aan de totstandkoming van dit proefschrift bijgedragen.

Published and distributed by:

Anton Bakker
Ternatestraat 179
2612 BC Delft, The Netherlands
Phone: +31-15-2138285
Fax: +31-15-2785755
E-mail: a.bakker@its.tudelft.nl

CIP-DATA KONINKLIJKE BIBLIOTHEEK DEN HAAG

Bakker, A.

High-Accuracy CMOS Smart Temperature Sensors / A. Bakker -
Thesis Delft University of Technology - With ref. - With summary in Dutch
ISBN 90-901-3643-6
Subject headings: Smart Sensors / Solid-State Circuits / Temperature Sensors / Sensor Interfaces / Analog Electronics

PRINTED IN THE NETHERLANDS

The front page shows our one-year old daughter sucking on a thermometer disguised as a dummy.

Table of contents

Introduction 1

This thesis describes the theory and design of high-accuracy CMOS smart temperature sensors. The major topic of the work is the realization of a smart temperature sensor that has an accuracy that is so high that it can be applied without any form of calibration. Integrated in a low-cost CMOS technology, this yields one of the most inexpensive intelligent general purpose temperature sensors in the world. This chapter introduces the reader to the general aspects of the design of CMOS smart temperature sensors. It explores the possibilities and detect the bottle-necks. It ends with the motivation and the organization of the thesis.

1.1 Temperature sensing

Even many centuries ago, people already had a great desire to know the temperature. Reference points could be found, such as the freezing point of water or the body temperature. However, to determine temperatures between those fixed points a temperature meter or *thermometer* was needed. This led to the invention of the mercury-in-glass thermometer, which is still widely used today. Another issue was the definition of the scale. In 1714, Fahrenheit proposed a scale that takes the minimal possible temperature he could imagine to appear in Central Europe as 0°F (-18°C) and the body temperature as 96°F. Celsius proposed a few years later in 1742 a similar scale but he took the freezing point of water as 0°C and its boiling point as 100°C. Both scales are still in use, but the Celsius scale is the most widely used scale. More than a century later, in 1851, Kelvin proposed his law of thermodynamics in which he proved that the

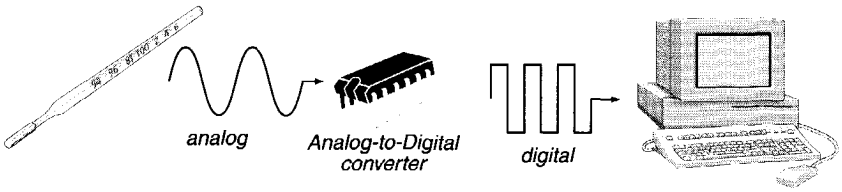

Fig. 1-1 *Communication between a temperature sensor and a computer through an analog-to-digital converter.*

theoretically minimal possible temperature is -273.15°C. Since that time, scientists have often used the Kelvin scale, which is a Celsius scale that is shifted by -273.15 degrees.

The early function of thermometers was to literally read the temperature. However, the industrial revolution asked for *automatic* temperature controllers. These so-called thermostats were used in central heating systems, ovens and engines. The thermometer needed to be adapted to control heaters, fuel valves or other so-called *actuators*. We still use these thermometers with mechanical output in thermostat taps and local temperature control in central heating systems.

In the second half of the 20th century, the temperature controllers became more and more intelligent. Switching a heater or ventilator *on* and *off* was in many applications not good enough anymore. For example, to reduce fuel consumption in central heating systems, it is necessary to monitor temperature *changes* and adapt the heating power to reduce unwanted temperature overshoots. The electronic controller was introduced and the thermometers had to be adapted to communicate with them. Thermometers needed to have an *electrical* output, which resulted in the development of resistance thermometers. The most widely used material for resistance thermometers nowadays is platinum because of its stability and reproducible temperature characteristics. The Pt-100, which is a platinum resistor with a nominal value of 100Ω, is now the standard resistance thermometer. Other widely used materials for resistance thermometers are (lightly-doped) semiconductors, such as silicon. These resistance thermometers are called *thermistors* and have the advantage of a much higher sensitivity than metal-based resistance thermometers.

To communicate with a computer, the resistance changes have to be converted into a digital signal. This is done with an analog-to-digital converter. Such a system is shown in figure 1-1. These kind of temperature sensor systems are still widely used. To reduce the cost of a system that consists of both a temperature sensor and a computer interface, integration of the temperature sensor on the

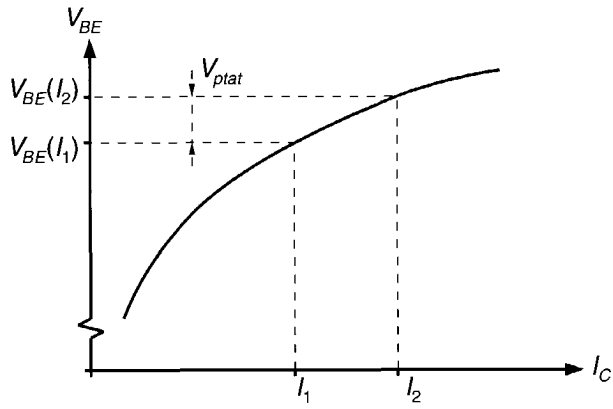

Fig. 1-2 *I_C-V_{BE} characteristic of a bipolar transistor, which can be used to make a temperature sensor that is almost perfectly Proportional To Absolute Temperature (PTAT).*

same chip as the A-to-D converter was attempted. This resulted in a whole new family called integrated *smart* temperature sensors.

1.2 CMOS smart temperature sensors

The research on smart or integrated temperature sensors started in the mid-seventies, together with the development of the Integrated Circuit (IC) technology. It was also the time of the development of a new temperature sensing device, the bipolar transistor. The bipolar transistor was completely compatible with the (bipolar) technology and therefore much easier to integrate than a platinum resistor. It was also much more accurate than a thermistor. Another great advantage was that the bipolar transistor delivered a voltage as output signal instead of a resistance variation, which is a much nicer signal for an A-to-D converter. However, the biggest advantage still is its extremely reproducible exponential characteristic. This made it possible to make a temperature sensor which has a very high intrinsic accuracy. This is explained by figure 1-2.

When a transistor is biased at a collector current of I_1, it will have a base-emitter voltage of

$$V_{BE}(I_1) = \frac{kT}{q}\ln\left(\frac{I_1}{I_s}\right) \qquad (1\text{-}1)$$

where k is Boltzmann's constant ($1.3807 \cdot 10^{-23}$ J/K), T the absolute temperature (in Kelvin), and q the electron charge ($1.6022 \cdot 10^{-19}$C). I_s is the saturation current of the transistor, which is very much dependent on process parameters. Because of the uncertainty in I_s it is not possible to make an accurate temperature sensor this way. However, if we take the *difference* of two base-emitter voltages of the same transistor but biased at two different collector currents I_1 and I_2, we get

$$V_{ptat} = V_{BE}(I_2) - V_{BE}(I_1) = \frac{kT}{q}\ln\left(\frac{I_2}{I_s}\right) - \frac{kT}{q}\ln\left(\frac{I_1}{I_s}\right) = \frac{kT}{q}\ln\left(\frac{I_2}{I_1}\right) \quad (1\text{-}2)$$

This equation shows that the voltage difference is, next to the constants k and q, only dependent on T and the ratio of I_2 and I_1. This means that there are no longer any parameters that are dependent on the process. The voltage difference is therefore very accurately proportional to the absolute temperature T and is called V_{ptat}.

The PTAT voltage can easily be generated on chip and forms the base of the family of smart temperature sensors. The first results were temperature sensors with a current or voltage output [1.1]. At the end of the eighties these temperature sensors were followed uo by a smart temperature sensor with a (semi-)digital output [1.2]. A problem occurred during the beginning of the nineties, when the bipolar technology started to become almost extinct in favour of CMOS technology. CMOS technology was developed very fast because of the great market demand for digital chips. CMOS, however, did not have a regular bipolar transistor. Fortunately, a solution was found in a parasitic bipolar substrate transistor. Although this transistor was not monitored nor wanted in the process, it was found to be very usable as a temperature sensing device. This discovery was made at the start of this thesis work (1995) and during this work the whole world-wide research on smart temperature sensors shifted from bipolar technology to CMOS technology. Nowadays, most commercial smart temperature sensors are made in CMOS technology.

1.3 Motivation and objectives

The main motivation of this thesis is definitely the fast growing demand for *smarter* and *cheaper* temperature sensors. Smarter temperature sensors are needed because the systems in which they are being applied are getting more and more complex. This implies that the designers of such systems will have less knowledge about subsystems such as temperature sensors. This lack of

knowledge needs to be compensated by an increase of intelligence in the smart temperature sensor.

Cheaper products are always needed when markets are growing. And markets are already large when you consider that there are at least five temperature sensors per car and two temperature sensors per PC. Cheaper temperature sensors will also further increase the number of applications.

The main objectives of this work are therefore firstly to increase the intelligence of the smart temperature sensor to such a level that it can be applied by, for example, a digital engineer, who does not know anything about analog circuit design nor physical temperature effects. The second objective is to reduce the costs as much as possible. This is first done by an analysis of the costs, which found that a major cost reduction can be achieved when it is possible to increase the accuracy to such a level that calibration is not necessary for most of the applications. This thesis will therefore focus on dedicated high-accuracy electronics that are needed to design high-accuracy CMOS smart temperature sensors.

1.4 Organization of the thesis

The accuracy of CMOS smart temperature sensors is limited by the offset of the read-out electronics. To increase the accuracy of CMOS smart temperature sensors to the highest possible accuracy level, it is therefore best to first consider techniques to reduce the offset of CMOS amplifiers. Chapter 2 describes the theory and design of dynamic offset-cancellation techniques. These techniques can reduce the offset by a factor of 100 to 1000 and do not need to be calibrated. At the end of this chapter a new technique is introduced that even further reduces the offset. A realization of this new technique and measurement results are shown.

Another important part of the design of a CMOS smart temperature sensor is the voltage reference. Such a reference is needed for the A-to-D converter. The accuracy of the total sensor system can never be higher than the accuracy of its reference. Chapter 3 therefore discusses the theory and design of CMOS bandgap references. It will discuss available bipolar devices in CMOS technology, thermal modelling of the bipolar transistor, and curvature correction techniques. Also the design and implementation of a novel high-accuracy CMOS bandgap reference is shown.

Chapter 4 describes general design aspects of CMOS bandgap references. It describes different kinds of analog-to-digital conversion, methods for Kelvin-to-Celsius conversion, curvature correction and bus interfaces. Also, an electronic interface for the read-out of a single remote bipolar transistor is proposed. The design aspects that are considered most important are high uncalibrated accuracy and low power consumption.

Chapter 5 describes the realizations of three different kinds of CMOS smart temperature sensors that have been designed within the framework of this thesis. The first realization finds its application in a tyre monitoring system, where it watches the temperature of the tyre and compensates thermal cross-sensitivities of a pressure sensor. The second realization is a general purpose ambient temperature sensor, where special attention has been paid to high uncalibrated accuracy. The last version can measure the temperature of a remote bipolar transistor and has its application in microprocessor thermal management.

References

[1.1] National Semiconductor, "LM135 - Precision Temperature Sensor", *http://www.national.com*, November 1999

[1.2] Smartec, "SMT160-30 Smart Temperature Sensor", *http://www.smartec.nl*, November 1999

Dynamic offset-cancellation techniques

2

This chapter describes the theory and design of the different kinds of dynamic offset-cancellation techniques. These techniques can reduce the offset of an amplifier by a factor of 100 to 1000 and do not need trimming. Knowledge of these techniques is necessary to improve the accuracy of CMOS smart temperature sensors. Also in this chapter, a new technique is proposed that can even further reduce the offset. This technique is called the "nested chopper technique". An implementation of this new technique is shown and measurement results are discussed.

2.1 Introduction

In many electronic systems, but especially in smart sensor interfaces, the overall performance is usually limited by the offset and noise of the input amplifiers. This problem has been growing in the past years, because of the shift from bipolar to CMOS IC processes, which have inherently higher noise and offset. Especially in low-frequency applications, like smart temperature sensors, the poor offset and $1/f$ noise performance of CMOS processes is a significant problem nowadays.

Laser trimming and other forms of calibration have been the most appropriate and cheapest solution for many years to reduce the offset, but they could not solve the $1/f$ noise problem. And with the ongoing reduction of processing costs, calibration is no longer an inexpensive solution.

A more fundamental approach to reduce offset and 1/*f* noise problems has already been explored in the late forties and was called chopper-stabilization [2.1]. At that time it was used to transport DC signals through vacuum tubes. Nowadays, all commercially available low-offset CMOS operational amplifiers use a dynamic offset-cancellation technique [2.2, 2.3, 2.4].

Comparing conventional trimming with dynamic offset-cancellation techniques, the latter has the advantages of:

- *Reduction of both offset and 1/f noise*
- *Excellent long-term stability*
- *No additional costs at test level*

Dynamic offset-cancellation techniques also have disadvantages. The first is the higher complexity of the circuit and the corresponding increase of chip area. A second, usually much more important disadvantage, is the unwanted mixing of input signals and modulation signals. This can be explained as follows. All dynamic offset-cancellation techniques perform a periodic sampling or switching of the input signal. The frequency of this control signal is usually in the order of 100Hz to 100kHz. If the input signal has components that are close to the frequency of the dynamic offset-cancellation control signals, a kind of aliasing occurs. This implies that the original input signal is distorted and can not be perfectly reconstructed anymore. The only good method to prevent this from happening is limiting the bandwidth of your input signal to half the frequency at which the dynamic offset-cancellation is performed.

In this chapter a systematic overview of all existing dynamic offset-cancellation techniques will be given. But before that, we will first analyse and model the offset and noise of CMOS amplifiers more accurately.

2.1.1 Offset and noise in CMOS amplifiers

If we perform a noise spectrum analysis on a standard CMOS operational amplifier and calculate it back to the input voltage of that amplifier, we retrieve an image as shown in figure 2-1.

We see that for high frequencies the input referred voltage noise is white. At these frequencies the noise is dominated by shot noise and/or thermal noise. This area is usually called the thermal noise floor. At lower frequencies, however, the noise is not white but dependent on the frequency. In this part of the spectrum another

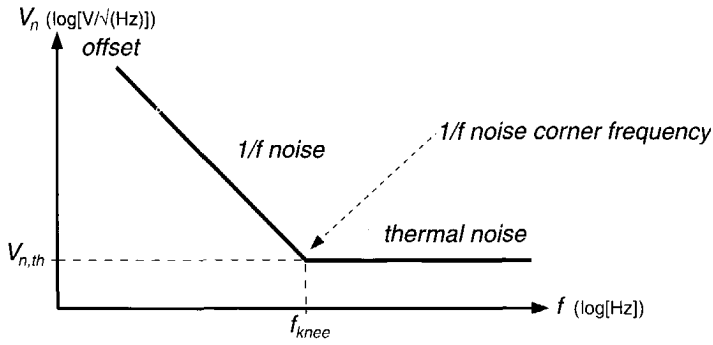

Fig. 2-1 *Noise power spectrum of standard CMOS operational amplifier*

noise source is dominant over the thermal noise, which is usually called flicker noise or $1/f$ noise. The reason for this latter name is that this noise decreases linearly with frequency. The point where the thermal noise becomes dominant over the $1/f$ noise is called the $1/f$ noise corner frequency or f_{knee}.

At very low frequencies close to DC, a third phenomenon appears that degrades the performance of our amplifier which is called offset. Offset is assumed to be time-invariant. However, offset changes due to aging or variations in temperature. This implies that it has a certain *bandwidth* and can therefore be shown in the same figure as the other noise sources. Another reason why we like to put it in the same picture is that dynamic offset-cancellation techniques reduce both the offset and the $1/f$ noise. It will appear to be much easier to explain these techniques if we consider offset as a form of ultra-low-frequency $1/f$ noise.

The three above determined parts of the noise spectrum (thermal noise floor, $1/f$ noise and offset) can be explained by the different noise sources in the CMOS transistor. These noise sources are modelled in figure 2-2.

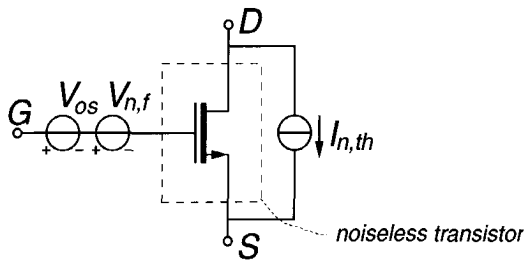

Fig. 2-2 *Offset and noise modelling of CMOS transistor*

Assuming that the transistor is operated in weak inversion, which is usually the case in low-power, low-noise applications, the following equations apply for the different noise sources in a normalized 1 Hz bandwidth [2.5]

$$I_{n,th} = \sqrt{2qI_D} = \sqrt{\frac{2kTg_m}{n}} \tag{2-1}$$

$$V_{n,f} = \sqrt{\frac{4kT\rho}{fWL}} \tag{2-2}$$

$$V_{os} = \frac{V_{os,norm}}{\sqrt{WL}} \tag{2-3}$$

where n is the sub-threshold slope factor ($n \approx 0.7$ for NMOS and $n \approx 0.5$ for PMOS), ρ is a process-dependent parameter, and $V_{os,norm}$ the offset voltage of a normalized sized transistor.

The shot noise is modelled as a current source $I_{n,th}$ because this matches better with its physical background. The flicker or $1/f$ noise $V_{n,f}$ and offset V_{os} can be modelled best as voltage sources [2.5].

For comparison reasons it is usually easier to also model the thermal noise as a voltage instead of a current. If we do that and add it to the flicker noise, we get:

$$V_n = \sqrt{\frac{4kT\rho}{fWL} + \frac{2k^2T^2}{n^2qI_D}} \tag{2-4}$$

Table 2-1 *Classification of dynamic offset-cancellation techniques*

autozeroing	**chopping**
correlated double-sampling	synchronous detection
ping-pong opamp	chopper amplifier
chopper-stabilization	chopper-stabilization
self-calibrating opamp	dynamic element matching
two or three-signal approach	

From these equations we can conclude that for a given process, the only *static* way to reduce offset and $1/f$ noise is to increase the area (WL) of the transistor, and the only way to reduce thermal noise is to increase the drain current I_D.

2.1.2 Naming conventions and classification

Due to historical reasons some confusion has been arisen in the naming conventions of the different dynamic offset-cancellation techniques. Nowadays, it is generally accepted to distinguish two main groups: autozeroing and chopping [2.6, 2.10, 2.16]. The fundamental difference between them is the handling of the offset. While the autozero principle first measures the offset and subtracts it in a next phase, chopping modulates the offset to higher frequencies, which will be explained later.

In data books and literature you can find many derivatives of these two basic offset-cancellation techniques like correlated double-sampling [2.6, 2.16], ping-pong opamps [2.8, 2.11], self-calibrating opamps [2.7], synchronous detection, the two- or three-signal approach [2.17] and dynamic element matching (D.E.M.). The name chopper-stabilization is even used for both autozeroing and chopping techniques [2.3, 2.6].

If we classify these techniques into the two main groups autozeroing and chopping, we get a result as shown in table 2-1. The next paragraphs will discuss the basics of these techniques. Practical applications will also be shown.

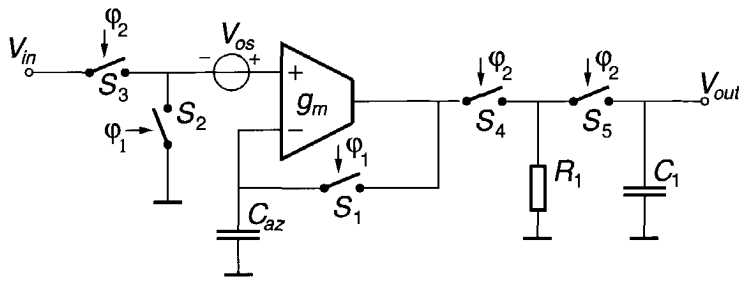

Fig. 2-3 *Autozero principle*

2.2 Autozero techniques

2.2.1 Principle

The basic principle of the autozero technique is shown in figure 2-3 and can be explained as follows [2.6, 2.9]. The offset cancellation is done in two phases, a sampling phase when φ_1 is high, and an amplification phase when φ_2 is high. In the sampling phase, the input signal V_{in} is disconnected from the transconductance amplifier g_m by switch S_3, while the non-inverting input of g_m is connected to ground by S_2. Switch S_1 is also closed and the offset V_{os} is sampled on capacitor C_{az}. In the second or amplification phase, when φ_2 is high, V_{in} is connected to the amplifier g_m. Switches S_4 and S_5 are also closed, which results in an output voltage at V_{out} of $g_m R_1$ times V_{in}. The offset V_{os} is in this phase compensated by the sampled offset voltage on C_{az}. A disadvantage of this circuit is that the output signal is not continuously available. This is solved by the addition of the sample-and-hold circuit formed by S_5 and C_1.

The residual offset of the autozero principle is determined by:

- *Charge injection on C_{az} at the opening of switch S_1*
- *Leakage on C_{az} during amplification phase φ_2*
- *Limited voltage gain of g_m*

This residual offset is minimized by choosing a large value of C_{az} and a high voltage gain of g_m. To reduce the influences of charge injection and leakage, low-sensitivity inputs are often used. Examples of this are shown in sections 2.2.5 and

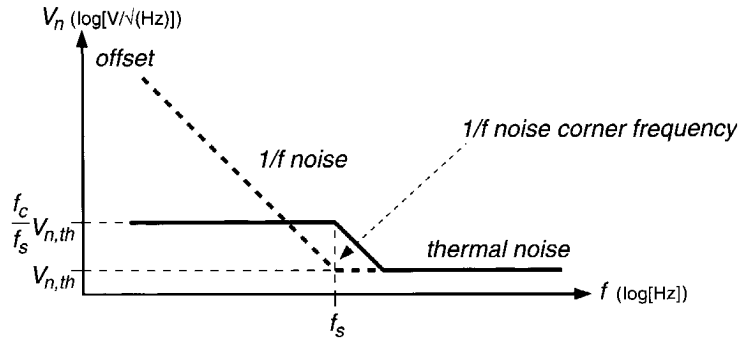

Fig. 2-4 *Noise power spectrum of autozeroed amplifier*

2.2.6. With these low-sensitivity inputs, residual offsets of 1-10μV can be achieved with capacitance values of 50pF.

2.2.2 Residual noise

As with all dynamic offset-cancellation techniques, autozeroing reduces the 1/*f* noise as well as the offset. This reduction depends on the sampling frequency and is maximal when this sampling frequency is chosen higher than the 1/*f* noise corner frequency. A quantitative calculation of this reduction is rather complicated, but has been done thoroughly by Enz et al [2.6]. In the framework of this book it is, however, not necessary to understand all details of the residual 1/*f* noise in autozeroed amplifiers. We will therefore limit ourselves to a more qualitative approach.

The voltage noise spectrum of an autozeroed amplifier is shown in figure 2-4. We see that the 1/*f* noise is removed at the cost of an increased thermal noise in the signal band. This can be explained as follows: the capacitor C_{az} and the switch S_1 in figure 2-3 act as a sample-and-hold filter. This sample-and-hold filter samples the offset and noise of the amplifier g_m, during the sampling phase φ_1. In the amplification phase φ_2, the sampled offset and noise are extracted from the offset and momentary noise of the amplifier. Because the sampling and extraction do not occur at the same time but subsequently, only the slowly changing noise components are cancelled. If we assume that the sampling frequency f_s is higher than the 1/*f* noise corner frequency, all 1/*f* noise frequency components are removed.

A disadvantage of this method is that not only the low-frequency noise is sampled on C_{az}, but the noise of the *full* bandwidth of the amplifier g_m. This results in an

increased thermal or white noise component. This effect always occurs in sampled systems and is often explained as folding of high-frequency components. How much the white noise component will increase depends on the unity-gain bandwidth of the amplifier f_c. The higher the bandwidth, the more the thermal noise will increase. In most practical cases, the increase in thermal noise is approximately equal to the amplifier's original broadband thermal noise $V_{n,th}$ multiplied by the ratio of the amplifier bandwidth f_c and the sampling frequency f_s [2.6]. This is shown in equation (2-5).

$$V_{n,res} = \frac{f_c}{f_s} V_{n,th} \qquad (2\text{-}5)$$

From equation (2-5) it can be seen that for optimal noise performance, the bandwidth of the amplifier f_c should be as low as possible. However, this bandwidth can not be chosen too small, because it needs enough speed to settle during the sampling period. In practical situations, f_c is chosen at least three times higher than f_s, which results in a residual thermal noise that is at least three times higher than the original wide-band thermal noise.

In the next paragraphs, several implementations of the autozero technique will be discussed.

2.2.3 Self-calibrating opamp

The principle of the self-calibrating opamp is shown in figure 2-5. This picture is directly copied from the Texas Instruments TLC4501 data sheet [2.7]. It is a genuine autozeroed opamp, but with the sample-and-hold circuit implemented by an A-to-D converter, a register, and a D-to-A converter. An advantage of this method is that there is no leakage on the sampling "capacitor" during the amplification phase. A disadvantage is obviously the increased chip area.

Although it is possible to run this offset correction loop at a higher frequency, which is done by Yu et al [2.8], Texas Instruments only performs an offset correction at start-up. This has the advantage that the amplifier does not suffer from spurious autozero-control signals in the signal band and does therefore not need an anti-aliasing filter.

The disadvantage is that the $1/f$ noise is not reduced and that the offset *drift* after start-up is not compensated for.

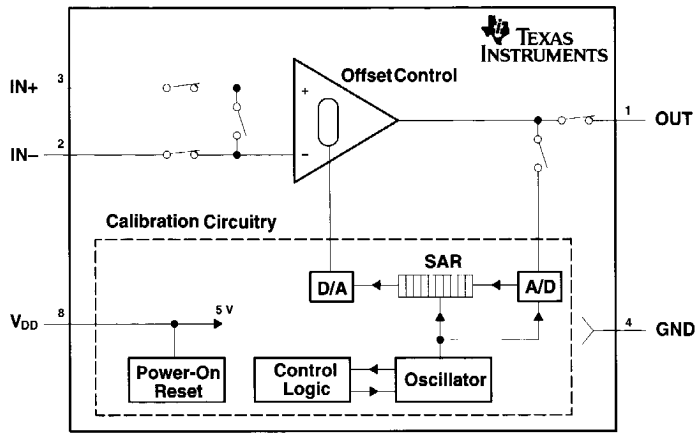

Fig. 2-5 *Self-calibrating opamp (Courtesy of Texas Instruments)*

2.2.4 Correlated double sampling

The correlated double sampling technique was originally introduced to reduce the noise in charge-coupled devices (CCDs) [2.12, 2.13]. It can be described as an autozero operation within a sample-and-hold circuit. Nowadays, it is widely used in sampled-data systems and particularly in switched-capacitor (SC) circuits. Many examples of high-accuracy SC sample-and-hold stages, voltage amplifiers, integrators, and filters can be found in [2.6].

The effect of the correlated double sampling technique on the amplifier offset and noise is equal to that of the autozero principle. However, because the input data is already sampled the white noise is not increased as is the case by a continuous-time input signal. This makes this technique widely used in sampled-data systems.

2.2.5 Ping-pong opamp

Some applications require bandwidths that are higher than the sampling frequency. In these cases a sample-and hold filter at the output can not be used. This problem can be circumvented by duplicating the autozeroed amplifier and using one amplifier in its amplification mode while the other is being autozeroed. This time-sharing principle is called "ping pong" technique in literature [2.8, 2.11] and is shown in figure 2-6.

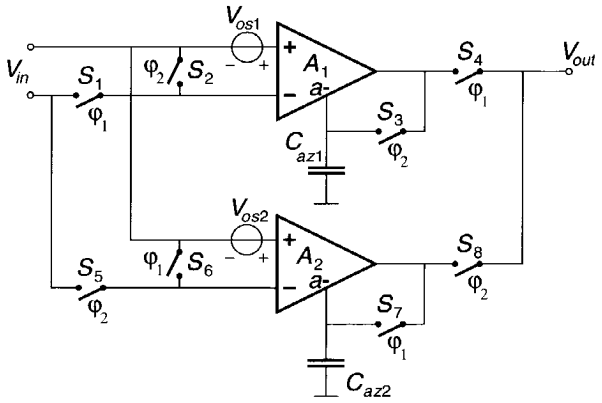

Fig. 2-6 *Ping-pong opamp principle*

The circuit consists of two amplifiers A_1 and A_2. In the first phase, when φ_1 is high, amplifier A_1 amplifies the input signal, while A_2 is being autozeroed. In the second phase, when φ_2 is high, A_2 amplifies the signal, while A_1 is being autozeroed. The amplifiers have special low-sensitivity inverting inputs *a*-. This has the advantage that the influence of charge injection on the autozeroing capacitors C_{az1} and C_{az2} is reduced and that the regular inverting input can be used for feedback circuitry. In practical situations, the inputs *a*- have approximately -40dB less sensitivity than the regular inverting inputs.

Disadvantages of this technique compared to regular autozeroing are the increased chip area and power consumption and the glitches at the output when switching the opamps. This latter problem can be reduced by the so-called chopper-stabilization technique, which is described in the next paragraph.

2.2.6 Chopper-stabilization

The chopper-stabilization technique is the industry standard nowadays for ultra-low offset CMOS opamps [2.2, 2.3, 2.4]. A better naming, however, would be continuous-time autozeroed amplifier as proposed by Enz et al. [2.6].

The principle of the chopper-stabilized opamp is shown in figure 2-7. It consists of a main opamp and a nulling opamp. The basic idea behind this circuit is to use an offset-free nulling amplifier to sense the main amplifier's offset and generate a correction voltage that is applied to the nulling input (*a*-, a low-sensitivity inverting input) of the main amplifier to cancel its own offset. To make the

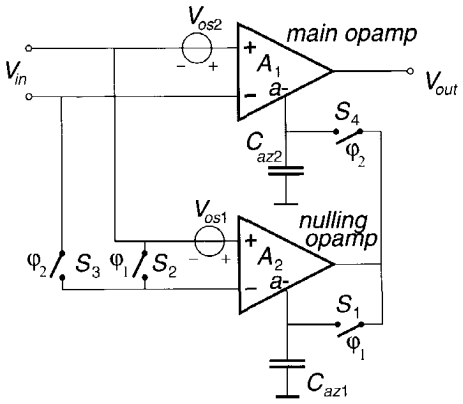

Fig. 2-7 *Chopper-stabilized opamp*

nulling opamp offset-free, it is autozeroed during phase φ_1, when switches S_1 and S_2 are closed. In the second phase φ_2, when switches S_3 and S_4 are closed, the main opamp is made offset-free. During the next φ_1 phase, the main opamp conserves its offset information on capacitor C_{az2}.

The main difference compared to basic autozeroing and ping-pong techniques is that the main opamp is not disconnected from the signal path, which reduces the glitches at the output. Disadvantages are again the higher power and chip area consumption.

2.2.7 Three-signal approach

In the previous paragraphs a number of different implementations and derivatives of the autozero principle have been discussed. All these principles cancel the offset and $1/f$ noise at the amplifier level. A more system-level approach is found in the two or three-signal approach [2.17]. In the two-signal approach, the offset is sampled during the first phase, converted to a digital format by an A-to-D converter, and stored in memory. In the second phase, the same is done with the input signal. Afterwards, the two signals are subtracted by a microcontroller. To also cancel amplification errors, the principle can be extended with a reference signal and is then called the three-signal approach. This principle is shown in figure 2-8.

In practical implementations [2.17], not a sampled but an integrating A-to-D converter is used (duty-cycle or V-f converter). As a result, the noise performance of the two and three-signal approach is not degraded by the sampling action. The

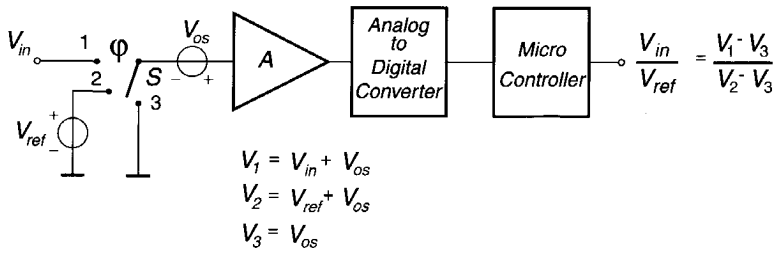

Fig. 2-8 *Three-signal approach*

noise will therefore not increase by at least a factor of five, but by a factor that is dependent on the ratio of the offset (and reference voltage) measurement time and the total time. For a two-signal approach with an offset measurement "duty cycle" of 50%, the increase in thermal noise is approximately a factor of two.

In general it can be stated that the two and three-signal approach is a very low-cost and easy offset and gain-cancellation technique if a microcontroller is already available, which is usually the case in modern data-acquisition systems.

2.3 Chopper techniques

2.3.1 Principle

As already discussed in paragraph 2.1.2, the main difference between the autozero technique and the chopper technique is the handling of the offset. While in the autozeroing principle the offset is sampled, stored and subtracted in the next phase, the chopping technique modulates the offset to higher frequencies and therefore needs an additional low-pass filter to remove the high-frequency offset components.

The chopping principle is shown in figure 2-9. The input signal V_{in} is multiplied by a square wave signal at a frequency f_{chop}. The modulated input signal is then amplified by an amplifier with gain A and demodulated back to the baseband by a second similar multiplier or chopper. The offset is, however, modulated only once and appears as frequency components around the odd harmonics of f_{chop}. These components need to be removed by a low-pass filter.

Next to the frequency domain, the chopping principle can also be explained in the time domain. The input signal V_{in} is in that case periodically inverted by the first

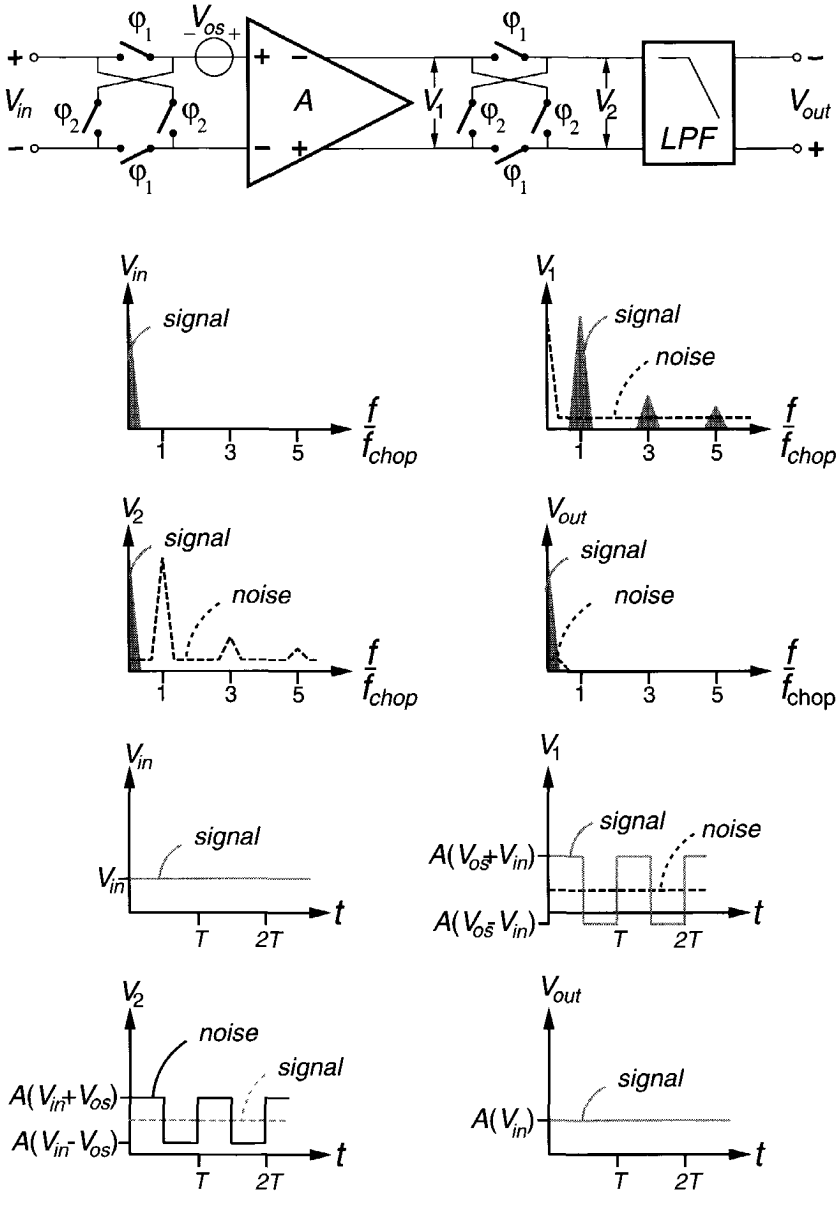

Fig. 2-9 *Amplifier using the chopper technique, including signals in the frequency and time domain*

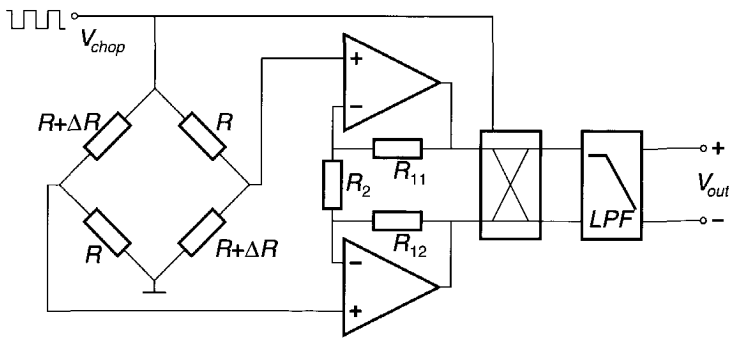

Fig. 2-10 *AC-driven Wheatstone bridge*

multiplier or chopper. After amplification, the inverted and amplified signal is inverted again, while the non-inverted part of the signal is left alone. The offset is periodically inverted only once and therefore appears as a square wave at the output.

The chopping technique is older than the autozeroing technique and was already applied with vacuum tubes in the late fifties. It is also called synchronous or coherent detection. A well-known example is the AC-driven Wheatstone bridge, which is shown in figure 2-10. The bias voltage on the bridge V_{chop} is inverted periodically, thus modulating the output signal of the bridge. This output signal is fed to the instrumentation amplifier. At the output of the amplifier the signal is demodulated again, while the offset is modulated. Finally, a low-pass filter removes the offset components.

Dynamic Element Matching (DEM) is an other derivative of the chopping technique. This technique periodically interchanges a number of elements and takes the average. The mismatch between the elements is cancelled this way. Fundamentally, the chopping principle can be seen as a DEM with two elements, for example the two input transistors which form the input stage of an amplifier. A very nice application has been described by de Jong et al [2.18].

2.3.2 Residual noise

In contrast to the increased white noise component of autozeroed amplifiers, the baseband white noise of chopped amplifiers is almost equal to the wideband thermal noise, assuming that the chopping frequency is higher than the $1/f$ noise corner frequency. The fundamental reason for this is that the input signal of a chopper amplifier is not sampled, which makes it impossible for wideband

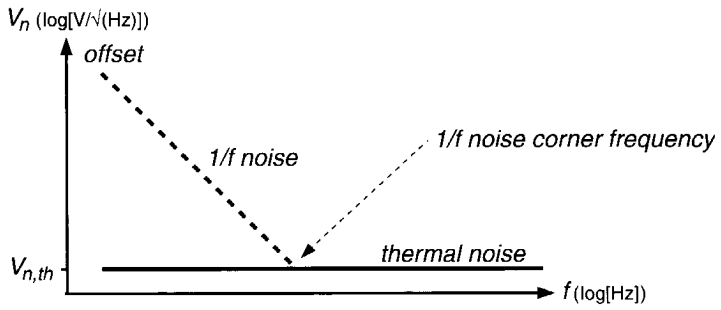

Fig. 2-11 *Noise power spectrum of chopper amplifier*

thermal noise to fold back in the baseband. The typical noise power spectrum of a chopped amplifier is shown in figure 2-11.

The superior noise performance of the continuous-time chopper technique over the sampled autozero technique, makes it the best choice to realize low-noise sensor interfaces. Since in low-bandwidth systems the power is mainly dictated by the required signal-to-noise ratio and not by the power-bandwidth product, the chopping principle is also the best choice to realize sensor interfaces with minimal power.

2.3.3 Residual offset

The residual offset of the chopper amplifier as shown in figure 2-9 is determined by the spikes of the input chopper [2.6, 2.10, 2.14]. This is schematically shown in figure 2-12. The spikes are generated around the odd harmonics of the modulating signal. They are demodulated to the baseband by the second modulator. The residual offset equals the area under the spike. Given a time-constant τ that is determined by the source impedance and the input capacitance, the residual offset can be written as

$$V_{os,\,res} = \frac{V_{spike}}{0.5T} \int_{0}^{0.5T} e^{-t/\tau} dt \qquad (2\text{-}6)$$

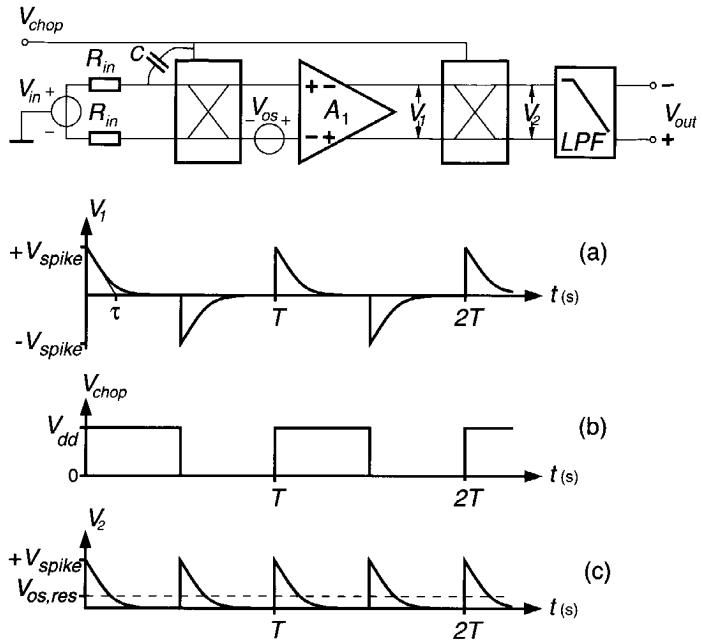

Fig. 2-12 *Residual offset caused by spikes*
(a) Spike signal (b) Demodulation signal
(c) Demodulated spike

for T >> τ this can be written as

$$V_{os,res} = \frac{2V_{spike}}{T}\int_0^{\infty} e^{-t/\tau}dt = \frac{2\tau}{T}V_{spike} \qquad (2\text{-}7)$$

Since most of the spike energy appears at high frequencies, the influence of these spikes on the residual offset can be reduced using a band- or low-pass filter between the modulator and demodulator. This is implemented by Menolfi et al [2.14, 2.15] and shown in figure 2-13. Because of the tuning error between the filter and the oscillator of 1%, the quality of the filter can not be set higher than five. With this quality factor, the residual offset can be improved by approximately a factor five. The analysis has been done very accurately by Menolfi et al [2.14, 2.15]. They report a chopper amplifier with bandpass filter,

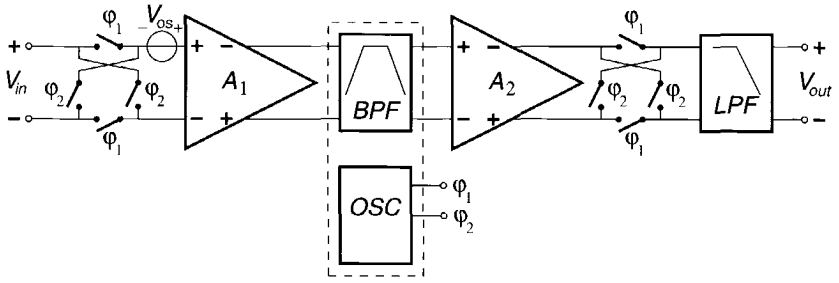

Fig. 2-13 *Chopper amplifier with bandpass filter to improve residual offset*

having a 500nV offset at a source impedance of 20kΩ and a chopping frequency of 5kHz.

2.3.4 Gain accuracy

The gain accuracy of the chopper amplifier principle, which is shown in figure 2-9, is limited by the bandwidth of the amplifier *A*. This is explained in figure 2-14. The output voltage for $\tau_c << T$ is given by:

$$V_{out} = \left(1 - \frac{4\tau_c}{T}\right) A V_{in} \qquad (2\text{-}8)$$

Where τ_c is inversely proportional to the bandwidth of the amplifier f_c.

The gain error can be reduced by delaying the demodulation signal with $0.7\tau_c$, see figure 2-14 (d). The output signal is then given by:

$$V_{out} = \left(1 - \frac{2\tau_c}{T}\right) A V_{in} \qquad (2\text{-}9)$$

If we consider an amplifier with a unity-gain bandwidth of 10Mhz and a dc-gain of 100, f_c is 100kHz. A chopping frequency of 5 kHz then already gives an error of 10%, even if we reduce the gain error by delaying the demodulation signal! It is shown from this example that for high gain accuracies, high bandwidth amplifiers are needed, which imply high power consumption. A more power-efficient solution can be found in the chopper opamp as discussed in the next paragraph.

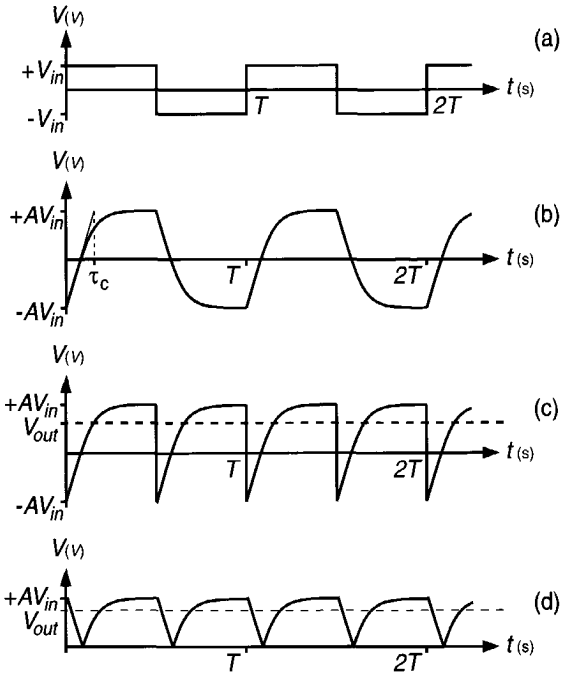

Fig. 2-14 *Gain errors due to limited amplifier bandwidth*
(a) Modulated input signal
(b) Amplified modulated input signal
(c) Output signal with in-phase demodulation signal
(d) Output signal with delayed demodulation signal

2.3.5 Chopper opamp

An interesting derivative of the chopper amplifier principle is the chopper *opamp* as shown in figure 2-15. Amplifier A_1 is a regular opamp with high dc-gain. Amplifier A_2 has a limited gain and serves mainly to generate a single-ended output.

In contrast to the above-mentioned chopper amplifier, the chopper opamp does not modulate the input signal, but only the *error* signal. This implies that all errors due to spikes and limited gain are independent of the input signal and thus do not contribute to the gain error but to the residual offset.

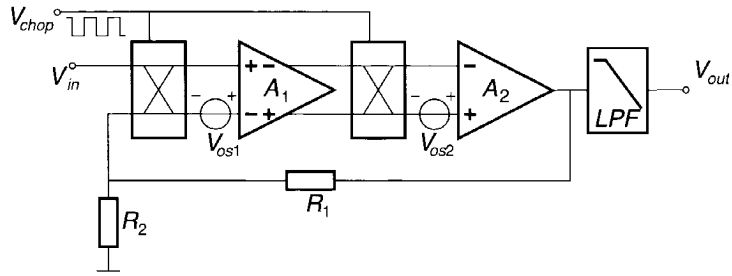

Fig. 2-15 *Chopper opamp*

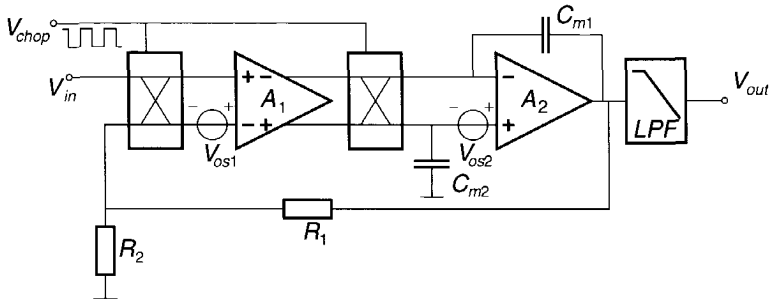

Fig. 2-16 *Chopper opamp with Miller pole-split*

Another source for residual offset is caused by the limited gain of amplifier A_1 and the second amplifiers offset V_{os2} and is given by:

$$V_{os,res} = \frac{2f_{chop}}{f_{0,A_1}} V_{os2} \qquad (2\text{-}10)$$

Assuming that the first amplifier A_1 has a unity-gain bandwidth ($f_{0,A1}$) of 10MHz, a chopping frequency of 5kHz and a second amplifier's offset of 5mV, the residual offset caused by limited gain is 5µV.

A method to increase the bandwidth of the first opamp is already reported by Hsieh et al [2.16] in 1981. Their principle is shown in figure 2-16. By adding a Miller capacitor across the second stage of the opamp, the pole of the first amplifier is shifted to higher frequencies. Assuming that the tail current of the input stage is already high for noise reasons, the bandwidth of the first stage can be increased to frequencies up to 100MHz, without any increase of power

consumption! This makes it possible to neglect the contribution of the limited gain to the residual offset as shown in equation (2-10).

The conclusion of this paragraph is that the chopper opamp with Miller split, as shown in figure 2-16, is the best implementation of the chopping principle regarding gain accuracy and power consumption. The offset performance is, however, worse than the chopper amplifier with bandpass filter.

In the next paragraph we will present a new technique to improve the offset performance of the chopper opamp, and bring it to a level that is higher than can be achieved by chopper amplifiers with bandpass filter.

2.4 Nested chopper technique

2.4.1 Principle

At first glance, equation (2-7) on page 24 does not leave much room to improve the residual offset of the chopper amplifier. The product τ times V_{spike} can hardly be reduced because it is dictated by the source impedance and by process parameters like the gate-drain overlap capacitance. Also the period T, which is inversely proportional to the chopping frequency f_{chop}, is dictated by the $1/f$ noise corner frequency. This $1/f$ noise corner frequency can only be reduced at the cost of higher thermal noise, or at the cost of larger input transistors, which are needed for lower $1/f$ noise.

However, considering a chopper amplifier as a regular amplifier with a small residual offset, but without $1/f$ noise, we can reduce this offset by doing another chopping action. This is shown in figure 2-17. The chopping frequency of this outer chopping pair is now not dictated by the $1/f$ noise corner frequency, but can be chosen as low as possible, usually two times the bandwidth of the input signal.

Since the residual offset caused by spikes is linearly dependent on f_{chop} according to equation (2-7), the influence of these spikes on the residual offset can be considerably reduced. If we look at a temperature sensor with a bandwidth of 10Hz for example, the outer chopper frequency can be as low as 20Hz, which is the minimal frequency at which no aliasing occurs. With an inner chopper frequency of 2kHz, which is determined by the $1/f$ noise corner frequency, the improvement can theoretically be a factor of 100.

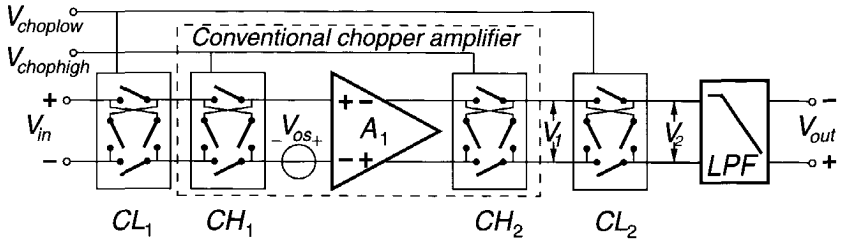

Fig. 2-17 *Nested chopper amplifier principle*

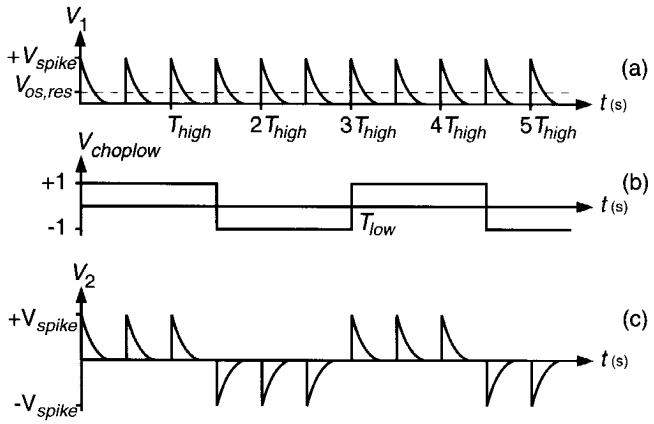

Fig. 2-18 *Reduction of residual offset by nested chopper technique*
(a) Spikes after first demodulator
(b)Low-frequency modulation signal
(c) Spikes after second demodulator

This short description of the nested chopper technique shows that a major improvement over the conventional chopper technique can be obtained with a rather simple extension. The extra chopper pair occupies negligible chip area and does not consume power.

2.4.2 Analysis

An analysis of the nested chopper technique on the residual offset caused by spikes is shown in figure 2-18. Figure2-18 (a) shows the result at the output of the second high-frequency chopper CH_2, which is the same as for a conventional chopper amplifier. Figure2-18 (c) shows the spikes at the output of the second low-frequency chopper CL_2. The spikes are again modulated but now at a much

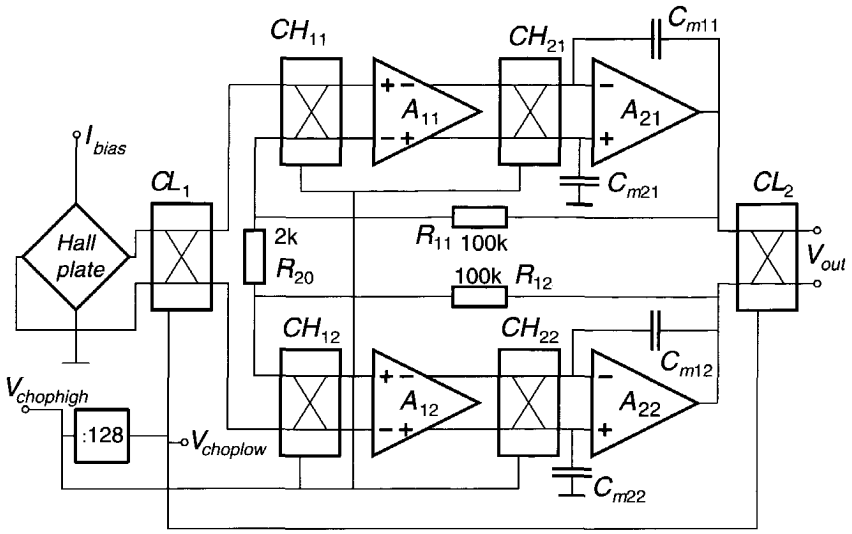

Fig. 2-19 *Nested chopper amplifier realization*

lower frequency with a period T_{low}. By this second modulation, the average value of the spikes becomes zero and therefore they no longer contribute to the residual offset. The extra spikes that are generated because of this low-frequency chopper pair can be neglected if $T_{low} >> T_{high}$.

In the figure it is suggested that the spikes are not influenced by the input low-frequency chopper CL_1. This is only true if the impedances of the input source at both inputs of the amplifier are completely matched, which is in practice usually not the case. This mismatch in source impedances will be the main source of residual offset of the nested chopper technique. Nevertheless, the residual offset of a nested chopper amplifier will be 10 to 50 times better than that of a conventional chopper amplifier, depending on the mismatch of the source impedances.

To examine the performance of the proposed nested chopper amplifier technique, a realization has been designed and tested. The results will be shown in the next paragraphs.

2.4.3 Realization

The realization of the nested chopper amplifier is shown in figure 2-19. A choice has been made for an instrumentation amplifier to have a good matching of input

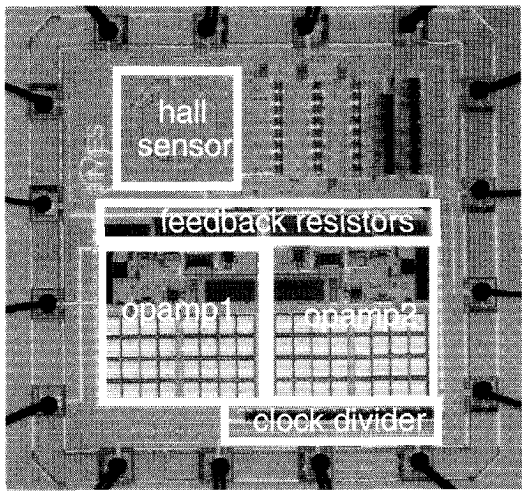

Fig. 2-20 *Chip photograph of nested chopper instrumentation amplifier*

impedances. The high-frequency choppers CH_{11} to CH_{22} are configured as chopper opamps with Miller pole-split as discussed in paragraph 2.3.5. The low-frequency pair (CL_1 and CL_2) modulates and demodulates the input signal as in a regular chopper amplifier (see paragraph 2.3.1). The gain is determined by R_{11}, R_{12} and R_{20} and is set to 100. The ratio between the higher chopping frequency and the lower chopping frequency is fixed and set to 128. The bandwidth of the amplifier is around 4kHz, power consumption is 200µA.

The choice for the input signal needs to be done very carefully. Applying an off-chip input signal will give rise to unwanted thermocouple effects and the protection diodes will add non-linearity and leakage currents. The best way is to add a Wheatstone bridge or a Hall plate on chip. We made the choice for a spinning-current Hall plate [2.19], because this directly shows an important application. The offset of the nested chopper instrumentation amplifier can be measured by switching off the bias current of the Hall plate.

The chip has been made in the in-house DIMES standard CMOS, 1.6µm, single poly process. A chip photograph is shown in figure 2-20. The chip measures 6mm^2 of which 1mm^2 is occupied by the instrumentation amplifier and another 1mm^2 by the high-linearity metal1-metal2 Miller capacitors.

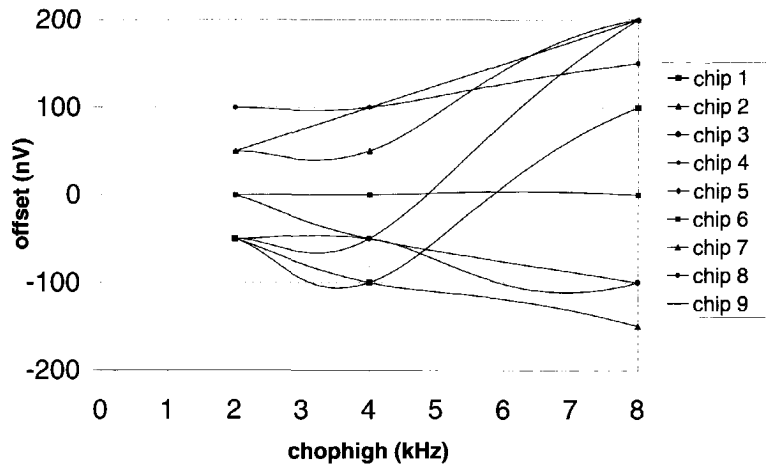

Fig. 2-21 *Input referred offset versus chophigh frequency, resolution of measurement is 50nV*

2.4.4 Measurement results

Nine samples of the chip have been tested. The output of the amplifier is connected to an external amplifier/filter, which has a 20-times amplification and a 3Hz bandwidth. For dc measurements, a multimeter with 500ms integration time has been used.

The offset as a function of the higher chopping frequency is shown in figure 2-21. It can be seen that the residual offset at a chophigh frequency of 2kHz is below 100nV for all samples. This offset becomes worse at higher frequencies, but stays below 200nV at 8kHz. The other specifications are listed in table 2-2.

These results show that the nested chopper amplifier principle works. It gives an improvement of 5 times over the best reported value so far by Menolfi [2.15]. Over regular chopper amplifiers it gives a 10 to 50 times improvement.

Given the fact that the charge injection is dependent on switch widths, even better results are to be expected in up-to-date processes.

Table 2-2 *Specifications of the nested chopper instrumentation amplifier*

	Typ	Unit
Supply voltage	5	V
Supply current	200	μA
High chopping frequency	2	kHz
Low chopping frequency	16	Hz
Offset	100	nV
Offset temperature coefficient (20-50°C)	3	nV/°C
Noise (input referred)	27	nV/√(Hz)
CMRR (0-3Hz)	140	dB

2.5 Conclusions

In this chapter it has been shown that, although many different names are being used, basically only two different dynamic offset-cancellation techniques exist, namely the autozeroing and the chopping technique. Autozeroing is a sampling technique, while chopping is a modulating technique.

The main difference in performance is the residual noise in the signal band. The chopping technique reduces the 1/f noise to the minimal thermal noise floor, while autozeroing has an increased thermal noise due to the sampling action.

An extension of the chopper technique is proposed, called the nested chopper technique. This new technique can reduce the residual offset to values below 100nV, while keeping the noise at the minimal thermal noise floor. The effectiveness of the nested chopper technique has been proved by a realization and measurement results.

Based on the above-mentioned results, the nested chopper technique is the best choice for realization of high-accuracy CMOS amplifiers and will therefore be used as a starting point for the rest of the design and implementation of our high-accuracy CMOS smart temperature sensor.

References

[2.1] A.J. Williams, R.E. Tarpley and W.R. Clark, "D-C amplifier stabilized for zero and gain", *Trans. AIEE*, vol.67, pp.47-57, 1948

[2.2] Maxim Integrated Products, "ICL7650, Chopper Stabilized Operational Amplifier", *http://www.maxim-ic.com*, June 1999

[2.3] Linear Technology, "LTC1050, Precision chopper stabilized operational amplifier with internal capacitors", *http://www.linear.com*, June 1999

[2.4] National Semiconductor, "LMC2001, High Precision, 6MHz Rail-To-Rail Output Operational Amplifier", *http://www.national.com*, June 1999

[2.5] J.F. Franca and Y.P.Tsividis, "Design of VLSI Circuits for Telecommunications and Signal Processing", *Prentice Hall*, 1991

[2.6] C.C. Enz and G.C. Temes, "Circuit techniques for reducing the effects of opamp imperfections: autozeroing, correlated double sampling and chopper stabilization", *Proc. of the IEEE*, vol.84, no.11, pp. 1584-1614, november 1996

[2.7] Texas Instruments, "TLC4501, Self-calibrating operational amplifier", *http://www.ti.com*, May 1999

[2.8] C.G. Yu and R.L. Geiger, "An automatic offset compensation scheme with ping-pong control for CMOS operational amplifiers", *IEEE Journal of Solid-State Circuits, vol.29,* pp.601-610, May 1994

[2.9] M. Degrauwe, E.A. Vittoz and I. Verbauwhede, "A micropower CMOS Instrumentation amplifier", *IEEE Journal of Solid-State Circuits, vol.20,* pp.805-807, June 1985

[2.10] C.C. Enz, E.A. Vittoz and F. Krummenacher, "A CMOS chopper amplifier," *IEEE Journal of Solid-State Circuits*, vol.22, pp.335-342, June 1987

[2.11] I.E. Opris and G.T.A. Kovacs, "A rail-to-rail ping-pong opamp", *IEEE Journal of Solid-State Circuits, vol.31,* pp.1320-1324, September 1996

[2.12] K.H. White, D.R. Lampe, F.C. Blaha and I.A. Mack, "Characterization of surface channel CCD image arrays at low light levels," *IEEE Journal of Solid-State Circuits,* vol.9, pp.1-14, February 1974

[2.13] R.W. Brodersen and S.P. Emmons, "Noise in buried channel charge-coupled devices", *IEEE Journal of Solid-State Circuits,* vol.11, pp.147-156, February 1976

[2.14] C. Menolfi and Q. Huang, "A low-noise CMOS instrumentation amplifier for thermoelectric infrared detectors, *IEEE Journal of Solid-State Circuits,* vol.32, pp.968-976, July 1997

[2.15] C. Menolfi and Q. Huang, "A fully integrated CMOS instrumentation amplifier with submicrovolt offset", *IEEE Journal of Solid-State Circuits,* vol.34, pp.415-420, March 1999

[2.16] K.C. Hsieh, P.R. Gray, D. Senderowicz and D.G. Messerschmitt, "A low-noise chopper stabilized differential switched-capacitor filtering technique, *IEEE Journal of Solid-State Circuits, vol.16,* pp.708-715, December 1981

[2.17] G.C.M. Meijer, "Concepts and focus point for intelligent sensor systems", *Sensors and Actuators,* vol.41, pp.183-191, 1994

[2.18] P.C. de Jong, G.C.M. Meijer and A.H.M. van Roermund, "A 300°C dynamic-feedback instrumentation amplifier", *IEEE Journal of Solid-State Circuits,* vol.33, pp.1999-2009, December 1998

[2.19] A. Bakker and J.H.Huijsing, "Low-offset, low-noise 3.5mW CMOS spinning-current Hall-effect sensor with integrated chopper amplifier", *Proc. Eurosensors XIII,* The Hague, The Netherlands, pp. 1045-1048, September 1999

CMOS bandgap references

3

Voltage references are key elements in analog and mixed-mode circuits. They determine the overall accuracy in many data-acquisition systems. The accuracy of our smart temperature sensor is also mainly dependent on the quality of its voltage reference.

This chapter describes bandgap voltage references with emphasis on CMOS technology. It discusses available bipolar devices in CMOS technology, thermal modelling of the bipolar transistor, and curvature correction techniques. Also the design and implementation of a novel high-accuracy CMOS bandgap reference is shown.

3.1 Introduction

Nowadays the bandgap reference is the most often used implementation of a voltage reference. It was invented in the early seventies [3.1, 3.2], when people started to explore integrated circuits. The popularity of the bandgap voltage reference is based on its outstanding stability and reproducibility. It is called a *bandgap* reference, because it exploits the bandgap voltage of the material it is made of. Because silicon is the most widely used semiconductor material, many commercial references have a value of approximately 1.2V or a multiple of that.

A typical bandgap reference circuit and its behaviour over temperature are shown in figure 3-1. It is explained as follows. The temperature coefficient of the base-

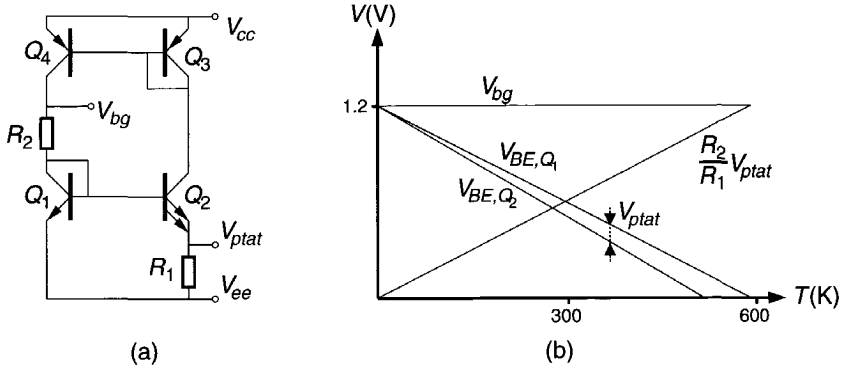

Fig. 3-1 *Typical bandgap reference circuit (a) and temperature curves (b)*

emitter voltage V_{BE} of a bipolar transistor is approximately -2 mV/°C. To obtain the desired temperature-stable voltage, a signal with an equal but positive temperature coefficient needs to be added. This voltage is derived by subtracting the base-emitter voltages of two bipolar transistors (Q_1 and Q_2) with different emitter areas but biased at the same current. Assuming an emitter area ratio of n, the voltage difference can be written as

$$V_{ptat} = \frac{kT}{q} \ln n \qquad (3\text{-}1)$$

where k is the Boltzmann's constant and q the electron charge. The ratio kT/q is called the thermal voltage and equals approximately 26 mV at room temperature (T=300K). This voltage difference is called V_{ptat} because it is proportional to the absolute temperature T. The temperature coefficient of the basic V_{ptat} signal yields approximately 200µV/°C for n is 10. To obtain the necessary temperature coefficient of 2 mV/°C, it needs to be amplified by approximately 10. This amplification factor is set by the ratio of R_2 and R_1.The bandgap reference circuit described above, fitted perfectly in bipolar technology, which was the standard technology of the seventies. When CMOS became the leading technology for digital and mixed-mode circuits, a lot of effort was put into implementing the bandgap reference principle in CMOS technology. The first description of a CMOS-compatible bandgap reference was published in 1978 [3.3].

The choice for the best device in CMOS technology is less easy than it is in bipolar technology. A conventional MOS transistor does not exhibit a very well defined and reproducible temperature characteristic. It also suffers from large

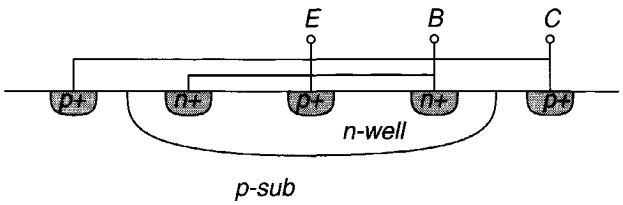

Fig. 3-2 *Substrate PNP transistor in n-well CMOS technology*

mismatch and spread in V_t. MOS transistors operated in weak inversion have a much better defined temperature dependence, which make them suitable for generating PTAT voltages, but the poor absolute accuracy remains. Vittoz et al [3.4] did a lot of research on CMOS compatible lateral bipolar transistors, which are available in all standard CMOS processes. These devices have the same temperature characteristics as regular bipolar transistors and some bandgap references have been reported using this device [3.4, 3.5]. However, the lateral bipolar suffers from very large spread in current gain (10-300) [3.4] and is very poorly described and monitored by the process suppliers. This makes it very difficult to make a commercial version of this type of CMOS bandgap reference.

The best device for implementation of a CMOS bandgap reference is by far the parasitic substrate bipolar transistor. This device is available in every CMOS process and is very well defined and highly reproducible. In addition, the monitoring by the various suppliers is improving, since it is being used more often. Most of the reported CMOS bandgap references are based on this substrate bipolar transistor [3.6, 3.7, 3.8, 3.9, 3.10, 3.11].

The substrate bipolar transistor is in a modern n-well CMOS process a pnp-type which is formed by the p+ source-drain implant, the n-well and the p-substrate. This is shown in figure 3-2. The main disadvantage, which becomes directly clear, is that the collector is formed by the substrate. Consequently, some people call it a diode. It is, however, a regular pnp-transistor of which the collector is connected to substrate. An important advantage of this device is its high accuracy. The base is rather large, which gives small spread in base-width. Compared to the small base-width in general frequency-optimized bipolar processes, the bipolar substrate transistor in CMOS technology has better absolute base-emitter voltage accuracy and also the mismatch is smaller. This is also concluded in [3.12, 3.13].

In this paragraph we only discussed the basic principles of bandgap references and possible implementations in CMOS technology. The next paragraph will focus more deeply on the temperature modelling of the substrate bipolar transistor and the error sources in CMOS bandgap references.

3.2 Temperature curves of the bipolar transistor

The temperature characteristics of the bipolar transistor are well known and show almost no dependency on process parameters. Meijer [3.14] made a very useful description of the bipolar transistor for application in temperature sensors, which we will follow in this paragraph.

For a bipolar transistor with a zero base-collector voltage, the collector current I_C as a function of the base-emitter voltage V_{BE} can be written as

$$I_C = I_s \exp\left\{\frac{qV_{BE}}{kT}\right\} \tag{3-2}$$

where I_s is the saturation current, q the electron charge, k Boltzmann's constant and T the absolute temperature. Expressing V_{BE} as a function of I_C, we get

$$V_{BE} = \frac{kT}{q}\ln\frac{I_C}{I_s} \tag{3-3}$$

This equation seems to imply that V_{BE} increases with temperature, which is however not true. The reason for this is that the saturation current I_s is very strongly dependent on temperature. Taking this into account, we can write equation (3-3) as

$$V_{BE} = V_{g0} + \frac{kT}{q}\ln\frac{I_C}{CT^{\eta}} \tag{3-4}$$

where V_{g0} is the extrapolated bandgap voltage at zero Kelvin and C and η are constants. Because $I_C/CT^{\eta} < 1$, the value of the ln function is negative, which gives the V_{BE} function a negative temperature coefficient.

Introducing a new constant $V_{BE}(T_r)$, which is the base-emitter voltage at a chosen reference temperature T_r, equation (3-4) can be written as

$$V_{BE}(T) = V_{g0}\left\{1-\frac{T}{T_r}\right\} + \frac{T}{T_r}V_{BE}(T_r) + \frac{kT}{q}\eta\ln\frac{T_r}{T} \tag{3-5}$$

This equation is derived in the appendix on page 117. In practice, I_C is usually not constant over temperature, but for example PTAT (proportional to absolute temperature). This can be introduced by writing I_C as

$$I_C = K_{IC}T^m \tag{3-6}$$

A constant I_C is then represented by m=0 and a PTAT I_C by m=1. Equation (3-4) then becomes

$$V_{BE} = V_{g0} + \frac{kT}{q}\ln\frac{K_{IC}}{CT^{(\eta - m)}} \tag{3-7}$$

and equation (3-5) will change into

$$V_{BE}(T) = V_{g0}\left\{1-\frac{T}{T_r}\right\} + \frac{T}{T_r}V_{BE}(T_r) + \frac{kT}{q}(\eta - m)\ln\frac{T_r}{T} \tag{3-8}$$

For further analysis and design, it is easier to write equation (3-8) as the sum of a constant term (V'_{g0}), a linear term (λT), and higher-order terms ($O(T^2)$). Equation (3-8) then becomes:

$$V_{BE}(T) = V'_{g0} - \lambda T + O(T^2) \tag{3-9}$$

where

$$V'_{g0} = V_{g0} + (\eta - m)\frac{kT_r}{q} \tag{3-10}$$

$$\lambda = \frac{1}{T_r}\{V'_{g0} - V_{BE}(T_r)\} \tag{3-11}$$

$$O(T^2) = (\eta - m)\frac{k}{q}\left\{T - T_r + T\ln\frac{T_r}{T}\right\} \tag{3-12}$$

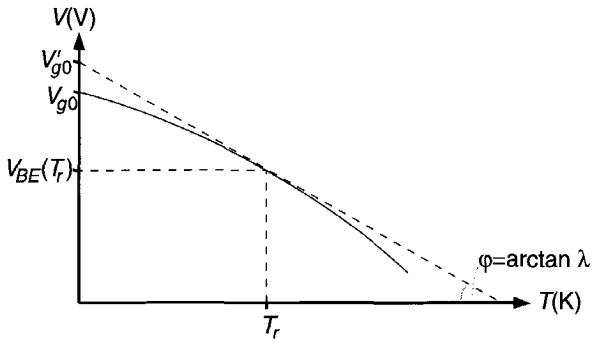

Fig. 3-3 *Plot of the base-emitter voltage $V_{BE}(T)$ versus temperature T*

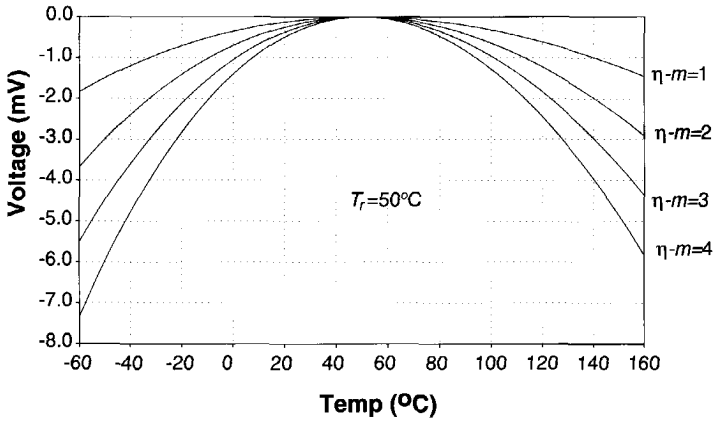

Fig. 3-4 *Plot of the non-linearity versus the temperature for different values of η-m*

A plot of the base-emitter voltage $V_{BE}(T)$ versus temperature T according to equation (3-9) is shown in figure 3-3. The non-linearity $O(T^2)$ is plotted in figure 3-4.

In the next paragraphs, design strategies for high-accuracy bandgap references are presented. A realization and measurement results are also shown.

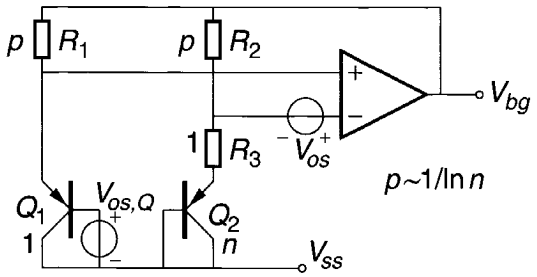

Fig. 3-5 *Traditional CMOS bandgap reference with non-idealities*

3.3 Design

3.3.1 Non-idealities in practical CMOS bandgap references

A widely used practical CMOS bandgap reference is shown in figure 3-5. The main difference of this circuit compared to the bipolar version, as shown in figure 3-1, is the presence of an operational amplifier. This opamp is needed because the collector of the substrate transistor can not be connected. It is therefore not possible to measure I_C. However, by means of the opamp the emitter-current I_E can be measured, which is a good alternative if the current-gain β is high or well known. The output voltage V_{bg} yields:

$$V_{bg} = V_{BE,Q_1} + p\frac{kT}{q}\ln n \qquad (3\text{-}13)$$

The parameters p and n are chosen in such a way that the second term in equation (3-13) exactly cancels the first-order temperature dependence of $V_{BE,Q1}$. According to equation (3-11) this yields:

$$p\frac{k}{q}\ln n = \frac{1}{T_r}\{V_{g0}' - V_{BE}(T_r)\} \qquad (3\text{-}14)$$

Practical values of n and p are 8 and 10 respectively. When the first-order temperature dependence is completely cancelled, the bandgap voltage equals V_{g0}' at T_r, which is approximately 1250 mV.

Table 3-1 *Errors in CMOS bandgap references at room temperature*

Kind of error	Typical value	Error contribution for typical value		Error contribution when applying dynamic offset-cancellation	
$\sigma(V_{os})$	0.4mV	4mV	0.33%	0.4mV	0.03%
$\sigma(V_{os,Q})$	0.1mV	1mV	0.08%	1mV	0.08%
$\sigma(\Delta R)$	0.1%	0.6mV	0.05%	0.6mV	0.05%
$\sigma(\Delta R_{sh})$	4%	1mV	0.08%	1mV	0.08%
$\sigma(\Delta V_{be})$	2mV	2mV	0.17%	2mV	0.17%
Total:		4.7mV	0.39%	2.5mV	0.21%

Errors uncorrelated, V_{be} = 600 mV, T = 300K, p=10, n=8

Looking at this CMOS bandgap reference, we can determine several non-idealities. The major one is the offset of the opamp V_{os}. Other non-idealities are the offset of the bipolar transistors $V_{os,Q}$, the mismatch of the resistors ΔR, the spread in the sheet resistance ΔR_{sh} and the spread in the base-emitter value ΔV_{BE}. Taking these errors into account, equation (3-13) becomes:

$$V_{bg} = V_{BE,Q_1} + \Delta V_{BE} + \frac{kT}{q}\Delta R_{sh} + p\,(1+\Delta R)\left\{\frac{kT}{q}\ln n + V_{os} + V_{os,Q}\right\} \quad (3\text{-}15)$$

The estimated values of the above-mentioned non-idealities and their contributions to the total error of the bandgap voltage at room temperature are shown in table 3-1. This table shows that a total typical error of 4.7mV or 0.39% can be expected. The largest contributor to this error is by far the offset of the opamp. This error can however be reduced by at least a factor of ten, if appropriate dynamic offset-cancellation techniques are applied (see chapter 2). Doing this, the total error in the bandgap reference can be reduced to typically 2.5mV or 0.21% as shown in table 3-1. This indicates that by removing the offset of the opamp, the overall accuracy of the bandgap reference circuit will be improved approximately two times. The largest contributor to the error then

becomes the spread in V_{be}. This error is however due to process variations and can not be reduced by circuit techniques.

Considering errors not only at room temperature but also for a larger temperature range, the non-linearity of V_{BE} will also add to the total error. Figure 3-4 shows that this can be up to 5mV or 0.4% over a temperature range from -50°C to 150°C. Comparing this to other error sources, this error will be the dominant one over an extended temperature range. Although this problem was encountered many years ago, techniques to solve this non-linearity problem are still being developed.

An overview of these so-called curvature correction techniques is given in the next paragraph. We will also propose a new class of techniques, which can be very useful in situations where trimming is not possible.

3.3.2 Curvature correction techniques

To obtain a high-accuracy bandgap reference, it is not sufficient to consider inaccuracies only at room temperature. Especially the non-linearity of the V_{BE}-curve over temperature needs more attention. Looking at figure 3-4 and introducing values η=4 ([3.6, 3.14]), m=1 (PTAT bias current) and T_r=50°C, the non-linearity can be up to 5mV at -50°C and +150°C. Comparing this to the total error at room temperature of 2.5mV, it is obvious that the non-linearity is the dominant error for temperatures below 0°C and above 100°C.

To reduce the inaccuracy due to the non-linearity, several so-called curvature correction techniques have been developed. The first ones appeared in the early eighties for bipolar references [3.15, 3.16], followed a few years later by CMOS [3.17]. Meijer describes a few of these early techniques in his Ph.D. thesis [3.18]. They were based on the application of resistors with different temperature coefficients. It was obvious that only poor accuracy could be obtained, due to the matching problems of the different types of resistors.

Meijer himself proposed a new technique that is based on the so-called linearized thermal behaviour of V_{BE}. This principle is still the starting point for many modern curvature correction techniques. Gunawan et al [3.19] adapted it for lower supply voltages. The principle of this curvature correction technique is shown in figure 3-6.

The non-linear correction voltage V_{nl} is generated by the extraction of two base-emitter voltages of which one (Q_3) is biased at a PTAT current and the other (Q_4) at a temperature-constant current . This differential voltage is amplified by η-1

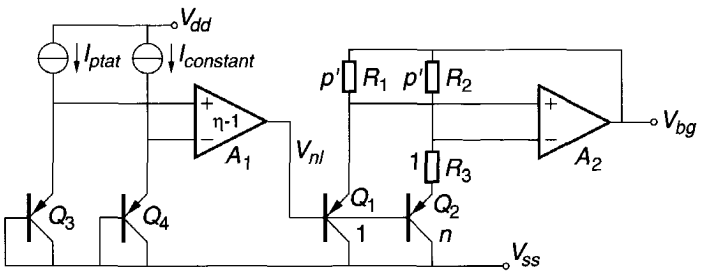

Fig. 3-6 *Curvature correction technique based on linearized V_{BE}*

and added to the conventional bandgap voltage. The emitter areas of Q_3 and Q_4 are chosen in such a way that at the chosen reference temperature T_r the corresponding base-emitter voltages are equal.

The correction voltage yields:

$$V_{nl} = (\eta - 1)\{V_{BE}(T_r, m=1) - V_{BE}(T_r, m=0)\} \tag{3-16}$$

Applying equation (3-9), this becomes:

$$V_{nl} = (\eta - 1)\left\{\frac{-kT_r}{q} + \frac{kT}{q} - \frac{k}{q}\left[T - T_r + T\ln\frac{T_r}{T}\right]\right\} \tag{3-17}$$

The last term in this equation is exactly opposite to the non-linearity term shown in equation (3-12). Equation (3-17) also shows a linear term $(\eta\text{-}1)kT/q$, which makes it necessary to slightly adapt the parameter p, resulting in a p'. The absolute value of the bandgap voltage is also decreased by a factor $(\eta\text{-}1)kT_r/q$, which makes it now exactly equal to V_{g0}.

A good implementation of the curvature correction circuit of figure 3-6 is not so straightforward. Amplifier A_1 has offset that is amplified by η-1 and added to the output voltage. Also the voltage V_{nl} can both be positive and negative, which requires a different method of adding it to the bandgap voltage. A solution in the current domain, as proposed by Gunawan et al [3.19], requires multiple current mirrors, which again add to the inaccuracy. Until now, all proposed curvature correction techniques use calibration to obtain the required accuracy.

Because our objective is to disregard calibration to reduce costs, we will introduce a new curvature correction technique that is based on a piece-wise-

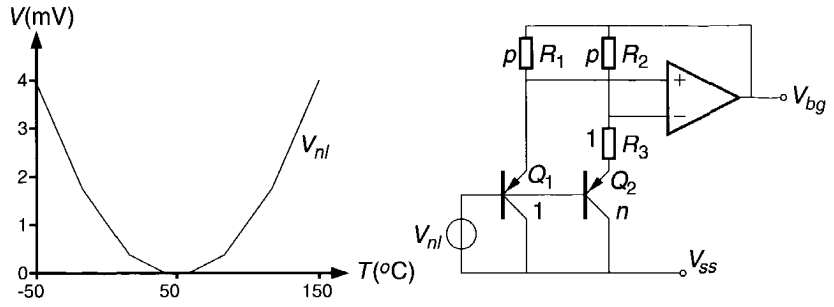

Fig. 3-7 *Piece-wise-linear curvature correction technique*

linear (PWL) principle. This technique is proposed by Rincon-Mora et al in basic form [3.20], but their implementation is still based on a non-linear approach.

The principle of the PWL curvature correction technique is shown in figure 3-7. Just as in the linearized V_{BE} technique, the PWL technique adds a correction voltage V_{nl} to a regular bandgap voltage. The difference is that V_{nl} is now built out of linear segments that follow the inverse curvature. These linear segments are made out of the V_{BE} signal and the amplified V_{ptat} signal, which have large opposite temperature coefficients of approximately $2\text{mV}/^{\circ}\text{C}$. It is therefore unnecessary to amplify them, which would introduce offset problems. The main advantage over the linearized V_{BE} technique is therefore that the implementation of the PWL technique is much easier. A designed and tested implementation example is shown in paragraph 3.4.3

Another important advantage is that the PWL technique can not degrade but only improve the accuracy, because the correction signal V_{nl} only has a non-linearity term, while in the linearized V_{BE} technique V_{nl} also has a significant constant and linear term. The obvious disadvantage of PWL is the worse approximation of the non-linearity curve. This makes PWL less suitable in ultra-high performance references, where multi-point calibration is not an issue. However, because our goal is only to reduce the error due to curvature below the next dominant error, an improvement of a factor of five is sufficient. This improvement can be obtained with both linearized V_{BE} and PWL techniques. A choice for PWL is then justified, because it has an easier implementation.

3.4 Realization

3.4.1 Filtering of modulated offset

To dynamically cancel the offset of the opamp in our bandgap reference circuit, a choice can be made between chopping and autozeroing. Because chopping has a superior noise performance (see chapter 2), our choice will be the chopping technique.

An important issue when using the chopping technique is the necessary filter to remove offset that is modulated by the output chopper. To make the chopping technique invisible to the outside world, we have to reduce the modulated offset signal below the noise floor. Doing this, we have to be aware that when filtering the output signal, the noise is also being filtered. Taking this into account, this gives the following condition

$$p V_{os} |H(f_{chop})| \leq V_{noise} \sqrt{f_{-3dB}} \tag{3-18}$$

where p is the amplification factor of the PTAT signal, V_{os} the initial offset voltage, $|H(f_{chop})|$ the filter attenuation at the chopping frequency, V_{noise} the output noise of the original bandgap reference circuit, and f_{-3dB} the -3dB bandwidth of the filter. The relation between $|H(f_{chop})|$ and f_{-3dB} for a first-order passive filter is given by:

$$|H(f_{chop})| = \frac{f_{-3dB}}{f_{chop}} \tag{3-19}$$

The required bandwidth for a first-order passive filter is then given by:

$$f_{-3dB} \leq \left(\frac{V_{noise} f_{chop}}{p V_{os}} \right)^2 \tag{3-20}$$

Filling in practical values $p=10$, V_{os}=1mV, f_{chop}=10kHz and V_{noise}=1μV/√(Hz), f_{-3dB} needs to be smaller than 1Hz. It is obvious that such a frequency requires large RC-constants, which consume a huge amount of chip area. Also the start-up time will become very large. A better approach is to use a second-order low-pass

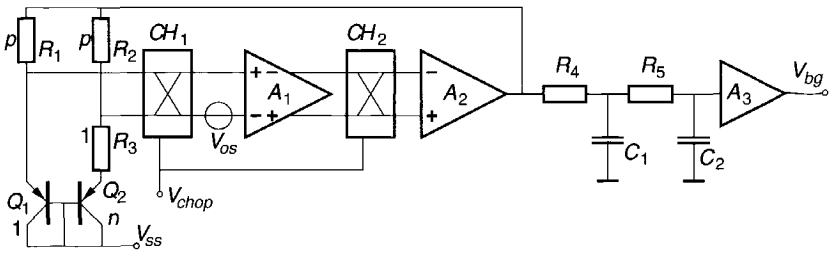

Fig. 3-8 *Filtering of modulated offset at output by a second-order continuous-time filter*

filter, which is shown in figure 3-8. The relation between $|H(f_{chop})|$ and f_{-3dB} for a second-order passive filter is given by:

$$|H(f_{chop})| = \left(\frac{f_{-3dB}}{f_{chop}}\right)^2 \tag{3-21}$$

and the corresponding bandwidth:

$$f_{-3dB} \leq \left(\frac{V_{noise} f_{chop}^2}{pV_{os}}\right)^{\frac{2}{3}} \tag{3-22}$$

This relaxes the requirements for the filter bandwidth to approximately 500Hz, which is possible to make but still requires large *RC*-constants.

An interesting different approach is to use a discrete-time filter, instead of a continuous-time one. Such an *SC*-filter implementation is shown in figure 3-9. It is explained as follows. The modulated offset of amplifier A_1 appears at the output of amplifier A_2 as a square wave, with an amplitude of pV_{os} and an average value of the desired V_{bg}. The switches S_1 and S_2 are closed with the non-overlapping signals φ_1 and φ_2 respectively. Doing this, the top values $V_{bg}+pV_{os}$ and $V_{bg}-pV_{os}$ of the square wave are sampled. Buffers A_{31}, A_{32} and the resistors R_{41} and R_{42} take care of the averaging of the two signals, resulting in the desired V_{bg} signal at the output.

The advantage of the *SC*-filter over the continuous-time filter is obviously the much smaller chip area. The resistors R_{41} and R_{42} can be quite small and also the size of the capacitors C_{11} and C_{12} is not determined by filter characteristics.

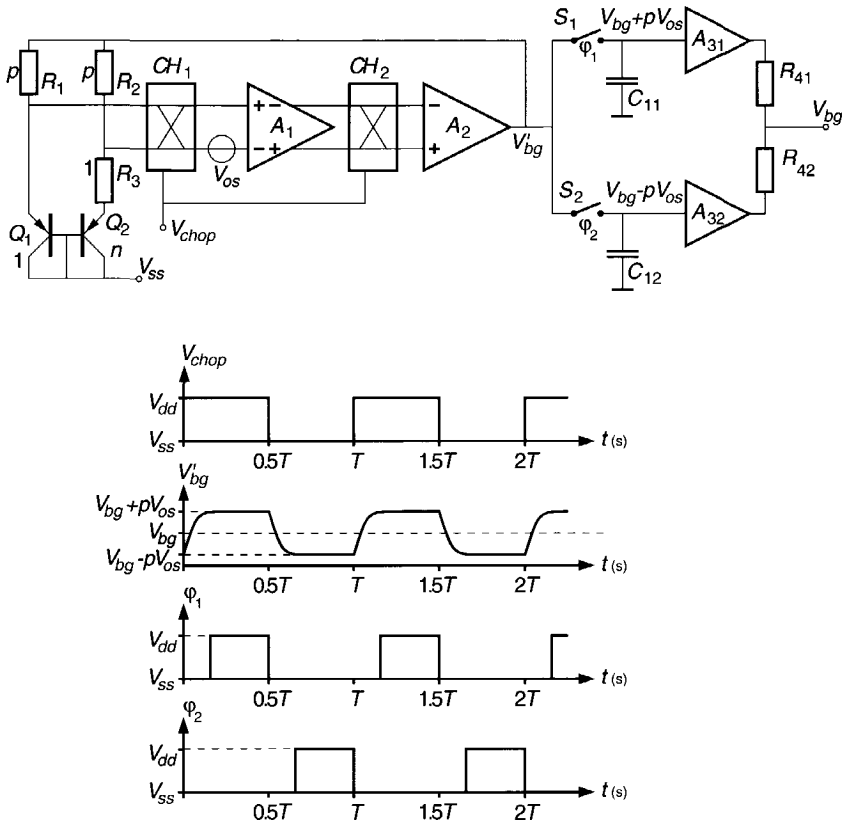

Fig. 3-9 *Filtering of modulated offset at output by SC-filter and corresponding signals*

However, charge injection caused by the switches S_1 and S_2 gives rise to spurious signals at the output. To reduce these signals below the noise floor, it is still necessary to use capacitors in the order of a few hundreds of pF.

A third method to implement the filter is shown in figure 3-10. It is again a continuous-time filter, but the filtering is now being done inside the loop. Compared to the two previous methods, a clear advantage of this approach is that no buffer is needed at the output. In addition to the lower power consumption and chip area, this has the advantage of not introducing additional offset at the output.

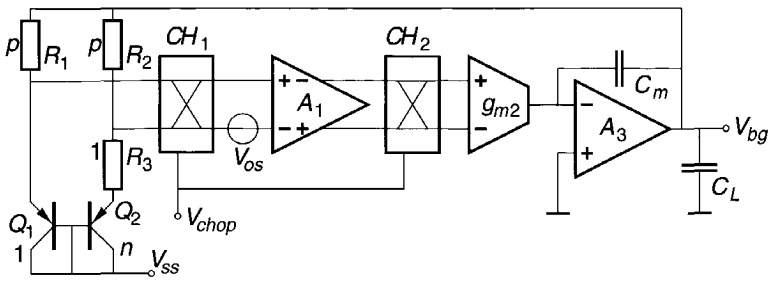

Fig. 3-10 *Filtering of modulated offset inside the loop*

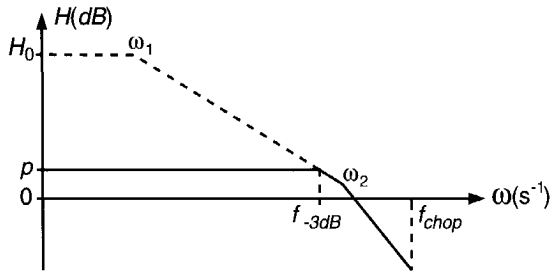

Fig. 3-11 *Bode plot of second-order in-loop filter*

Another great advantage is that the bandwidth of the filter is now determined by the Miller capacitance C_m and the transconductance g_{m2} of amplifier A_2. Compared to the conventional continuous-time filter, this reduces the chip area.

When implementing a second-order low-pass filter within the loop, attention should be paid to stability. Figure 3-11 shows a Bode-plot of the second-order in-loop filter. The dominant pole ω_1 is formed by g_{m2} and C_m. The second pole ω_2 can be formed by adding a second capacitor or by the load capacitance C_L. We decided to let it be determined by C_L, which gives the circuit better stability performance for heavy load capacitances. The dashed line indicates the amplifier loop without feedback. The total amplification at 0 Hz is given by H_0. Further information about stability analysis can be found in [3.21].

It is also possible to implement an *SC*-filter inside the loop. We showed that in [3.22]. This has approximately the same advantages and disadvantages as the regular *SC*-filter. We decided not to use this implementation here, because it requires too much capacitance to reduce the spurious frequencies below the noise floor.

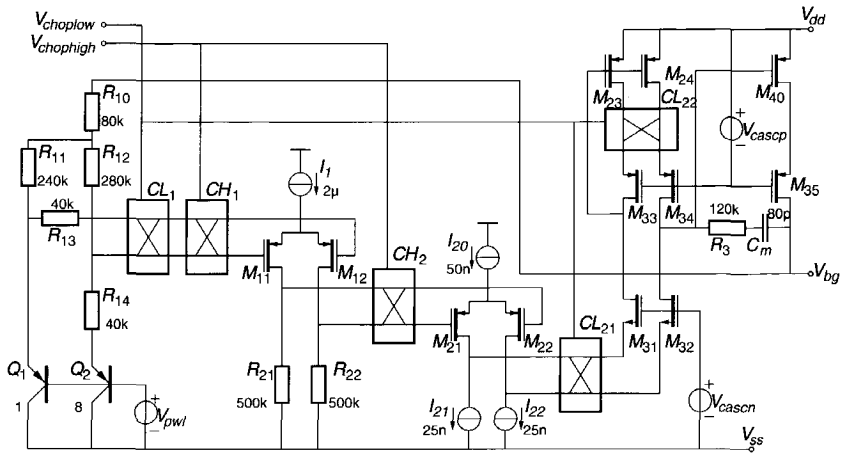

Fig. 3-12 *Bandgap reference with nested chopper technique*

3.4.2 Nested chopper technique

Until now, we have always chosen a value of 10kHz as an example for the chopping frequency. It would be much appreciated if this value could be higher, because that would relax our filter requirements. This is, however, not possible with a regular chopping technique, because higher chopping frequencies cause higher residual offset. Measurements on implementations of chopper amplifiers show a linear increase of 5-10µV per kHz. Keeping the residual offset well below 50µV, the chopping frequency can therefore not be higher than 10kHz.

A method to increase the chopping frequency, while still maintaining a low residual offset is found in applying the nested chopper technique as described in chapter 2. The residual offset is then determined by the lower chopping frequency, while the filter requirements can be lowered by the higher chopping frequency. In our implementation, we chose a high chopping frequency of 50kHz and a low chopping frequency of 6.25kHz. The implementation including the nested chopper technique is shown in figure 3-12.

The high-frequency chopper pair CH_1 and CH_2 is located across the first amplifier stage formed by M_{11}, M_{12}, R_{21} and R_{22}. The amplification of this stage is approximately ten. The amplification is kept relatively low to assure a high bandwidth, which is necessary to reduce errors due to the high-frequency

chopping (see chapter 2). The low-frequency pair CL_1, CL_{21} and CL_{22} is located across the first stage and a transconductance stage formed by M_{21} and M_{22}. This transconductance stage is biased at a relatively low bias current of only 50nA to obtain with the Miller capacitor C_m a very low-frequency dominant pole. The bandwidth between the low-frequency chopper pair, however, is relatively high, because the chopping is done in the current domain, before the cascodes. The second chopper is split into two choppers CL_{21} and CL_{22}, to be able to also remove the offset of the mirror formed by M_{23} and M_{24}. Resistor R_3 is added to remove the zero in the right-hand plane. The amplification of the fed back amplifier is approximately equal to $1+(R_{12}+2R_{10})/R_{14}$ which is twelve. Because the amplification of the first stage is also of the same order, the bandwidth of the amplifier is approximately $g_m/2\pi C_m$. Having a g_m of the second stage of 500nA/V and a C_m of 80pF, this yields a bandwidth of approximately 1 kHz. Having a high chopping frequency of 50kHz, this gives for a first-order filter a suppression of 34dB and for a second-order 68dB. According to equation (3-18) on page 48, this should be at least 50dB to reduce the modulated offset below the noise floor. This condition can be met easily by introducing an extra pole at the output by adding a load capacitor.

The substrate bipolar transistors Q_1 and Q_2 have an emitter area ratio of eight. They are biased at slightly different currents determined by R_{11} and the sum of R_{12} and R_{14}. Resistor R_{10} is formed out of a part of the parallel resistors R_{11} and R_{12} and is only introduced to reduce some chip area. The values of the resistors R_{11} through R_{14} are mainly determined by power consumption. The chip area is of less importance, because for values below 200kΩ the size of the resistors is not determined by their value but by their matching constraints. The reason for this unusually high "break-even point" is the availability of high-ohmic polysilicon resistors with values of 2kΩ/square.

The reference temperature T_r, at which the temperature coefficient of the bandgap reference circuit is exactly zero, is chosen at 45°C. This is exactly in the middle of our target temperature range of -40°C to +130°C. This value can be set by changing the unity-emitter size of the bipolar transistors Q_1 and Q_2, or by the absolute values of the emitter current controlling resistors R_{11} through R_{14}.

Resistor R_{13} is a very interesting one, because it has no influence on the DC behaviour of the circuit. It is introduced to match the impedances for the input choppers. This is necessary because spikes due to charge injection of the input choppers are cancelled out if these input impedances are exactly matched. Without R_{13} the impedance at the upper connection of CL_1 would be approximately $1/g_m$ of transistor Q_1. At the lower connection this is approximately R_{14} plus $1/g_m$ of transistor Q_2. Since transistors Q_1 and Q_2 are

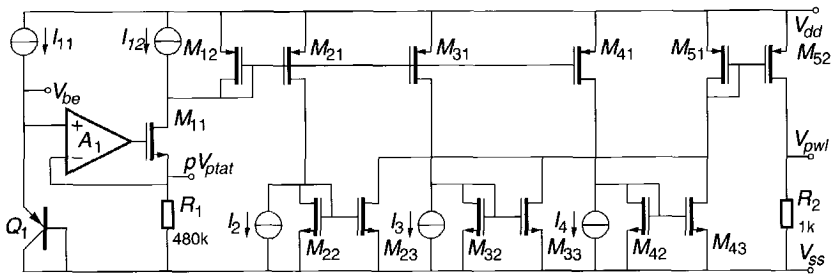

Fig. 3-13 *Piece-wise-linear voltage generator for $T<T_r$*

biased at approximately the same currents, they have equal g_ms. The impedance at the input of CL_1 can therefore be made approximately equal by adding a resistor with the same value as R_{14}.

3.4.3 Piece-wise-linear circuit implementation

In figure 3-12, the voltage source V_{pwl} is not implemented yet. This voltage source is meant to compensate the non-linearity of the V_{BE} curve, by means of a piece-wise-linear method as described in paragraph 3.3.2. To obtain enough accuracy, we decided to make a PWL source with six segments. The implementation of this circuit for temperatures smaller than the reference temperature T_r, which is 45°C, is shown in figure 3-13. For simplicity cascodes are omitted. All current sources in this circuit are PTAT and derived from the main bandgap circuit. The left part of the circuit generates a current that is equal to

$$I_{D,M12} = \frac{V_{BE} - pV_{ptat}}{R_1} \approx \frac{4mV(T_r - T)}{R_1} \qquad (3\text{-}23)$$

V_{be} and pV_{ptat} are chosen in such a way that their values are equal at the reference temperature T_r. This equation is only valid for $T<T_r$, because M_{12} can not drain a negative current. The current through M_{12} is copied three times and compared with different current sources I_1 through I_3. The differences are added and transformed to a voltage by R_2. The value of R_2 is kept at a low value of 1kΩ to reduce output resistance, which is necessary to reduce the influence of the base currents of Q_1 and Q_2 in figure 3-12.

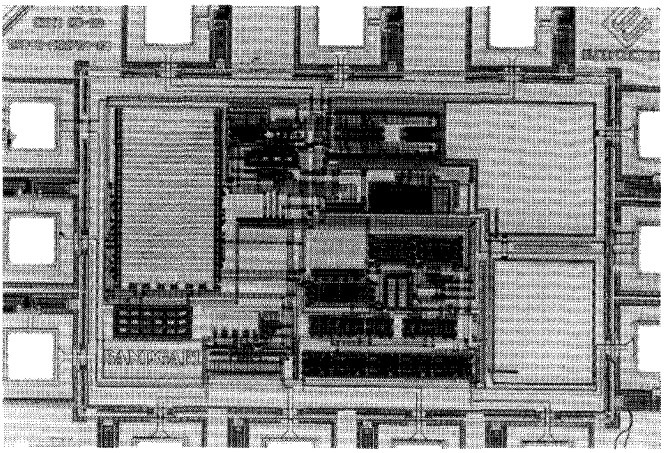

Fig. 3-14 *Chip micrograph of the fabricated circuit*

For $T>T_r$ a circuit with an extra NMOS current mirror is added. The current generated by this circuit is added to the circuit described above at M_{51}.

3.4.4 Measurement results

The bandgap reference circuit is fabricated in a standard 0.7μm CMOS process having high-ohmic polysilicon resistors and high-density linear capacitors. A chip micrograph is shown in figure 3-14. The total chip measures 1500μm x 1100μm. Active area is 1000μm x 600μm, of which 25% is occupied by capacitors and 10% by high-accuracy resistors. All circuitry is on the chip except for the oscillator.

The measurement results of nine samples of the circuit without curvature correction are shown in figure 3-15. The same measurements, but now with curvature correction are shown in figure 3-16. The standard deviation at 45°C is 1.1mV. This is in very good accordance with table3-1 on page 44 if we take into account that the spread in V_{be} within one batch is much smaller than the given value of 2 mV.

Over the full temperature range from -40°C to +130°C, the 3σ-spread of the bandgap references without curvature correction is 5.6mV. The same measurement for the circuits with curvature correction gives a 3σ-spread of 4.7mV.

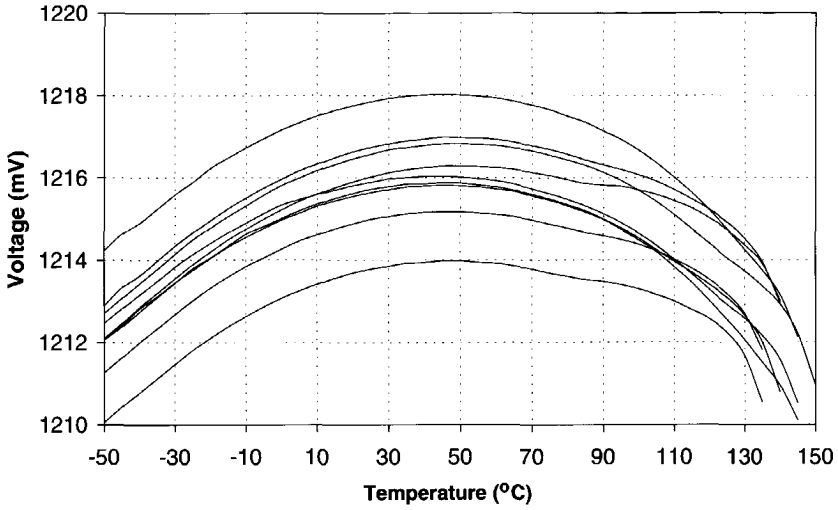

Fig. 3-15 *Measurement results of nine samples without curvature correction*

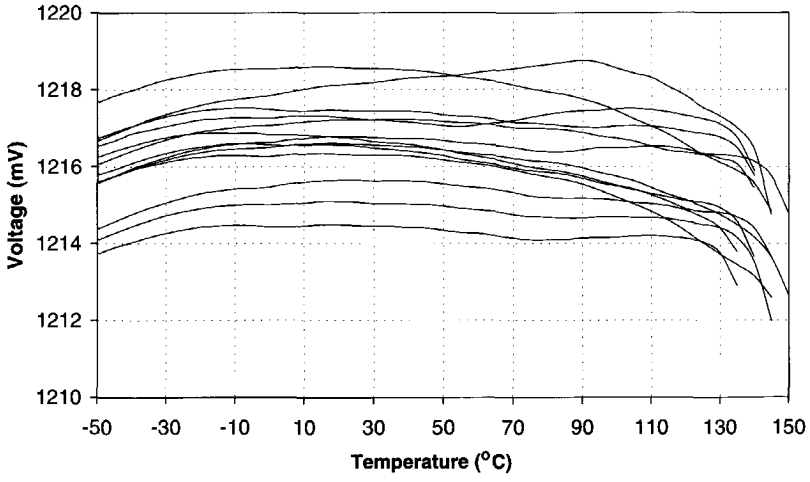

Fig. 3-16 *Measurement results of 14 samples with curvature correction*

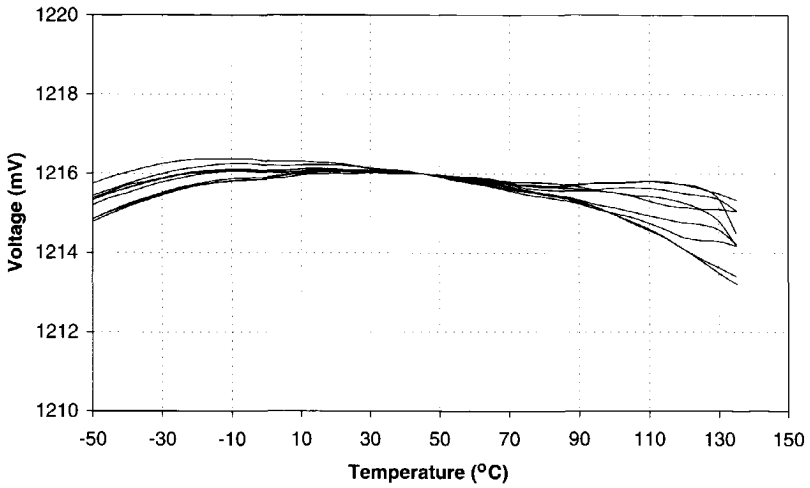

Fig. 3-17 *Measurement results of 14 samples with curvature correction, normalized at 45°C to determine temperature coefficients*

Looking at temperature coefficients, the difference between the measurements with and without curvature correction can become larger. Defining the temperature coefficient as the maximum slope of the line that starts at the reference temperature T_r at 45°C and follows the curve, we find for the nine circuits without curvature correction a maximum of 45µV/°C or 37ppm/°C, and for the fourteen measurements with curvature correction a maximum of 28µV/°C or 23ppm/°C. This difference is larger if we compensate for the small gain error in the curvature correction which shifts the reference temperature to lower values. This is clearly shown if we do a normalization of the circuit at 45°C, which is shown in figure 3-17. If we, for example, redefine the reference temperature for both measurements to 23°C and consider a temperature range from -40°C to +85°C, the temperature coefficients for the measurements without correction reduce to 24ppm/°C, but for the circuits with correction to 12ppm/°C. These values correspond to very low typical temperature coefficients of 16ppm/°C and 8ppm/°C respectively.

Another conclusion that can be drawn from figure 3-17 is that the curvature is slightly undercompensated, which indicates that further improvements can be made by applying a stronger curvature correction. If this is done in a redesign, temperature coefficients of typically less than 6ppm/°C can be obtained.

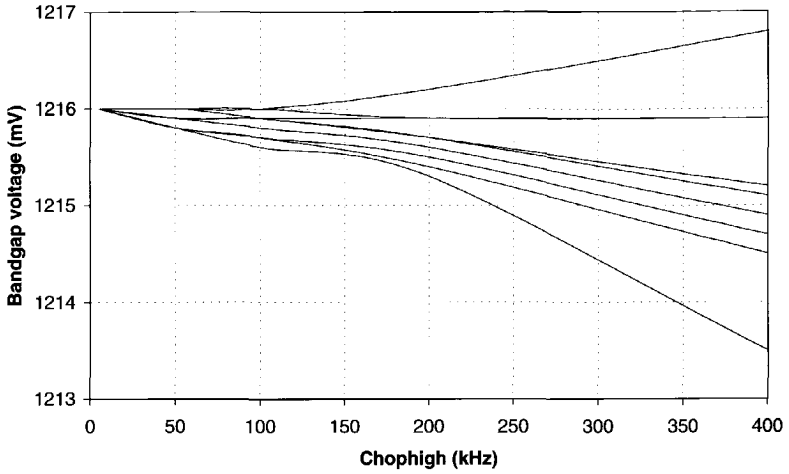

Fig. 3-18 *Measurement results of the accuracy versus chophigh frequency for nine samples, normalized at 10kHz*

Measurements on the accuracy versus chopping frequency have also been done. A plot of the bandgap voltage at room temperature for different values of chophigh and choplow are shown in figure 3-18. The curves are normalized at 1216mV for chophigh is 10kHz. The ratio between chophigh and choplow is fixed at eight. The output voltage dependency on chophigh has an average value of -2μV/kHz and a spread of 2.4μV/kHz. For the nominal chophigh frequency of 50kHz this gives a worst-case 3σ-spread of approximately -500μV or 0.04% at room temperature. This is around 10% of the maximum spread of 0.3%.

A list of specifications is shown in table 3-2.

3.5 Conclusions

It has been shown that the initial accuracy of CMOS bandgap references is mainly dictated by the offset of the opamp. This offset can be removed by applying a dynamic offset-cancellation technique. Using the nested chopper technique, spurious signals at the output can be reduced below the noise floor, using a total capacitor value of only 80pF.

Table 3-2 *Specifications of PWL-corrected bandgap reference*

	Min	Typ	Max	Unit
Supply voltage	2.4	3.0	5.5	V
Supply current		30	60	uA
High chopping frequency	10	50	400	kHz
Low chopping frequency	1.25	6.25	50	kHz
Nominal voltage @27°C		1216		mV
Initial inaccuracy@27°C		0.1	0.3	%
Temperature coefficient (-40 to +85°C) (-40 to +130°C)		 8 15	 16 30	ppm/°C
Noise (white)		1		$\mu V/\sqrt{(Hz)}$
Power supply regulation (2.4-3.5V) (3.5-5.5V)		 100 2000		ppm/V

Accuracy over temperature can be improved by applying a sixth-order piece-wise-linear curvature correction technique. This technique has the advantage over other techniques that it is easier to implement.

The measurement results show a high initial accuracy of typically 0.1% . This is the highest value ever reported for an uncalibrated bandgap reference. Also the temperature drift of typically 8ppm/°C is the best ever reported without calibration.

After applying all the above-mentioned techniques, the dominant error in the CMOS bandgap reference becomes the bipolar substrate transistor. The spread in base-emitter voltage of this transistor can, however, be much smaller than in comparable devices in bipolar technology. This work shows, therefore, that although bandgap references are based on bipolar transistors, the better choice for the implementation technology is CMOS.

References

[3.1] R.J. Widlar, "New developments in IC voltage regulators", *IEEE Journal of Solid-State Circuits*, vol.6, pp. 2-7, Feb. 1971

[3.2] K.E. Kuijk, "A precision reference voltage source", *IEEE Journal of Solid-State Circuits*, vol.8, pp. 222-26, June 1973

[3.3] Y.P. Tsividis and R.W. Ulmer, "A CMOS voltage reference", *IEEE Journal of Solid-State Circuits*, vol.13, pp. 774-778, Dec. 1978

[3.4] E.A. Vittoz, "Mos transistors operated in the lateral bipolar mode and their application in CMOS technology" *IEEE Journal of Solid-State Circuits*, vol.18, pp. 273-279, June 1983

[3.5] M.G.R. Degrauwe, O.N. Leuthold, E.A. Vitoz, H.J. Oguey and A. Descombes, "CMOS voltage references using lateral bipolar transistors", *IEEE Journal of Solid-State Circuits*, vol.20, pp. 1151-1157, Dec. 1985

[3.6] A.-J. Annema, "Low-power bandgap references featuring DTMOST's", *IEEE Journal of Solid-State Circuits*, vol.34, pp. 949-955, July 1999

[3.7] K.-M. Tham and K. Nagaraj, "A low supply voltage high PSRR voltage reference in CMOS Process", *IEEE Journal of Solid-State Circuits*, vol.30, pp. 586-590, May 1995

[3.8] E.A. Vittoz, and O. Neyroud, "A low-voltage CMOS bandgap reference", *IEEE Journal of Solid-State Circuits*, vol.14, pp. 573-577, June 1979

[3.9] G. Nicollini and D. Senderowicz, "A CMOS bandgap reference for differential signal processing", *IEEE Journal of Solid-State Circuits*, vol.26, pp. 41-50, Jan. 1991

[3.10] J. Michejda and S.K. Kim, "A precision CMOS bandgap reference", *IEEE Journal of Solid-State Circuits*, vol.19, pp. 1014-1021, Dec. 1984

[3.11] M. Ferro, F. Salerno and R. Castello, "A floating CMOS bandgap voltage reference for differential applications", *IEEE Journal of Solid-State Circuits*, vol.24, pp. 690-697, June 1989

[3.12] E. Sanchez-Sinencio and A.G. Andreou, "Low-voltage, low-power integrated circuits and systems: low-voltage mixed-signal circuits", *Piscataway*, ISBN 0-7803-3446-9, p.61, 1998

[3.13] G. Wang and G.C.M. Meijer, "The Temperature Characteristics of Bipolar Transistors for CMOS Temperature Sensors", *Proc. of Eurosensors XIII*, pp. 553-556, September 1999

[3.14] G.C.M. Meijer, "Thermal sensors based on transistors" *Sensor and Actuators*, vol.10, pp.103-125, 1986

[3.15] C.R. Palmer and R.C. Dobkin, "A curvature-corrected micropower voltage reference", *ISSCC Dig. Tech. Papers*, pp.58-59, Feb. 1981

[3.16] G.C.M. Meijer, P.C. Schmale and K. van Zalinge, "A new curvature-corrected bandgap reference", *IEEE Journal of Solid-State Circuits*, vol.17, pp. 1139-1143, Dec. 1982

[3.17] B.S. Song and P.R. Gray, "A precision curvature compensated CMOS bandgap reference", *IEEE Journal of Solid-State Circuits*, vol.18, pp.634-643, Dec. 1983

[3.18] G.C.M. Meijer, "Integrated circuits and components for bandgap references and temperature transducers" *Ph.D.-thesis, Delft University of Technology*, Delft, The Netherlands, March 1982

[3.19] M. Gunawan, G.C.M. Meijer, J. Fonderie and J.H. Huijsing, "A curvature-corrected low-voltage bandgap reference", *IEEE Journal of Solid-State Circuits*, vol.28, pp. 667-670, June 1993

[3.20] G.A. Rincon-Mora and P.E. Allen, "A 1.1V current-mode and piecewise-linear curvature-corrected bandgap reference", *IEEE Journal of Solid-State Circuits*, vol.33, pp. 1551-1554, Oct. 1998

[3.21] K.-J. de Langen, "Advanced Low-Voltage and High-Speed Techniques for BiCMOS, CMOS and Bipolar Operational Amplifiers", *Ph.D. Thesis*, ISBN 90-407-1846-6, Delft University Press, March 1999

[3.22] A. Bakker and J.H. Huijsing, "A CMOS chopper opamp with integrated low-pass filter", *Proc. ESSCIRC'97*, Southampton, UK, September 1997, pp. 200-203.

Design of CMOS Smart Temperature Sensors

4

This chapter discusses the various design aspects of CMOS smart temperature sensors. It describes different kinds of analog-to-digital conversion, methods for Kelvin-to-Celsius conversion, curvature correction and bus interfaces. Also an electronic interface for the temperature read-out of a single remote bipolar transistor is proposed. The design aspects that are considered most important are high uncalibrated accuracy and low power consumption.

4.1 Introduction

The research on smart or integrated circuit temperature sensors started in the mid-seventies and was mainly driven by the need for better interfaces. Until that time, the market was dominated by resistance thermometers, such as the Pt-100 and the thermistor. The development of IC-technology gave researchers the possibility to integrate bipolar transistor temperature sensors together with interface electronics. Meijer describes in his Ph.D. thesis [4.1] the development of monolithic temperature transducers with current, voltage and logic output.

The increasing level of complexity of integrated circuits in the eighties and nineties paved the way for the addition of analog-to-digital conversion and bus interfaces on the temperature sensor chip. The company Smartec was the first one in 1992 to produce a temperature sensor with duty-cycle output (SMT160) [4.2]. National Semiconductor's LM75 [4.3] was, in 1995, the first "fully" integrated smart temperature sensor. This circuit has become the standard for smart temperature sensors.

The real breakthrough of the smart temperature sensor came in the second part of the nineties, when Intel decided to add thermal management to their microprocessors and motherboards. Costs became a much more important issue. All large chip manufacturers now have several kinds of smart temperature sensors in their catalogue, making for a heavy competition and price erosion. The price for an LM75 compatible circuit is already significantly below $1 in 2000.

The work described in this thesis focuses mainly on the application of electronic thermometers in the above-mentioned microprocessor systems. The main objective is to enhance the accuracy, to be able to reduce the costs by omitting the need for calibration. Another important design aspect is power consumption, because the applications for battery-supplied systems are expected to grow fast in the near future.

4.1.1 Accuracy

The largest market for smart temperature sensors is currently found in thermal management systems. Especially the thermal management in personal computers yields a market of over 5 million pieces a year. The accuracy specification for this application is 2°C in a limited temperature range from -20°C to +100°C, and 3°C for an extended temperature range from -40°C to +125°C. All existing temperature sensors need at least one calibration to meet this accuracy. This calibration can usually be done at room temperature at the wafer level test. The long-term stability of the silicon temperature sensors is usually sufficient to keep its specification during its lifetime.

Analysis of CMOS bandgap references that has been done in chapter 3 showed that when applying dynamic offset-cancellation techniques and curvature correction, it is possible to make a voltage reference with a typical uncalibrated accuracy of 0.1%. Taking into account that a bandgap reference has all the necessary signals to make a smart temperature sensor (e.g. PTAT voltage and reference voltage), it should be possible to make a temperature sensor with an uncalibrated accuracy of better than 0.1%. Since the PTAT voltage gives a temperature signal in Kelvin scale, 0.1% corresponds to approximately 0.3°C accuracy at room temperature (T=300K). For a production environment where is calculated with 3σ spreads, this indicates that a maximum inaccuracy of approximately 1°C can be met at room temperature. Assuming that errors introduced by the A-to-D converter can be neglected, it is theoretically possible to make a CMOS smart temperature sensor with 1°C accuracy without calibration. This would be a major cost breakthrough in the market for thermal management because omitting the calibration means reduced manufacturing costs, firstly

because no chip area is needed for EPROM or OTP (one time programmable), secondly no additional processing steps like double poly are necessary, and thirdly no temperature control and programming has to be done for calibration itself.

In this chapter we will show that it is possible to approach this theoretical limit. We will look at various aspects like A-to-D conversion, curvature correction, Kelvin-to-Celsius conversion and we will do research on the best implementation to meet the goal of uncalibrated 1°C accuracy.

4.1.2 Power consumption

Power consumption in smart temperature sensors has been an important issue since the first introduction in the mid-seventies. At that time, the main reason for reducing power consumption was self-heating. Having a thermal resistance of 100°C/W and drawing a few mili Amps from a 12V supply already gives rise to a temperature error of several degrees [4.1]. More modern sensors like the LM75 draw only a few hundreds of micro Amps from a 5V supply, thus reducing the power consumption below 1mW and the self-heating in the order of 0.1°C, which is negligible compared to the overall accuracy of 2°C.

For battery-powered systems, a power consumption of 1mW is usually not negligible. These applications require a power consumption that approaches the theoretical limit. An interesting question therefore is: what is the theoretical minimal power consumption of a smart temperature sensor? This question can only be answered if the bandwidth of the input signal and the required noise performance are known.

The required bandwidth of a temperature sensor for application in thermal monitoring systems is dictated by the thermal capacitance or mass of the monitoring object, the (parasitic) heating source and the thermal resistance. For example, in a microprocessor system the bandwidth is determined by the thermal mass of the package, the maximum power dissipation of the microprocessor and the worst-case thermal resistance when the fan is malfunctioning. This worst-case approach gives maximum temperature slopes of, for example, 10°C per second. The customer can then specify, for example, a maximum error of 2°C, giving a required bandwidth of 5 samples/s. In most applications the required bandwidth is limited to a few samples/s.

The maximum allowed noise is normally specified to be below 0.5 LSB (Least Significant Bit). For the industry-standard LM75 the resolution is 9 bits or 0.5°C, thus giving a maximum allowed noise of 0.25°C. Other sensors, like those from

Dallas Semiconductor [4.4] have resolutions up to 12 bits or 0.0625°C, which enables the measurement of very small temperature changes.

Tuthill showed that the power consumption of a smart temperature sensor with a bandwidth of 10 samples/s can be reduced to 1μW [4.5]. He made a circuit with a successive-approximation A-to-D converter, which powers-up, does a conversion and goes back to sleep again. Although the circuit dissipates approximately 1.25mA from a 3V supply, the short on-time of only 25μs reduces the average current consumption at 10 samples/s to below 0.3μA.

The noise of a smart temperature sensor is almost completely determined by the noise of the PTAT voltage generator, because this circuit has the smallest signal. According to equation 3-1 in chapter 3, the signal from a PTAT generator with an emitter-current density ratio of 10, is approximately 200μV/°C. To obtain temperature noise below 0.25°C, the voltage noise of the read-out circuit should therefore be less than 50μV. With a current consumption of the first amplifier stage of 1μA, thermal noise can be kept below 100nV/√(Hz). With a bandwidth of a few Hz, it is quite easy to keep the thermal noise below 50μV. Even the increased specification for the Dallas Thermometer of 0.03°C, which requires a noise performance of less than 5μV is easy to meet with a power consumption of 1μA. Flicker noise can be a greater problem, but this is solved using appropriate dynamic offset-cancellation techniques, as shown in chapter 2.

The answer to the question on the theoretical minimal noise is that it is neither determined by bandwidth nor by noise. These requirements only need less than 1μW of power consumption. In practice the total power consumption will be dominated by the digital interface and especially the buffers. This shifts the challenges on power reduction to the system level, where choices are made on the kind of data transfer. However, our goal will remain to reduce power consumption to the μW level, to assure a negligible addition to the total power consumption.

4.2 Analog-to-Digital conversion

The requirements for the analog-to-digital converter in a smart temperature sensor are quite relaxed regarding bandwidth. A bandwidth of 10 samples/s is usually sufficient. The accuracy is somewhere between 9 and 12 bits, depending on the application. These specifications give possibilities to choose an A-to-D converter principle that is optimized for, for example, simplicity or robustness. Literature shows implementations of smart temperature sensors with converters

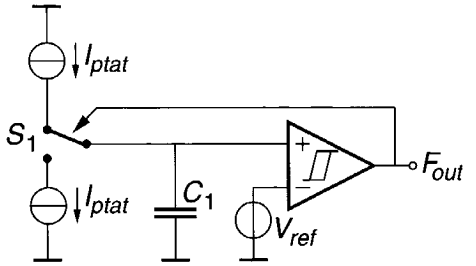

Fig. 4-1 *Principle of temperature to frequency conversion*

based on successive-approximation [4.5], frequency conversion [4.6, 4.7], duty-cycle modulation [4.2, 4.8], and sigma-delta modulation [4.3, 4.9, 4.10].

Successive approximation A-to-D converters are clearly the fastest converters among the mentioned ones. They can do an n-bit conversion in n clock cycles, which makes them easily achieve 10-100 ksamples/s. Their accuracy is however limited to the matching properties of passive components on chip. Assuming a best matching of capacitors or resistors of 0.1%, accuracy is limited to 10 bits. This accuracy is good enough for medium accuracy sensors like the 9-bit LM75, but lacks possibilities to extend it to more sophisticated smart temperature sensors. The author already mentioned in his paper [4.5] that the choice for the successive approximation principle was not determined by the temperature sensor, but by the necessity to use the same A-to-D converter for other higher frequency input signals.

The other mentioned principles are all interesting for our smart temperature sensor and will be discussed in the next paragraphs.

4.2.1 Frequency conversion

The principle of the temperature to frequency converter is shown in figure 4-1. The circuit is a relaxation oscillator of which the frequency is determined by the temperature-dependent input current I_{ptat}. The circuit has two phases. In the first phase, the capacitor C_1 is charged by I_{ptat}. When the voltage on C_1 reaches a level equal to V_{ref} plus a certain hysteresis voltage V_{hyst}, the output F_{out} switches to zero. The switch S_1 is set to the sinking version of I_{ptat} and C_1 starts to discharge. When the voltage on C_1 is lower than V_{ref} minus V_{hyst}, F_{out} switches again to one, and the whole cycle starts over.

The output frequency is equal to

$$f_{out} = \frac{2I_{ptat}}{C_1 V_{hyst}} \tag{4-1}$$

The great advantage of this circuit is its simplicity. A simple PTAT generator, a capacitor and a Schmitt-trigger are sufficient. The circuit also has only three terminals (V_{dd}, V_{ss} and F_{out}), which implies a very cheap package. The output signal is in the frequency domain, which makes it easy to interpret by a simple microcontroller. The signal is also quite robust to interfering signals, so it can be easily transported over long wires. This frequency signal is a so-called "semi-digital" signal, which means that it is digitized in amplitude, but not in time.

The great disadvantage of the circuit is its poor accuracy. The output frequency f_{out} is not only determined by I_{ptat} but also by C_1 and V_{hyst}. Especially C_1 can have a poor absolute accuracy. Also the temperature dependencies of C_1, V_{hyst} and the resistor R_{ptat} which is needed to generate I_{ptat} out of V_{ptat}, will have a bad influence on the accuracy of the total circuit. Also the reference oscillator, which is used by the microcontroller needs to be taken into account.

From a designer's point of view, one can say that this circuit has a problem with its reference. The reference of the temperature-dependent PTAT voltage V_{ptat} is formed by a resistor R_{ptat}, a capacitor C_1, another voltage (V_{hyst}) and an oscillator (of the microcontroller). Most of these devices are badly matched and therefore introduce errors.

4.2.2 Duty-cycle modulation

A method to overcome the problem with the reference is found in the duty-cycle modulation. The principle is shown in figure 4-2. The difference compared to the frequency converter is that the capacitor is now charged with a current I_{bg} - I_{ptat} and discharged with I_{ptat}. Another difference is that the temperature information is now not found in the frequency but in the duty-cycle. The duty-cycle at the output can be calculated as follows. The charge on capacitor C_1 at the end of each period is always the same. That means that during exactly one period the charging current times the charging time is equal to the discharging current times the discharging time. This yields

$$T_1 \cdot I_{bg} = T \cdot I_{ptat} \tag{4-2}$$

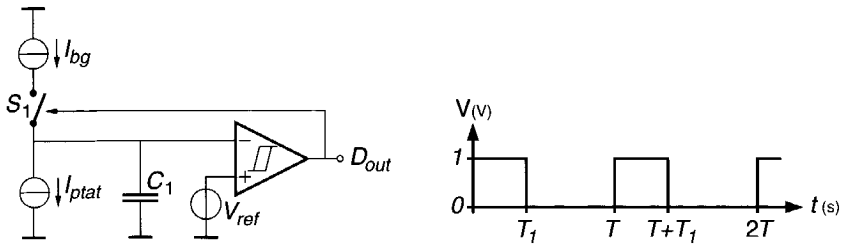

Fig. 4-2 *Principle of duty-cycle modulation*

where I_{bg} is derived from a bandgap voltage V_{bg}. Equation 4-2 can also be written as

$$D_{out} = \frac{T_1}{T} = \frac{I_{ptat}}{I_{bg}} \tag{4-3}$$

This type of converter is also called "charge balancing". It can be clearly seen from equation (4-3) that the duty-cycle output is now not determined by a capacitor or external oscillator, but only by I_{ptat} and I_{bg}. Also, the resistor's absolute value, which is used to derive these currents from their corresponding voltages, is now not important anymore, because both I_{ptat} and I_{bg} can be derived by the same type of resistor.

4.2.3 Sigma-delta A-to-D conversion

In systems where only one microcontroller is used to perform different tasks, such as in a thermal management system for PCs, the microcontroller has usually not enough time to process a duty-cycle signal during several tens or hundreds of milli seconds. To reduce the tasks of the microcontroller, it is therefore necessary to do a full digitalisation on the chip, instead of doing the time-discretization off-chip in the microcontroller.

One of the problems when doing a time-discretization of a duty-cycle modulated signal on-chip is the fact that the signal is not synchronized with the system clock. This is disadvantageous because it can cause locking between the two oscillators due to parasitic coupling. This will reduce the accuracy and resolution [4.9]. To avoid this, the duty-cycle signal is synchronized with the system clock by adding a flip-flop after the comparator. The result is shown in figure 4-3 and is called a sigma-delta modulator. The difference compared to the duty-cycle converter is

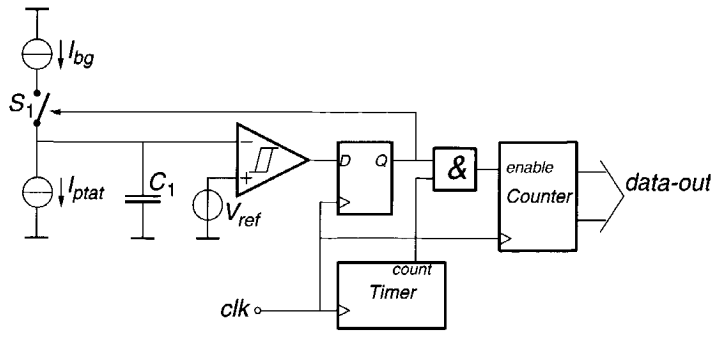

Fig. 4-3 *Principle of sigma-delta A-to-D converter*

that the output signal of the Schmitt-trigger is now sampled by the clocked flip-flop and is therefore synchronized with the system clock. For one period T, this gives an obviously large error, because the free-running duty-cycle frequency is not equal to the frequency of the system clock. However, this error is stored (integrated) on the capacitor C_1 and extracted in the next period. When a lot of periods are taken into account, the error is averaged and becomes theoretically zero when an infinite number of periods are taken. This strategy is called "oversampling" and is the strength of the family of sigma-delta modulators. More theory about sigma-delta modulators can be found in [4.11].

The pulses at the output of the flip-flop are counted for a fixed time defined by a timer. The result is a digital word, which can be sent directly to the outside world.

One reason why this synchronization technique is usually not used in systems where the time-discretization is done in the microcontroller is the fact that one more communication line is necessary for a clock signal. Another reason is that the parasitic coupling between the microcontroller oscillator and the duty-cycle modulator is much smaller because they are not on the same substrate.

The sigma-delta modulator is the best trade-off between accuracy, simplicity and power consumption. It has better accuracy than frequency converters or non-integrating converters like successive-approximation and consumes less power than duty-cycle modulators. In future, it may be worthwhile to switch to higher-order sigma-delta converters, because this reduces sampling time. A shorter sampling time reduces power consumption when a power-down mode is applied. It also opens ways to increase the resolution, while still keeping the same bandwidth.

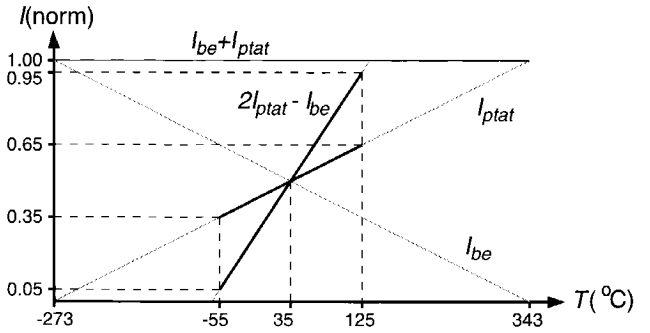

Fig. 4-4 *Kelvin-to-Celsius converter principle with normalized currents*

4.3 Kelvin-to-Celsius conversion

As already discussed in chapter 3, the most accurate temperature-dependent signal on-chip is the PTAT voltage. It is made by taking the difference of two base-emitter voltages which have a different base-emitter current density. Having a base-emitter current density ratio of p, the PTAT voltage can be written as

$$V_{ptat} = m\frac{kT}{q}\ln p \qquad (4\text{-}4)$$

where k is the Boltzmann's constant and q the electron charge. The factor m is called the non-ideality factor, which is caused by base-width modulation [4.12]. This non-ideality factor is slightly process-dependent, but is quite stable between batches of the same process. Its value is usually somewhere between 1.003 and 1.010.

As can be seen from equation (4-4), the PTAT signal is related to the temperature in Kelvin. Because we are usually only interested in a temperature range from -55°C to +130°C or roughly 220K to 400K, a large part of the dynamic range is not being used. This problem has already been encountered by Meijer [4.1, 4.8] and was solved by subtracting a V_{BE} signal from the basic PTAT signal. Something similar is done in our smart temperature sensor. Figure 4-4 shows the basic PTAT (I_{ptat}) and reference ($I_{ptat} + I_{be}$) signals of a smart temperature sensor. The straightforward implementation using I_{ptat} as input signal shows a dynamic range usage of only 30% in a temperature range from -55°C to +125°C. By not

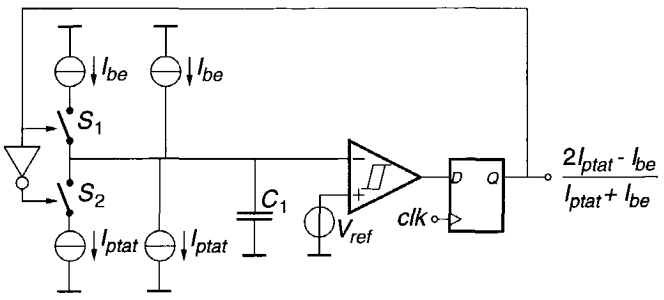

Fig. 4-5 *Kelvin-to-Celsius converter implementation in combination with a sigma-delta modulator*

using the standard I_{ptat} signal, but a combination of $2I_{ptat}$ minus I_{be}, the dynamic range usage can be increased to 90%.

An implementation of the Kelvin-to-Celsius converter for our sigma-delta based thermometer is shown in figure 4-5. The output is calculated as follows. Just as in the duty-cycle modulator in paragraph 4.2.2, the charge-balancing principle can be applied. Defining the total number of clock cycles in one sample as M and the number of ones as N, this yields

$$M \cdot I_{ptat} + (M - N) \cdot I_{ptat} = M \cdot I_{be} + N \cdot I_{be} \tag{4-5}$$

This can also be written as

$$M \cdot (2I_{ptat} - I_{be}) = N \cdot (I_{ptat} + I_{be}) \tag{4-6}$$

The measured duty-cycle is the ratio of N and M and equals

$$\frac{N}{M} = \frac{2I_{ptat} - I_{be}}{I_{ptat} + I_{be}} \tag{4-7}$$

Compared to the conventional approach when only the I_{ptat} signal is used as input signal, the efficiency of the sigma-delta converter is increased with a factor of three. This implies a three times smaller sampling time and a three times lower power consumption at the cost of only one switch, an inverter and an extra PTAT current source.

4.4 Curvature correction

In contrast to the PTAT voltage, which is a highly linear function of temperature, the signal related to the base-emitter voltage shows a slight non-linearity. This non-linearity is already discussed in paragraph 3.2 and yields approximately 5 mV in a temperature range from -55 °C to +125 °C. This corresponds to approximately 1.3 °C.

Chapter 3 also discusses techniques to reduce the effects of this non-linearity in bandgap references. These curvature correction techniques, that are based on linearization of V_{BE} or piece-wise-linear techniques, can also be applied to smart temperature sensors, with approximately the same results.

Meijer et al describe in [4.8] a different but very elegant technique to reduce the effects of non-linearity in their three-terminal temperature sensor. They show that the non-linearity in a smart temperature sensor can be reduced by making the reference signal ($I_{ptat}+I_{be}$) slightly temperature-dependent. Introducing a factor α which is the temperature dependence of the reference current and applying this to our smart temperature sensor, equation (4-7) becomes

$$\frac{N}{M}=\frac{2I_{ptat}-I_{be}}{(I_{ptat}+I_{be})\left(1+\alpha\frac{T}{T_r}\right)}\approx\frac{2I_{ptat}-I_{be}}{(I_{ptat}+I_{be})}\left\{1-\alpha\frac{T}{T_r}+\alpha^2\left(\frac{T}{T_r}\right)^2-\alpha^3\left(\frac{T}{T_r}\right)^3\right\} \quad (4\text{-}8)$$

This equation shows that by applying a slight first-order temperature dependence in the denominator, quadratic and higher-order components appear at the output. The quadratic component can be used to compensate the curvature. The higher-order components can be neglected for small values of α. In practice, values for α around 0.04 are used, which suppress third-order components by approximately 25. Equation 4-8 also shows a linear component ($\alpha T/T_r$) at the output. This component is unwanted and needs to be compensated by a slight modification of I_{ptat} or I_{be}. Figure 4-6 shows typical curves for this kind of curvature correction.

Meijer et al [4.8] showed that by applying this method, a reduction of the curvature can be achieved by a factor of approximately 5.

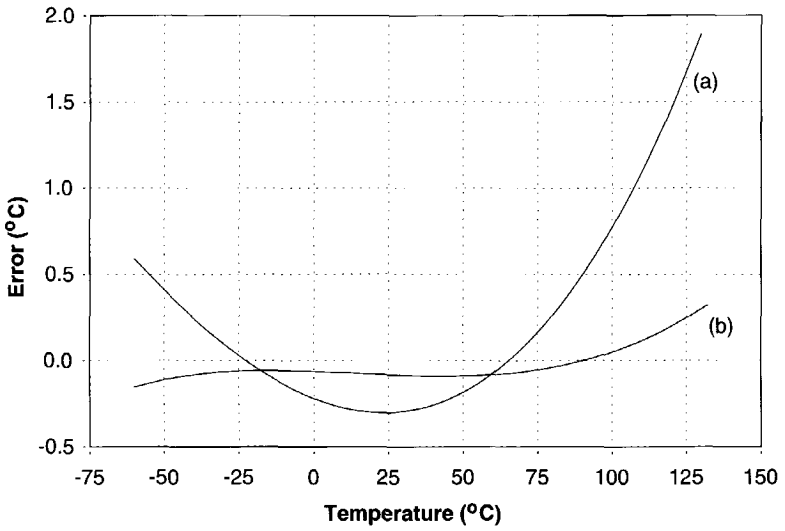

Fig. 4-6 *Non-linearity error of smart temperature sensor before (a) and after (b) correction by means of non-constant reference current*

4.5 Single transistor temperature sensors

Until now, the PTAT signal, which is used as the temperature signal, is derived from the difference of the base-emitter voltages of two separate transistors. It is, however, also possible to use only one transistor and multiplex the measurements in time. This has the advantage that no problems can arise with possible mismatch between the bipolar transistors, because for both measurements the same transistor is used. This advantage is however quite small, because the error due to this mismatch is relatively small (typically 100μV or 0.5°C) compared to errors due to amplifier offset and resistor mismatch. Table 3-1 in paragraph 3.3.1 shows that the error in the bandgap reference is already much greater. The disadvantages are obviously higher complexity which yields more chip area and higher power consumption.

A much more important reason to study single transistor temperature sensors is the situation when the transistor or p-n junction is on a different substrate. This is, for a very important example, the case in many microprocessors. Because the processes in which they are made are fully optimized for digital design and are

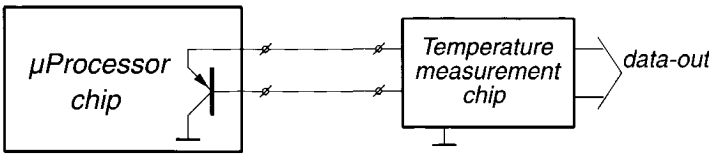

Fig. 4-7 *Principle of remote microprocessor temperature sensing*

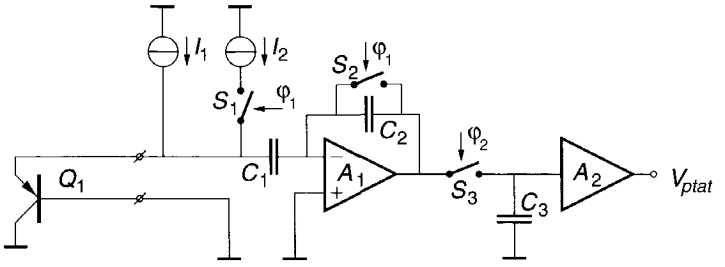

Fig. 4-8 *PTAT-voltage generation by means of SC-amplifier*

updated very often, it is almost impossible to design a temperature sensor circuit on the same substrate as the processor itself. To still be able to measure the real temperature of the silicon itself and not of the package, the microprocessor suppliers incorporate a single transistor on the microprocessor die. This single transistor temperature sensor is then read out by an external temperature measurement chip. The principle is shown in figure 4-7.

A straightforward way to generate the desired PTAT signal out of the microprocessor transistor is by a regular switched-capacitor technique. The principle is shown in figure 4-8. In the first phase, when φ_1 is high, the voltage on C_2 is reset and the voltage on C_1 is set to the base-emitter voltage of Q_1, which is biased at the sum of the currents I_1 and I_2. In the second phase, when φ_2 is high, the base-emitter voltage of Q_1 is lowered, because it is now biased only by I_1. The corresponding charge difference on C_1 is dumped on C_2 and sampled on C_3. The output voltage V_{ptat} is equal to the voltage on C_3 and yields

$$V_{ptat} = \frac{C_1}{C_2} \cdot \frac{kT}{q} \ln\left(\frac{I_1 + I_2}{I_1}\right) \tag{4-9}$$

This circuit has some potential drawbacks. First is the effect of the offset of amplifier A_1. This offset is sampled on C_1 during φ_1 and appears amplified at the output. Because our input signal has a sensitivity of only 200µV/°C, this offset can cause errors of several degrees Celsius. This offset can however be removed

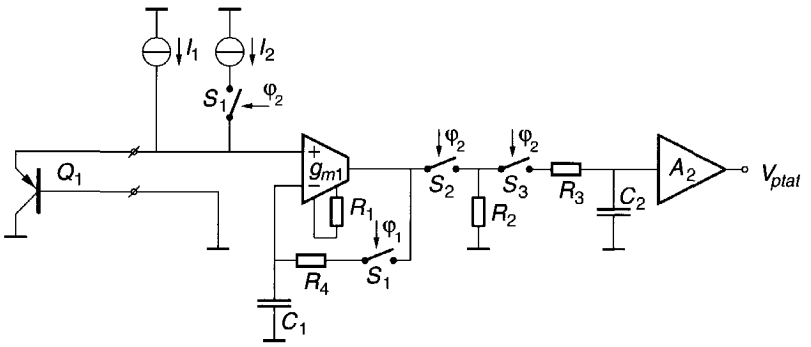

Fig. 4-9 *PTAT-voltage generator with offset cancellation and improved interference suppression*

by an autozeroing or double-correlated sampling technique as described in chapter 2.

Another, more fundamental, drawback of this circuit is its sensitivity to interference. Transistor Q_1 is situated on a microprocessor chip with high clock frequencies and steep edges. Spikes on base and emitter can appear of several hundreds of mili Volts. To reduce their influence, the temperature measurement chip should integrate the input signal. This is not done in the circuit of figure 4-8 where the whole bandwidth is sampled on C_1.

An improvement of the circuit of figure 4-8 is shown in figure 4-9. In the first phase φ_1, the base-emitter voltage of Q_1 for the bias current I_1 is sampled on C_1 together with the offset of transconductance amplifier g_{m1}. In the second phase φ_2, the base-emitter voltage of Q_1 is raised by the PTAT voltage. This voltage difference is amplified by g_{m1} in combination with R_2 and sampled on C_3. The output voltage yields

$$V_{ptat} = \frac{R_2}{R_1} \cdot \frac{kT}{q} \ln\left(\frac{I_1 + I_2}{I_1}\right) \qquad (4\text{-}10)$$

Advantages of this circuit over the regular SC-implementation is that the offset is auto-zeroed and that the interference rejection is much better. The latter is achieved by low-pass filtering in the autozeroing phase (φ_1) by means of R_4 and C_1 and in the amplification phase (φ_2) by R_3 and C_2.

The interference suppression is maximal when the filter time constants are chosen as large as possible. In practical realizations, as will be shown in chapter 5, in

terms of chip area it is better to implement those large time-constants in the digital domain, by using a combination of a comparator, counter and D-to-A-converter instead of a capacitor which is basically an analog storage device.

4.6 Bus interfaces

One of the questions after digitization of the temperature signal is how to transport it to a computer or microcontroller. For more general signals, this question has been asked and answered already many times. It has led to the development of several bus interfaces, each one having its own specific application or optimization. The most widely used for distances between 10 and 1000 meters are the field buses, such as Fieldbus, RS-485 and CAN-bus. They are characterized by an elaborate transmission protocol and high data transmission rates.

For most smart sensor systems, where the demands for distance and transmission rates are much more relaxed, these high-level buses are too expensive in terms of chip area and power consumption. For these applications, very basic protocols have been developed. Examples of these are the Microwire protocol from National Semiconductor or the IS^2 bus interface from Riedijk [4.14].

One of the major problems with these low-level bus interfaces has been the standardization. No specific protocol could gain enough market share yet to become a real standard. At this moment, the I^2C protocol from Philips, that has been designed over twenty years ago for communication in audio and video equipment, seems to have become the leading standard. The reason for this is *not* that it is the best available protocol in technical terms, but many peripheral chips like microprocessors and RAMs can be bought with such an interface. Also the smart temperature sensor standard, the LM75, has an I^2C interface. Another very important reason is that all modern laptops have a System Management bus (SM-bus), which is upwards compatible with I^2C [4.13].

Based on the above-mentioned considerations, the choice for the bus interface of our smart temperature sensor will be the SM-bus. This has the advantages that it will be completely compatible with the standard LM75, it can be easily interfaced by several microcontrollers, and it can be directly applied in laptops. The latter is becoming more and more important, and thermal management in laptops is at the moment already the most important application for smart temperature sensors.

References

[4.1] G.C.M. Meijer, "Integrated Circuits and Components for Bandgap References and Temperature Transducers", *PH.D. Thesis*, Delft University of Technology, The Netherlands, March 1982

[4.2] Smartec, "SMT160-30 Smart Temperature Sensor", *http:// www.smartec.nl*, November 1999

[4.3] National Semiconductor, "LM75 - Digital Temperature Sensor and Thermal WATCHDOG with Two-Wire Interface", *http:// www.national.com*, November 1999

[4.4] Dallas Semiconductor, "DS75 - Digital Thermometer and Thermostat", *http://www.dalsemi.com*, November 1999

[4.5] Mike Tuthill, "A Switched-Current, Switched-Capacitor Temperature Sensor in 0.6-μm CMOS", *IEEE Journal of Solid-State Circuits*, vol.33, pp. 1117-1122, July 1998

[4.6] P. Krummenacher and H. Oguey, "Smart Temperature Sensor in CMOS Technology", *Sensors and Actuators*, A21-A23, pp.636-638, 1990

[4.7] D. van Maaren, J. Klijn and G.C.M. Meijer, "An Integrated Micropower Low-Voltage Temperature-Controlled Oscillator", *IEEE Journal of Solid-State Circuits*, vol.17, pp. 1197-1201, 1982

[4.8] G.C.M. Meijer, R. van Gelder, V. Nooder, J. van Drecht and H. Kerkvliet, "A Three-Terminal Integrated Temperature Transducer with Microcomputer Interfacing", *Sensors and Actuators*, A-18, pp.195-206, 1989

[4.9] F.R. Riedijk and J.H. Huijsing, "An Integrated Absolute Temperature Sensor with Sigma-Delta A-D Conversion" *Sensors and Actuators*, A-34, pp.249-256, 1992

[4.10] A. Bakker and J.H. Huijsing, "Micropower CMOS Temperature Sensor with Digital Output", *IEEE Journal of Solid-State Circuits*, vol.31, pp. 933-937, July 1996

[4.11] S.R. Norsworthy, R. Schreier and G.C. Temes, "Delta-Sigma Data Converters", *IEEE Press*, ISBN 0-7803-1045-4, 1997

[4.12] G.C.M. Meijer and K. Vingerling, "Measurement of the Temperature Dependence of the $I_C(V_{BE})$ Characteristics of Integrated Bipolar Transistors", *IEEE Journal of Solid-State Circuits*, vol.15, pp. 237-240, April 1980

[4.13] Smart Battery System-Implementers Forum, "Smart Battery System Specification", *http:// www.sbs-forum.org*, rev.1.1, December 11, 1998

[4.14] F.R. Riedijk, "Integrated Smart Sensors with Digital Bus Interface", *PH.D. Thesis*, Delft University of Technology, The Netherlands, November 1993

5 Realizations of CMOS Smart Temperature Sensors

This chapter describes the realizations of three different CMOS smart temperature sensors that have been designed within the framework of this thesis. The first realization finds its application in a tyre monitoring system, where it watches the temperature of the tyre and compensates thermal cross-sensitivities of a pressure sensor. The second realization is a general-purpose ambient temperature sensor, where special attention has been paid to high uncalibrated accuracy. The last version can measure the temperature of a remote bipolar transistor and has its application in microprocessor thermal management.

5.1 Tyre monitoring system

5.1.1 Motivation

A tyre monitoring system has been desired by transportation companies for several years. The reasons for that are quite clear. Independent investigations show that more than half of all road vehicles have under-inflated tyres. Under pressures of only 0.2 bar result in a significant increase in fuel consumption and decrease of tyre life-time. Furthermore, safety is increased when tyres are at the correct pressure.

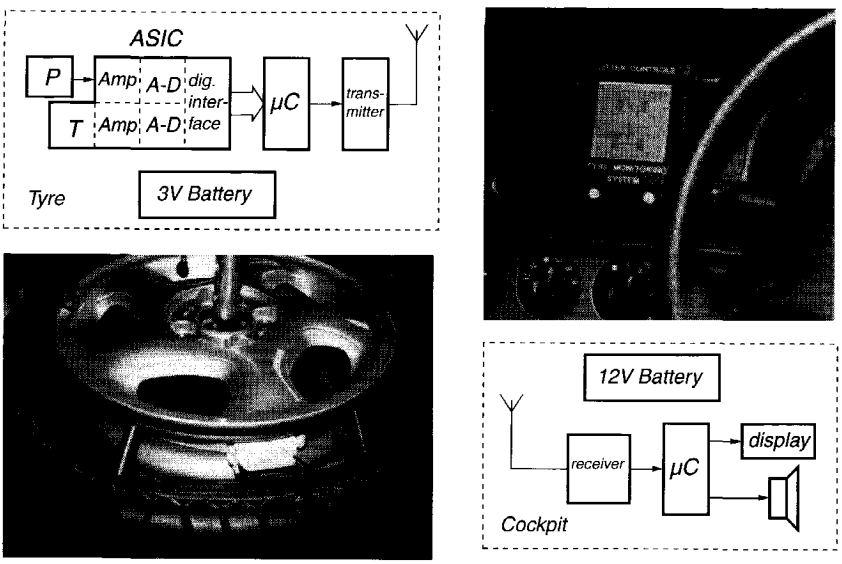

Fig. 5-1 *Principle of the tyre monitoring system*

At this time, many companies already offer a tyre monitoring system. Such systems consist of pressure and temperature sensors within the different tyres and a display in the cockpit of the vehicle. Data transmission between sensors and display is done by radio signals. The frequency of these radio signals lies in the reserved "low-power devices" bands on 434MHz, 869MHz or 2.456GHz, which are the same frequencies used for cordless phones, baby phones and garage openers [5.1].The principle of the tyre monitoring system is shown in figure 5-1.

Within all the tyres of the vehicle, which can be a car or a truck, a module is connected that consists of a pressure sensor, an ASIC (Application-Specific Integrated Circuit) containing a smart temperature sensor, a microcontroller, a transmitter, an antenna and a power supplying device, usually a battery. The ASIC has several functions. It can read out the pressure sensor and perform an analog-to-digital conversion of the pressure signal. It can also generate a digital code from its integrated smart temperature sensor. These two digital codes are sent to a microcontroller by the digital interface. The microcontroller analyses the results and sends it via a transmitter to the receiver unit in the cockpit of the vehicle. The receiver unit transports the signal further to another microcontroller,

which drives the display in the cockpit and can sound an alarm bell in emergency cases.

One of the major problems in these systems is the supply of power to the sensors in the tyre. Batteries seem to be the most straightforward choice, but they should be very small to avoid imbalances in the wheel. However, the batteries should have enough energy to assure the specified life-time. This life-time should be at least equal to the life-time of the tyre to avoid extra maintenance. In practical situations, this can be up to a few years. This trade-off requires extremely low-power circuitry for the sensor systems within the tyre.

Another possibility to supply power to the sensors is by electromagnetic radiation. This principle is also used for non-contact read-out of ID cards or transponders. Although this principle is feasible, it would be too costly for application in a low-cost market such as the automotive industry.

The need for temperature information from the tyre may not be clear at first sight. Still, almost all tyre monitoring systems measure temperature besides pressure. The reason for this is that the temperature of the tyre can give a lot of extra information. For example, on a cold winter day, when starting your engine the tyre temperature can be -10°C. After driving a quarter of an hour your tyres easily warm up to +40°C. During this warming-up, the pressure in your tyre will increase by approximately 50/300 or 16%. A 16% under-inflation normally sounds alarm bells, but at temperatures below zero, this can be quite normal.

Another reason for needing the temperature information is to be able to compensate for the pressure sensor's temperature dependence. Especially low-cost pressure sensors (automotive!) can have large thermal cross sensitivities. If this sensitivity can be compensated for, the price for the pressure sensor can be much lower.

A final reason for needing the temperature information is in cases where the pressure sensor is broken or out of specification. When a tyre is under-inflated or punctured, it will become hotter than usual. This information can warn the driver, even when the pressure sensor transmits a normal condition.

The project to make a tyre monitoring system started in early 1994 and was supported by the European Committee. Several companies were asked to participate because of the multi-disciplinary character. Delft university was asked to develop the integrated temperature sensor because of their experience in this area. The pressure sensor interface has been developed by CSEM in Switzerland. Results have been published by both TU Delft [5.3] and CSEM [5.4].

Table 5-1 *Target specifications for the tyre temperature sensor*

	Minimal	Typical	Maximal	Unit
Supply voltage	2.4	3.0	3.6	V
Average supply current (3 V, 2 samples/s)			3	μA
Temperature range	-40		130	°C
Inaccuracy			±1	°C
Bandwidth	2			samples/s
Resolution		8		bits

5.1.2 Specification

As already discussed in the previous paragraph, one of the major issues of the tyre monitoring system is the power consumption. After discussions with battery suppliers, it was found that for the maximum allowed size and weight, the (lithium) battery has an energy content of approximately 2 Ah at 3 Volt. This indicates that for a minimal system life-time of six years (which is the maximum life-time of a tyre), the maximum average current consumption of the whole tyre monitoring system should be below 2 Ah/6 years times 365 days times 24 hours, which is approximately 40 μA. This 40 μA has to be divided between the pressure sensor plus interface electronics, the integrated smart temperature sensor, the microcontroller and the transmitter. Because the microcontroller and the transmitter needed most of the energy, only 3 μA was left for the integrated smart temperature sensor. The bandwidth for the smart temperature sensor was set to 2 samples/s, which is expected to be enough to follow the temperature changes within the tyre with an accuracy of 1°C. The nominal supply voltage is set to 3 V, but the circuit should stay functional at supply voltages down to 2.4 V, which is the minimal voltage of the battery at the end of its life-time. The temperature range of the circuit is set from -40°C to +130°C. The maximum temperature of +130°C is not expected to occur too often (tyres are usually guaranteed to approximately 100°C), but investigations on tyre temperatures showed that temperatures up to +130°C can occur when heavy and lasting braking are needed during mountain descents. A summary of the main specifications is shown in table 5-1.

5.1.3 Design of the tyre temperature sensor

A-to-D conversion

One of the design issues that has not been discussed yet in the specification is the noisy environment in cars. This noisy environment is due to the many interference signals, especially from the engine ignition. The tyre monitoring system should be immune to these interfering signals, which indicates the necessity of integration of the pressure and temperature sensor signals or averaging of multiple samples.

From a power consumption point of view, it is better to use a slow, low-power integrating A-to-D converter, than a fast, high-power, sampling one and take the average of multiple samples. Although faster A-to-D converters may consume less power for single samples [5.2], obtaining the same noise immunity will always take more power than an integrating converter.

From the available integrating A-to-D converters, it is shown in chapter 4 that the best choice is the first-order continuous-time sigma-delta A-D converter. Higher-order sigma-delta A-D converters consume more power in exchange for shorter sampling times. These shorter sampling times are however unwanted, because of the worse interference rejection. Sampling sigma-delta A-D converters also have less interference suppression, because of the sampling action.

Technology

The choice for the implementation technology for the ASIC has been made after discussions with the different partners in the project. For the integrated temperature sensor, a bipolar or BiCMOS technology would be advantageous, because of the availability of high-performance bipolar transistors. However, a bipolar technology was not wanted by our partner who was responsible to read out the pressure sensor. The pressure sensor is based on a capacitive principle and can not be read out without the availability of high-performance switches. A BiCMOS process was not wanted by the commercial partner because it would increase the price for the ASIC too much. A choice was therefore made to implement the ASIC in standard CMOS.

Block diagram

The block diagram of the integrated tyre temperature sensor is shown in figure 5-2. On the left is shown the temperature sensor and reference block, which generates the temperature-dependent current I_{temp} and reference current I_{ref} for

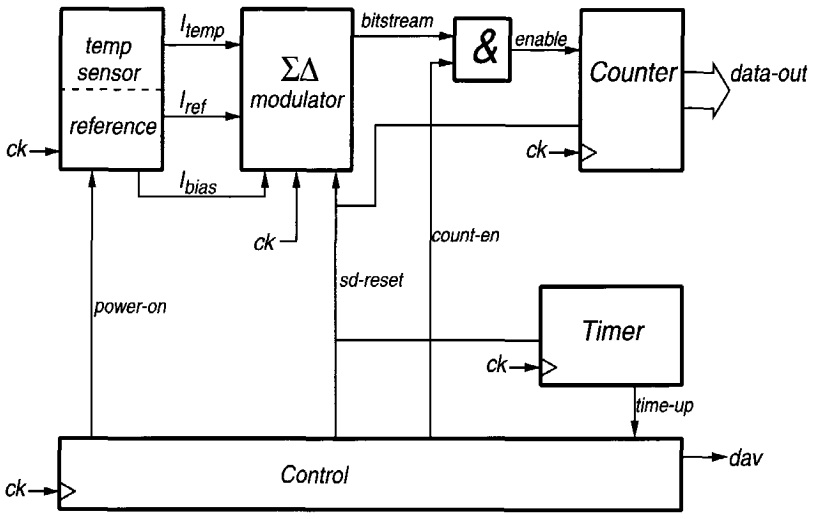

Fig. 5-2 *Block diagram of the integrated tyre temperature sensor*

the sigma-delta modulator. This block also generates the bias current I_{bias}, which is necessary to bias the different parts of the sigma-delta modulator. The bitstream output of the sigma-delta modulator is counted during a time that is specified by the timer. All the blocks are controlled by the control block.

At the start of each sample, the sigma-delta modulator, the counter and the timer are reset by the signal "*sd-reset*". The signal "*count-en*" becomes *high*, which enables the counter to count the number of *ones* in the bitstream. When the timer says "*time-up*", the "*count-en*" signal becomes *low*, which disables the counting and the "*dav*" signal becomes *high*, showing that the data at "*data-out*" is valid.

The control block also has the ability to power down the whole circuit. When it is in this mode, the signals "*power-on*" and "*dav*" are *low* while "*sd-reset*" is set *high*. By doing this, the analog blocks are powered down and the digital blocks consume the least amount of power, because there are no signal transitions, except for the clock. This power-down facility is necessary, because at this stage of the design it is expected that it is not possible to have all the blocks functioning at a total current consumption of less than 3μA.

The only way to meet the power consumption requirement is by reducing the sampling time and switching off the complete circuit between the samples. However, making the sampling time too short will result in worse interference

rejection. Therefore, the sampling time should be taken as large as possible, but small enough to have an average supply current of 3μA. The exact sampling time and ratio between power-off and power-on time will be decided at the end of the design track.

The temperature dependent current I_{temp} and the reference current I_{ref} will both be derived from a current that is PTAT (Proportional To Absolute Temperature) and a current that is proportional to the base-emitter voltage of a substrate bipolar transistor. Of those two, the PTAT current needs special attention because it is derived from a very small PTAT voltage, which has a temperature coefficient of typically 200μV/°C. To be able to reduce inaccuracies due to offsets or mismatch by dynamic offset-cancellation techniques (see chapter 2), a clock signal is also fed to the "temperature sensor and reference" block.

Calibration

It is expected that the integrated temperature sensor has to be calibrated to meet the accuracy specification of ±1°C over the whole temperature range from -40°C to +130°C. This calibration can be performed by laser trimming, programmable fuses or E(E)PROM. Another possibility is to do the calibration in the microcontroller's E(E)PROM. This has the advantage that the technology to implement the ASIC does not need to have facilities for the above-mentioned on-chip calibration techniques. A disadvantage is that the software engineer who writes the program for the microcontroller has to deal with calibration figures from the temperature sensor. This may complicate the high-volume production, because each microcontroller E(E)PROM should be programmed with different temperature calibration figures. However, because the pressure sensor interface also had to be calibrated for mismatches between the different pressure sensors [5.4], it was already impossible to calibrate the ASIC without a pressure sensor. Therefore, the decision has been made to calibrate both the temperature sensor and the pressure sensor interface in the microcontroller. This implies that in the design of the temperature sensor no attention will be paid to calibration.

5.1.4 Detailed design of the tyre temperature sensor

Low-power, high-accuracy PTAT current generator

The schematic of the low-power high-accuracy PTAT current generator is shown in figure 5-3. Cascodes are not shown because of simplicity, but are present at each PMOS transistor that is connected to V_{dd} and each NMOS transistor that is connected to V_{ss}. Compared to a conventional PTAT current generator, the accuracy is improved by the chopping of the first amplifier stage, which is formed

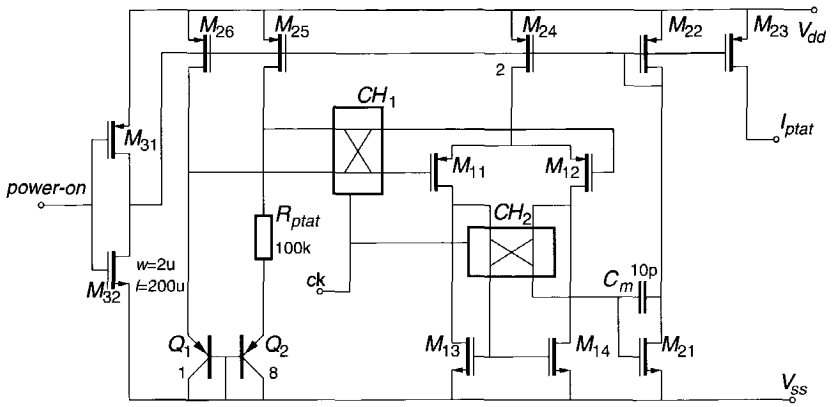

Fig. 5-3 *Schematic of the low-power, high-accuracy PTAT current generator. Cascodes are omitted for simplicity.*

by M_{11} through M_{14}. The offset of both the input stage and the mirror are modulated to the chopping frequency, which is chosen equal to the clock frequency. The filtering of the modulated offset is partly done by the Miller capacitor C_m, but the main filtering is done in the ΣΔ A-to-D converter. The ΣΔ A-to-D converter acts as a digital low-pass filter with notches at the clock frequency and higher harmonics (see paragraph 4.2).

The current of the PTAT generator is determined by the emitter area ratio of Q_2 and Q_1 and the resistor R_{ptat}. For the chosen emitter area ratio of 8 this yields:

$$I_{ptat} = \frac{1}{R_{ptat}} \cdot \frac{kT}{q} \ln 8 \qquad (5\text{-}1)$$

For a nominal value of R_{ptat} of 100kΩ, I_{ptat} equals approximately 500nA at T=300K. This nominal PTAT current is a trade-off between power consumption and accuracy. The accuracy reason can be explained as follows. A nominal current of 500nA at 300K yields a sensitivity of approximately 1.7nA/°C. Especially at temperatures above 100°C and up to 130°C, leakage currents can easily rise close to the nA level. Calibration of these errors is possible, but because leakage currents can change greatly after a certain time of operation, major aging problems can occur.

The reason for using a two-stage amplifier is to have enough gain at the lowest amount of power. This can be explained as follows. To have enough accuracy, it is

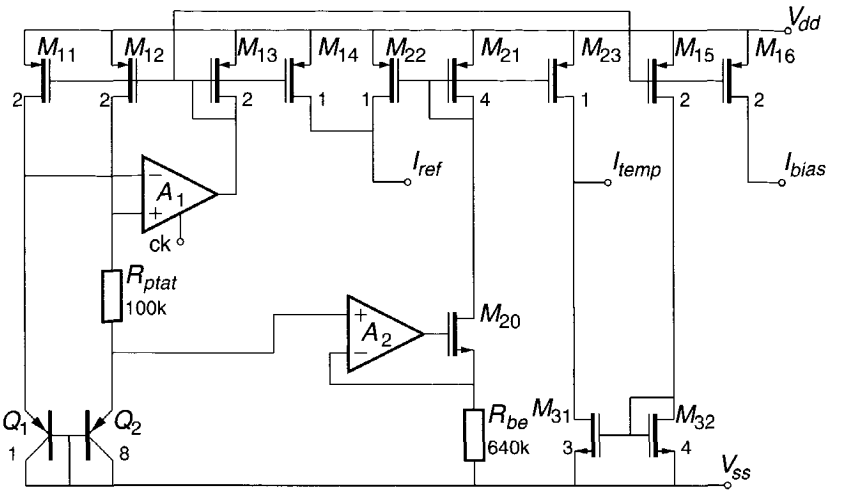

Fig. 5-4 *Simplified schematic of the circuit generating I_{temp}, I_{ref} and I_{bias}.*

necessary to have a gain within the loop of at least a factor of 50. Higher is not really necessary, because only offsets and mismatches need to be compensated. To obtain a voltage gain of 50 with only one stage and an output resistor R_{ptat} of 100kΩ, would require a g_m of 500μA/V. This would require a drain-current, even in weak inversion, of at least 20μA, which is far above our available budget. Our two-stage amplifier with long input transistors M_{11} and M_{12} and cascodes on all the mirrors has a voltage gain at room temperature of approximately 200. The total current consumption is only 3μA, which we believe is the lowest possible power consumption for the required accuracy.

The power-down circuit requires special attention. If the signal *power-on* is *high*, a narrow and long transistor M_{32} is switched on, resulting in a very small current through the transistors M_{22} to M_{26}. This is enough to start up the circuit. When *power-on* is switched to *low* again, M_{31} is switched on, forcing the gates of M_{22} to M_{26} to V_{dd}, resulting in a power-down of the circuit. Although not shown in figure 5-3 for simplicity, the current supplied by M_{32} is switched off after a few clock pulses to increase the accuracy. Also the start-up behaviour of the cascodes, which is not shown nor explained here, needs special attention.

Generation of I_{temp}, I_{ref} and I_{bias}

A simplified schematic of the circuit that generates the currents I_{temp}, I_{ref} and I_{bias} is shown in figure 5-4. On the left the PTAT current generator can be

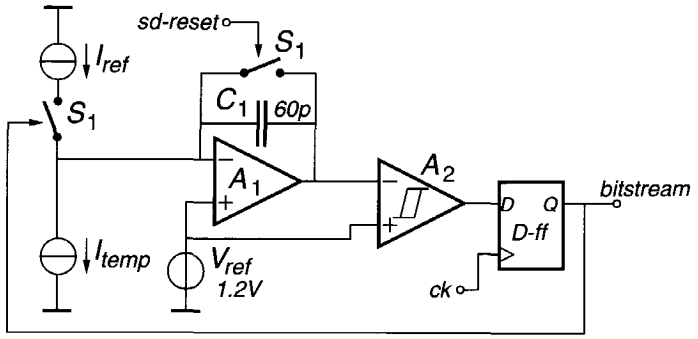

Fig. 5-5 *Schematic of the ΣΔ modulator*

distinguished. In the centre of the figure, the base-emitter voltage of Q_2 is transformed into a current by means of A_2, M_{20} and R_{be}. The value of this current is at room temperature approximately 1μA, which is a trade-off between power consumption and chip area of the resistor R_{be}. The I_{be} current is divided by four and split into two parts by the current mirrors formed by M_{21}, M_{22} and M_{23}. The left part of this current is added to a PTAT current generated by M_{14}, which is half the nominal PTAT current of 500nA. The sum of these two currents forms the reference current I_{ref}, which is independent of temperature (if we do not take into account the temperature dependencies of the resistors R_{ptat} and R_{be}). I_{temp} is formed by a PTAT current that is 0.75 times the nominal I_{ptat} current minus the right part of the I_{be} current from transistor M_{23}. The bias current I_{bias} is exactly equal to the nominal PTAT current I_{ptat}. The reason for choosing a temperature signal current I_{temp} as the difference between a certain amount of I_{ptat} and I_{be} is to be able to fully use the dynamic range of the ΣΔ modulator. This so-called Kelvin-to-Celsius conversion technique is explained in paragraph 4.3.

The total power consumption of this circuit is kept as low as possible and yields approximately 7μA.

Sigma-Delta modulator

The schematic of the ΣΔ modulator is shown in figure 5-5. The bitstream output of the modulator is given by

$$bitstream = \frac{I_{temp}}{I_{ref}} = \frac{1.5I_{ptat} - I_{be}}{I_{ptat} + I_{be}} \qquad (5\text{-}2)$$

The currents I_{temp} and I_{ref} are chosen in such a way that at the low end of the temperature range, which is -40°C, this ratio is 0.1 and at the high end, which is +130°C, the ratio is 0.9. The margins of around 10% at each corner are necessary to be able to perform the calibration. The choice for the size of the integrating capacitor C_1 of 60pF is a trade-off between chip area and accuracy. The larger the capacitance, the smaller the problems with charge injection from switch S_1. A large capacitance also offers lower possible clock frequencies. This is advantageous because this implies that the frequency of the chopper amplifier in the PTAT-current generator can also be lower. This results in better accuracy in the chopper amplifier, because the residual error due to chopping is linearly dependent on the chopping frequency. The clock frequency has been chosen 16kHz. This frequency can be easily derived from a low-cost 32kHz-watch crystal. With a required resolution of 8 bits, this clock frequency implies a maximum sampling speed of $16\text{kHz}/2^8$, which is approximately equal to 60 samples/s.

The reference voltage V_{ref} is set to 1.2V to be sure that both current sources I_{ref} and I_{temp} have their cascodes working in the saturation domain, where they have the highest possible output impedance. At the start of each sample the integrator is reset by S_2 to assure that the initial voltage at the beginning of each sample is the same.

The power consumption of the complete ΣΔ modulator block is approximately equal to10μA. Together with the temperature sensor and current reference block, this yields 17μA. Adding a few micro Amps for the digital part, the total power consumption of the circuit should be around 20μA at continuous operation at 60 samples/s. For the required sampling rate of 2 samples/s this can theoretically be reduced to less than 1μA. This will, however, not be met in practice because of the leakage currents that can be up to a few micro Amps.

Layout

The ASIC containing the circuitry for the tyre temperature sensor and the read-out of the pressure sensor is implemented in a 2μm standard CMOS process. A micrograph is shown in figure 5-6. Special attention has been paid to the matching of transistors and resistors in the circuit that generates the temperature and reference currents. The total chip measures 12mm^2 of which 25% is occupied by digital cells. The analog part of the tyre temperature sensor is marked with a white rectangle and measures 1.5mm^2. Around 20% of this area is occupied by the resistors R_{be} and R_{ptat} of 640k and 100k resp. The integrating capacitor C_1 in the ΣΔ modulator of 60pF occupies approximately 10% of the chip area.

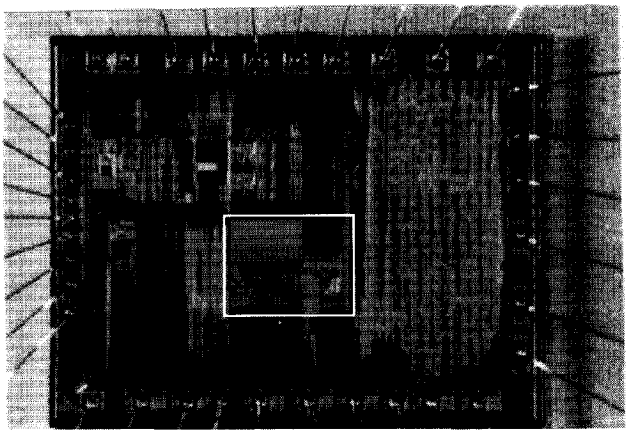

Fig. 5-6 *Micrograph of the integrated circuit for the tyre monitoring system. The white rectangle shows the analog part of the tyre temperature sensor.*

To reduce costs, the circuit has afterwards also been implemented in a 0.7µm CMOS process, with a reduced chip area of $5mm^2$. The measurements did not show significant differences and will therefore not be discussed here.

5.1.5 Measurement results

The chip has been encapsulated in a ceramic IC package and has been exposed to temperatures in the range from -40°C to +120°C. The circuit showed full functionality for supply voltages between 2.2V and 5.0V. The total current consumption at continuous operation is 25µA. At stand-by, the current consumption drops to 2µA. For the nominal sampling rate of 2 samples/s the current consumption is 3µA, which exactly meets our target specification. The full specifications are listed in table 5-2.

Three samples have been measured more elaborately to analyse the temperature inaccuracies over the temperature range. The results of the uncalibrated parts are shown in figure 5-7. These samples show an error of ±5°C at room temperature and ±7°C over the full temperature range. To achieve the specified accuracy of ±1°C over the full temperature range, the circuit has to be calibrated at least two temperatures. The results for a proposed calibration at -20°C and +100°C are shown in figure 5-8.

Table 5-2 *Specifications of the measured tyre temperature sensor*

	Minimal	**Typical**	**Maximal**	**Unit**
Supply voltage	2.2	3.0	5.0	V
Supply current Nominal (3V, 2 samples/s) During sample		3 25		µA
Temperature range	-40		120	°C
Inaccuracy (2 calibrations)			±1	°C
Bandwidth			50	samples/s
Resolution		8		bits
Clock frequency	25		40	kHz
Power supply regulation		0.1		°C/V
Noise		0.1		°C

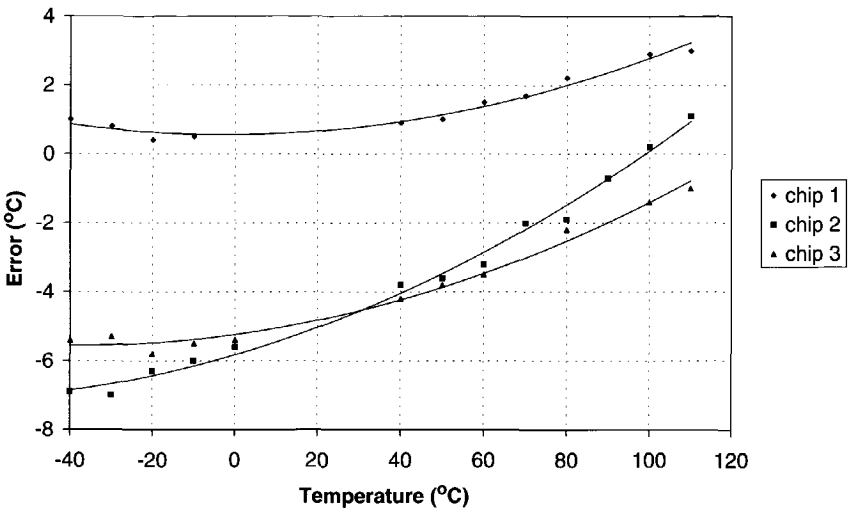

Fig. 5-7 *Temperature errors of three samples before calibration*

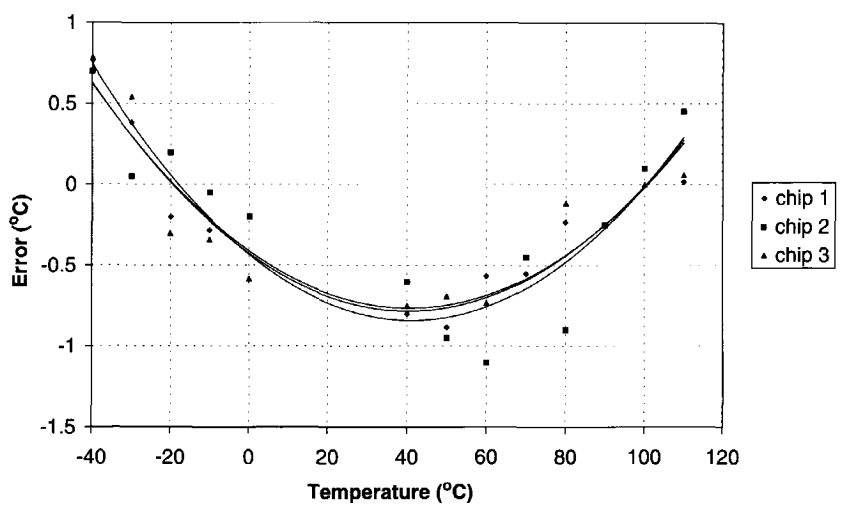

Fig. 5-8 *Temperature errors after calibration at -20°C and +100°C*

Table 5-3 *Manufacturing costs of smart temperature sensors per piece*

Description	Cost
Design	10¢
Processing	30¢
Test	15¢
Sawing/Bonding/Packaging	25¢
Calibration	20¢
Total	$1.00

5.2 High-accuracy temperature sensor

5.2.1 Motivation

One of the major issues in the tyre monitoring system project was the size of the battery. This implied that the most important specification of the tyre monitoring system was the extreme low power consumption. We succeeded in that by reducing the amount of circuitry to the minimum, at the cost of a very simple digital interface and medium uncalibrated temperature accuracy. The market for tyre temperature sensors is however quite small. The largest markets at this moment are found in thermal management applications, like personal computers and domestic appliances. In these kinds of applications, power consumption is not one of the most important issues. The magic word here is low-cost: How can I make my device cheaper than that of my competitor?

Although the author does not pretend to have much experience in sales, a short overview of the costs of smart temperature sensor manufacturing is considered useful here. In this overview, which is shown in table 5-3, it is assumed that the average sales are over 1 million per year. The calculation is also based on a 1K+ selling price of $1, which is quite arbitrary, but reasonably matches the current price of an LM75 equivalent. The costs have been divided into five parts. In the next paragraphs the possibility of reducing costs related to each of these five parts is discussed.

Design

The design costs are built up of costs related to the complete design process, varying from salaries of design engineers to processing of first silicon and writing the datasheets. These investment costs can be rather high and are estimated to be somewhere between $500,000 and $1,000,000. However, divided by the total expected sales of over 10 million pieces, the costs per piece can still be kept below 10¢. The costs for design are quite hard to reduce. A simple design can, however, reduce the risks and the additional costs of a redesign.

Processing

The processing costs for large volume production are mainly dictated by the size of the chip and the complexity of the process, i.e. the number of processing steps. The cheapest available process that can fulfil the specifications of a general smart temperature sensor is obviously CMOS. The standard CMOS process, however, may need to be extended with options for high-density capacitors, thin-film resistors or calibration facilities like E(E)PROM or OTP (One Time PROM). The choice for these extensions is usually a trade-off between costs and performance. It is a challenge for the designer to reduce the processing costs by meeting the specifications without the need for process extensions and without increasing the chip size significantly.

Test

The costs for testing are almost fully dictated by the testing time. The testing time can be reduced by adding special test features in the chip. However, these additions should not increase the chip size too much.

Sawing/Bonding/Packaging

The cost for sawing, bonding and packaging are mainly dictated by the number of pins of the chip. The number of pins is usually dictated by the customer or by competitors and can therefore hardly be optimized.

Calibration

The costs for calibration depend mainly on the time the chip is in the tester. Time is needed to assure temperature stability, to obtain the temperature information and to do the programming. These costs can offcourse be removed if the sensor has the desired accuracy without calibration. However, this seems hardly feasible because datasheets from the different suppliers show that all currently available smart temperature sensors need to be calibrated to meet the accuracy specification.

Table 5-4 *Specifications of the high-accuracy temperature sensor*

	Minimal	Typical	Maximal	Unit
Supply voltage	3.0	3.3	5.5	V
Supply current I^2C inactive I^2C active Shutdown mode		250 1	1000	µA
Temperature range	-55		125	°C
Inaccuracy T_A= -25°C to +100°C T_A= -55°C to +125°C			±2.0 ±3.0	°C
Conversion time		100		ms
Resolution		9 0.5		bits °C

The conclusion from the above-mentioned discussion is that the costs can be reduced by using a standard CMOS process, by minimizing the chip size and by increasing the accuracy to be able to omit the calibration. If we succeed in removing the need for calibration, around 20% of the costs can be directly reduced from the calibration phase, according to table 5-3. Moreover, the processing costs will be further reduced because of less chip area and less processing steps. Adding this, it can be expected that the total manufacturing costs can be reduced by approximately 25%.

The next paragraphs will describe a smart temperature sensor which is optimized for high accuracy before calibration. The project has been done in cooperation with Philips Semiconductors in Sunnyvale, California.

5.2.2 Specification

The specifications of our smart temperature sensor follow those of the industry-standard LM75 from National Semiconductor [5.5]. These are listed in table 5-4. The most challenging on this specification is, as already mentioned, to meet the temperature accuracy without any calibration. National specifies a temperature

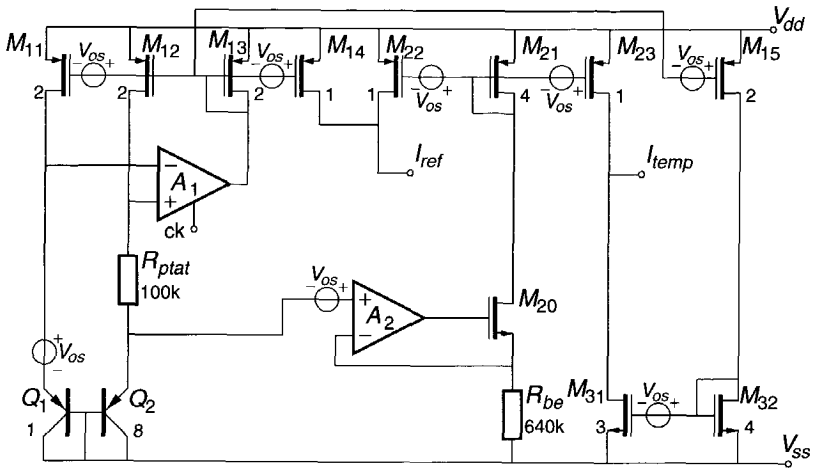

Fig. 5-9 *Copy of the part of the tyre temperature sensor that generates I_{temp} and I_{ref} to explore the sources of error.*

inaccuracy of ±2°C in a limited range from -25°C to +100°C and ±3°C in the total specified range from -55°C until +125°C. The major question for us is: Is it possible to meet this accuracy specification without the need for trimming?

Expected accuracy before trimming

To estimate the uncalibrated accuracy of our smart temperature sensor, the most important sources of error should be explored and modelled. If we consider the circuit structure of the tyre temperature sensor, we can assume that all dominant errors are made before the ΣΔ modulator. This is highly feasible, because the ΣΔ modulator is known for its very high linearity which implies that resolutions up to 9 bits can be achieved without any problem. The major errors will therefore be generated in the circuit generating I_{temp} and I_{ref}. This part of the circuit of the tyre monitoring system is shown again in figure 5-9, together with the most important sources of error. A list of these errors and their contributions are shown in table 5-5. The calculation of the contributions of the different error sources is not so straightforward. For example, to calculate the error due to mismatch between M_{12} and M_{14}, one should know the transfer function between I_{ptat} and the output digital code. For the circuit of figure 5-9, the output is given by:

$$Tempdata = \frac{I_{temp}}{I_{ref}} = \frac{1.5I_{ptat} - I_{be}}{I_{ptat} + I_{be}} \qquad (5\text{-}3)$$

Table 5-5 *Typical errors in tyre temperature sensor according to* figure 5-9

Kind of error		Typical value	Error contribution for typical value
Description	**Symbol**		
Offset Q_1 and Q_2	$\sigma(V_{os,Q})$	0.1mV	0.25°C
Mismatch M_{12}, M_{11}	$\sigma(\Delta M_{12,11})$	1.0%	1.3°C
Mismatch M_{12}, M_{14}	$\sigma(\Delta M_{12,14})$	1.0%	0.3°C
Mismatch M_{12}, M_{15}	$\sigma(\Delta M_{12,15})$	1.0%	1.8°C
Mismatch M_{32}, M_{31}	$\sigma(\Delta M_{32,31})$	1.0%	1.8°C
Mismatch M_{21}, M_{22}	$\sigma(\Delta M_{21,22})$	1.0%	0.3°C
Mismatch M_{21}, M_{23}	$\sigma(\Delta M_{21,23})$	1.0%	1.2°C
Offset A_2	$\sigma(V_{os,A2})$	1.0mV	0.25°C
Spread absolute value R_{ptat}	$\sigma(\Delta R_{sh})$	4.0%	0.25°C
Mismatch R_{ptat}, R_{be}	$\sigma(\Delta R_{ptat,be})$	0.1%	0.15°C
Spread absolute value V_{BE}	$\sigma(\Delta V_{BE})$	2.0mV	0.5°C
Total:			3.2°C

Errors uncorrelated, V_{BE} = 600 mV, T = 300K

It can be clearly seen from equation (5-3) that errors in the signal I_{ptat} change both the numerator and the denominator. If we assume that at room temperature I_{ptat} and I_{be} are approximately equal, an error of Δ in I_{ptat} results in:

$$Tempdata = \frac{1.5(1+\Delta)-1}{(1+\Delta)+1} = \frac{0.5+1.5\Delta}{2+\Delta} \approx 0.25 + 0.62\Delta \qquad (5\text{-}4)$$

Due to the Kelvin-to-Celsius conversion (see paragraph 4.3), the temperature range is limited to approximately 600K/2.5 which equals 240K or 240°C. A Δ of 1% therefore yields a temperature error of 0.62*1%*240°C or 1.5°C.

From table 5-5 it can be seen that the largest contributors to the total error of 3.2°C are the mismatches in the various current mirrors. When we cancel these with dynamic offset-cancellation techniques, the residual error will be greatly reduced to approximately 0.7°C. For a 3-σ production bandwidth, this feasible accuracy is around 2.1°C which comes very close to the LM75 specification, which specifies ±2°C in a limited temperature range and ±3°C over the full range. The major difference, however, will be that we can achieve this accuracy without one single calibration.

The next paragraphs will describe the detailed design of the high-accuracy smart temperature sensor.

5.2.3 Design

As was seen in the previous paragraph, the largest contributors to the error in the tyre temperature sensor were the various current mirrors. To improve the accuracy, we can do a chopping action on all the current mirrors. It is, however, better to first consider the necessity of these current mirrors and find out if we can omit some of them. If we look in detail, we can see from table 5-5 that the largest contributors to the error are the mismatches between M_{12} and M_{15} and between M_{32} and M_{31}. The reason for this is that these current mirrors transport the I_{ptat} signal, which is the most important temperature signal. The need for these current mirrors is explained by the fact that firstly the ΣΔ modulator needs both a *sourcing* and a *sinking* current to achieve a charge balance and, secondly, that the I_{ptat} signal is needed both as a *sourcing* and a *sinking* current because it appears as a positive signal, both in the numerator and the denominator of the *tempdata* output.[1]

The conclusion of the above-mentioned reasoning is that the accuracy of the smart temperature sensor can be much improved if it is possible to 1) make a ΣΔ modulator which only needs a *sourcing* I_{be} and a *sinking* I_{ptat} and, 2) make a *high-accuracy sinking* PTAT-current generator without mirrors in the signal path.

ΣΔ modulator with sourcing I_{be} and sinking I_{ptat}

A trick to move the I_{ptat} current source from the sourcing side to the sinking side of the ΣΔ modulator is to use a switch with an *inverted* bitstream signal. This principle is shown in figure 5-10. When the switch S_1 in figure 5-10 (a) is closed,

[1] The signal I_{be} is only needed as a *sourcing* current because it appears as a *negative* signal in the numerator

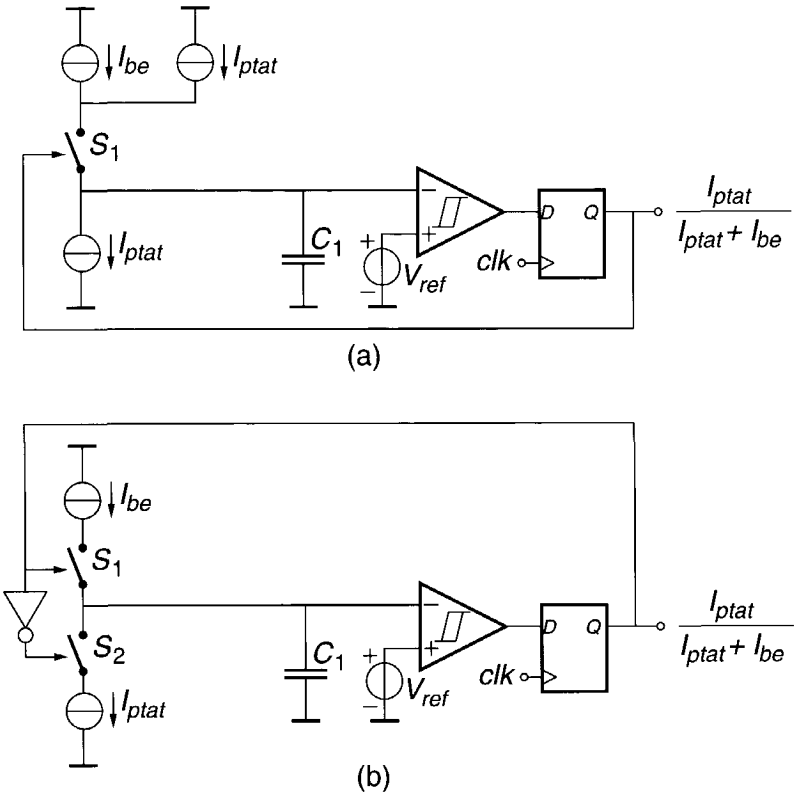

Fig. 5-10 *Method to move the I_{ptat} current source from the sourcing side to the sinking side in the ΣΔ modulator. Circuits (a) and (b) both have the same output.*

the sourcing I_{ptat} cancels the sinking I_{ptat}, so the net result is zero. When switch S_1 is open, I_{ptat} is sinked out of C_1. The same action can however also be achieved by using the inverse bitstream signal to control switch S_1 and omitting the sourcing I_{ptat}. This is shown in figure 5-10 (b).

The implementation of the ΣΔ modulator with only *sourcing* I_{be} current sources and *sinking* I_{ptat} current sources is shown in figure 5-11 The output yields

$$bitstream = \frac{2I_{ptat} - I_{be}}{I_{ptat} + I_{be}} \tag{5-5}$$

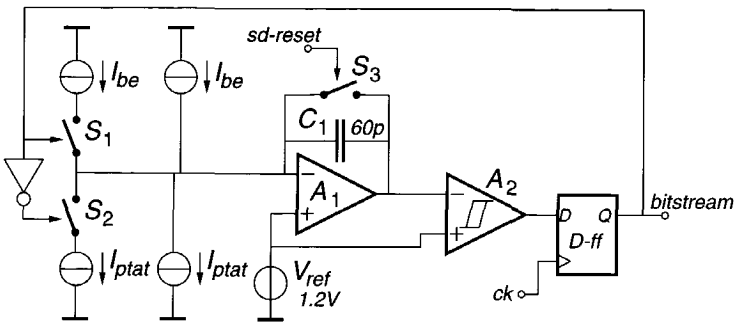

Fig. 5-11 *ΣΔ modulator implementation which only needs sourcing I_{be} current sources and sinking I_{ptat} current sources.*

More information about the derivation of this result can be found in paragraph 4.3. Compared to the tyre temperature system, the weight of I_{ptat} in the numerator has been slightly increased from 1.5 to 2. This results in an improved usage of dynamic range of the ΣΔ A-to-D converter. Compared to the tyre temperature sensor, the temperature range of this new high-accuracy smart temperature sensor is slightly reduced from 240°C to 200°C. However, this reduced temperature range still gives enough headroom to meet the minimal range specification of the LM75 of 180°C (-55°C to 125°C).

High-accuracy PTAT-current generator

The principle of the sinking PTAT-current generator without current mirrors is shown in figure 5-12. The bipolar transistors Q_1 and Q_2 are biased with two bias currents I_1 and I_2 that have a ratio of n. The emitters of Q_1 and Q_2 are kept at the same voltage by the loop involving amplifier A_1. The bases of Q_1 and Q_2 therefore have a voltage difference that is equal to kT/q*ln(n). The PTAT current that is flowing through R_2 is split into two equal parts by M_1 and M_2, thus generating the two I_{ptat} currents that are needed by the ΣΔ modulator.

A problem with this circuit is that the base current of Q_2 also flows through the resistor R_2, which can significantly reduce the accuracy. To compensate for this problem, a resistor R_1 is added to the base of Q_1, which is n times smaller than R_2. The voltage drop across this resistor due to the base current of Q_1 exactly compensates for this error.

The major sources of inaccuracy in this circuit are the offset of the opamp A_1, the mismatch between M_1 and M_2, and the mismatch between the bias currents I_1 and

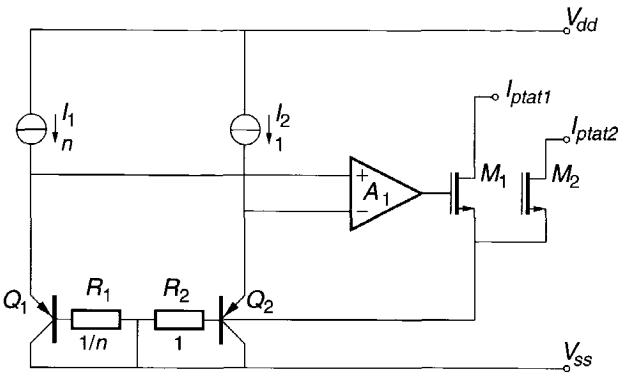

Fig. 5-12 *PTAT-current generator principle with sinking output without current mirrors in the signal path.*

I_2. These inaccuracies will be cancelled by a chopping technique. The choice for the chopping frequency should at first sight be as low as possible, because the magnitude of residual errors increase linearly with the chopping frequency (see paragraph 2.3). The lowest possible chopping frequency equals the sampling frequency. However, we should take care that, at that frequency, large errors in the denominator are not completely cancelled. This can be explained as follows. Suppose we have an error in I_{ptat} of Δ. If we choose the chopping frequency equal to the sampling frequency, the output of the $\Sigma\Delta$ A-to-D converter will be equal to the average of one sample with $I_{ptat}(1+\Delta)$ and one with $I_{ptat}(1-\Delta)$,

$$Bitstream = \frac{1}{2}\left\{\frac{2I_{ptat}(1+\Delta)-I_{be}}{I_{ptat}(1+\Delta)+I_{be}} + \frac{2I_{ptat}(1-\Delta)-I_{be}}{I_{ptat}(1-\Delta)+I_{be}}\right\} \qquad (5\text{-}6)$$

This equation is quite hard to solve for general values of I_{ptat} and I_{be}. However, for $T=T_{ref}$, I_{ptat} is equal to I_{be}. Equation 5-6 is then reduced to

$$Bitstream = \frac{1}{2}\left\{\frac{1+2\Delta}{2+\Delta} + \frac{1-2\Delta}{2-\Delta}\right\} = \frac{1}{2}\left\{\frac{4-4\Delta^2}{4-\Delta^2}\right\} \approx \frac{1}{2}(1-0.75\Delta^2) \qquad (5\text{-}7)$$

Equation 5-7 shows that with the lowest possible chopping frequency the errors are approximately squared. This implies that, for example, an error of Δ of 1% is reduced to 0.0075%, which equals approximately 0.0075%*200K is 0.015°C. For our specification with a maximum inaccuracy of ±2°C this can be assumed negligible. However, errors of Δ of 10% are reduced to 0.75%, which equals

approximately 0.75%*200K is 1.5°C. This latter error is definitely not negligible. If we define that temperature errors of typically 0.1°C are acceptable, only errors in I_{ptat} that are smaller than approximately 2.5% can be cancelled with a chopping frequency that is equal to the sampling frequency. The question is therefore: Are there errors in the PTAT-current generator of figure 5-12 that are larger than 2.5%? The mismatch in the bias currents I_1 and I_2 will be approximately 1%. Also the mismatch in M_1 and M_2 will not be larger than 1%. However, the offset of the operational amplifier A_1 can be up to 1 mV, which will, depending on the choice of the bias current ratio n, give an error between 2 and 3%.

The offset of the amplifier after low-frequency chopping will therefore be on the edge of acceptable errors. Increasing the chopping frequency to the clock frequency will completely cancel the errors, because the chopping moments will fall between the decision moments of the comparator, so they are not "seen" by the ΣΔ modulator. Another advantage of increasing the chopping frequency of the amplifier is that no special attention has to be given to minimize the offset. Finally, the $1/f$ noise of A_1 will also be reduced.

A drawback of increasing the chopping frequency is the larger residual error. This can be up to a few hundreds of micro Volts for frequencies up to 10kHz, based on experiences with regular chopper amplifiers. However, a residual error of 100μV gives a temperature error of 1 to 2°C, depending on the choice of the bias current ratio n. A solution is therefore found in applying the nested chopper technique as explained in paragraph 2.4. This results in the high-accuracy PTAT-current generator with nested chopper amplifier as shown in figure 5-13. The amplifier is split into three parts. The first part A_1 is chopped at the clock frequency which is called in this figure *chophigh*. The second one A_2 is chopped together with the first one A_1 at the sampling frequency, which is called *choplow*. The third part of the amplifier, A_3, has no amplification but serves only to make the output single-ended. The output transistors M_1 and M_2 are chopped by CH_6. The bipolar transistors Q_1 and Q_2 are also chopped by S_2 and S_3.

The ratio n of the bias currents is chosen three. This eases the Dynamic Element Matching because this can now be done in four phases, needing two control signals. These signals are already available as *choplow* and *chophigh*. If we would have chosen a bias current ratio n of for example seven, we would have needed at least eight phases and three control signals to dynamically match these currents.

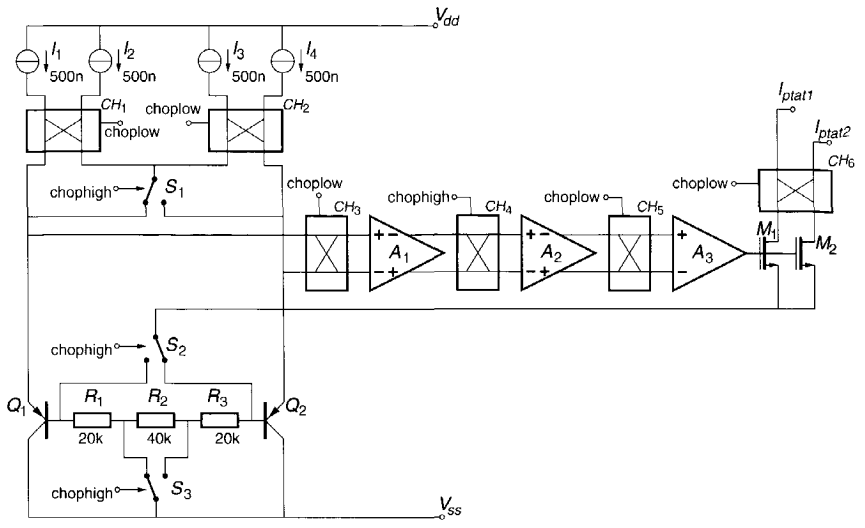

Fig. 5-13 *High-accuracy PTAT-current generator with nested chopper amplifier.*

High-accuracy I_{be} generator

The errors of the I_{be} generator are all smaller than 1%. This implies that all mismatches can be chopped at the low sampling frequency and no nested chopper technique is necessary. The implementation is shown in figure 5-14. The V_{BE} voltage at the input of the generator is made by a bias current I_{bias} and a bipolar transistor. This bias current is derived from a bias circuit and has an accuracy that is approximately equal to the spread in sheet resistance, which is assumed to be approximately 4%. According to table 5-5, this equals an inaccuracy of 0.25°C. The current mirror formed by M_{31} and M_{32} is chopped by CH_3. The current splitter formed by M_{21} and M_{22} is chopped by CH_4. All the choppers are controlled by the low frequency sampling signal *choplow*.

I^2C bus interface

The I^2C bus interface is completely designed by Philips. Specifications for this interface can be found in [5.6].

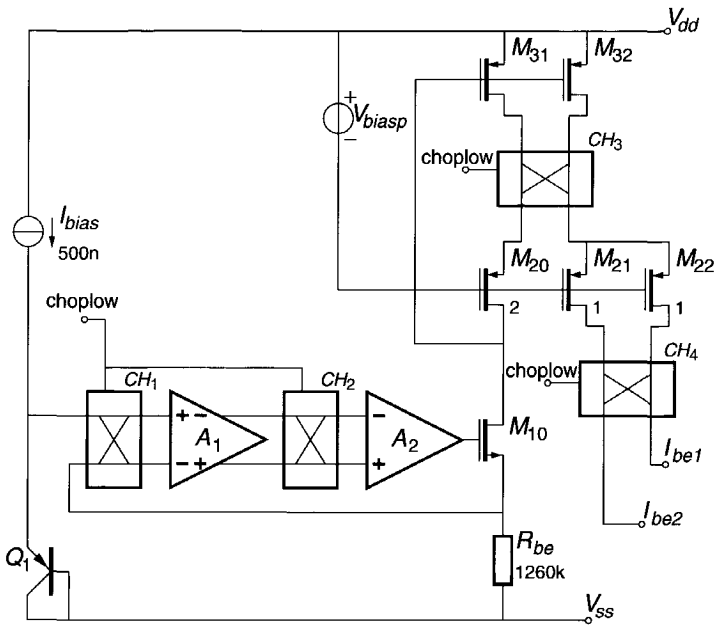

Fig. 5-14 *High-accuracy I_{be} generator*

Control

The control block is comparable to the tyre temperature sensor as compared to handling of the power down. It has some additional features like over-temperature and under-temperature watchdogs that are listed in the LM75 specification.

Layout

A micrograph of the layout is shown in figure 5-15. The circuit is made in a standard 0.7µm CMOS process. Next to the high-accuracy temperature sensor, an interface for an external bipolar temperature sensor has also been integrated. This interface will be described in paragraph 5.3. The total chip size is 4.5mm^2 of which approximately 50% is occupied by digital circuitry. The analog building blocks for the high-accuracy temperature sensor are marked by a white rectangle and occupy approximately 0.7mm^2. A stand-alone version with LM75 specifications is expected to have a chip size of 2mm^2.

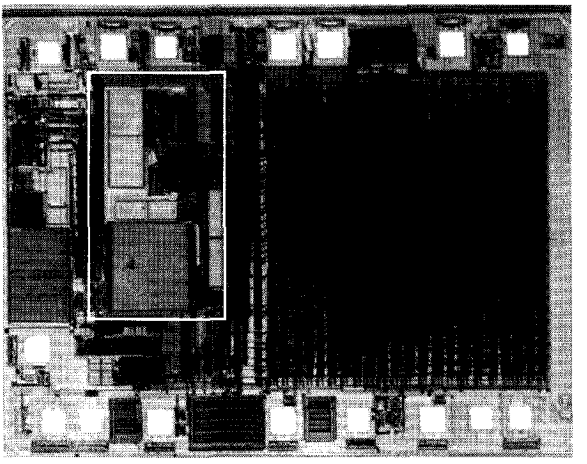

Fig. 5-15 *Micrograph of the high-accuracy temperature sensor chip*

5.2.4 Measurement results

The chip has been encapsulated in a standard plastic IC package and has been exposed to temperatures in the range from -60°C to +130°C. The circuit shows full functionality for supply voltages between 2.9V and 5.5V.

We tested 24 samples of the circuit from one batch over a temperature range from -40°C to +120°C. The results are shown in figure 5-16. All circuits are within ±1°C in a temperature range from -40°C to +80°C. The typical spread at room temperature is ±0.4°C. Based on the error analysis from the tyre temperature sensor on page 96, we expected ±0.7°C between batches. The largest error, however, is the spread in V_{BE} which is definitely much lower within one batch. Until now, several hundreds of chips have been tested from different batches. Errors were all within ±1°C at room temperature. This proves that it is possible to make a smart CMOS temperature sensor with a 3-σ uncalibrated accuracy at room temperature of better than ±2°C.

From the results shown in figure 5-16, it can also be seen that the LM75 specification of an accuracy of ±2°C in a limited temperature range from -25°C to +100°C is met for these 24 samples. However, if we take into account the full 3-σ production bandwidth, this is not completely sure. We solved this problem by doing a curvature correction in a subsequent version that is currently being processed. This curvature correction will lower the high-temperature curves with

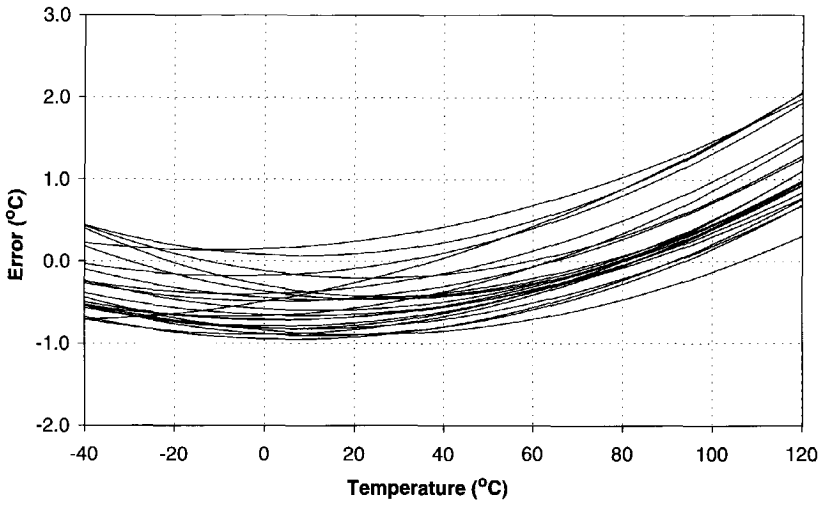

Fig. 5-16 *Temperature errors of 24 samples without calibration.*

approximately 1°C. With this new version the LM75 specification will be completely met, without any calibration.

A summary of the measured specifications is listed in table 5-6.

The results of this work are published in [5.7].

5.3 Remote microprocessor temperature sensor

5.3.1 Motivation

One of the most important issues in laptop design is the thermal management. Laptops are optimized for small size and low-weight. This implies that only a minimum of heat sink mass will be used to cool high-power dissipating parts like the microprocessor. To still maintain a high reliability, modern laptops have variable microprocessor clock speeds. In situations where the microprocessor is in danger because of a too high temperature, its power dissipation is reduced by lowering the clock frequency.

Table 5-6 *Measured specifications of the high-accuracy temperature sensor*

	Minimal	Typical	Maximal	Unit
Supply voltage	2.9	3.3	5.5	V
Supply current I^2C inactive I^2C active Shutdown mode		50 1	 500 	µA
Temperature range	-55		125	°C
Inaccuracy T_A= -25°C to +100°C T_A= -55°C to +125°C			 ±2.0 ±3.0	°C
Conversion time		100		ms
Resolution		10 0.25		bits °C
Power supply regulation		0.2		°C/V
Noise		0.05		°C

A block diagram of the above-mentioned thermal control system is shown in figure 5-17. Laptops, but also desktops, have a system management that takes care of the system functions like the different power supply voltages. Communication between the different parts of the system is done over a two-wire SM-bus interface that is based on I^2C [5.8]. The junction temperature of the microprocessor is continuously measured by a temperature measurement chip, which also measures the ambient temperature. This information is sent to the system management. When the temperature becomes too high, the system management decides whether or not to increase the fan speed or decrease the microprocessor clock speed.

Our role in the project, that has been done in cooperation with Philips Semiconductors in Sunnyvale, was the design of the temperature measurement chip. For the measurement of the ambient temperature, the design of the high-accuracy temperature sensor, as described in paragraph 5.2, has been used. The

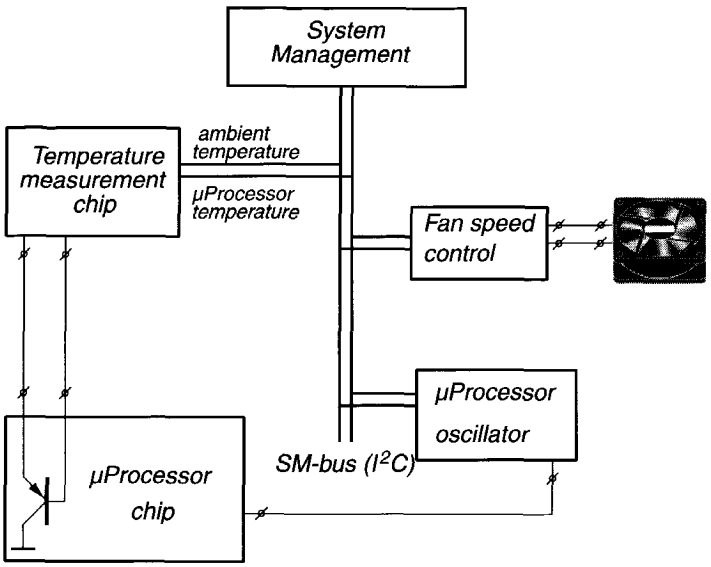

Fig. 5-17 *Thermal control system in laptops and other PCs.*

design of the remote microprocessor temperature sensor will be described in this paragraph.

5.3.2 Specification

The most challenging part of the project is to reduce the influence of interference of the microprocessor signals on the temperature measurement. Due to the high clock speeds of up to 450MHz and the consequent steep edges, there is a lot of noise that is coupled to the bipolar temperature sensor through the microprocessor substrate. Measurements show spikes on the base-emitter voltage of the bipolar transistor of more than 300mV. If we compare this with the sensor signal of approximately 200µV/°C, it is obvious that this interference can cause great problems.

The high-frequency interference problem implies that the sensor signal should be integrated. However, due to the non-linearity of the bipolar transistor, some errors will remain even when the integration action is ideal. This can be explained as follows. Figure 5-18 shows the typical situation of the bipolar transistor on the microprocessor chip. For simplicity it is assumed that the base of the transistor is grounded and that the interference source is a triangular signal which is coupled to the emitter via a capacitance C_I. In practice, the effect of the parasitic coupling

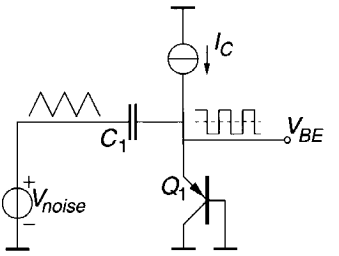

Fig. 5-18 *Influence of high-frequency interference on the DC-behaviour of the bipolar transistor temperature sensor.*

can somewhat be reduced by a floating measurement of the base-emitter voltage. However, due to mismatches in impedances between the base and the emitter, the effect will always be existent. The interference source V_{noise} will add a small current ΔI_C to the bias current I_C. If the shape of V_{noise} is triangular, the current ΔI_C is a square wave. If we state that the collector current of Q_1 is equal to the emitter current, the average output voltage V_{BE} yields:

$$V_{BE} = \frac{1}{2}\left\{\frac{kT}{q}\ln\left(\frac{I_C(1+\Delta)}{I_s}\right) + \frac{kT}{q}\ln\left(\frac{I_C(1-\Delta)}{I_s}\right)\right\} \tag{5-8}$$

which equals:

$$V_{BE} = \frac{1}{2}\frac{kT}{q}\left\{2\ln\left(\frac{I_C}{I_s}\right) + \ln\{(1+\Delta)(1-\Delta)\}\right\} \tag{5-9}$$

or:

$$V_{BE} = \frac{kT}{q}\ln\left(\frac{I_C\sqrt{1-\Delta^2}}{I_s}\right) \tag{5-10}$$

From equation (5-10) it can be seen that the DC voltage of a bipolar transistor is changed due to high-frequency interference. To have a negligible influence on V_{BE}, the interference current should be small compared to I_C. If the interference can not be changed by the designer, which is usually the case, a possible way to reduce the influence is to increase I_C. However, this can not be done too much because of self-heating and power consumption limitations. In practice, the

interference source and the coupling is not very well known. However, the conclusions derived from equation (5-10) will still be applicable.

After deliberation with the microprocessor supplier, it was decided that the current through the bipolar transistor is limited to 100μA. Simulations and measurements, however, showed that the temperature read-out error, even with a bias current of 100μA was several degrees Celsius. To reduce this error, it was decided that an external capacitor between base and emitter of maximum 2.2nF would be added.

The accuracy specification was set to ±2°C for temperatures between +70°C and +100°C and ±3°C for the total temperature range from -20°C to +125°C. The resolution was somewhat lowered to 8 bits, which equals 1°C. The other specifications did not differ very much from the specifications of the high-accuracy temperature sensor described in paragraph 5.2.

5.3.3 Design of the single-transistor temperature sensor interface

The most accurate way to generate temperature information from a bipolar transistor is to perform a differential measurement of two base-emitter voltages biased at two currents with a fixed ratio. To generate such a PTAT signal from a single transistor, it is always necessary to have a storage device to store the value of the first measured base-emitter voltage. The most straightforward way is to use a capacitor in a sample-and-hold circuit. This is also shown in figure 4-9 in chapter 4. A major disadvantage of capacitors as storage devices is leakage. Especially in our design, where large signal integration times are needed to reject interference, this can be a major problem. The PTAT voltage will have a sensitivity of approximately 200μV/°C. To achieve leakages smaller than 100μV during an integration period of several milliseconds would require large storage capacitors, especially at temperatures above 100°C where leakage becomes very high.

A different approach, that does not have the leakage problem, is to store the base-emitter voltage digitally. This can be done in two ways. The first method is to perform just two samples of the base-emitter voltage and subtract them digitally. However, this requires a great improvement of performance of the A-to-D converter. Normally, the A-to-D converter has to convert a signal with a maximal value of approximately 80mV (kT/q*ln10 at T=400K) with an accuracy of 200μV, which requires a dynamic range of 52 dB or 9 bits. In the proposed method, the maximum input signal is increased to approximately 700mV which is

the nominal base-emitter voltage, while the maximal inaccuracy is decreased to 100μV, because a factor two of resolution is lost with digital extraction. This implies that an A-to-D converter is needed with a dynamic range of at least 77dB or 13 bits.

The second method is to make a local digital sample-and-hold including an A-to-D converter, a register and a D-to-A converter. Although this A-to-D converter and D-to-A converter also need a resolution of 13 bits, the accuracy is of no importance because it is used in a feed-back loop. However, to make an extra A-to-D and D-to-A converter with a resolution of 13 bits still requires a lot of chip area.

We propose a third method that does a digital storage of the base-emitter voltage but does not have the disadvantages of the two above-mentioned methods. The principle is shown in figure 5-19. We introduce an extra sample-and-hold phase where a rough sampling is done of the base-emitter voltage at the high bias current of 100μA, which is called V_{be10}. This sampling has a resolution of only 6 bits, which corresponds to approximately 10mV. In the next phases, a small offset voltage V_1 of 20mV is subtracted from this sampled voltage, to assure that the voltage V_s always lays between V_{be10} and V_{be1}, which is the base-emitter voltage at a bias current of 10μA. In the second phase, the voltage difference of V_{be10} and V_s is converted into a digital number. In the third phase, the voltage difference of V_s and V_{be1} is converted. Both the conversions result in a positive number, because V_s lies between V_{be10} and V_{be1}. The two samples are *added* digitally, which results in the value V_{be10} minus V_{be1}, which is the desired PTAT signal.

The advantage of this method over the two others is that the maximum signal at the input of the A-to-D converter is only 80mV and the required accuracy 100μV, which corresponds to a dynamic range of 58dB or 10 bits. This is much lower than the required dynamic range of 13 bits, which is needed when doing a digitization of the full base-emitter voltage. This is achieved by using a digital sample-and-hold that has a resolution of only 6 bits.

5.3.4 Detailed design of the remote temperature sensor

ΣΔ modulator

The first-order ΣΔ modulator from the high-accuracy temperature sensor is also used to convert the remote microprocessor temperature sensor signal to the digital domain. The conversion of the ambient and microprocessor temperature is done subsequently. In the phase when the microprocessor temperature is converted, the

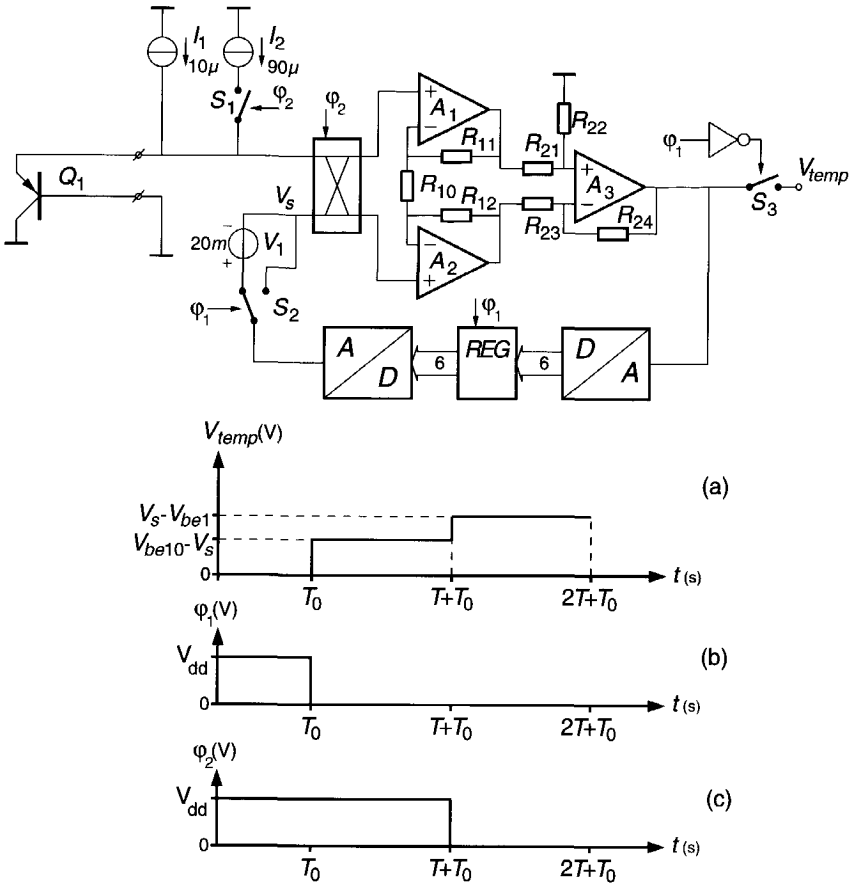

Fig. 5-19 *Principle of low-resolution sample-and-hold circuit in single-transistor temperature sensor interface*

I_{ptat} and I_{be} signals from the ambient temperature measurement are added to obtain a reference signal.

Single-transistor temperature sensor interface

Because the first-order ΣΔ modulator needs a current input, the instrumentation amplifier from figure 5-19 is changed to have a current output. The bias currents I_1 and I_2 are dynamically matched to obtain higher matching properties. The digital sample-and-hold consists of a comparator, a counter and a resistor-string-type 6-bit D-A converter. The counter has a size of 8 bits to perform some extra

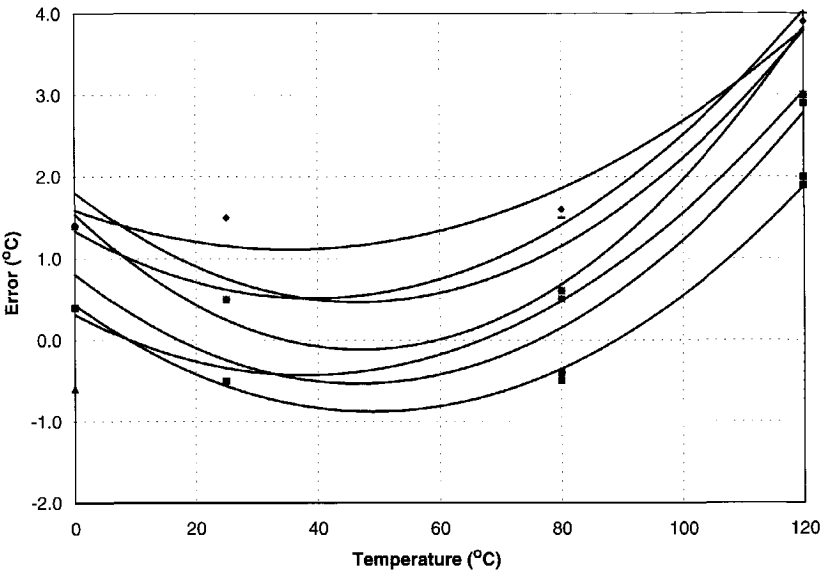

Fig. 5-20 *Temperature errors of 30 samples without calibration.*

digital filtering. The 20mV offset is implemented in the D-A converter and can be switched on with an additional control line.

Layout

The layout is already shown in figure 5-15. The single-transistor temperature sensor interface occupies approximately 0.3 mm^2. The $\Sigma\Delta$ modulator is shared with the high-accuracy ambient temperature sensor and therefore needs no additional chip area. The control block is a little bit enlarged.

5.3.5 Measurement results

The chip has been encapsulated in a standard plastic IC package and has been exposed to temperatures in the range from -60°C to +130°C. The circuit shows full functionality for supply voltages between 2.9V and 5.5 V.

The accuracy of 30 samples of the circuit from one batch have been tested at 0°C, +25°C, +80°C and +120°C. The results are shown in figure 5-20. The resolution of measurement is 1°C. The reason why only a small number of the 30 tested samples can be seen is because many samples have exactly the same inaccuracy because of the limited resolution.

Table 5-7 *Measured specifications of the remote microprocessor temperature sensor*

	Minimal	Typical	Maximal	Unit
Supply voltage	2.9	3.3	5.5	V
Supply current I^2C inactive I^2C active Shutdown mode		100 1	500	µA
Temperature range	-20		125	°C
Inaccuracy T_A= 70°C to +100°C T_A= -25°C to +125°C			±3.0 ±5.0	°C
Conversion time		100		ms
Resolution		8 1		bits °C
Power supply regulation		0.2		°C/V
Noise		0.1		°C

All circuits are within ±2°C in a temperature range from 0°C to +80°C. The typical spread at room temperature is ±1°C. The accuracy specification of ±2°C for temperatures between +70°C and +100°C and ±3°C for the total temperature range between -20°C and +125°C are not completely met. However, this can be realized by doing a hard-wired shift of -1°C for the whole range and a curvature correction.

A summary of the measured specifications is listed in table 5-7.

5.4 Conclusions

In this chapter three different CMOS smart temperature sensors have been described, each one with its own optimization and application. It is shown that it is possible to make a smart temperature sensor with a power consumption of less

than 10μW at a sampling rate of 2 samples/s. A circuit is also shown that can measure the temperature of a remote transistor, independent of the process in which it is made.

However, the best result from the author's point of view, is the realization of the first smart temperature sensor that meets the standard accuracy specifications without any form of calibration. This implies a major cost breakthrough and will further enlarge the number of smart temperature sensor applications.

References

[5.1] IQ Mobil, "Why use 2.45GHz for the RDKS system", *http://www.iqmobil.de*, December 1999

[5.2] M. Tuthill, "A Switched-Current, Switched-Capacitor Temperature Sensor in 0.6-µm CMOS", *IEEE Journal of Solid-State Circuits*, vol.33, pp. 1117-1122, July 1998

[5.3] A. Bakker and J.H. Huijsing, "Micropower CMOS Temperature Sensor with Digital Output", *IEEE Journal of Solid-State Circuits*, vol.31, pp. 933-937, July 1996

[5.4] B. de Geeter, O. Nys and J.-P. Bardyn, "A Wide Temperature Range Micropower Sensor Interface Circuit", *Proc. ESSCIRC'96*, Neuchâtel, Switzerland, pp. 136-139, September 1996

[5.5] National Semiconductor, "LM75 - Digital Temperature Sensor and Thermal WATCHDOG with Two-Wire Interface", *http://www.national.com*, November 1999

[5.6] Philips Semiconductors, "The I^2C-bus and how to use it (including specifications)", *http://www.philips.com*, April 1995

[5.7] A. Bakker and J.H. Huijsing, "A Low-Cost, High-Accuracy CMOS Smart Temperature Sensor", *Proc. ESSCIRC'99*, Duisburg, Germany, pp. 302-305, September 1999

[5.8] Smart Battery System-Implementers Forum, "Smart Battery System Specification", *http:// www.sbs-forum.org*, rev.1.1, December 11, 1998

Appendix

Derivation of $V_{BE}(T)$ by importing the constant $V_{BE}(T_r)$. This derivation is referenced in chapter 3 on page 41.

$$V_{BE} = V_{g0} + \frac{kT}{q}\ln\frac{I_C}{CT^{\eta}} \tag{2-1}$$

$$V_{BE} = V_{g0} + \frac{kT}{q}\ln\frac{I_C T_r^{\eta}}{CT_r^{\eta}T^{\eta}} \tag{2-2}$$

$$V_{BE} = V_{g0} + \frac{kT}{q}\left(\ln\frac{I_C}{CT_r^{\eta}} + \ln\frac{T_r^{\eta}}{T^{\eta}}\right) \tag{2-3}$$

$$V_{BE} = V_{g0} + \frac{kT}{q}\left(\ln\frac{I_C}{CT_r^{\eta}} + \eta\ln\frac{T_r}{T}\right) \tag{2-4}$$

$$V_{BE} = V_{g0} + \frac{kT}{q}\ln\frac{I_C}{CT_r^{\eta}} + \frac{kT}{q}\eta\ln\frac{T_r}{T} \tag{2-5}$$

$$V_{BE} = V_{g0} + \frac{T}{T_r}\left(\frac{kT_r}{q}\ln\frac{I_C}{CT_r^{\eta}}\right) + \frac{kT}{q}\eta\ln\frac{T_r}{T} \tag{2-6}$$

$$V_{BE} = V_{g0} - \frac{T}{T_r}V_{g0} + \frac{T}{T_r}\left(V_{g0} + \frac{kT_r}{q}\ln\frac{I_C}{CT_r^{\eta}}\right) + \frac{kT}{q}\eta\ln\frac{T_r}{T} \tag{2-7}$$

$$V_{BE} = V_{g0}\left(1-\frac{T}{T_r}\right) + \frac{T}{T_r}V_{BE}(T_r) + \frac{kT}{q}\eta\ln\frac{T_r}{T} \tag{2-8}$$

Summary

This thesis describes the theory and design of high-accuracy CMOS smart temperature sensors. The major topic of the work is the realizzation of a smart temperature sensor that has an accuracy that is so high that it can be applied without any form of calibration. Integrated in a low-cost CMOS technology, this yields one of the most inexpensive intelligent general purpose temperature sensors in the world.

Chapter 1. Introduction

The first thermometers were based on mercury-in-glass and could only be read by the human eye. The industrial revolution and the following computerization asked for more intelligent sensors, which could easily communicate to digital computers. This led to the development of integrated temperature sensors that combine a bipolar temperature sensor and an A-to-D converter on the same chip. The implementation in CMOS technology reduces the processing costs to a minimum while having the best-suited technology to increase the (digital) intelligence. Another major cost reduction is achieved by increasing the accuracy to such a level that for general-purpose applications the sensor does not require any calibration.

Chapter 2. Dynamic offset-cancellation techniques

The accuracy of conventional CMOS smart temperature sensors is degraded by the offset of the read-out electronics. Calibration of these errors is quite expensive. Dynamic offset-cancellation techniques can reduce the offset of amplifiers by a factor 100 to 1000 and do not need trimming.

Dynamic offset-cancellation techniques can be classified into two groups. The first group make use of the autozero technique, which is a sampling technique. A disadvantage of this technique is that high-frequency noise is folded back to the baseband, which results in a residual noise that is higher than the thermal noise

floor. The second group makes use of the chopper technique, which is a modulation technique. This technique lacks the problem of the folding back of high-frequency noise components and is therefore a fundamentally better technique for high-accuracy data-acquisition systems. A disadvantage of the chopper technique is that the residual offset is slightly higher than that which can be achieved with the autozero technique. In this chapter a new technique is proposed which utilizes an additional low-frequency chopper pair to further reduce the residual offset. This new technique is called the nested-chopper technique. An implementation of a CMOS nested-chopper instrumentation amplifier shows a residual offset of less than 100nV, which is the best result ever reported.

Chapter 3. CMOS bandgap references

Voltage references are the key elements in analog and mixed-mode circuits. The accuracy of a CMOS smart temperature sensor is mainly dependent on the quality of its reference.

The (silicon) bandgap voltage reference is the most favourable voltage reference for integrated circuits nowadays. The main reasons for this are its high accuracy, stability and reproducibility. Although this reference is based on a bipolar transistor, it can be implemented in standard CMOS technology by utilizing the (parasitic) substrate PNP transistor. In high-accuracy bandgap references the accuracy is limited by non-linearity. Because this non-linearity is highly reproducible it can be corrected by so-called curvature correction techniques. A well-known curvature correction technique is based on the linearized V_{BE} principle. This technique, however, requires highly-matched components, which can usually only be retrieved after trimming. A different approach is to use a piece-wise-linear (PWL) technique. This new technique is slightly less accurate, but does not require trimmed components. A realization of a CMOS bandgap reference utilizing the PWL technique shows a 3σ-spread before calibration of only 4.7mV in a temperature range between -40 to +130°C. The 3σ temperature coefficient is 23ppm/°C, which is almost twice as good as for a bandgap reference without curvature correction. These results are the best ever reported for an untrimmed bandgap reference.

Chapter 4. Design of CMOS smart temperature sensors

The driving force behind the development of smart temperature sensors is both compatibility with the digital world of computers and cost reduction. The current

standard for smart temperature sensors is the LM75, which is a thermometer with an on-board Analog-to-Digital converter and an I^2C bus-interface. The largest market for smart temperature sensors nowadays is multimedia, where it is the main component in thermal management of personal computers. Modern smart temperature sensors are obliged to be cheap and have low-power consumption. The accuracy should be high to reduce calibration costs. These requirements are addressed in the different design aspects.

Analog-to-Digital conversion. A-to-D conversion can be done at low-frequency, but should be at low-cost and low-power. Oversampling converters are definitely the best choice. A first-order Sigma-Delta A-to-D converter is chosen to be the best compromise for chip area and power consumption. In future, a second-order Sigma-Delta A-to-D converter may be favourable to further reduce power consumption.

Kelvin-to-Celsius conversion. Kelvin-to-Celsius conversion improves the usage of the dynamic range of the Sigma-Delta modulator. This implies higher accuracy and less power consumption.

Curvature correction. The curvature that is due to the non-linearity of the base-emitter voltage can be compensated by various techniques. Adding a small temperature dependence to the voltage reference can reduce the curvature and enhance the accuracy.

Single-transistor temperature sensors. Single-transistor temperature sensors are useful for remote temperature sensing, for example on a different die. The interface to these kind of temperature sensors should be very robust to interference and insensitive to process variations. Continuous-time interfaces are highly favourable over time-discrete (switched-cap) ones.

Bus interfaces. The standard temperature sensor bus interface is the I^2C bus. There is no reason to deviate from that.

Chapter 5. Realizations of CMOS smart temperature sensors

Three different CMOS smart temperature sensors have been realized in the frame of this Ph.D. work.

Tyre monitoring system. A tyre monitoring system consists of a smart pressure sensor, a smart temperature sensor, a transmitter and receiver, and a display unit. The main issue is the power supply within the tyre. The cheapest solution is a battery. However, this puts heavy power requirements on the in-tyre electronics.

The power budget for the smart temperature sensor was limited to 7μW for a data-rate of two 8-bit samples per second. This ultra low-power consumption has been achieved by reducing the bias current of each single transistor to the absolute minimum and by incorporating a power-down facility to reduce the power consumption between the samples.

High-accuracy temperature sensor. Approximately 20 percent of the total manufacturing costs of a CMOS smart temperature sensor originate from the calibration. By improving the sensor's initial accuracy, calibration procedures can be relaxed and, in many general-purpose applications, completely omitted. By applying dynamic offset-cancellation techniques, the accuracy of a CMOS smart temperature sensor can be increased to typically better than 0.7°C. For a 3-σ spread production bandwidth this becomes 2.1°C, which is very close to the 2°C specification of the industry-standard LM75. A realized chip shows that these accuracies are feasible.

Remote microprocessor temperature sensor. The largest challenge in the design of a remote microprocessor temperature sensor is to minimize the influence of the microprocessor interference on the accuracy. This is achieved by using a continuous-time input amplifier. The sampling of the base-emitter voltage is done with a low-resolution (5-bit) fully digital sample-and-hold filter. Quantization errors that are introduced because of this low-resolution are removed by a chopping technique. Three-sigma accuracies of better than 3°C are achieved.

Samenvatting

Dit proefschrift beschrijft de theorie en het ontwerp van zeer nauwkeurige, intelligente temperatuur sensoren in CMOS technologie. Het belangrijkste onderwerp van dit werk is de realisatie van een intelligente temperatuur sensor die een nauwkeurigheid heeft die zo hoog is dat hij toegepast kan worden zonder enige vorm van calibratie. Geïntegreerd in een goedkope CMOS technologie levert dit één van de goedkoopste intelligente temperatuur sensoren op in de wereld.

Hoofdstuk 1. Inleiding

De eerste thermometers waren gebaseerd op een kwikkolom in glas en konden alleen uitgelezen worden met het oog. De industriële revolutie en de daarop volgende computerisatie vroegen om intelligentere sensoren die eenvoudig konden communiceren met computers. Dit leidde tot de ontwikkeling van geïntegreerde temperatuur sensoren die een bipolaire temperatuur sensor combineren met een A-naar-D omzetter op de zelfde chip. De implementatie in CMOS technologie reduceert de productiekosten tot een minimum, terwijl dit tegelijkertijd de meest geschikte technologie is om de (digitale) intelligentie te verhogen. Een andere belangrijke kostenverlaging wordt bereikt door de nauwkeurigheid zodanig te verhogen dat voor algemene toepassingen er geen calibratie meer nodig is.

Hoofdstuk 2. Dynamische offset-corrigerende technieken

De nauwkeurigheid van conventionele intelligente temperatuur sensoren in CMOS-technologie wordt beperkt door nulpuntsfouten (offset) van de uitlees-elektronica. Calibratie van deze fouten is tamelijk duur. Dynamische offset-corrigerende technieken kunnen de nulpuntsfout van versterkers reduceren met een factor 100 tot 1000 terwijl ze niet getrimd hoeven te worden.

Dynamische offset-corrigerende technieken kunnen in twee groepen worden ingedeeld. De eerste groep gebruikt de *autozero* techniek, wat een techniek is die werkt met *samples*. Een nadeel van deze techniek is dat hoogfrequente ruis wordt teruggevouwen naar de basisband, wat resulteert in een residu-ruis die hoger is dan de thermische ruis vloer. De tweede groep maakt gebruik van de *chopper* techniek, wat een modulerende techniek is. Deze techiek heeft geen last van terugvouwende hoogfrequente ruiscomponenten en is daarom fundamenteel een betere techniek voor hoge nauwkeurigheids data-acquisitie-systemen. Een nadeel van de *chopper* techniek is dat de residu-offset iets hoger is dan wat behaald zou kunnen worden met de *autozero* techniek. In dit hoofdstuk wordt een nieuwe techniek voorgesteld, die gebruikt maakt van een extra laag-frequent *chopper* paar om de residu-offset verder te verlagen. Deze nieuwe techniek wordt de *nested-chopper* techniek genoemd. Een implementatie van een nested-chopper instrumentatie versterker in CMOS technologie laat een residu-offset zien van minder dan 100nV, wat het beste resultaat is dat ooit is gerapporteerd.

Hoofdstuk 3. Bandgap referenties in CMOS technologie

Spanningsreferenties zijn sleutel elementen in analoge en gemengd analoog-digitale circuits. Ook de nauwkeurigheid van een intelligente temperatuur sensor is voornamelijk afhankelijk van de kwaliteit van zijn referentie.

De spanningsreferentie die afhankelijk is van de energieband-afstand (*bandgap*) van silicium is tegenwoordig de favoriete spanningsreferentie voor geïntegreerde schakelingen. De belangrijkste redenen hiervoor zijn zijn hoge nauwkeurigheid, stabiliteit en reproduceerbaarheid. Hoewel deze referentie gebaseerd is op een bipolaire transistor kan hij toch worden geïmplementeerd in een standaard CMOS technologie door gebruik te maken van de (parasitaire) substraat PNP transistor. In zeer nauwkeurige bandgap referenties is de nauwkeurigheid begrensd door niet-lineariteit. Omdat deze niet-lineariteit zeer goed reproduceerbaar is, kan hij gecorrigeerd worden door zogenaamde krommings-correctie technieken. Een bekende krommingscorrectie techniek is gebaseerd op de linearisatie van V_{BE}. Deze techniek heeft echter componenten nodig die zeer goed aan elkaar gelijk zijn, wat meestal alleen voor elkaar te krijgen is door ze te trimmen. Een andere aanpak is om de lineaire-gedeelten techniek (piece-wise-linear, PWL) te gebruiken. Deze, voor deze toepassing nieuwe, techniek is iets minder nauwkeurig, maar heeft geen getrimde componenten nodig. Een realisatie van een bandgap referentie in CMOS technologie, die gebruik maakt van de PWL techniek heeft een 3σ spreiding voor calibratie van slechts 4.7mV voor temperaturen tussen -40 en +130°C. De 3σ temperatuurafhankelijkheid is 23ppm/°C, wat bijna twee keer zo goed is als van een bandgap referentie zonder

krommings-correctie. Deze resultaten zijn de beste die ooit gerapporteerd zijn voor een ongecalibreerde bandgap referentie.

Hoofdstuk 4. Ontwerp van intelligente temperatuur sensoren in CMOS technologie

De drijvende krachten achter de ontwikkeling van intelligente temperatuur sensoren zijn de compatibiliteit met de digitale wereld van de computers en de lagere kosten. De huidige standaard voor intelligente temperatuur sensoren is de LM75. Dit is een thermometer met een geïntegreerde Analoog-naar-Digitaal omzetter en een I^2C bus interface. De grootste markt voor intelligente temperatuur sensoren is momenteel die van de multimedia, waar het een belangrijke component is in het thermisch management van personal computers. Moderne intelligente temperatuur sensoren moeten goedkoop zijn en weinig vermogen verbruiken. De nauwkeurigheid moet hoog genoeg zijn om de calibratiekosten te beperken. Deze noodzakelijke eigenschappen worden besproken in de verschillende ontwerpaspecten.

Analoog-naar-Digitaal conversie. A-naar-D conversie kan gedaan worden op een lage frequentie, maar dient goedkoop te zijn en weinig stroom te verbruiken. Omzetters die op een hogere frequentie werken dan strikt noodzakelijk (oversampling) zijn zonder twijfel de beste keus. Een eerste-orde Sigma-Delta A-naar-D omzetter is het beste compromis tussen weinig chipoppervlak en laag stroomverbruik. In de toekomst kan een tweede-orde Sigma-Delta A-naar-D omzetter de voorkeur krijgen in verband met het lagere energieverbruik.

Kelvin-naar-Celsius conversie. Kelvin-naar-Celsius conversie verbetert het gebruik van het dynamisch bereik van de Sigma-Delta modulator. Dit zorgt voor hogere nauwkeurigheid en lager stroomverbruik.

Krommings-correctie. De kromming die ontstaat door de niet-lineariteit van de basis-emitter spanning kan gecompenseerd worden met verschillende technieken. Het toevoegen van een kleine temperatuurafhankelijkheid aan de spannings-referentie kan de kromming verminderen en de nauwkeurigheid verbeteren.

Temperatuur sensoren met één transistor. Temperatuur sensoren met slechts één transistor zijn handig voor temperatuur metingen op afstand, bijvoorbeeld op een ander substraat. De uitlezing van dit soort temperatuur sensoren moet ongevoelig zijn voor storingen en proces variaties. Uitlezingen die continu meten hebben verreweg de voorkeur boven degenen die tijd-discreet meten (zoals in *switched-cap* circuits).

Bus interfaces. De standaard bus interface voor temperatuur sensoren is de I^2C bus. Er is geen goede reden om hiervan af te wijken.

Hoofdstuk 5. Realisaties van intelligente temperatuur sensoren in CMOS technologie

Drie verschillende intelligente temperatuur sensoren in CMOS technologie zijn gerealiseerd in het kader van dit promotiewerk.

Autobanden controle systeem. Een systeem om autobanden continu te controleren bestaat uit een intelligente druk sensor, een intelligente temperatuur sensor, een zender en ontvanger en een display. Het belangrijkste probleem ligt in het voeden van het systeem in de band. De goedkoopste oplossing is een batterij. Dit stelt echter hoge eisen aan het energieverbruik van de elektronica in de band. Het vermogen dat maximaal gebruikt mocht worden voor de intelligente temperatuur sensor was gesteld op 7μW. Hiervoor moesten per seconde twee metingen gedaan worden met een resolutie van 8 bits. Dit extreem lage energieverbruik is behaald door de instelstroom van iedere transistor tot het absolute minimum te beperken en door toepassing van een uitschakelmogelijkheid om het verbruik tussen de metingen door te minimaliseren.

Hoge nauwkeurigheids temperatuur sensor. Ongeveer 20% van de totale productiekosten van een intelligente temperatuur sensor in CMOS technologie bestaat uit kosten voor calibratie. Door de initiële nauwkeurigheid van de sensor te verbeteren kunnen calibratie procedures vereenvoudigd worden en in veel algemene toepassingen zelfs worden weggelaten. Door dynamische offset-corrigerende technieken toe te passen kan de nauwkeurigheid van een intelligente temperatuur sensor worden verhoogd tot typisch 0.7°C. Voor een 3σ productie spreiding wordt dit 2.1°C, wat zeer dicht komt bij de 2°C specificatie van de LM75 die de industriële standaard is. Er is een chip gerealiseerd die aantoont dat deze nauwkeurigheid goed haalbaar zijn.

Temperatuur sensor voor microprocessoren. De grootste uitdaging in het ontwerp van een temperatuur sensor voor microprocessoren is de minimalisatie van de invloed van de storing van de microprocessor op de nauwkeurigheid. Dit is behaald door het gebruik van een ingangsversterker die continu werkt. Het opslaan van de basis-emitter spanning is gedaan met een lage-resolutie (5 bit) volledig digitaal *sample-and-hold* filter. Quantisatie fouten die onstaan door deze lage resolutie worden gecorrigeerd door een chopper techniek. Drie-sigma nauwkeurigheden van beter dan 3°C zijn behaald.

Acknowledgments

Although this thesis has only one author's name, it has in fact been realized by many people. By writing acknowledgments one always runs the risk of forgetting somebody. To avoid this, I would like to start by thanking everybody who has in any way contributed to this thesis. I have had and still have the good fortune of being supported by so many people who are willing to share their knowledge and experience with me. Thank you all!

However, there are still many people who deserve specific thanks for their outstanding support.

Firstly, I would like to thank all the people from the Electronic Instrumentation laboratory for providing me with a stimulating environment. The most important person in this matter is Prof. Han Huijsing, who supervised me not only as a boss, but also as a colleague and as a friend. I would also like to thank my room-mates Frank Riedijk, Klaas-Jan de Langen, Lucien Breems, Kofi Makinwa and Ovidiu Bajdechi. I also consider them not only as colleagues, but also as friends. Here I should also mention Maureen Meekel, who was not a room-mate, although it seemed most of the time that the only difference was the location of her desk. Special thanks to Michiel Pertijs, who added to this work and did proof reading of the manuscript. Special thanks also to Wim van der Sluys, Evelyn Sharabi and Inge Egmond. Without them there would be no computers, no paper, no pencils, no coffee. In fact, everything you need to be able to work.

Other people from the Delft University of Technology that should be mentioned are Gerard Meijer with whom I have had very stimulating discussions. Also thanks to many people from DIMES Technology Center, especially Wim van der Vlist for processing and bonding the chips and John Slabbekoorn for making chip photographs.

Thanks also go to many people from Philips Semiconductors in Sunnyvale, California. Especially Kevin Thiele, who was my technical supervisor while designing the chips for Philips. But also Tony Bull and Don Remsen, who were

willing to invest money in my project. Hung Nguyen, who did very precise and accurate measurements on my temperature sensor chips.

Here I would also like to thank the people from STW, the dutch technology foundation, for their financial support. Without their investments, this project wouldn't have even started.

At home, there has been my loving wife, Miriam, who supplied tea and beverages when I was writing. She helped me through when there were times that I thought I would never make it in time. Miriam, thanks for everything, I love you!

Also thanks to Judith, our 1.5 year old daughter, for being on the front page of my thesis. I hope you will still like it when you are grown up.

I would also like to thank my parents, who stimulated me to go to school and do my homework. That is in fact the time when it all started. Thanks also for stimulating me to go to university.

Finally, I would like to say something about God. Although some people may think he does not exist, I believe he is the one who is supporting me the most. He created the world and electronics. He showed his perfect love by sending his son. I should thank him the most.

List of Publications

A. Bakker, K. Thiele and J.H. Huijsing, "A CMOS Nested-Chopper Instrumentation Amplifier with 100nV Offset", *Proc. ISSCC'00*, San Francisco, USA, February 2000, pp. 156-157

A. Bakker and J.H. Huijsing, "A CMOS Spinning-Current Hall Effect Sensor with Integrated Submicrovolt Offset Instrumentation Amplifier", *Proc. ProRISC'99*, Mierlo, The Netherlands, November 1999, pp. 17-20

A. Bakker and J.H. Huijsing, "A low-cost High-Accuracy CMOS Smart Temperature Sensor", *Proc. ESSCIRC'99*, Duisburg, Germany, September 1999, pp. 302-305

A. Bakker and J.H. Huijsing, "Low-offset low-noise 3.5mW CMOS spinning-current Hall effect sensor with integrated chopper amplifier", *Proc. Eurosensors'99*, The Hague, The Netherlands, September 1999, pp.1045-1048

A. Bakker and J.H. Huijsing, "Low-cost temperature monitoring IC", *Proc. ICMP'99*, Campinas, Brazil, August 1999, pp.244-248

A. Bakker and J.H. Huijsing, "Low-cost CMOS smart temperature sensor with digital bus interface", *Journal A*, vol. 40, no.1, March 1999

A. Bakker and J.H. Huijsing, "CMOS smart temperature sensor with uncalibrated 1 degree C accuracy", *Proc. ProRISC'98*, Mierlo, The Netherlands, November 1998

A. Bakker and J.H. Huijsing, "A CMOS Chopper Opamp with Integrated Low-Pass Filter", *Proc. ProRISC'97*, Mierlo, The Netherlands, November 1997, pp.31-34

A. Bakker and J.H. Huijsing, "A CMOS Chopper Opamp with Integrated Low-Pass Filter", *Proc. ESSCIRC'97*, Southampton, UK, September 1997, pp. 200-203

A. Bakker and J.H. Huijsing, "Micropower CMOS Temperature Sensor with Digital Output", *IEEE J. Solid-State Circuits*, vol. 31, no.7, July 1996, pp.933-937

A. Bakker and J.H. Huijsing, "Offset Reduction in Current-Mode Microsystems", *Proc. ISCAS'96*, Atlanta, USA, May 1996, vol.4, pp. 344-347.

A. Bakker and J.H. Huijsing, "Micropower CMOS Smart Temperature Sensor", *Proc. ESSCIRC'95*, Lille, France, September 1995, pp. 238-241.

M. van Doornik, E. van Hoeven and A. Bakker, "3D Sensor Packaging for an Electronic Compass", *Proc. ProRISC'99*, Mierlo, The Netherlands, November 1999, pp. 125-128

M.J. Meekel, A. Bakker, G. Vega, J.H. Huijsing, "The I2C-bus as smart sensor bus", Proc. ProRISC '98, Mierlo, The Netherlands, November 1998

G.J.A. van Dijk, A. Bakker and J.H. Huijsing, "Low-Power Transadmittance Amplifier with Extended Common-Mode Input Range for a Smart Thermocouple Interface", *Proc. ProRISC '97*, Mierlo, The Netherlands, November 1997, pp. 143-148

A.A. Bellekom, S. Yin and A. Bakker, "Spinning-Current Hall Plate with Integrated Switches", *Proc. ProRISC '97*, Mierlo, The Netherlands, November 1997, pp. 43-48

Patents

A. Bakker and J.H. Huijsing, "Circuit comprising means for reducing the DC-offset and the noise produced by an amplifier", *European Patent Application* #98203827.5, filed November 12, 1998

A. Bakker and J.H. Huijsing, "A current generator for delivering a reference current of which the value is proportional to the absolute temperature", *European Patent Application* #98203826.7, filed November 12, 1998

J.H. Huijsing and A. Bakker, "Chopper amplifier technique with internal filter", *European Patent Application* #1001231, March 20, 1997

About the author

Anton Bakker was born in Amsterdam, the Netherlands, on 11 July 1968. After finishing his pre-university education (VWO) at the "Sweelinck College" in Amsterdam, he started his study in Electrical Engineering at the Delft University of Technology in 1986. In 1991 he received the M.Sc. degree. In the same year he joined the "Werkgroep Elektrotechnisch Practicum" where he developed a laboratory course for second-year students on the design of a complex integrated circuit. In 1993 he joined the Electronic Instrumentation Laboratory where he was involved in the European Project (ESPRIT) on the design of a ultra low-power tempersature sensor. During this project he spent three months at CSEM, Neuchâtel, Switzerland to implement his design. In 1996 he started his Ph.D. project on CMOS smart temperature sensors, which was funded by STW (Dutch Technology Foundation). During this research period he designed a number of temperature sensors for Philips, Sunnyvale, USA. From 1 May 2000, he will start as a senior design engineer at Philips Semiconductors, Sunnyvale.