ERROR MODELING, SELF-CALIBRATION AND DESIGN OF PIPELINED ANALOG TO DIGITAL CONVERTERS

Volume II

A Dissertation

by

ERIC GEORGES SOENEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 1992

Major Subject: Electrical Engineering

ERROR MODELING, SELF-CALIBRATION AND DESIGN OF PIPELINED ANALOG TO DIGITAL CONVERTERS

Volume II

A Dissertation
by
ERIC GEORGES SOENEN

Approved as to style and content by:

Randall Geiger
(Co -Chair of Committee)

Edgar Sanchez-Sinencio
(Co-Chair of Committee)

Richard Hester
(Member)

Maciej Styblinski
(Member)

Donald Friesen
(Member)

Jo Howze
(Head of Department)

May 1992

TABLE OF CONTENTS

CHAPTER Page

Volume I

TABLE OF CONTENTS (Continued)

CHAPTER Page

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

CHAPTER Page

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

CHAPTER Page

LIST OF FIGURES

FIGURE Page

Volume I

LIST OF FIGURES (Continued)

FIGURE Page

LIST OF FIGURES (Continued)

FIGURE Page

LIST OF FIGURES (Continued)

FIGURE Page

Volume II

LIST OF FIGURES (Continued)

FIGURE Page

LIST OF FIGURES (Continued)

FIGURE Page

LIST OF FIGURES (Continued)

FIGURE Page

CHAPTER IX

ANALYSIS OF SWITCHED-CAPACITOR GAIN STAGES

9.1. Introduction

Many analog discrete-time systems like filters and A/D converters require the use of fast, linear sample/hold amplifiers. Sample/hold amplifiers must be able to realize a fixed gain, as well as to temporarily store a certain signal level. The use of such amplifiers in pipelined or other analog-to-digital converters has been discussed extensively in earlier chapters.

This chapter discusses a possible circuit implementation, in switched-capacitor (SC) technology. Similar configurations have traditionally been used in many different practical filtering and conversion applications. Their use is well understood and well documented [66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 30, 34, 7]. Another, new implementation will be discussed in the next chapter.

SC amplifiers are based on a high-gain, high-linearity operational amplifier (opamp), a number of MOSFET switches and a feed-back circuit consisting of two capacitors. The structure is fairly simple, and the design of operational amplifiers is well understood. As long as the opamp gain can be made high enough, the voltage gain of a SC stage will be determined by the ratio of input to feed-back capacitor.

This structure has some drawbacks. There is an inherent trade-off between the speed of an SC amplifier, its accuracy and its power consumption, which tends to be high. Since the gain is set by a ratio of capacitors, it is difficult to adjust after fabrication. This makes it impractical for example to tune an SC filter. The structure is subject to parasitic effects which can corrupt its performance. A well-known but hard to avoid example is clock feed-through. Finally, due to the feed-back, instability can occur in the basic SC structure and the dimensioning of active devices has to be done very carefully.

9.2. Basic Operation

Two well-known SC amplifier schemes are shown in figure 91 and 92. The first one realizes a positive voltage gain, approximately equal to C_{in}/C_f. The second

configuration realizes a negative voltage gain of $-C_{in}/C_f$. Both configurations are called stray-insensitive, because the ultimate voltage gain is not affected by parasitic capacitance between the capacitor electrodes and ground (the substrate, when realized in IC technology).

In both cases, the input voltage is sampled onto C_{in} and then transferred onto C_f using the opamp, thus realizing a certain voltage gain. Important performance considerations of these schemes are accuracy of the gain, linearity, speed and to a lesser extent power consumption. These parameters are primarily determined by the characteristics of the opamp. However, additional non-idealities may be introduced by the switches.

In the first scheme, during clock phase Φ_1, the input electrode of capacitor C_{in} is connected to V_{in} while its output electrode is grounded. C_{in} is charged to a voltage V_{in}, which normally is provided by a low output impedance voltage source. If the output impedance is finite rather than zero, the voltage on C_{in} will increase exponentially with a certain time constant, but given sufficient time, the final situation will be the same. The finite ON resistance of the switches has a similar effect and does not affect the final voltage across the capacitor terminals. The total capacitor charge on C_{in} after settling, is $V_{in}\ C_{in}$.

There is a parasitic capacitance to ground at the input side of C_{in}, C_{p1}. This capacitance is due to stray capacitance of C_{in} (capacitance to the substrate), S_1 and S_3. C_{p1} is charged to the same voltage V_{in}.

Also during Φ_1, C_f is discharged by S_5. The opamp operates in a unity-gain configuration, and its input as well as output voltage approximately settles to ground potential.

After Φ_1, S_1 and S_3 are opened, C_{in} is left floating and holds its charge for a short period of time. For the moment being, we will not consider any clock feed-through effects and assume that the charge on C_{in} is not influenced by the switching.

During Φ_2, the input electrode of C_{in} is grounded and the output electrode connected to the inverting opamp input. The parasitic charge on C_{p1} flows to ground, while the charge on C_{in} is inverted and redistributed between C_{p2} (stray capacitance of output electrode, stray capacitance of S_3 and S_4), C_{p3} (stray

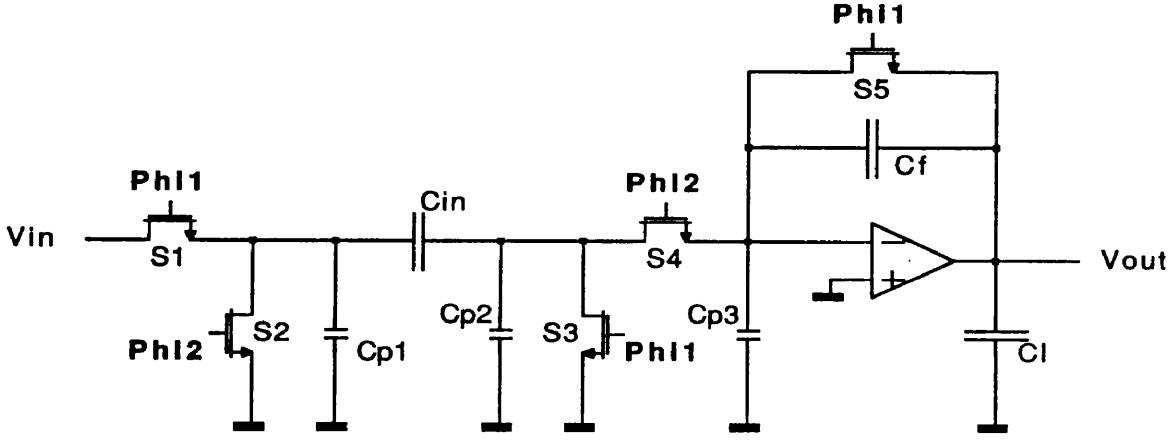


Fig. 91: Non-Inverting Switched-Capacitor Stage

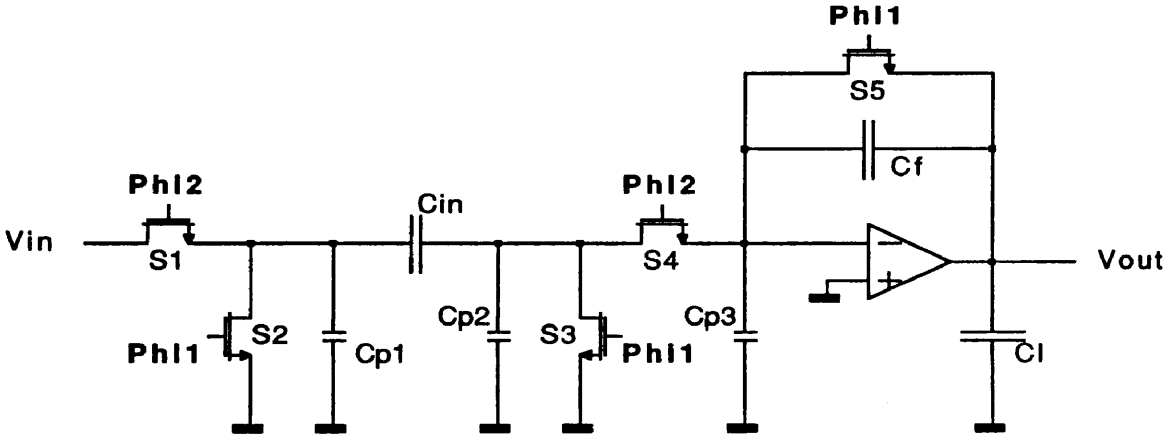


Fig. 92: Inverting Switched-Capacitor Stage

capacitance of S_4 and S_5 and input capacitance of the opamp) and C_f (feed-back capacitor).

This situation is maintained for a short time only, because the opamp immediately reacts to the non-zero input signal and produces an output which will alter the charge balance. Eventually, the feed-back system will force the inverting input node of the opamp to a final value close to 0 (virtual ground principle). This implies that the total charge on C_{in}, C_{p2}, C_{p3} and C_f (the initial charge $-C_{in}\ V_{in}$), has to flow to C_f. By equating initial and final charges (assuming that the inverting opamp node actually goes to ground, which would be the case if the opamp gain were infinite), the voltage across C_f can be found to be $-V_{in}\ C_{in}/C_f$. This also is the output voltage of the system, irrespective of the load capacitance C_L.

The delay between input signal and output signal is one half clock period and hence the transfer function of the system in the Z domain can be written as

$$H(z) = \frac{C_{in}}{C_f}\ z^{-\frac{1}{2}} \tag{105}$$

A similar analysis can be performed for the inverting configuration, which is identical except for the clocking scheme. During Φ_1, S_2, S_3 and S_5 are closed and both C_{in} and C_f are discharged.

During Φ_2, only S_1 and S_2 are closed. The output electrode of C_{in} is forced to 0 (virtual ground), while the input electrode is pulled to V_{in}. A charge $V_{in}\ C_{in}$ is accumulated onto C_{in}. This charge is provided by the input voltage source on the one hand, while the opposite charge flows from the inverting input node of the opamp. For practical purposes, this is equivalent to a charge $V_{in}\ C_{in}$ flowing through the capacitor, to the inverting input node. This results (again assuming infinite gain and perfect virtual ground) in an output voltage $-V_{in}\ Cin/C_f$.

One can easily verify that the scheme of figure 92 is insensitive to stray capacitance (C_{p1}, C_{p2} and C_{p3}) as well.

Since for this configuration, the output voltage becomes available during the same clock phase as in which the input voltage is sampled, the transfer function can be written as

$$H(z) = -\frac{C_{in}}{C_f} \tag{106}$$

9.3. Gain and Speed

Some important conclusions can be drawn concerning the speed of an SC stage, using a simple opamp model. The opamp is modeled as an ideal amplifier, with a certain open-loop gain A (assumed to be frequency-independent and linear for the moment being) and a certain *series* output impedance R_o. The very same situation can alternatively be described by a transconductance element with transconductance $g_m = A/R_0$ and *parallel* output conductance $g_0 = 1/R_0$ (Thevenin's Theorem).

The distinction between an opamp and an operational transconductance amplifier (OTA) is somewhat artificial anyway. An opamp is a high-gain, low output impedance device. In practice, high DC gain is often realized in an OTA stage ($R_0 \approx 1k\Omega \cdots 10M\Omega$). The OTA output is then buffered by a unity gain, low output impedance ($1k\Omega \cdots 10k\Omega$) stage, resulting in a much higher overall transconductance A/R_0. In theory however, an SC gain stage could operate using an OTA rather than an opamp.

Speed considerations are best analyzed in the frequency domain. For this analysis, it will be assumed that the effect of finite switch resistance can be neglected. In both SC configurations, the input capacitor, charged to a certain level, is eventually switched between the inverting input node of the opamp and ground. This can functionally be modeled as if a certain voltage V_{in} were suddenly (in a step-like fashion) applied in series with a zero-charge input capacitor. The simplified model is shown in figure 93.

The equations describing the system are

$$(v_{in} - v_1)sC_i - v_1 sC_p + (v_{out} - v_1)sC_f = 0 \tag{107}$$

and

$$(v_1 - v_{out})sC_f - v_{out}sC_l + \frac{(-Av_1 - v_{out})}{R_o} = 0 \tag{108}$$

The last equation can be rewritten as

$$v_1(sC_f - \frac{A}{R_o}) = v_{out}(sC_f + sC_l + \frac{1}{R_o}) \tag{109}$$

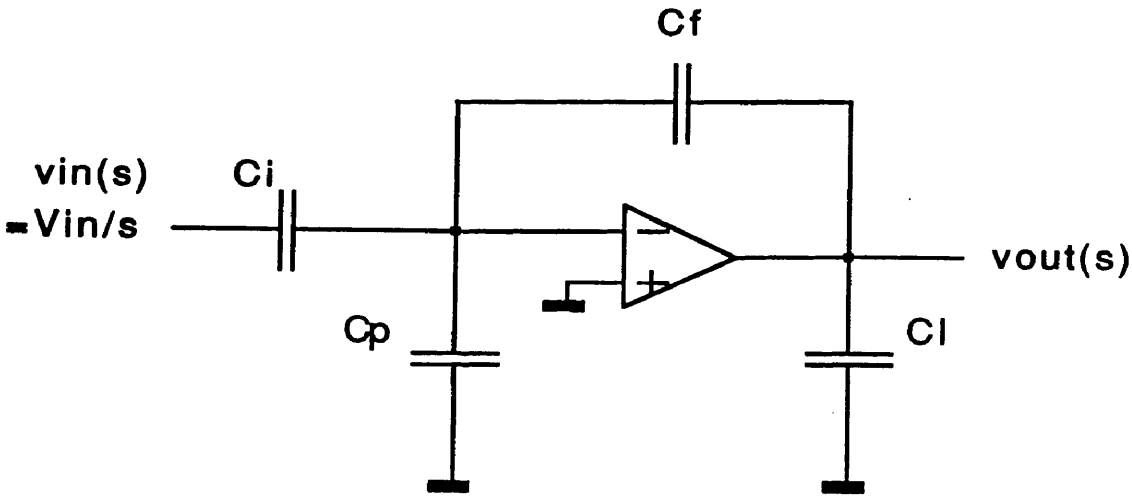


Fig. 93: Simplified Model of the Stage

Substitution into equation (107) yields

$$v_{out}\ (sC_f - \frac{s(C_f + C_l) + \frac{1}{R_o}}{sC_f - \frac{A}{R_o}}\ s(C_i + C_p + C_f)) = -v_{in} sC_i \qquad (110)$$

or

$$v_{out} = -v_{in}\ \frac{sC_i(sC_f - \frac{A}{R_o})}{sC_f(sC_f - \frac{A}{R_o}) - (s(C_f + C_l) + \frac{1}{R_o})(s(C_i + C_p + C_f))} \qquad (111)$$

or

$$v_{out} = -v_{in}\ \frac{C_i}{C_f}\ \frac{sC_f - \frac{A}{R_o}}{s(C_f - \frac{C_f + C_i}{C_f}(C_i + C_p + C_f)) - (\frac{C_i + C_p + C_f}{R_o C_f} + \frac{A}{R_o})} \qquad (112)$$

or finally

$$v_{out} = -v_{in}\ \frac{C_i}{C_f}\ (1 - \frac{C_i + C_p + C_f}{C_f A})\ \frac{1 - s\frac{C_f R_o}{A}}{1 + s(\frac{(1 + \frac{C_l}{C_f})(C_i + C_p + C_f) - C_f}{\frac{A}{R_o}(1 - \frac{C_i + C_p + C_f}{C_f A})})} \qquad (113)$$

This represents a first order transfer function with one real pole in the left-half complex plane and one real zero in the right-half complex plane.

Since the switches are assumed to be infinitely fast (zero ON resistance), the input signal v_{in} can be modeled as a step function. In the actual situation, the input capacitor holds an initial voltage V_{in} before the switching. In the S domain model, the input capacitor is assumed to hold zero initial voltage, but the *input voltage* is represented by the step function

$$v_{in}(s) = \frac{V_{in}}{s} \qquad (114)$$

The transfer function of the SC gain stage has one pole and one zero, as indicated by equation (113). All capacitors in the system are initially discharged (C_{in} is modeled as being discharged before a step is applied). Right after the switching (beginning of Φ_2, or application of the step function), the system output will instantaneously jump to a finite value V_{init}, as indicated by the presence of

the pole and the zero. Physically, the effect is due to capacitive feed-through from input to output.

This can be confirmed through Fourier analysis. The Fourier transform of the output signal is

$$v_{out}(j\omega) = v_{in}(j\omega)\ H(j\omega) = \frac{V_{in}}{j\omega}\ H(j\omega) \tag{115}$$

The time-domain value for $t = 0^+$ (right after switching) can be found directly by multiplying this expression by $j\omega$ and taking the limit for $\omega \rightarrow \infty$ (initial value theorem). This corresponds to setting s to ∞ in the original transfer function.

$$V_{init} = V_{in}\ \frac{C_i}{C_f}\ (1 - \frac{C_i + C_p + C_f}{C_f A})\ \frac{\frac{C_f R_o}{A}}{\frac{(1+\frac{C_l}{C_f})(C_i+C_p+C_f)-C_f}{\frac{A}{R_o}(1-\frac{C_i+C_p+C_f}{C_f A})}} \tag{116}$$

Since the voltage gain A of an opamp or an OTA can normally be considered large, this simplifies to

$$V_{init} \approx V_{in}\ \frac{C_i}{(1 + \frac{C_l}{C_f})\ (C_i + C_p + C_f) - C_f} \tag{117}$$

For a typical situation where $C_i = C_l$ and $C_f = C_i/2$ (gain of 2), V_{init} would be $\frac{3}{8}V_{in}$.

The presence of a pole in the system means that the output voltage will exponentially reach a finite final value, with a time constant equal to the inverse of the pole. If the system is given infinite time to settle, the final output voltage V_∞ can be found by setting $s = 0$ in the transfer function (final value theorem of Fourier analysis)

$$V_\infty = -V_{in}\ \frac{C_i}{C_f}\ (1 - \frac{C_i + C_p + C_f}{C_f A}) \tag{118}$$

If A is large enough, this expression reduces to

$$V_\infty \approx -V_{in}\ \frac{C_i}{C_f} \tag{119}$$

This confirms that the gain of the stage can be accurately set by a ratio of capacitors, provided A is large enough *and* the system is given enough time to

settle. The time constant τ for this settling is given by

$$\tau = \frac{(1+\frac{C_l}{C_f})(C_i+C_p+C_f)-C_f}{\frac{A}{R_o}(1-\frac{C_i+C_p+C_f}{C_f A})} \tag{120}$$

or for large A

$$\tau \approx \frac{(1+\frac{C_l}{C_f})(C_i+C_p+C_f)-C_f}{g_m} \tag{121}$$

This shows how the SC gain stage can be made faster (smaller time constant τ) by either increasing the effective transconductance ($g_m = \frac{A}{R_o}$) of the opamp, or by decreasing the value of input and feed-back capacitors (C_p is small anyway). The first solution usually requires increased power consumption in the opamp, while the second one tends to decrease reproduceability and linearity of the gain, since small capacitances are harder to match and more subject to clock feed-through.

9.4. Accuracy

The term accuracy is used here to refer to the predictability and reproduceability of the overall gain in the SC stage. Ideally, the gain is determined by the ratio of C_{in} and C_f. In practice, two factors mainly affect the accuracy of the gain: the DC open-loop gain of the opamp (A) and the total allowable settling time (the duration of Φ_2 with respect to the intrinsic time constant τ of the system).

The gain of the system, after infinite settling time, can easily be derived from equation (118).

$$A_\infty = -\frac{C_i}{C_f}\,(1-\frac{C_i+C_p+C_f}{C_f A}) \tag{122}$$

The factor $-\frac{C_i}{C_f}$ is the desired value, which can be set accurately by carefully matching the capacitors in the design. The second factor can be considered a parasitic factor, due to the fact that the opamp gain cannot be made infinite. Due to finite opamp gain, the overall gain will be slightly less in absolute value than the desired value. The relative error is given by

$$\epsilon_A = \frac{C_i+C_p+C_f}{C_f A} \approx \frac{C_i+C_f}{C_f A} \tag{123}$$

This equation makes it possible to determine the minimum open-loop DC opamp gain in order to obtain a certain overall relative accuracy.

$$A_{min} \approx \frac{C_i + C_f}{C_f \ \epsilon_A} \tag{124}$$

The minimum required settling time for the system can be determined using equations (120) or (121), which give a value for the time constant τ. Since the SC gain stage is a one-pole system, the output voltage (time domain) will exponentially evolve from its initial value V_{init} to its final value V_∞, during clock phase Φ_2. This exponential settling is described by

$$v_{out}(t) = V_\infty + (V_{init} - V_\infty) \ e^{-\frac{t}{\tau}} \tag{125}$$

After a time t_1, the relative settling error is

$$\epsilon = \frac{V_\infty - V_{init}}{V_\infty} \ e^{-\frac{t_1}{\tau}} \tag{126}$$

Which often can be coarsely approximated by (V_{init} smaller than V_∞)

$$\epsilon \approx e^{-\frac{t_1}{\tau}} \tag{127}$$

This yields an expression for the minimum settling time required to obtain a certain relative accuracy

$$t_{min} \approx -\tau \ \ln(\epsilon) \tag{128}$$

In A/D converter applications, the relative error is often related to the effective number of bits of the system

$$nbits = -\log_2(\epsilon) \approx -\log_2(e^{\frac{-t_1}{\tau}}) = \frac{t_1}{\tau} \ \log_2(\epsilon) = \frac{t_1}{\tau} \ \frac{\ln(\epsilon)}{\ln(2)} \tag{129}$$

and hence

$$nbits \approx \frac{t_1}{\tau \ \ln(2)} \approx \frac{t_1}{.693 \ \tau} \tag{130}$$

9.5. Slew Rate

The analysis of the exponential settling and the resulting estimate for the relative error was based on a linear model for the opamp. This is justified as long as all signals are small. For larger signals, additional effects can come into play. One of them is slewing, due to the fact that a practical opamp can only provide a finite maximum current at its output. As a result, only finite current is available to charge capacitors and the maximum rate of voltage change at the output is hard-limited.

The current required from the opamp reaches its maximum at the beginning of the exponential settling, right after the input switch S_5 is closed, since this is when the output voltage changes at the fastest rate. This is also when the input voltage of the opamp (inverting node) is largest. This voltage can easily be calculated from the capacitors in the system (C_p can be neglected).

At $t = 0^+$, a step of size V_{in} is applied to the series combination of C_i, C_f and C_l. The input voltage of the opamp is the voltage across the series combination of C_f and C_l, due to this capacitive voltage divider (figure 93).

$$V_{in}^{0+} = V_{in} \frac{\frac{1}{C_f} + \frac{1}{C_l}}{\frac{1}{C_i} + \frac{1}{C_f} + \frac{1}{C_l}} = V_{in} \frac{C_i(C_l + C_f)}{C_f C_l + C_i C_l + C_i C_f} \tag{131}$$

For a typical situation where $C_i = C_l$ and $C_f = C_i/2$ (gain of 2), V_{in}^{0+} would be $\frac{3}{4}V_{in}$, which is *not* insignificant.

The initial current provided by the opamp is given by

$$I^{0+} = V_{in}^{0+}\, g_m \tag{132}$$

When dimensioning the opamp, the configuration (e.g. class A, class AB...) and bias currents of the output stage have to be chosen such that this current can indeed be provided without significant distortion, for all possible values of V_{in}.

9.6. Linearity

The model introduced so far is linear, since all components of the system are represented by simple linear elements. In practice, the gain of the SC stage could become non-linear due to several effects: non-linearity of the opamp (A

is input-dependent), non-linearity (voltage dependency) of C_i and C_f and clock-feed-through. Non-linearity of the input parasitic capacitor C_p, or other parasitic elements of the circuit, is less significant, since the configuration is essentially stray-insensitive.

The non-linearity of C_i and C_f can be a critical issue in standard, general-purpose IC technologies, which are often optimized for digital rather than precision analog applications. In such technologies, the linearity of capacitors is often limited to about 10 bits or less. Specialized technologies can provide quality capacitors with linearities up to 10-14 bits (.01% to .006%) over a wide voltage range. In that case, gain non-linearity due to the capacitors is limited. However, non-linearity of the open-loop opamp gain A can be harder to control.

The opamp has modeled been by a simple linear relationship between input and output voltage, as shown in figure 94.

$$v_{out} = A\ v_{in} \tag{133}$$

More accurate modeling is possible by expressing v_{out} as a higher-order function of v_{in}, like illustrated in the opamp transfer characteristic of figure 95.

$$v_{out} = A_0 + A_1 V_{in} + A_2 V_{in}^2 + A_3 V_{in}^3 + \cdots \tag{134}$$

The term A_0 is often omitted, since it represents an offset term, which is irrelevant in many applications.

This expression describing the opamp could theoretically be used instead of the original linear relationship when deriving an expression for the output voltage of the SC stage after settling. However, this would lead to non-linear mathematical expressions. It is more practical to determine some worst-case values for the opamp gain over its input voltage range. These values can then be plugged into equation (122) in order to calculate the worst-case error on the closed-loop gain after settling, A_∞.

A simple, conservative method is to estimate the worst-case (smallest) small-signal value of the opamp gain, being the derivative of output voltage with respect to input voltage, considered over the whole potential signal range (figure 95)

$$A_{wc} = min(\frac{\delta v_{out}}{\delta v_{in}}) \tag{135}$$

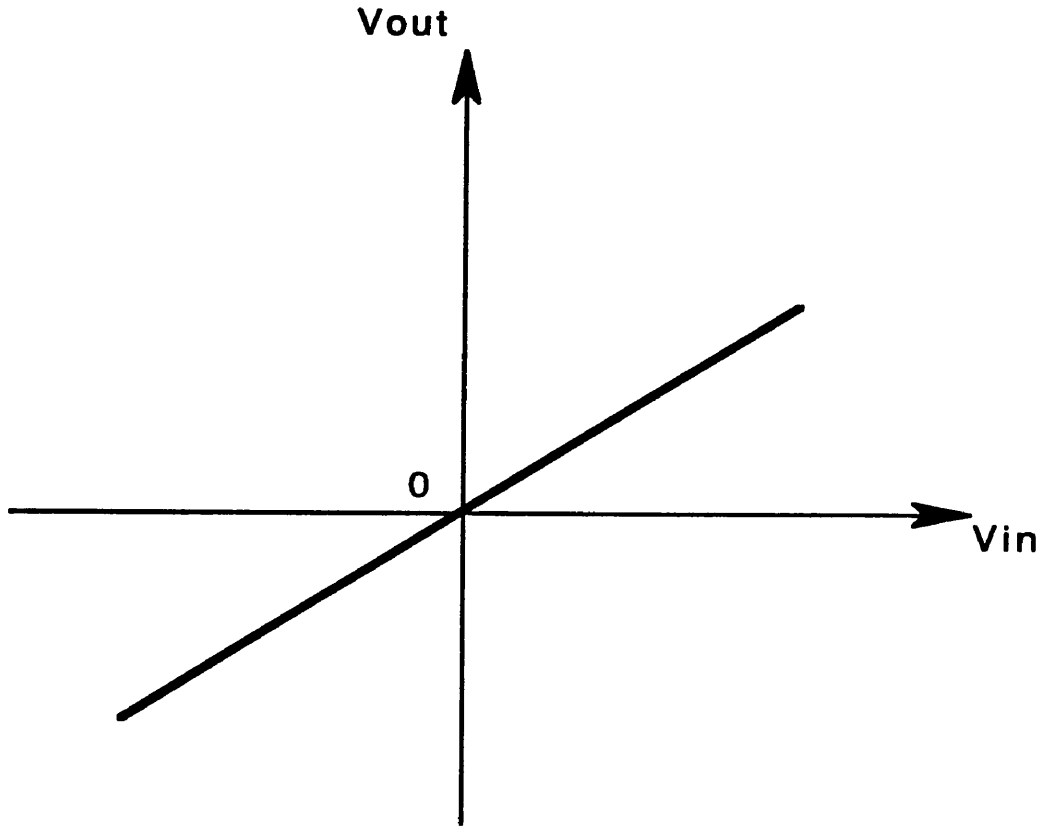


Fig. 94: Linear Gain

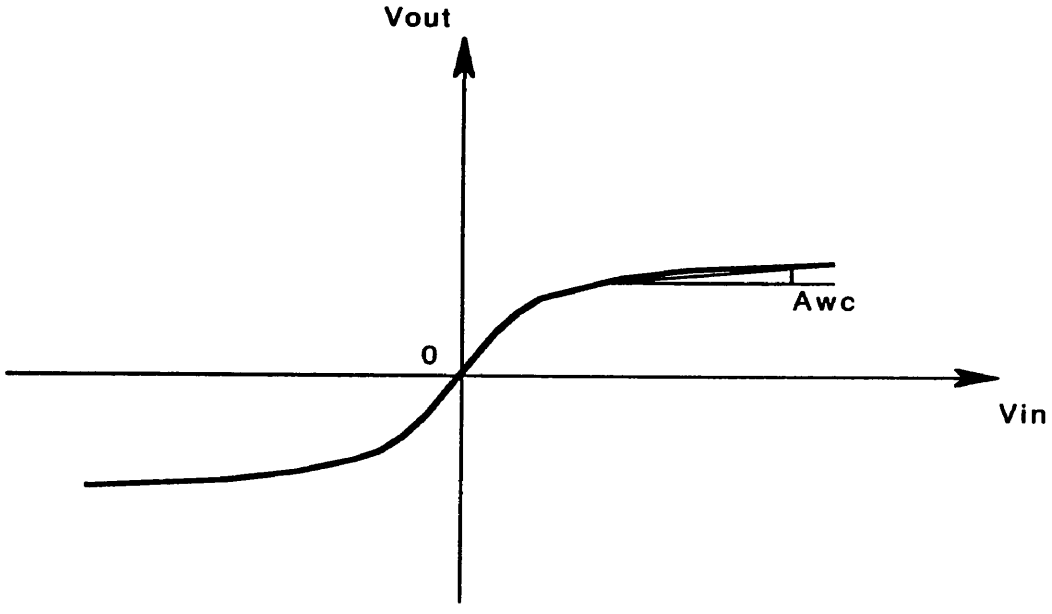


Fig. 95: Non-Linear Gain

Plugging this value into equation (122) will establish an upper limit for the potential relative overall gain error of the SC stage, due to opamp non-linearity.

$$\epsilon_A = \frac{C_i + C_p + C_f}{C_f A} \approx \frac{C_i + C_f}{C_f A_{wc}} \tag{136}$$

A less conservative worst-case value for A is the slope of the line between the origin of the opamp transfer curve and one of its extremities (figure 96). Either worst case value can easily be determined from a standard circuit simulation of the opamp (e.g. SPICE). The transfer curve usually exhibits a certain flattening, and hence lower A_{wc}, for larger (positive or negative) values of the input voltage. As a result, the larger the range of the opamp that is used by the signal, the lower the worst-case gain and the higher the overall non-linearity error will be.

9.7. Switching Speed

The switches $S_1 \cdots S_5$ usually are implemented using MOS transistors. Each switch can consist of either one transistor (an N-channel MOSFET), controlled by a single clock phase, or two transistors (N-channel and P-channel) in parallel, controlled by two opposite clock phases. The second approach has advantages as far as dynamic range is concerned, but is not fundamentally different from the single-transistor approach. For simplicity, only the single-transistor approach will be considered.

The switches are mainly used to connect a capacitor to a fixed voltage source. When they are on, they essentially act as a resistor (MOSFET in ohmic region) during most of the settling process.

The drain current of a MOSFET in the ohmic region can be approximated by the equation

$$I_D = \frac{K'W}{L} \left(V_{GS} - V_T - \frac{V_{DS}}{2}\right) V_{DS} \tag{137}$$

When V_{DS} is small, the equivalent ON resistance can be calculated as

$$R_{eq} = \frac{V_{DS}}{I_D} \approx \frac{1}{\frac{K'W}{L}(V_{GS} - V_T)} \tag{138}$$

This expression, together with the known value of the capacitance, can be used to calculate the worst-case (i.e. for the least favorable input voltage) RC

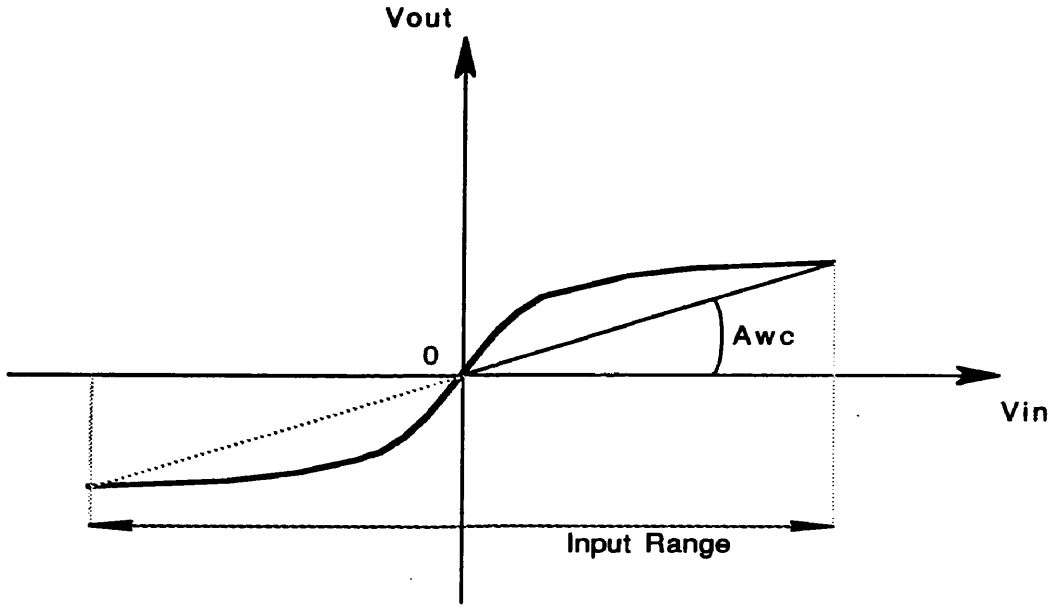


Fig. 96: Linearized Transfer Curve

time constant of the switch/capacitor combination. This time constant can be reduced by making the aspect ratio W/L of the switch large, and the capacitance small. Unfortunately, these are the exact conditions that will generate significant clock feed-through.

9.8. Noise

An advantage of switched-capacitor gain stages, is that noise is relatively easy to model and to control [79, 80, 81]. Noise will be discussed in more detail in subsequent chapters (chapters XI through XIV). However, at this point, it should be pointed out that noise can limit the accuracy (dynamic range) of a switched-capacitor based gain stage.

In electronic systems, noise is mainly introduced by resistors and/or active devices (transistors, either MOSFET or bipolar). It can also be introduced by junction diodes. Capacitors usually do not introduce significant noise, but their frequency response strongly influences the power spectrum (discussed in detail below) of the noise from the other devices.

When a signal is amplified by a switched-capacitor gain stage, noise is normally added in two ways: by the switches and by the opamp. The mechanism by which the (MOSFET) switches introduce noise can best be described by the fact that when the switches are closed, their channels essentially act as low-impedance resistors. It can be shown that the expected variance (σ^2) of the capacitor voltage, after switching, can be approximated by the expression

$$\sigma^2 \approx \frac{kT}{C} \tag{139}$$

With k Boltzmann's constant, T the absolute temperature and C the capacitance. It can be verified that for "normal" values of the capacitance (pF range), this kind of noise is rather limited (microvolts).

The second kind of noise however, is more important. During normal operation, under constant bias conditions, the opamp will generate noise. The parasitic noise current of each semiconductor device can be calculated from the bias conditions. The influences from several devices can be combined in order to calculate the equivalent output-referred noise voltage. This is normally

accomplished using a frequency domain analysis, since noise can be described by a deterministic power spectrum in the frequency domain (this will be treated in detail in chapter XI).

From the value of the output-referred noise, an equivalent input-referred noise voltage can be calculated, by dividing the output-referred value by the open-loop gain of the opamp. For practical purposes, the noisy opamp can be modeled as a noise-free opamp, with an equivalent noise source at the input, of which the voltage is added to the useful signal. Depending on the design of the opamp, the effective input-referred noise voltage can be as low as $10\mu V$, or as high as several mV. These values should be related to the voltage range of the input signal in order to determine the dynamic range of the gain stage.

9.9. Clock Feed-Through

Clock feed-through is the variation of a capacitor charge when a switch connecting that capacitor to some other node is opened or closed [82, 83, 84, 78, 85, 86, 7]. The effect is normally worst when the switch is opened, since the parasitic charge that is introduced does not have a path back, and stays stored on the capacitor.

Clock feed-through is mainly due to two effects: parasitic capacitance from the controlling switch terminal (gate) to the main capacitor, and channel charge redistribution.

The parasitic capacitance at the gate creates feed-through from the gate (when it is being turned off) to the signal capacitor [7]. As long as the switch is still ON (gate voltage above threshold voltage V_T for an N-channel MOSFET), the feed-through charge can escape through the switch to the driving node. However, as soon as the transistor is cut off, the charge can no longer escape, and the parasitic capacitance forms a capacitive voltage divider with the signal capacitor. Since the moment at which the transistor cuts off is dependent on the input voltage, the amount of feed-through will be roughly proportional to the input voltage (figure 97). To the first order, this effect creates a combination of offset and gain error on the SC stage.

It is claimed that this aspect of clock feed-through can be reduced by the use of complementary switches, or dummy switches controlled by a complementary gate

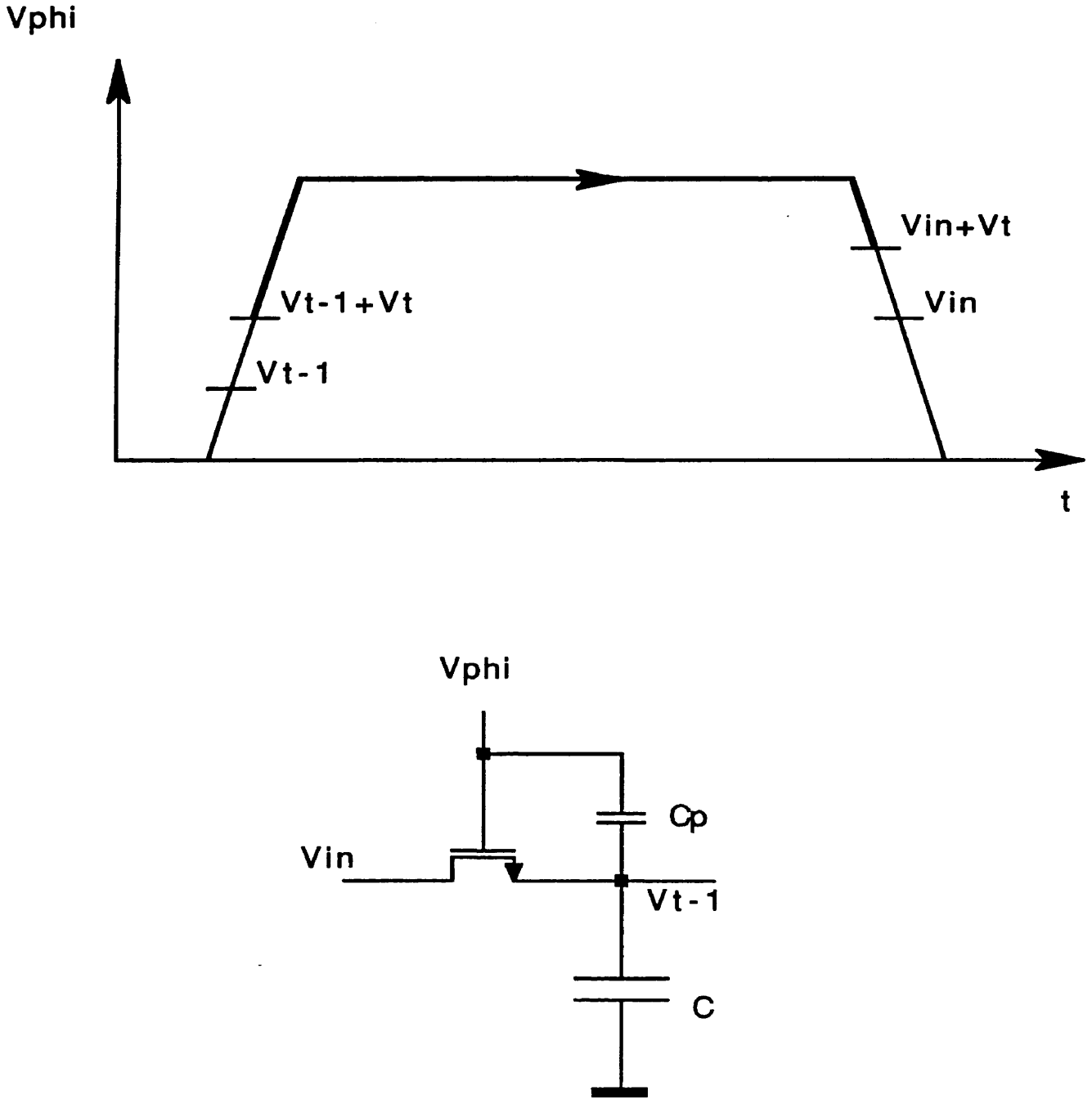


Fig. 97: Effect of Clock Feed-Through

signal.as described in [7]. In some cases, a certain improvement can be achieved. In others however, the mismatch between different switches and clock signals limits the benefits and the additional complexity of the technique may outweigh the advantages. Other schemes switch the gate to a voltage that is higher than the input voltage by a fixed amount (bootstrapped gate signal). This ideally makes the clock feed-trough repeatable, but like the previous class of techniques, the additional complexity not always produces spectacular improvements.

Channel charge redistribution is due to the fact that the conductive channel of a MOSFET consists of an inversion layer with finite electrical charge [78]. When an N-channel MOSFET is turned ON, electrons are attracted by the gate from the source and drain substrate junctions. When the MOSFET operates in the ohmic region, the channel charge is a function of the gate oxide capacitance density, the width and length of the transistor and the gate-source voltage. An approximative expression is

$$Q_c = C_{ox}\ W\ L\ V_{GS} \tag{140}$$

The gate voltage is usually clocked to a fixed high level, while source and drain voltage are equal to the input voltage of the system. As a result, the channel charge is dependent on, and roughly proportional to the input voltage. When the switch is opened however, the channel charge will be redistributed between source and drain terminals. The way the charge is redistributed can be very hard to predict, since the channel momentarily acts as a distributed-line system. The amount of charge that ends up in each terminal depends on the voltage of each terminal, the impedance seen by each terminal and the rate of change (dV/dt) of the gate voltage.

If switching conditions are not very well controlled or known, equation (140) will have to be used as a worst-case estimate for the clock feed-through due to channel charge injection on the input capacitor. Since the charge injection into one particular node depends on the input voltage in a not necessarily linear way, channel charge injection will create a combination of offset, gain and non-linearity errors. Depending on the size of the capacitor and of the switch itself, the maximum relative error can be fairly high (0.1% to 1%). Accuracy can be improved by using minimum-sized switches and relatively large capacitors (pF

range), but this will degrade the speed of the scheme.

The unpredictability of the charge redistribution can be reduced somewhat by bottom-plate switching (figure 98). The input capacitor of the SC stage of figure 91 is always connected to ground by one switch, and to another node by a second switch. If the switch to ground is opened first, the effect of charge redistribution can be, if not predicted, at least made repeatable. When the ground-side (bottom-plate) switch is opened, its channel always sees the same voltage and the same impedance, and its charge will redistribute in a repeatable way. When the second switch is opened, the first one is already open, leaving only one path for the charge. If bottom-plate switching could be implemented perfectly, the error due to channel charge redistribution would be limited to an offset error.

9.10. Stability

So far, it has been assumed that the opamp gain was frequency-independent. This was a simplification, since any opamp internally consists of active devices with associated parasitics. A better, frequency-dependent gain model would express A (or g_m) as a function of s and would include a number of poles and/or zeros in the transfer function.

Simple mathematical expressions can be obtained if the opamp is modeled as a one-pole system, with the transconductance given as

$$g_m = \frac{g_{m0}}{1 + \frac{s}{\omega_p}} \tag{141}$$

From equation (113), an approximate expression for the output voltage of the SC stage can be derived.

$$v_{out} \approx -v_{in} \quad \frac{C_i}{C_f} \quad \frac{1 - s\frac{C_f}{g_m}}{1 + s\frac{C_{eq}}{g_m}} \tag{142}$$

with

$$C_{eq} = (1 + \frac{C_l}{C_f})(C_i + C_p + C_f) - C_f \tag{143}$$

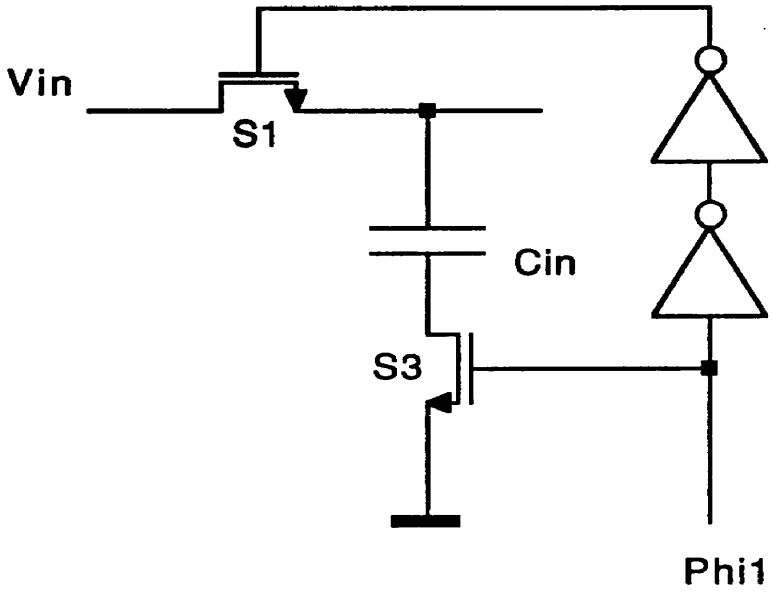


Fig. 98: Bottom-Plate Switching

Replacing g_m by expression (141) yields

$$v_{out} \approx -v_{in} \frac{C_i}{C_f} \frac{1 - s\frac{C_f}{g_{m0}} \frac{1}{1+\frac{s}{\omega_p}}}{1 + s\frac{C_{eq}}{g_{m0}} \frac{1}{1+\frac{s}{\omega_p}}} \tag{144}$$

or

$$v_{out} \approx -v_{in} \frac{C_i}{C_f} \frac{\frac{C_f}{\omega_p} s^2 - C_f\ s + g_{m0}}{\frac{C_{eq}}{\omega_p} s^2 + C_{eq}\ s + g_{m0}} \tag{145}$$

The poles of this system are

$$p_{1,2} = \frac{\omega_p}{2C_{eq}}(-C_{eq} \pm \sqrt{C_{eq}^2 - 4\frac{C_{eq} g_{m0}}{\omega_p}}) \tag{146}$$

In this simplified second order system, strictly speaking, instability cannot occur. According to expression (146), the two poles are either real and negative, or complex conjugate with negative real part. Yet despite its simplicity, this simple model shows the conditions under which the system has complex conjugate poles and ringing (underdamped oscillatory settling) occurs at the SC output. Ringing can be avoided if the dominant pole of g_m is made high enough with respect to the pole of the feed-back system.

$$\omega_p > \frac{4 g_{m0}}{C_{eq}} \tag{147}$$

For a typical situation where $C_i = C_l$ and $C_f = C_i/2$ (gain of 2), the minimum ω_p would be $\approx g_{m0}/C_i$.

When the DC value of g_m is increased, its dominant internal pole must be increased accordingly. When the capacitances in the SC stage are decreased (and hence C_{eq} is decreased), the same rule applies. In other words, the whole SC stage can only be made faster if the dominant internal pole of the opamp is simultaneously moved up in frequency.

In practice, the assumption of a first order transfer function for g_m is not entirely justified and instability is still possible. During Φ_1, the combination of

opamp, C_in and C_f forms a feed-back system. If R_0 is small or C_l is small, the feed-back factor would be approximately

$$\beta \approx C_f/C_i \tag{148}$$

During Φ_1, the opamp is configured as a unity gain feed-back system (C_f shorted by S_5 and $\beta = 1$). Stability in that particular situation requires that all the higher poles and zeros of the opamp, which were not considered in the previous formulas, would be significantly higher in frequency than the open-loop gain-bandwidth product $A_0\ \omega_p$ of the opamp. If this condition is met, the opamp is unity-gain compensated, and chances are it will be stable during Φ_2 as well (β real and ¡1). However, if R_0 and C_l are significant, beta will become frequency-dependent and a more detailed circuit analysis may be needed.

Yet it is clear that for high values of A_0 (DC open-loop gain), which are necessary in order to achieve accurate closed-loop gain, the required opamp gain-bandwidth product can be very high. Moving non-dominant poles and zeros even beyond that frequency, which is necessary to guarantee stability, can be difficult and usually requires significantly increased power consumption.

9.11. Design Methodology

For a particular application, the specifications of an SC stage are usually given by the maximum allowable gain error, the maximum allowable non-linearity and the minimum overall speed (clock frequency) of the stage. Some of these specs are affected by the switches, others by the opamp. Designing and dimensioning the SC stage consists of carefully balancing the different requirements in order to get satisfactory overall performance.

It usually makes more sense to first dimension the switches, as well as the input capacitor of the stage, in order to obtain a switching structure that is fast enough, but not subject to excessive clock feed-through. The difficulty is that faster switching has to be achieved with wide switches and small capacitances, while the same conditions favor distortion due to clock feed-through. In addition, the design of the opamp is easier if C_l is small. If fast, accurate switching is not possible for a wide signal range, choosing a smaller input signal range may be

helpful. Once the input capacitor is determined, the feed-back capacitor can be chosen so as to realize the required overall gain.

Next, the desired overall time constant of the system can be determined using equation (130), with t_1 the duration of one clock phase. Equation (121) makes it possible to determine the minimum DC transconductance g_m of the opamp in order to obtain the desired speed. Equation (147) sets a minimum value for the dominant *internal* pole of the opamp, to be determined with shorted output (opamp operating in transconductance mode).

The opamp design will have to be such that the dominant pole is located at about the frequency predicted by (147), or slightly higher. If for a given initial opamp design, the dominant pole is lower, it will have to be increased by increasing the bias currents of the circuit. If the dominant pole is higher, it may have to be brought down by "compensation", or addition of internal capacitance in order to slow down the system. A high-frequency dominant pole is not bad per se, but makes it harder to assure stability.

Next come additional open-loop gain and stability considerations. The minimum open-loop gain is predicted by equation (124). In order for the opamp to be unity-gain stable, non-dominant poles and zeros must be moved to frequencies well beyond (factor 5-10) the gain-bandwidth product $A\ \omega_p$. In general, satisfying this condition may require several design iterations and associated circuit simulations. The location of poles and zeros of a system is roughly proportional to the square root of the bias currents involved. It is inversely proportional to the square root of the width and inversely proportional to the power 3/2 of the length of critical transistors. The DC gain is proportional to the length of critical transistors and inversely proportional to their DC bias currents.

If an opamp structure can be found that satisfies all the previous requirements, the slew rate must be checked. The opamp must be able to provide the current predicted by equations (131) and (132), without excessive distortion. Finally, opamp linearity must be checked. The minimum small-signal gain, defined in equation (135) can be used as a criterion.

9.12. Conclusion

Gain accuracy, linearity and speed are all desirable, but mutually contradic-

tive attributes of an SC stage. Improving the speed while maintaining performance in other areas, must necessarily be done at the price of significantly larger power consumption within the opamp. Increased power consumption is the only way to increase transconductance, while at the same time moving internal poles and zeros to higher frequencies in order to maintain stability. An additional negative effect of higher power consumption, can be the need for larger devices (in order to accommodate larger currents) and hence increased silicon area. Beside parasitic effects introduced by the opamp, clock feed-through introduced by the switches can also corrupt performance.

It will be shown in the next chapter, how an alternative sample/hold amplifier design avoids many of the problems associated with SC stages. The new approach has been used in a prototype high-speed pipelined converter, which will be discussed later.

CHAPTER X

THE CHARGE AMPLIFIER

10.1. Introduction

This chapter introduces a new discrete-time (sample/hold) amplifier scheme that is based on charge rather than voltage amplification. The charge is temporarily held on a capacitor and amplified using a combination of simple comparators and current sources. The gain can be pre-set by choosing the ratio of bias currents, but unlike an SC amplifier, it can also be adjusted through an external current. Although the scheme was originally developed for pipelined converter applications, this peculiarity could open the road to tunable or programmable discrete-time analog filters as well.

The scheme does not necessarily require MOSFET switches, which makes it possible to integrate it using either MOS or bipolar technologies. It is not subject to instability, which eases design constraints. Finally, due to its inherent robustness, this scheme makes it possible to realize systems that operate reliably in harsh or radiation-filled environments [87]. The radiation hardness can even be improved if the scheme is implemented in bipolar technologies.

A draw-back of the scheme is that in many cases a specialized input stage is needed to convert a voltage signal to a charge stored on a capacitor. However, in some applications, like optical sensing using photodiodes or charge-coupled devices (CCD's), the stimulus appears as a charge and the charge amplifier provides a direct method to accomplish analog signal processing or analog to digital conversion.

10.2. Basic Operation

Unlike in other amplifiers, the signal-carrying variable of this scheme is a charge rather than a voltage. In practice, this distinction is minor, since the charge will be temporarily stored on a capacitor. As long as the capacitor has an approximately linear Q/V relationship, the capacitor voltage will be proportional to the signal.

The amplification relies on a ratio of currents, which can be made very accurate, even when the capacitors in the system are not accurately matched or entirely linear. The amplification occurs by discharging the signal charge of one capacitor, while copying it onto the capacitor of the next stage. The discharge operation is realized by a highly constant (high output impedance) current source, which is turned ON until the voltage across the capacitor reaches a pre-determined reference level (figure 99). During this operation, the capacitor voltage decreases in a linear way, until the current source is turned OFF. (figure 100).

The output of a fast comparator is used to control (gate) the current source. Before the reference level is reached, the comparator output is logically high and the current source ON. As soon as the reference level is detected, the comparator output switches and the current is turned OFF. The same comparator drives another, similar current source, connected to the capacitor of the next stage.

Since both current sources are ON during the same amount of time, the charge ($Q = \int_0^t i(t)\ dt \approx I\ t$) displaced by each of them will stand in a fixed ratio, determined by the ratio of currents. In other words, the current ratio of the two current sources determines the charge gain of the system. By carefully choosing the reference levels involved and the direction of the currents, an inverting or non-inverting charge amplifier can be realized, with any desired level of systematic offset.

In this scheme, capacitor linearity or matching is essentially irrelevant in order to achieve a constant, linear charge gain. The gain is set by a ratio of precision current sources, which can be set very accurately [88]. (However, it should be mentioned that better matching may require sacrificing speed: very precise current ratios are usually obtained using fairly large transistors, which have a tendency to switch slower than small ones. In the analog to digital converter that is described at the end of this text, this is of little concern, since gain errors will be calibrated out digitally, using accuracy bootstrapping.)

10.3. Clocking Scheme

Just like a switched-capacitor amplifier, the charge amplifier requires a clock. In its most elementary form, the scheme needs three distinct, non-overlapping phases. Cancellation of parasitic offsets and low-frequency noise, if

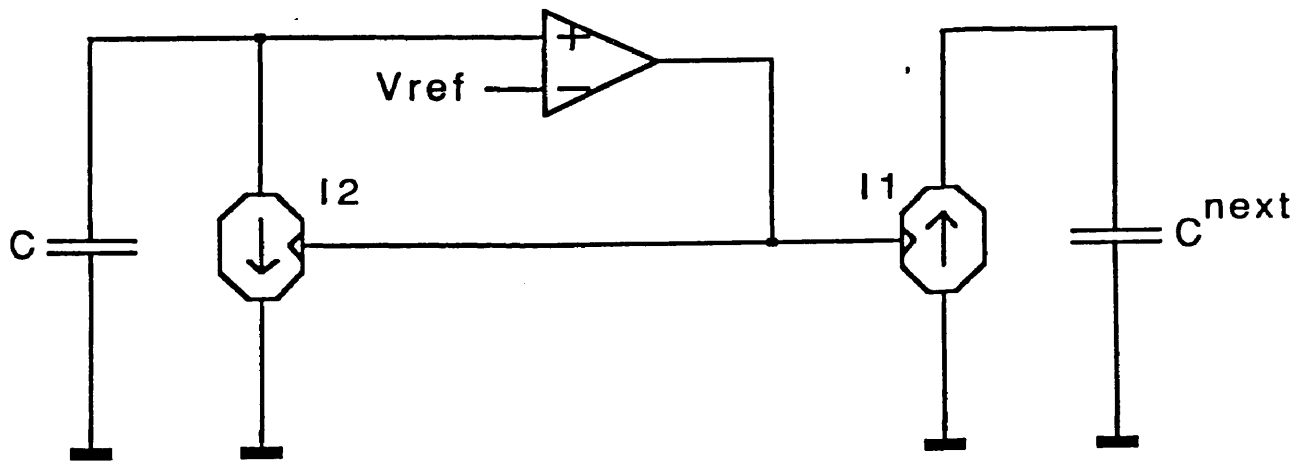


Fig. 99: Basic Charge Amplifier Components

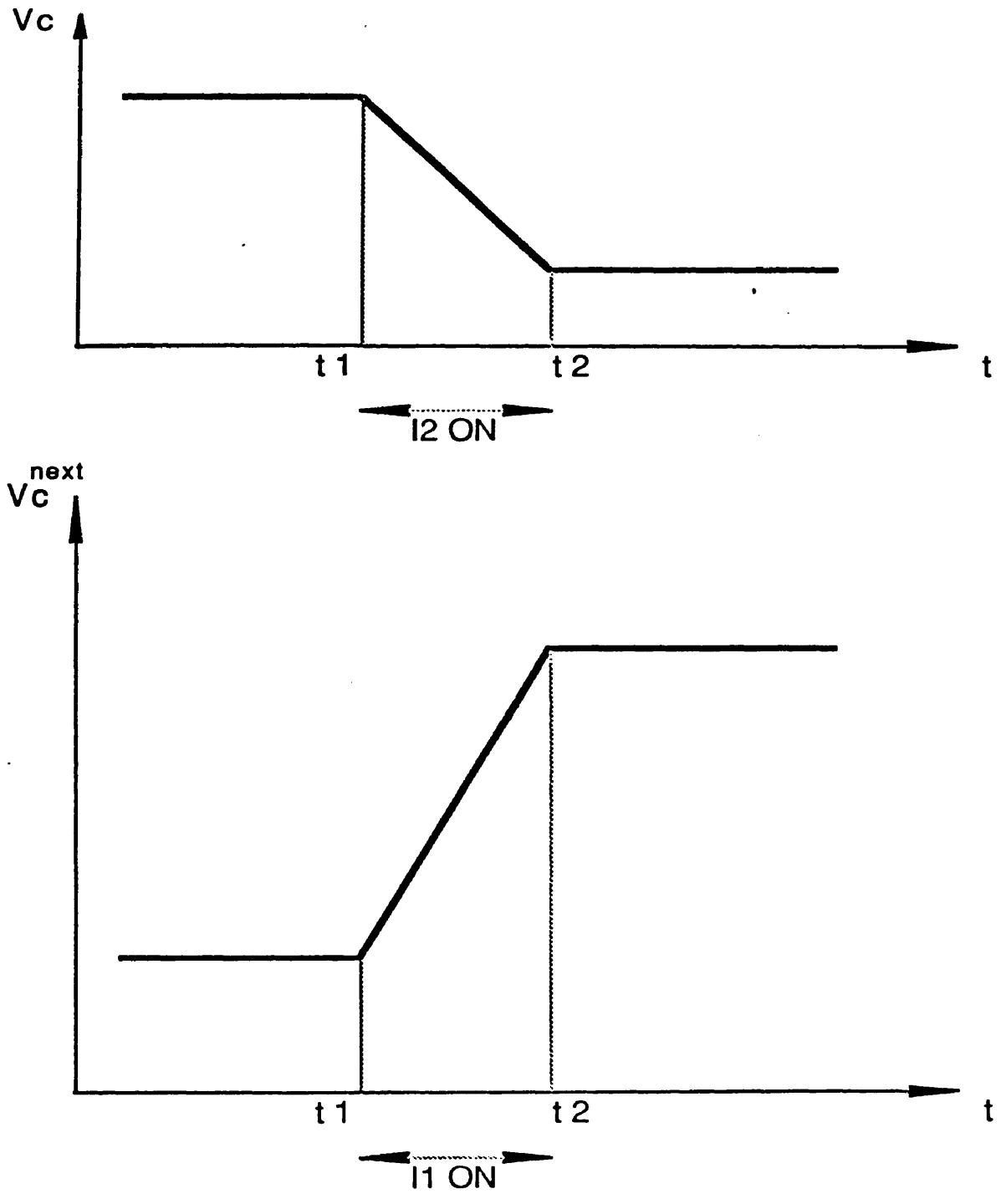


Fig. 100: Basic Capacitor Discharge Waveform

desired, requires at least one additional clock phase. When used as part of an analog/digital converter (ADC), the charge amplifier may need additional phases, bringing the total up to six.

During clock phase Φ_1, the signal capacitor C is precharged to a fixed level (the precharge level). Φ_1 will be referred to as the *precharge phase.*

Clock phase Φ_2 is reserved for offset cancellation, which will be described in detail below. During this phase, the initial charge is adjusted with the help of the comparator. The capacitor is partially discharged to a first reference level (the initial discharge level), slightly below the precharge level. Φ_2 will be called the *initial discharge phase.*

During clock phase Φ_3, the signal charge is applied to the capacitor with the help of a constant current source. The signal charge can be assumed to be copied from a previous stage. The comparator of the previous stage drives the current source of the present stage. The difference in charge on the capacitor between the beginning of Φ_3 and the end of Φ_3 represents the analog signal. Depending on the direction of the current source, the capacitor can be either charged or discharged. This distinction determines the sign of the gain between the current stage and the next one. Φ_3 will be called the *load phase.*

Clock phase Φ_4 is reserved for ADC applications and will be described in more detail below. During this phase, the signal is compared against a number of reference levels using comparator/latch combinations. The operation is equivalent to the operation of a "flash" ADC. Φ_4 will be called the *flash phase.*

Phase Φ_5 is also reserved for ADC applications. During this phase, a given charge is subtracted from the main capacitor. The magnitude of the subtracted charge can be any one value out of a set of possible charges, depending on the result of the comparisons during the flash phase. The purpose is to decrement the signal level in order to calculate a residue, which is the remainder of the local analog to digital conversion. The residue calculation is equivalent to the subtraction of the output of an elementary digital/analog converter (DAC) from the input signal. The remaining signal (residue) is amplified and passed on to the next stage. Φ_5 will be called the *DAC phase.*

During Φ_6, the capacitor is further discharged to a fixed zero state (final

discharge level), using the current source. The same comparator as in Φ_2 is used to detect when the appropriate level is reached. While the capacitor is being discharged, the charge is copied (and multiplied) onto a similar capacitor in the next stage. The copying is realized by a second current source, connected to the output of the same comparator. The ratio of the two currents involved, determines the charge gain. Obviously, in order to maintain synchronization, phase Φ_6 of the present stage must correspond to the load phase (Φ_3) of the next stage. In other words, there is a three phase delay between one stage and the next one. Phase Φ_6 will be called the *final discharge phase.*

10.4. Analysis of the Precharge Phase

The capacitor C is precharged to a fixed precharge voltage V_p, through switch S_1 (figure 101). S_1 is normally a single MOSFET, but implementation in bipolar technology is possible as well. Depending on the size of the switch, the value of C and the duration of Φ_1, the capacitor voltage may or may not have time to fully settle to a final value (the reference voltage $V_{ref,p}$). When the switch is opened, a certain amount of parasitic charge may also be injected onto the capacitor, due to clock feed-through (parasitic capacitance between clock and switch terminal, and channel charge redistribution).

Despite these parasitic effects, which cause the actual value of V_p to be slightly different from $V_{ref,p}$, V_p is still perfectly deterministic (repeatable). C is always charged starting from the same initial condition (fully discharged), controlled by the same clock phase Φ_1. The symbol Q_p will be used to refer to the capacitor charge after completion of the precharge operation.

$$Q_p = \int_0^{V_p} C \, dv \tag{149}$$

10.5. Analysis of the Initial Discharge Phase

The precharge voltage V_p, though deterministic, may be hard to predict with a high degree of accuracy, due to imperfect settling and/or clock feed-through. This is one reason, though not the main one, why an initial discharge phase is included. A more important reason is the performance of an auto-zero (or

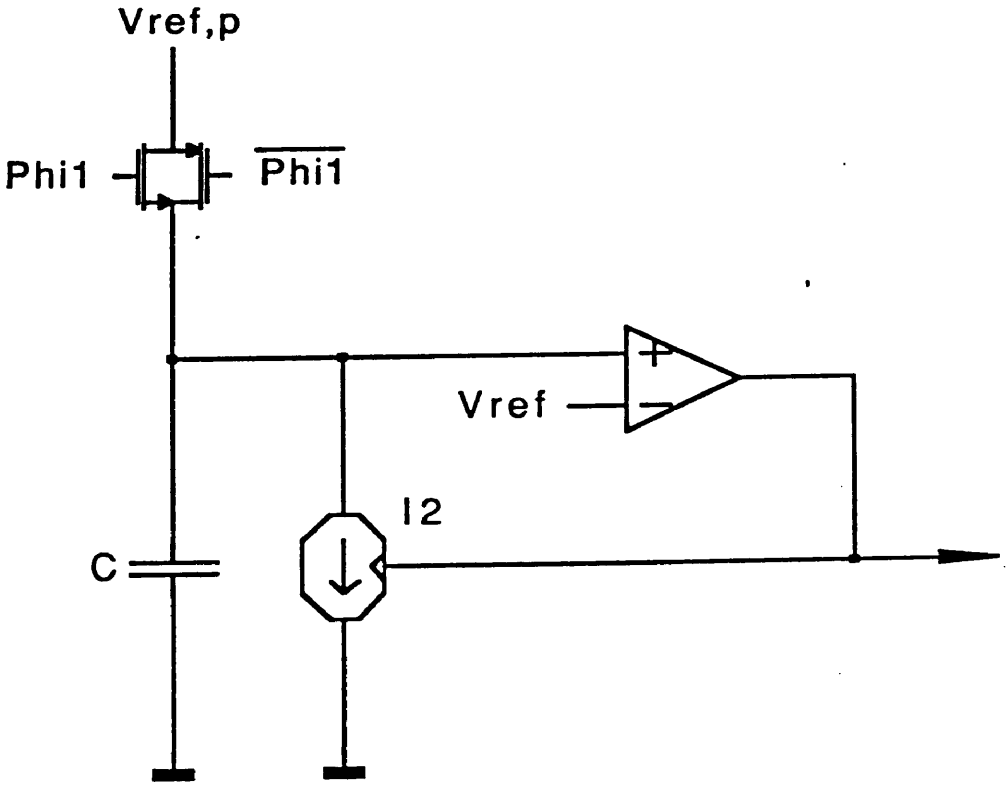


Fig. 101: Precharge Phase with Switch

offset-cancellation) operation, in order to compensate for imperfections in the fast comparator. Such imperfections are finite switching time, random offset voltage and low-frequency (mainly $1/f$) noise.

The technique is functionally equivalent to correlated double sampling (CDS). CDS consists of sampling a zero (or reference) level before sampling the actual input signal. By forming the difference of the corresponding two output signals, the effect of any systematic offset errors is cancelled. The effect of low-frequency random variations is greatly reduced, due to the fact that these noise components cannot change much within one clock period.

During Φ_2, the fast comparator of the charge amplifier stage is used to discharge the capacitor C to some voltage below V_p (figure 102). For this purpose, a reference voltage $V_{ref,i}$ is applied to the inverting input of the comparator. Later on, during phase Φ_6 (after the signal charge is applied), the capacitor will be discharged further, to a final state. At that time, another reference level, $V_{ref,d}$ is applied to the inverting comparator input.

Ideally, assuming that each time the capacitor is discharged *exactly* down to the specific reference level, the difference in capacitor charge (after Φ_2 and after Φ_6) will be

$$Q_{26}^{id} = Q_i^{id} - Q_d^{id} = Q(V_{ref,i}) - Q(V_{ref,d}) \tag{150}$$

or

$$Q_{26}^{id} = \int_0^{V_{ref,i}} C(v)\, dv - \int_0^{V_{ref,d}} C(v)\, dv = \int_{V_{ref,d}}^{V_{ref,i}} C(v)\, dv \tag{151}$$

In practice, after Φ_2 the capacitor is discharged to a voltage V_i, slightly below $V_{ref,i}$. After Φ_6, the capacitor is discharged to a voltage V_d, slightly below $V_{ref,d}$. The actual value of Q_{26} is

$$Q_{26} = \int_0^{V_i} C(v)\, dv - \int_0^{V_d} C(v)\, dv = \int_{V_i}^{V_d} C(v)\, dv \tag{152}$$

If no offset cancellation phase (Φ_2) were applied, the comparator would only be used during the final discharge phase (Φ_6), in order to determine when the reference level $V_{ref,d}$ would be reached and when the current source used during the discharge operation could be turned off. Due to random offset, the actual

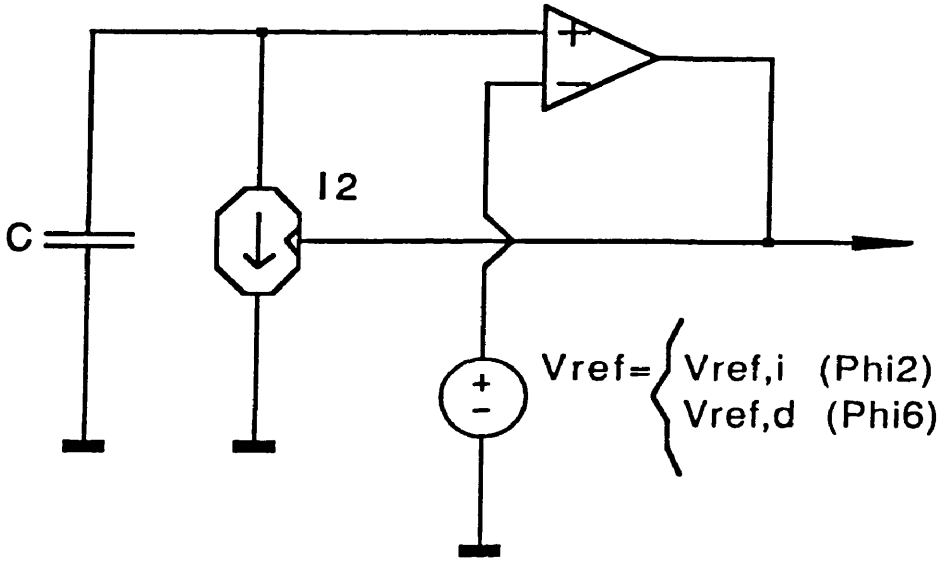


Fig. 102: Initial Discharge Phase and Offset Cancellation

voltage at which the comparator flips would not be accurately known. More importantly, due to finite switching speed, the current source would be switched off with a certain delay, and the capacitor voltage would undershoot with respect to $V_{ref,d}$. This undershoot might not even be entirely constant, since low frequency noise within the comparator would create a certain variability in the switching delay (time jitter).

Eventually, in the fully discharged state (after Φ_6), the voltage across C would have a certain value V_d, different from $V_{ref,d}$. The corresponding difference in charge on C, $Q(V_{ref,d}) - Q(V_d)$ will be called the undershoot charge Q_u.

This undershoot is introduced by the comparator, and entirely determined by the comparator characteristics. Assuming that the current through I_2 is perfectly constant, the discharging of C will occur in a perfectly predictable and repeatable way. As a result, the voltage as well as its derivative ($dV/dt = I_2/C$), seen by the non-inverting comparator input, will be perfectly predictable any time the capacitor voltage reaches the reference level $V_{ref,d}$. Since the comparator characteristics (offset and switching delay) are constant, the undershoot charge Q_u will be the same every time the comparator switches at the end of the discharge phase.

The initial discharge phase is designed to cancel the effect of the undershoot charge Q_u. The idea is to use the same comparator as in the *final* discharge phase, to create a similar undershoot on the *initial* discharge level. The reference level $V_{ref,i}$, slightly lower than $V_{ref,p}$, is applied to the inverting input of the comparator during Φ_2 and the current source I_2 is used to discharge C to this level. Since comparator offset and switching delay normally are fairly constant within a clock phase, even if a different reference voltage is applied, C will be discharged to a voltage V_i such that

$$Q(V_{ref,i}) - Q(V_i) \approx Q_u = Q(V_{ref,d}) - Q(V_d) \tag{153}$$

As a result, the actual difference in capacitor charge (after Φ_2 and after Φ_6) will be

$$Q_{26} \approx (Q_i^{id} - Q_u) - (Q_d^{id} - Q_u) \tag{154}$$

or

$$Q_{26} \approx [Q(V_{ref,i}) - Q_u] - [Q(V_{ref,d}) - Q_u] = Q(V_{ref,i}) - Q(V_{ref,d}) \tag{155}$$

This is the same expression as in equation (151). In other words, as far as Q_{26} is concerned, the effect of comparator offset and/or switching delay has been eliminated.

10.6. Analysis of the Load Phase

During Φ_3, a certain charge is added to or subtracted off the capacitor through current source I_1 (figure 103). The direction of I_1 determines if addition or subtraction occurs. This charge represents the input signal. In algebraic terms, a certain signal charge Q_s is added onto C, Q_s being positive for a non-inverting configuration, negative for an inverting configuration.

The signal that controls I_1 is a digital signal, produced by the comparator of a previous stage (or an input stage, to be described below). Since I_1 is a constant current source, the control signal must necessarily have a variable pulse width, proportional to the signal charge Q_s. This implies that the signal comes into a stage in pulse width modulated (PWM) form. Within the stage, the PWM signal is transformed into a charge by the current source I_1 and temporarily stored. Later on, the charge will be converted to PWM again and transmitted to the next stage. Charge amplification occurs during this process.

To our knowledge, the use of PWM digital coding to represent a signal in a complex signal processing device has never been reported before. One application that bears a remote resemblence is the digitally programmable filter presented by Vallancourt and Tsvidis [89]. However, their scheme uses an analog integrator which is activated for a predetermined time in order to achieve programmable gain. The signal itself is represented by a voltage.

The charge amplification scheme relies on an implicit change of signal-carrying variable from charge to pulse width and back, through the relation

$$Q = \int_0^T I(t)\, dt \tag{156}$$

Which for constant current simplifies to

$$Q = T\ I \tag{157}$$

After Φ_3, the charge on C is the sum of Q_i (total charge after initial discharge) and Q_s, which was added through current source I_1.

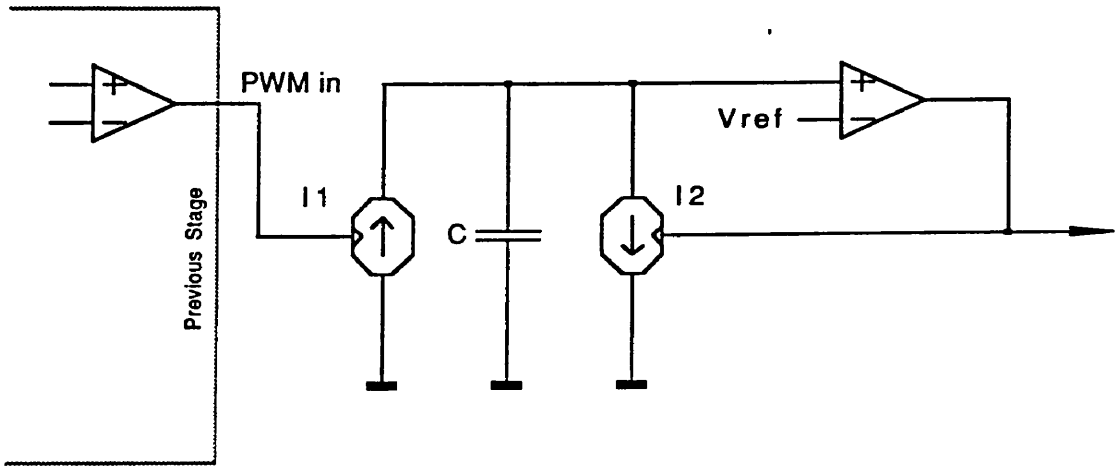


Fig. 103: Load Phase

10.7. Analysis of the Flash Phase

The flash phase, as well as the DAC phase, are needed to turn the charge amplifier stage into one stage of a pipelined analog/digital converter. A simplified schematic of such a converter stage is given in figure 104. The input signal is compared against a number of reference voltages, using comparators. The digital output pattern of the comparators forms a "thermometer" code, i.e. the number of ones in the binary word is proportional to the input signal. The code is latched, and if needed, converted to a regular binary number using an encoder. Finally, a certain quantity is subtracted from the input signal, depending on the output code determined by the comparators. The result of the subtraction is the residue, which is amplified and fed to the next stage.

During Φ_3, the signal charge Q_s was added to capacitor C. The resulting charge at the end of Φ_3 is

$$Q_3 = Q_i + Q_s = Q(V_i) + Q_s = Q(V_{ref,i}) - Q_u + Q_s \tag{158}$$

This charge determines the capacitor voltage at the end of Φ_3, through its C/V characteristic. If this characteristic is more or less linear, the voltage on C will be proportional to Q_s.

$$V_{C,3} \approx V_{ref,i} - \frac{Q_u}{C} + \frac{Q_s}{C} \tag{159}$$

This makes it possible to compare Q_s to a number of fixed reference levels using an elementary flash converter with a reference voltage string (usually a resistive divider string) and as many comparators as there are reference levels (figure 105). Each comparator is followed by a digital latch, which is activated during Φ_4. At the end of Φ_4, the bit pattern at the output of the latches forms a thermometer-coded local conversion code, proportional to the value of Q_s.

If Q_s is to be compared against N fixed reference charge levels $Q_1 \cdots Q_N$, the reference voltage string has to be dimensioned so as to generate the voltages $V_1 \cdots V_N$, with

$$V_i = V_{ref,i} - \frac{Q_u}{C} + \frac{Q_i}{C} \tag{160}$$

If the undershoot charge Q_u is small (negligible), this can easily be realized. However, if Q_u is significant, some kind of offset compensation scheme will have

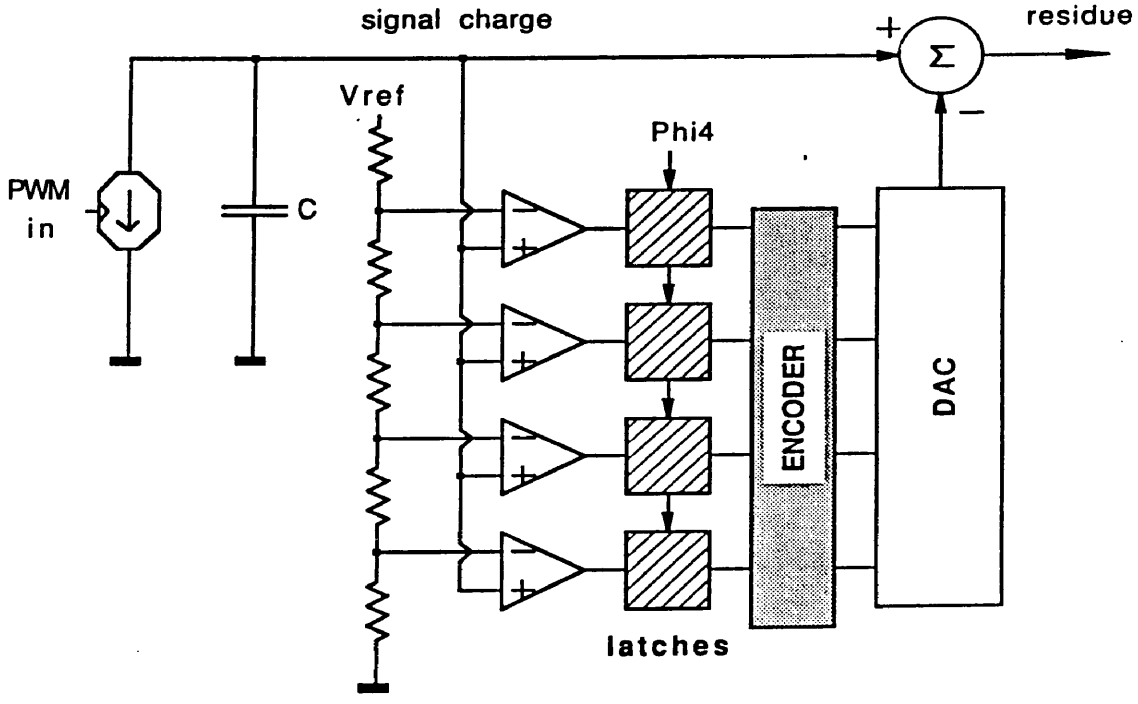


Fig. 104: A/D Converter Stage

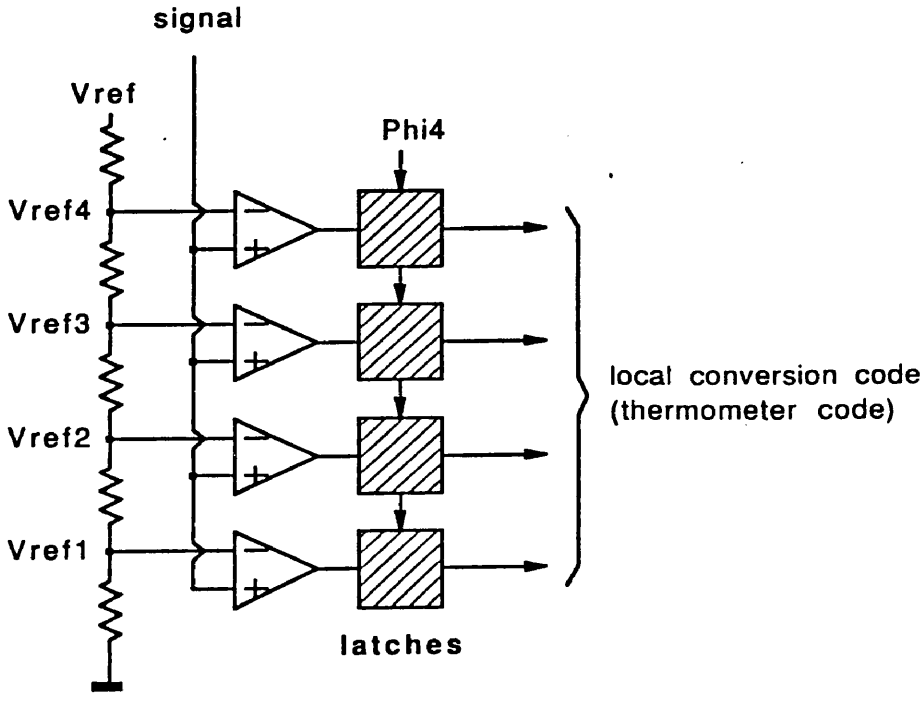


Fig. 105: Detail of the A/D Converter Stage

to be implemented in order to cancel the effect of Q_u on the reference string. This will be described below.

The approach assumed that the capacitance value of C was known, and that C was linear. If this is not actually the case, equation (160) will not accurately predict the reference voltages to be generated by the reference string. Still, the equation can be used, since the variation will be fairly small. Most A/D converters apply digital error correction (or redundancy, as described in chapters V and VI), which makes the whole structure completely insensitive to small variations on the reference voltages of the flash comparators. In many cases, the effect of the undershoot charge Q_u can be neglected for the same reason.

If Q_u cannot be neglected, this charge will have to be compensated for, using some offset cancellation technique. This usually is the case in fast systems, in which the capacitor is discharged by a large current. The slew rate (dV/dt) on the capacitor is high and for a given comparator delay, the undershoot created by the initial and final discharge operations will be proportionally larger. In equations (159) and (160), the parasitic term Q_u/C will have to be taken into account.

It would be possible to estimate this parasitic term through circuit simulations, and then adjust the reference levels V_i accordingly. However, this would imply that the operation of the charge stage could be predicted very accurately, and it would also strongly reduce the robustness of the system against drifts in device or technology parameters.

A safer approach is to build a self-adjusting scheme. Again, this can be accomplished using a pre-charge / initial discharge approach. This does not go without additional hardware however. Each reference level will need its own capacitor, current source and additional comparator, identical to the ones used for the main initial discharge operation (phase Φ_2). This is shown in figure 106 and 107 (detail). The reference levels are generated in two steps: precharge of the corresponding capacitor to a level slightly above the desired level, and discharge down to the desired reference level using the current source / comparator combination.

To the extent that all components match the main components of the system, the effect of undershoot will be eliminated. However, random offset on the

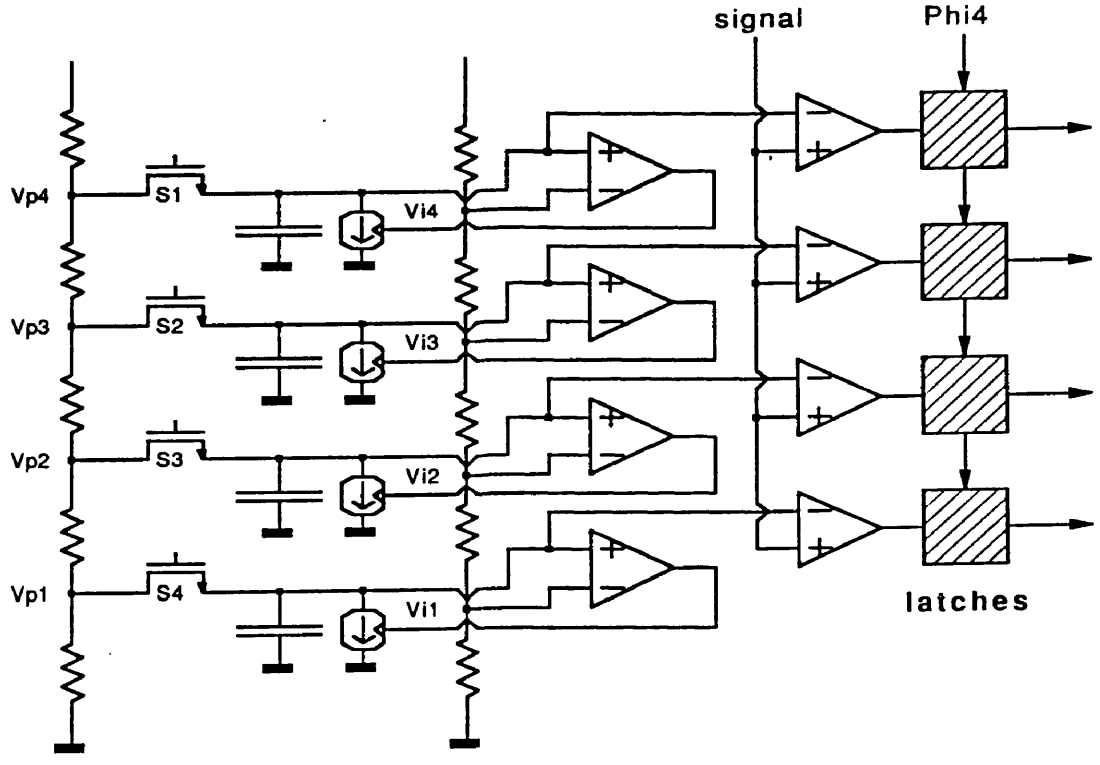


Fig. 106: Cancellation of Undershoot on the Reference Levels

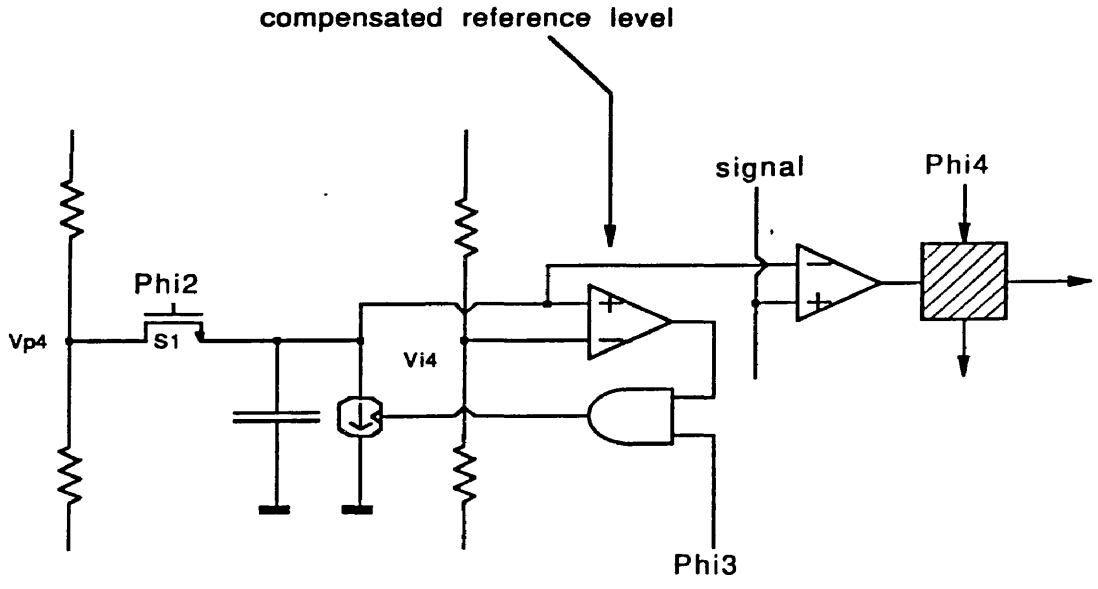


Fig. 107: Detail of Undershoot Cancellation

flash comparators themselves, will still be present. Another technique, using a combination of a resistive and capacitive voltage divider, may prove itself useful (figure 108).

The main disadvantage of these techniques is the increased complexity in terms of current sources, capacitors and comparators. In turn, this complexity implies a significantly higher power consumption. As a result, the approach normally only pays off for very fast ADC systems (large undershoot), with very few bits (and hence few reference levels) per stage.

10.8. Analysis of the DAC Phase

The purpose of the DAC phase is to use the local conversion code produced by the flash comparators and latched during Φ_4, in order to subtract a certain charge from the input signal. The amount of charge subtracted can be any one out of a number of pre-set values. Which charge is subtracted, is determined by the local conversion code. As a result, it can be said that the local conversion code acts as a digital input to an elementary charge-based digital/analog converter (DAC), of which the output charge is subtracted from the signal charge. The charge values themselves are determined so as to be proportional to the input signal, but increasing in discrete increments (figure 109 and 110).

If there are N comparators, there will be $N + 1$ possible local conversion codes, with each bit corresponding to the output of one comparator. The codes range from $00\cdots0$, $10\cdots0$, $11\cdots0$ through $11\cdots1$. Every time the input signal exceeds the reference voltage of one comparator, an additional comparator will have an output of 1.

Normally, the reference levels of the comparators are spaced apart in equal voltage intervals ΔV. The associated charge intervals are $\Delta Q = C\ \Delta V$. Ideally, the charge increments at the output of the DAC are spaced apart by the same amount ΔQ.

There are several ways to implement the subtraction. A straight-forward method is to use a number of auxiliary current sources, pulsed on for a certain time T_1, and controlled by the individual bits of the local conversion (thermometer) code (figure 111). The currents I^{aux} are determined so that

$$T_1\ I^{aux} = \Delta Q = C\ \Delta V \tag{161}$$

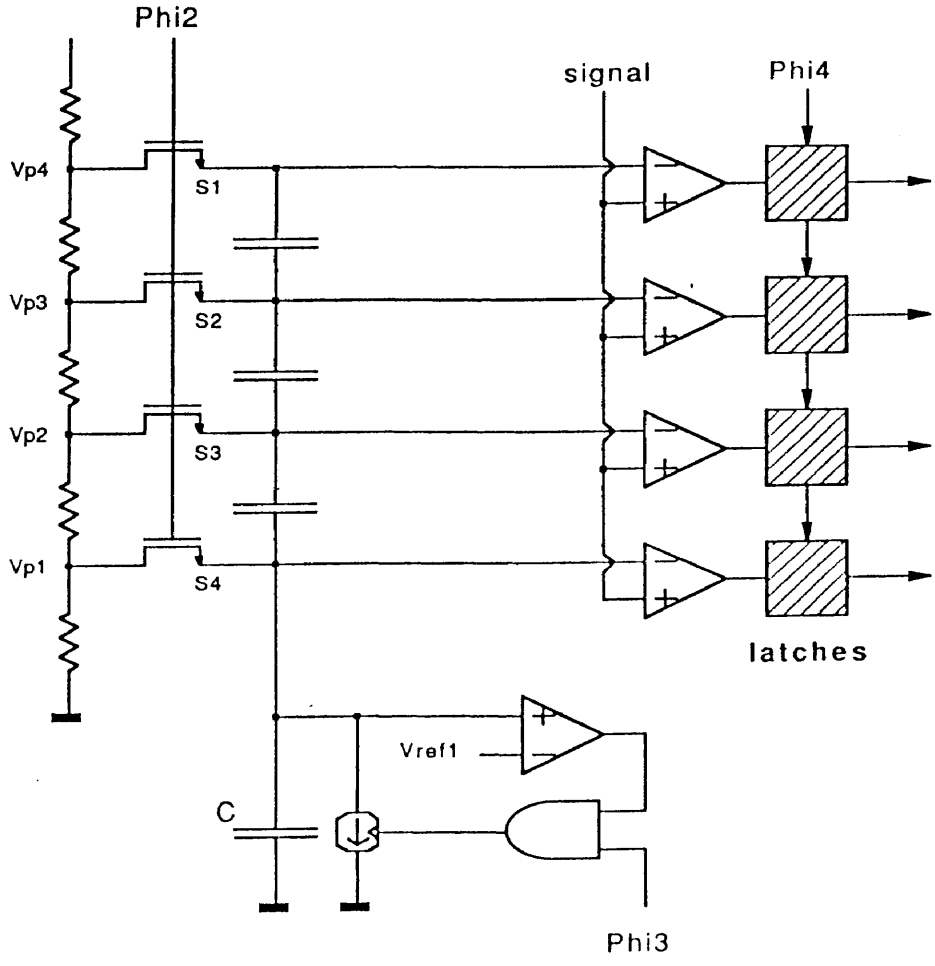


Fig. 108: Undershoot Cancellation Using Capacitive Divider

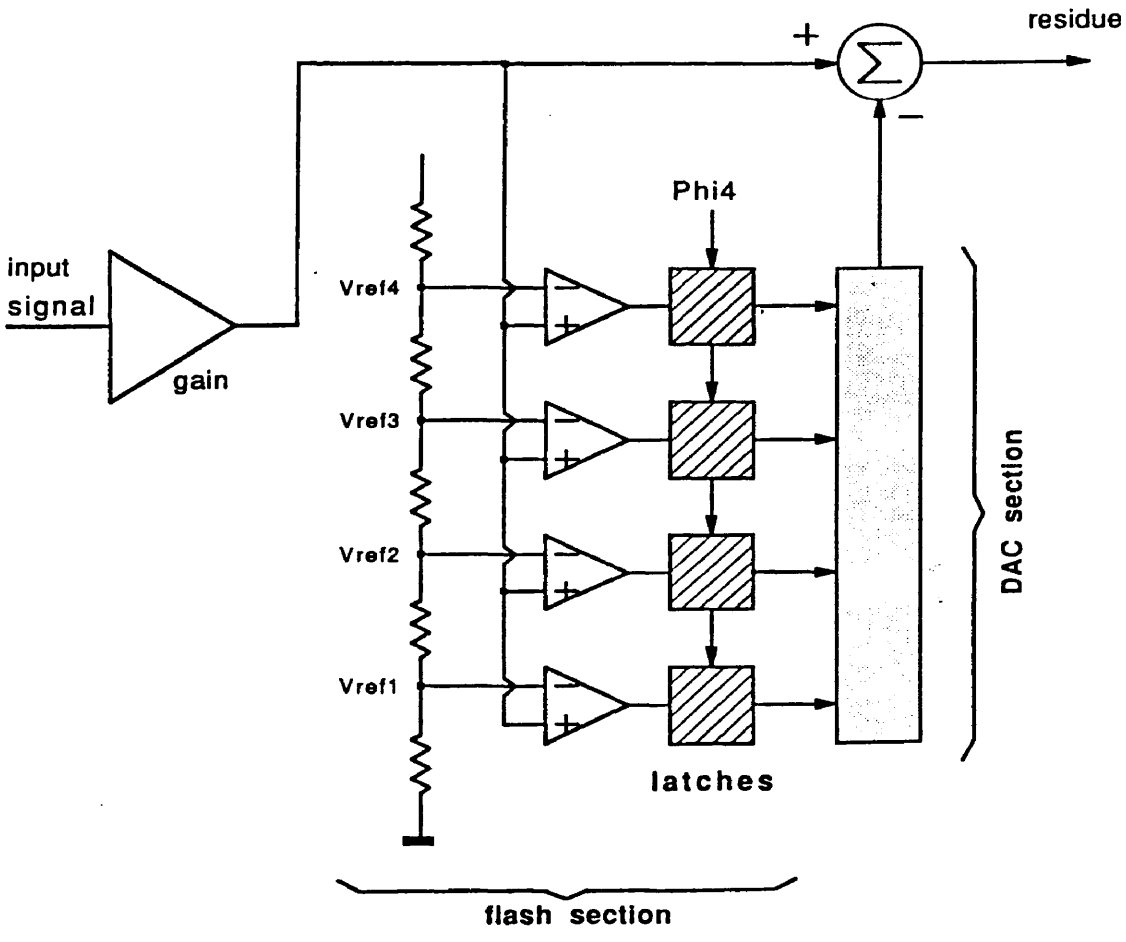


Fig. 109: A/D Converter Stage, Main Components

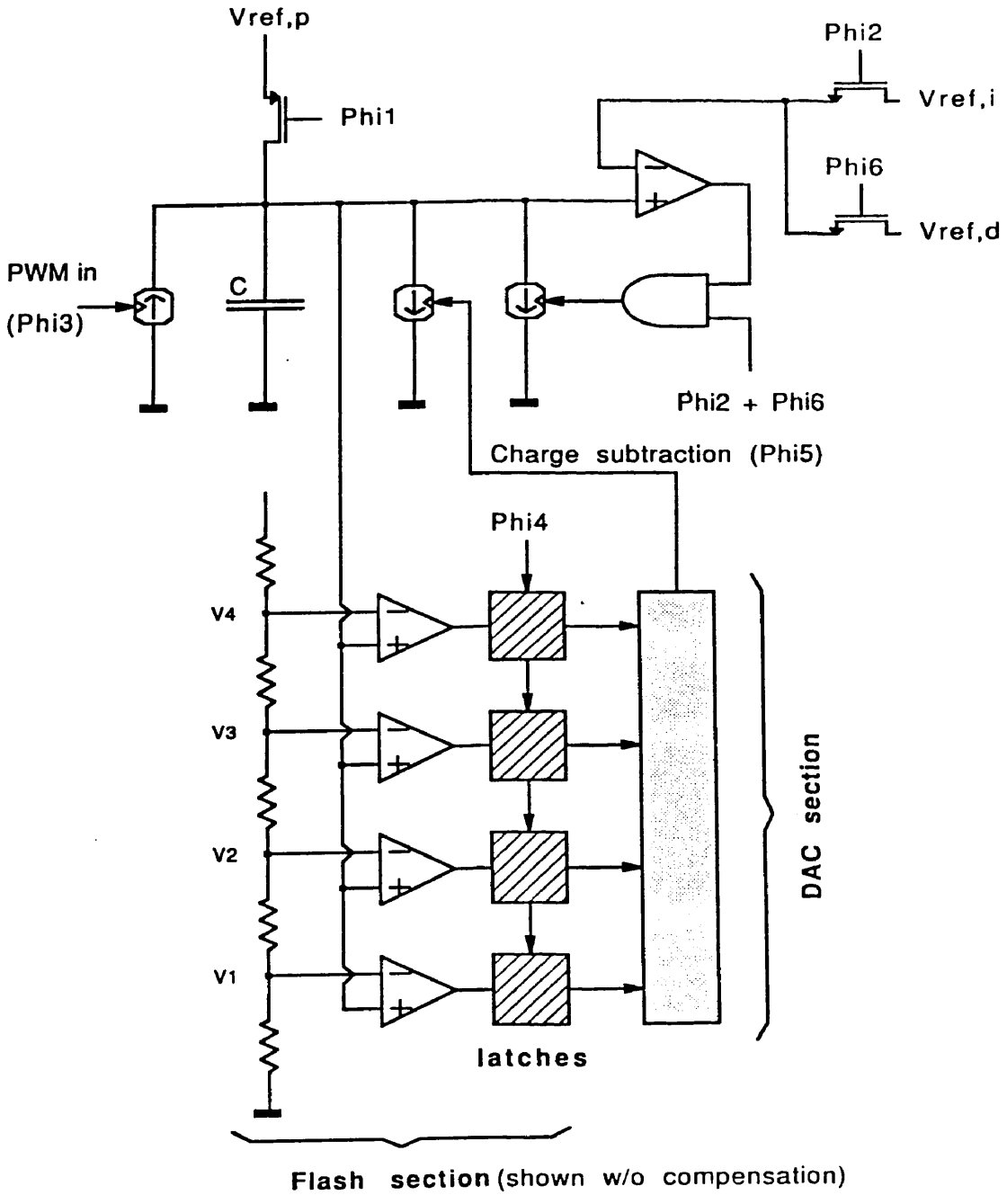


Fig. 110: A/D Converter Stage (Detailed Schematic)

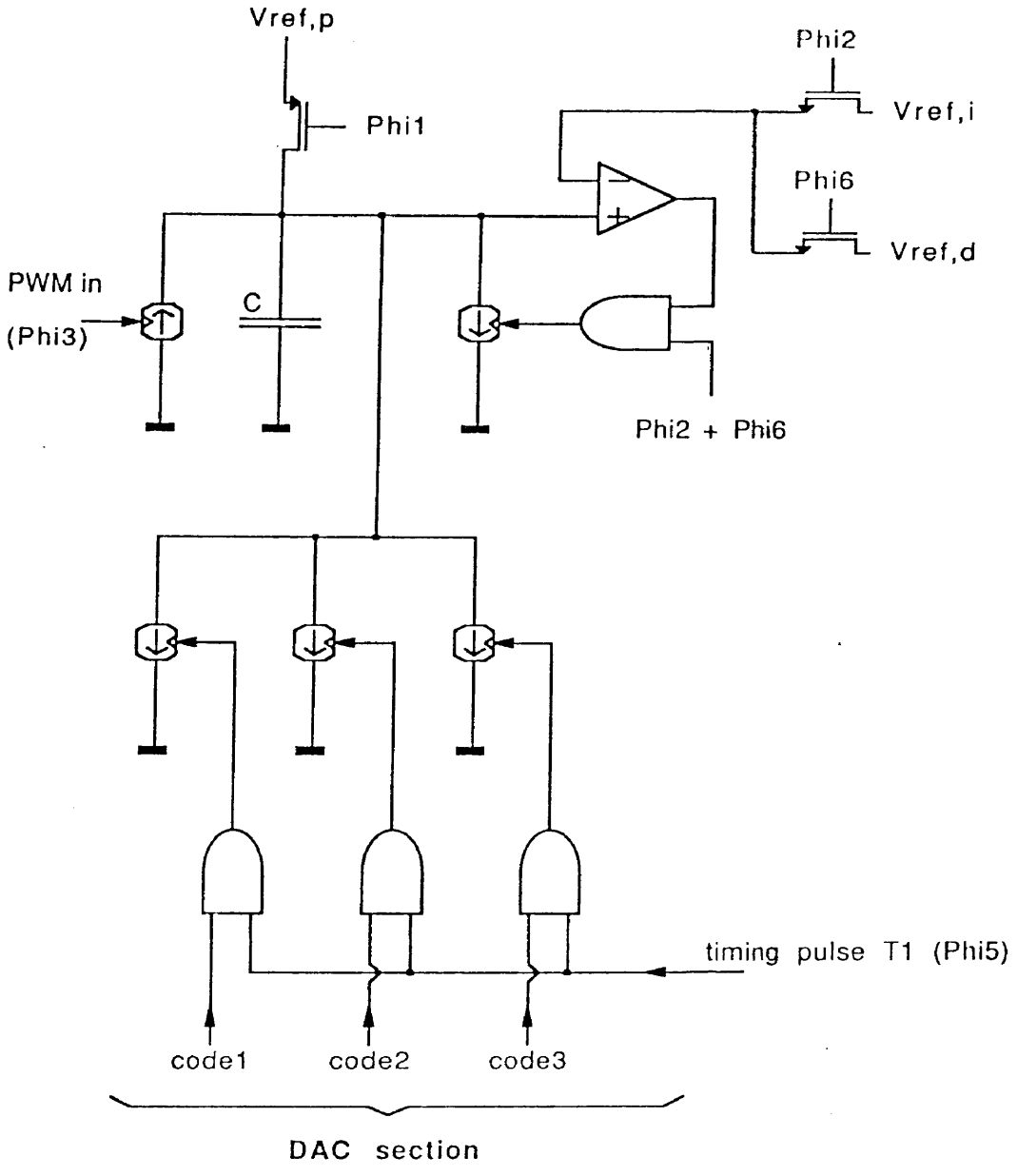


Fig. 111: DAC, Using Auxiliary Current Sources

Since T_1 is the same for all of the auxiliary current sources (equal charge increments), a common signal can be used for the timing. Actually, the whole DAC section can be implemented as N independent, identical current sources, connected to the common node of capacitor C. The gate signal of each current source is generated by ANDing together the appropriate bit of the local code and a common timing signal with pulse width T_1.

The timing pulse can be generated using an auxiliary capacitor C^{aux} and another fast comparator, similar to the one used during the precharge and final discharge phase (figure 112). During Φ_3, C^{aux} is precharged to a precharge level V_p^{aux}. During Φ_4, it is partially discharged to an initial discharge voltage V_i^{aux}. Finally, during Φ_5 a final discharge operation is performed down to a voltage V_f^{aux}. The whole procedure is very similar to the operation of the main part of the charge amplifier stage, and one can easily see that the width of the resulting pulse at the output of the comparator is not subject to any systematic offset, switching delay or low-frequency time jitter, potentially introduced by the comparator.

If the reference current source connected to C^{aux} carries a current I^{aux} as well, then the charge increments at the output of the DAC will be

$$\Delta Q = T_1 \; I^{aux} = \frac{(V_{ref,i}^{aux} - V_{ref,f}^{aux}) \; C^{aux}}{I^{aux}} I^{aux} = (V_{ref,i}^{aux} - V_{ref,f}^{aux}) \; C^{aux} \qquad (162)$$

By setting $V_{ref,i}^{aux}$ and $V_{ref,f}^{aux}$ so that their difference is equal to the increment of the resistive divider, the DAC stage is automatically correctly dimensioned. Actually, these voltages can be derived directly from the same resistive divider as is used for the reference levels of the flash ADC. Device mismatches may cause variations from the ideal values, but the characteristics of the system will not be signal-dependent. Reducing the effect of the variations is the task of the self-calibration algorithm to be used with the A/D converter and will not be discussed here.

However, it has to be mentioned that by externally controlling the auxiliary current sources, instead of controlling them by the flash comparator outputs, each charge increment can be applied individually. This is necessary in order to apply self-calibration through accuracy bootstrapping.

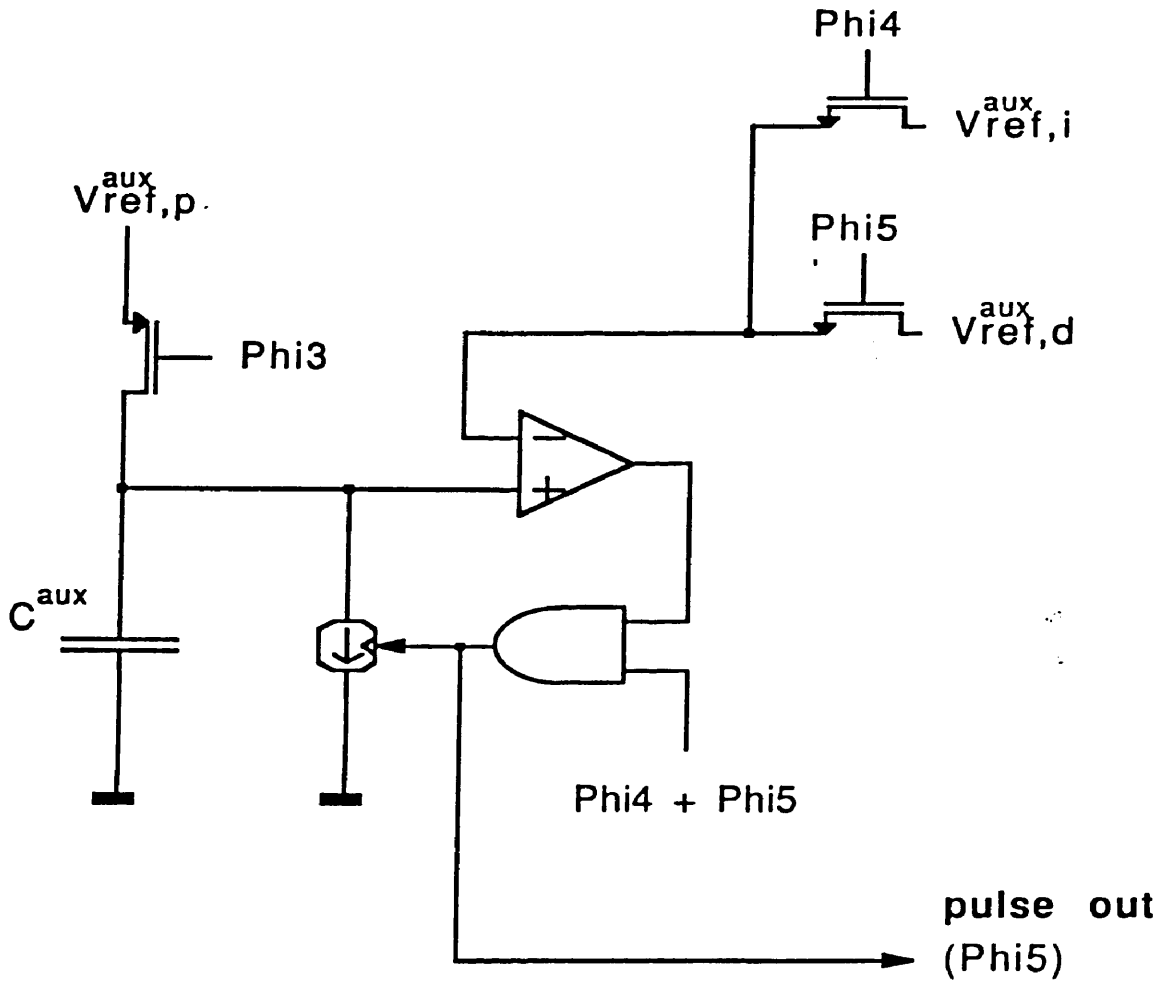


Fig. 112: Generation of the Timing Pulse

10.9. Analysis of the Final Discharge Phase

During Φ_6, capacitor C is further discharged to the voltage V_d, using current source I_2. The current source is switched ON at the beginning of the clock phase and switched OFF when the output of the comparator flips from a logic high to a logic low. Before Φ_6, the charge on C is $Q_i + Q_s$ (we do not consider the effect of the DAC phase, for simplicity). After Φ_6, the charge is Q_d, like described in an earlier paragraph. Hence, the charge removed by I_2 during Φ_6 is

$$Q_6 = Q_i + Q_s - Q_d = Q(V_{ref,i}) - Q_u + Q_s - [Q(V_{ref,d}) - Q_u] = Q_{26} + Q_s \quad (163)$$

This means that I_2 is ON during an effective time equal to (*effective* time means not taking into account switching transients).

$$T = \frac{Q_6}{I_2} = \frac{Q_{26} + Q_s}{I_2} \quad (164)$$

The comparator output is a PWM signal with pulse width T. This same signal controls the current source I_1 of the next stage (figure 113). As a result, the charge added to the capacitor of the next stage is

$$Q_s^{next} = \pm T\ I_1^{next} = \pm \frac{Q_{26} + Q_s}{I_2}\ I_1^{next} \quad (165)$$

The plus sign refers to the non-inverting configuration, the minus sign to the inverting configuration. The net result of the operation is charge amplification between the present stage and the next one. The charge gain is

$$A_q = \pm \frac{\delta Q_s^{next}}{\delta Q_s} = \pm \frac{I_1^{next}}{I_2} \quad (166)$$

The output-referred offset is

$$Q_{off}^{out} = \pm Q_{26}\ \frac{I_1^{next}}{I_2} = \pm [Q_i - Q_d]\ \frac{I_1^{next}}{I_2} = \pm [Q(V_{ref,i}) - Q_{ref,d}]\ \frac{I_1^{next}}{I_2} \quad (167)$$

The input-referred offset is

$$Q_{off}^{in} = \pm Q_{26} = \pm [Q_i - Q_d] = \pm [Q(V_{ref,i}) - Q_{ref,d}] \quad (168)$$

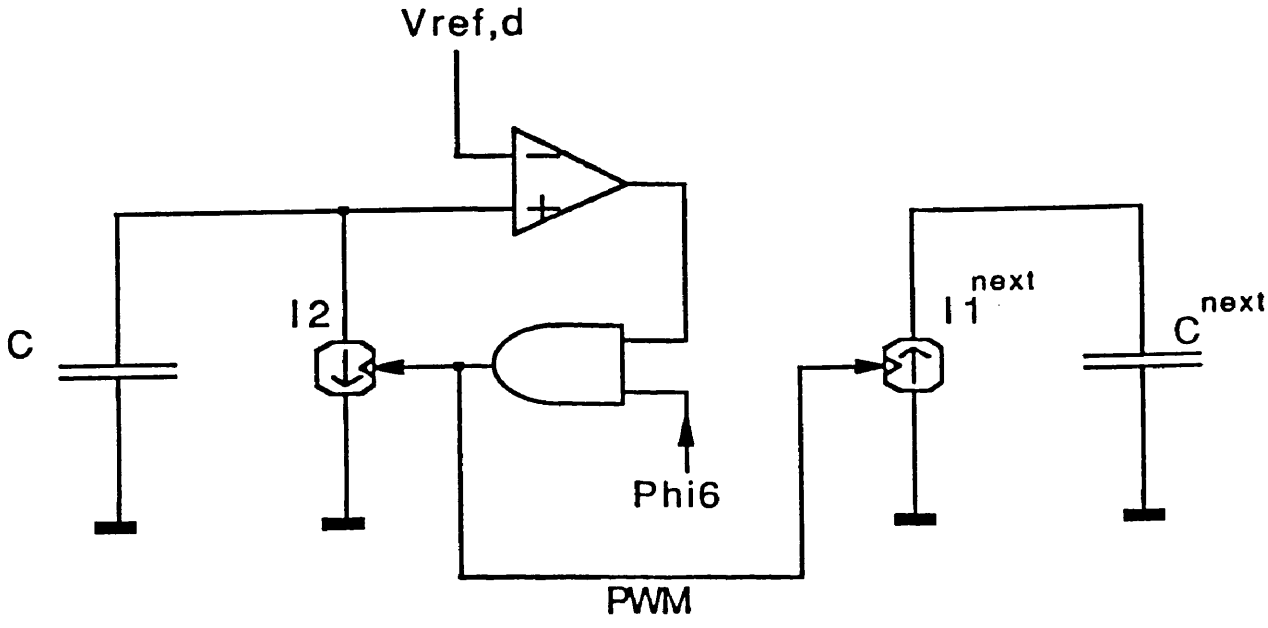


Fig. 113: Realization of the Charge Gain

These equations show that by choosing the currents I_1 and I_2 of two consecutive stages, any gain can be realized, positive as well as negative. By choosing $V_{ref,i}$ and $V_{ref,d}$, the input-referred offset of the system can independently be set to any desired value. These equations only apply if offset compensation is built in (Φ_2).

The charge amplification is actually realized using a current-mirroring scheme. The two current sources form a current mirror. However, the change of variable is somewhat unusual. The two sources are coupled by a common pulse duration rather than by a common voltage, like in more conventional mirrors (figure 114). The currents can be realized in any fixed ratio, by duplicating a number of unity current sources, or they can be made adjustable, by using at least one current source that is controlled through an external terminal.

The latter is a fairly unique feature of the charge amplifier, which cannot be realized easily in switched-capacitor technology, where the gain is set by a *fixed* ratio of two capacitors. This feature may prove crucial for the realization of high-speed, tunable, self-calibrating or even programmable discrete-time analog filters, which are not common at the time of this writing.

It may seem difficult to realize a fixed ratio of currents in two current sources that are part of a charge amplifier scheme, even when using nominally identical components. In the application we had in mind, this was not a significant disadvantage, since digital error correction was applied to the converter that uses the charge amplifier stages. The associated Accuracy Bootstrapping algorithm is insensitive to fairly large gain errors. Had this not been the case, relatively good matching could still be obtained by using large-area transistors in the current sources [88], although this might cause speed degradation. Another alternative could be to periodically recalibrate the current sources with respect to a common reference [90, 64, 91].

10.10. Transient Waveforms

The equations describing gain and offset of the charge amplifier have been derived through a simplified model. They rely on the assumption that the PWM output of the comparator can be modeled as a digital signal that is either perfectly ON or perfectly OFF. All current sources controlled by the comparator were

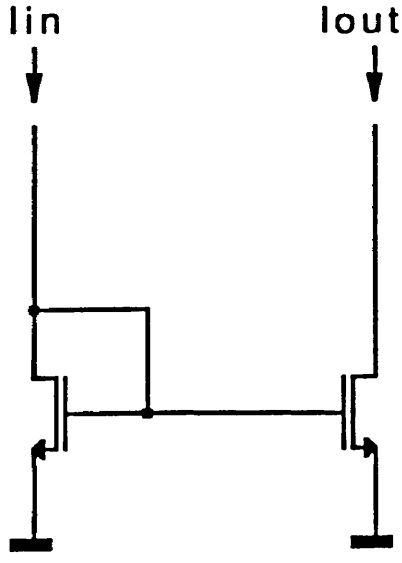


Fig. 114: Conventional Current Mirror

assumed to switch infinitely fast. As a result, the pulse width T was uniquely defined.

Obviously, this is not a realistic situation. Every time a current source is switched ON or OFF, the current exhibits some transient behavior, which has not been taken into account so far. However, it will be shown that this effect, if given some attention, will not affect the gain or the linearity of the scheme. At most, it may slightly change the offset, which for many applications is not critical anyway.

There are two main current sources per stage. I_1 is used to apply the input charge (the signal) during Φ_3. I_2 is used to discharge the capacitor to some reference level, either $V_{ref,i}$ (initial discharge, Φ_2) or $V_{ref,d}$ (final discharge, Φ_6). In analog/digital converter applications, auxiliary current sources may be used to subtract fixed amounts of charge during the DAC phase (Φ_5). Those current sources will not be explicitly discussed here, though similar considerations apply to them.

The comparator's output does not exhibit infinitely sharp transitions. Even if it did, the current sources that are controlled by that signal still could not react infinitely fast. Both effects together make that the current will exhibit a certain transient behavior when switched ON.

The duration of that transient will be called the *lead time*, t_{lead} (figure 115). The precise duration of t_{lead} has to be defined conventionally, since the current may settle exponentially to its steady-state value. We will define t_{lead} as the time between the instant at which the clock signal associated with the particular phase becomes active and the instant at which the current reaches its desired steady-state value, within a certain accuracy. This accuracy is such that the charge error associated with assuming that the rest of the waveform represents a constant current, would be negligible compared to the full signal range. The range is defined as the difference between the maximum and minimum possible signal charge. "Negligible" is defined as small enough to obtain a specified linearity or dynamic range within the system.

Associated with t_{lead} is a certain charge which is the time integral of the switch-ON transient current. We will call this charge the *lead charge*, Q_{lead}.

The current exhibits similar transient behavior when it is switched OFF. This

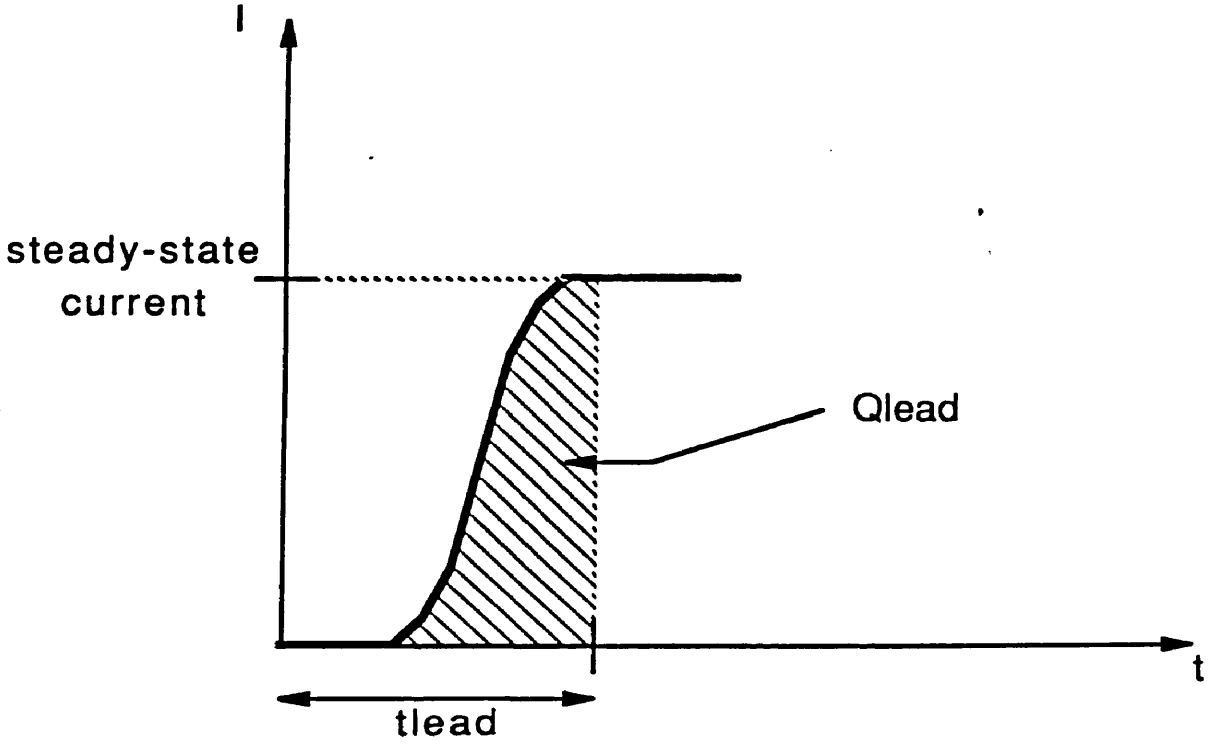


Fig. 115: OFF-ON Current Transient

occurs when the output of the controlling comparator flips from logic high to logic low, after a capacitor voltage reaches a pre-set reference level. We will define the *trail time*, t_{trail} as the time between the instant at which the capacitor voltage reaches the comparator's reference level and the instant at which the current has decayed "sufficiently" (again as compared to the signal range) close to zero (figure 116).

Associated with t_{trail} is a certain charge which is the time integral of the switch-off transient current. We will call this charge the *trail charge*, Q_{trail}. From its very definition, it is clear that the trail charge associated with current source I_2 is the same as the undershoot charge Q_u discussed earlier.

The presence of a lead and trail charge has two major implications. First, the offset of the charge amplifier may not have exactly the value predicted by equation (167). Secondly, the clocking scheme, as well as the offsets in the system, have to be dimensioned carefully, in order to guarantee that for *any* input signal, all currents always get enough time to rise to their steady-state value, *and* decay back to zero. Failure to do so would introduce severe non-linearity into the scheme.

The effect of switching transients may not be very significant in low-frequency or low-accuracy applications. In low-frequency charge stages, fairly small current values can be chosen, since a lot of time is available to charge or discharge the capacitors. In that case, the lead and trail times may be small compared to one clock phase, and the associated lead and trail charges may be relatively insignificant compared to the total useful signal range. However, in high-speed applications, the duration of one clock phase has to be minimized to the extreme (possibly down to a few tens of nanoseconds). In such cases, lead and trail times can take up to one half of a total clock phase. If proper care is taken to accommodate their effect, very accurate performance (up to about 14 bits linearity) is possible, at clock rates that can reach several tens of MHz.

10.11. Repeatability of the Switching

The effect of switching transients and their associated lead and trail charges can be accommodated. However, the method relies on the assumption that these effects are repeatable. In other words: every time a current source is turned ON or OFF, the transient current waveform must be exactly the same.

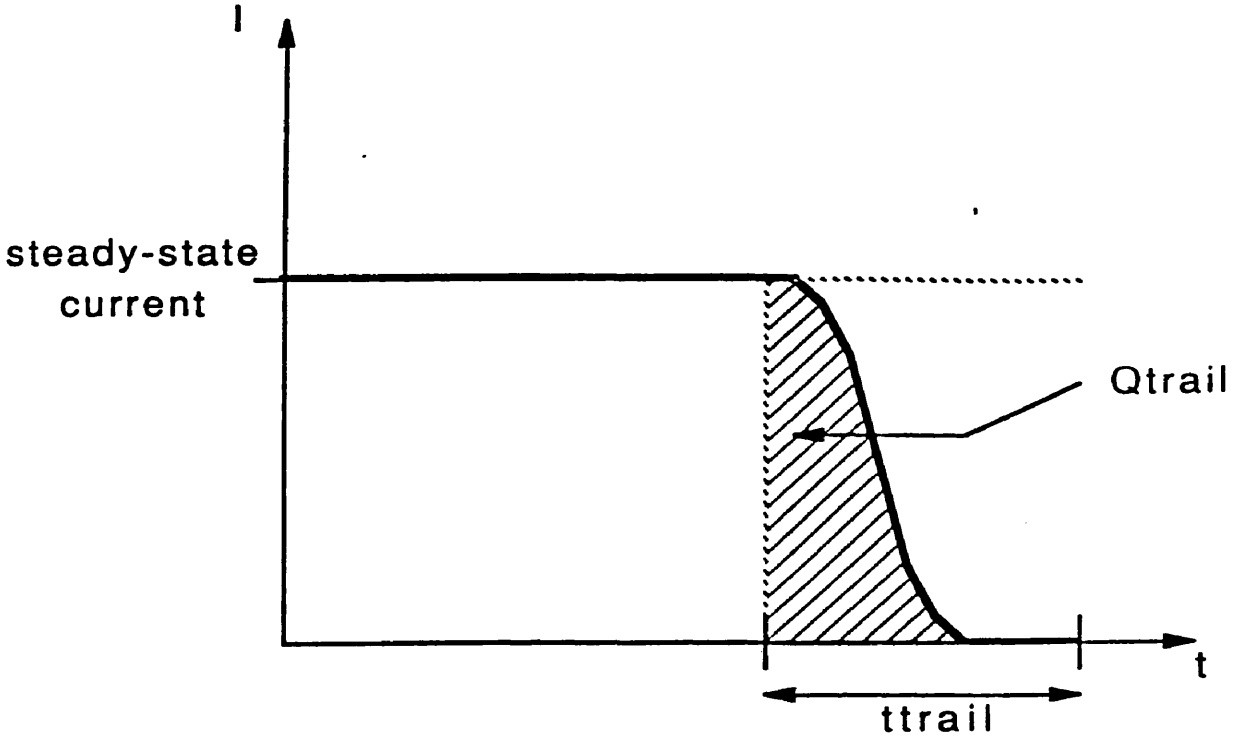


Fig. 116: ON-OFF Current Transient

Fortunately, this assumption is justified. Every current source is turned ON at the beginning of a clock phase and turned OFF when a specific comparator detects that a capacitor voltage falls below its pre-set reference value.

This condition can easily be realized by combining the clock signal and the digital comparator output through an AND gate (figure 117). Before the clock phase starts, the comparator output is logically high, while the clock signal is low. The AND gate output is low and all current sources that are connected to it are OFF. When the clock signal goes high, the AND output goes high as well, and the current sources are turned ON. As a result, the OFF/ON transients of the current sources *are* repeatable, since the clock signal is deterministic and the comparator output remains high during the whole transition.

When the capacitor is discharged to the desired reference level, the comparator output switches to a low level. The clock signal remains high during this operation. The output of the AND gate goes low and the current sources are turned OFF. The conditions under which this happens are perfectly repeatable. The reference voltage of the comparator is held constant, while the capacitor voltage decreases according to the formula

$$Q(t) = Q_{init} - \int_0^t I_2(t)\ dt \tag{169}$$

With the relationship between V and Q given by

$$Q(V) = \int_0^V C(v)\ dv \tag{170}$$

Before the capacitor voltage reaches the comparator reference level, I_2 is approximately constant and equal to its *steady-state* value (we assume that the current source has been on for a time longer than t_{lead}). The discharge of the capacitor is approximately linear, with $dV/dt \approx I_2/C$. All higher derivatives of V, if not entirely zero, are deterministic as well, since I_2 is constant. Hence, the conditions at the comparator input, just before switching, are perfectly deterministic.

During the switching (when the capacitor voltage exceeds the reference level), the clock signal normally remains high and the AND gate merely acts as a buffer.

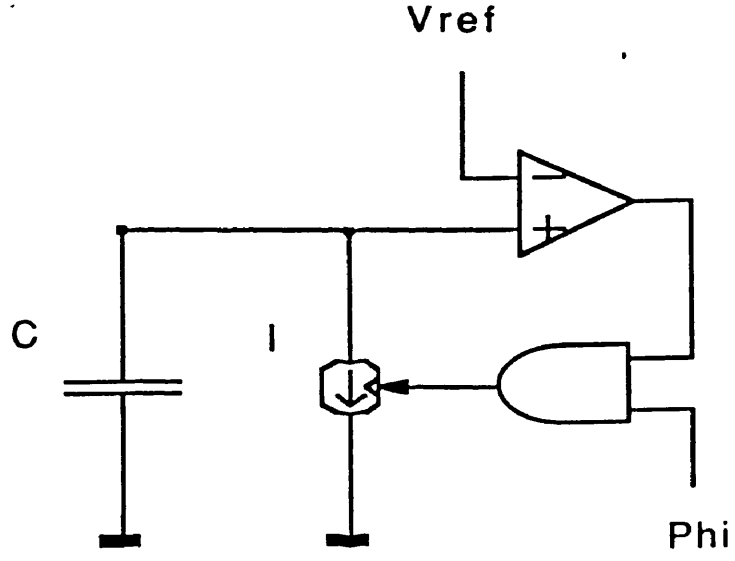


Fig. 117: Combination of Comparator Output and Clock Pulse

The combination of capacitor, comparator, AND gate and current source I_2 forms a system with feed-back. The capacitor charge (and voltage) eventually drops below the reference level and the current is switched OFF. This system works in an entirely deterministic way, and the final undershoot charge Q_u is fixed.

The feed-back loop has a certain parasitic delay associated with it. At the same time, the current is *very* sensitive to the comparator input voltage (ideally an infinitesimal voltage variation turns the current either ON or OFF). These conditions would normally make the system subject to strong instability. If a frequency domain analysis of the feed-back loop were performed, poles would be found well in the right-hand complex plane. However, this is of no concern here, since the system is strongly non-linear and one-directional. Eventually, the situation will evolve to an equilibrium state with zero current and a fixed capacitor voltage. Actually, the more unstable the system is, the faster it will evolve to its final state.

The only restriction, in order to guarantee repeatability of the ON/OFF transition, is that the clock signal remain at a steady high level for the whole duration of the transition. In practice, this means that the clock signal must not go low before at least t_{trail} *after* the capacitor voltage first reaches the reference level.

It has to be noted that asymmetric delay times in the AND gate do not affect the operation. The delay time associated with an OFF/ON transition is essentially irrelevant as far as the total displaced amount of charge is concerned. The delay associated with the ON/OFF transition will affect the trail charge (or undershoot), but this effect can be neutralized by the offset compensation scheme. In some cases, the presence of the AND gate can help to shape the PWM signal, since the transition at the comparator output may be fairly slow. The AND gate then makes the transition sharper.

To summarize, one could say that any time a current source is turned ON, the transition occurs in a deterministic way. Any time a current source is turned OFF, it occurs in a deterministic way as well, provided the current first had a chance to reach its steady-state value. The latter condition means that a current source *must not* be switched OFF unless it was first ON for at least t_{lead}. At the same time, the clock signal *must not* go low earlier than t_{trail} after the reference

level is exceeded.

This discussion assumed that the current transients are *only* dependent on the switching characteristics of the controlling signal, and not on the actual voltage or slew rate (dV/dt) at the output of the current source. For the current sources that will be described below, this assumption is justified to a great extent.

10.12. Accommodation of Lead and Trail Times

It is relatively easy to guarantee enough time for each current source to be turned OFF. It is sufficient to dimension the duration of Φ_6 so that the total available time would exceed the sum of the ideal time needed to remove the maximum possible charge on the capacitor, plus the lead time of current source I_2 (due to the delay associated with turning I_2 on), plus the trail time of I_2.

It is slightly more involved to accommodate the lead time of a current source under *any* conditions. It is clear that the PWM signal that controls the current sources cannot be allowed to get a pulse width of less than t_{lead}. This implies that a systematic offset will have to be built into the system, in such a way as to guarantee a certain minimum pulse width, and hence a minimum signal charge greater than Q_{lead}.

A formula was derived earlier (equation (168)) for the input-referred offset of a charge amplifier stage.

$$Q_{off}^{in} = \pm Q_{26} = \pm[Q_i - Q_d] = \pm[Q(V_{ref,i}) - Q_{ref,d}] \qquad (171)$$

This offset can be adjusted to any value by a proper choice of $V_{ref,i}$ and $V_{ref,d}$. In order to guarantee a certain minimum pulse width at the comparator output of one stage, it is sufficient to choose $V_{ref,i}$ and $V_{ref,d}$ so that Q_{off}^{in} would be at least equal to Q_{lead}. However, this systematic offset gets amplified and generates an output-referred offset Q_{off}^{in} (I_1^{next}/I_2), which is equivalent to an additional input-referred offset in the next stage. This offset must then be canceled, e.g. by subtracting a fixed charge from the capacitor of the next stage, or by correspondingly increasing or decreasing the desired systematic offset of the next stage.

10.13. Mismatches in Transient Currents

It has been shown that any time a current source is switched ON or OFF, a parasitic lead or trail charge will be associated with the transient current. This charge can be accommodated as outlined above, by building some extra systematic offset into each stage.

However, while the gain of the resulting system will not change (the gain is determined only by the ratio of steady-state currents), the random offset of the system may become harder to predict. This is due to the fact that the transients in the current sources of one stage may not exactly be matched by the transients in the next stage. Fortunately, this effect can be minimized to a certain extent by using identical current sources in two consecutive stages. The assumption is that identical current sources, controlled by the same gate signal, will have very similar transient behavior, even if the voltages at their output are not identical. The gain of the charge amplifier can still be set to any value by choosing an appropriate integer ratio of currents.

This strategy leads to sufficiently small random offset levels for many applications (a fraction of a percent of the full signal range). Further offset reduction would require some external adjustment.

10.14. Multiplexing Comparator References

In order to cancel random offsets (correlated double sampling), a different reference voltage has to be applied to the comparator during Φ_2 (initial discharge) and Φ_6 (final discharge). A similar situation arises during the generation of the DAC charges. Switching from one reference level to another can be accomplished in several ways.

In CMOS technology, the comparator reference input can easily be multiplexed between two fixed reference voltages. An analog multiplexer can be made out of two MOS transmission gates, controlled by complementary control signals. The transmission gates can consist of either a single MOSFET, or a parallel combination of an N-channel and a P-channel MOSFET, depending on the range of the required signals (figure 118).

Either way, when the transmission gate (switch) is ON, it forms a low resistance path between the reference voltage and the reference input of the comparator, which normally is the gate of another MOSFET. Since the comparator does

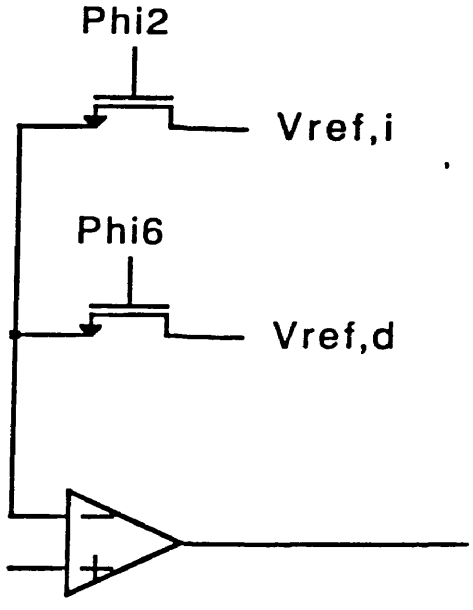


Fig. 118: Multiplexing of Comparator References (CMOS)

not draw any DC input current, there is no voltage drop across the switch. Since the input capacitance of the comparator is small and the ON resistance of the switch is low, settling will be extremely fast. As a result, the reference voltage is passed nearly perfectly by the switch and no additional error is introduced.

In bipolar technology, such ideal switches are not available. However, other ways exist to change the reference level of the comparator between two fixed values. As an example, one could think of two current sources connected to a resistor (figure 119). Depending on which current source is turned on, a different voltage drop can be generated across the resistor.

10.15. Summing Nodes

The charge amplifier scheme lends itself well to applications where several signals have to be summed together. An input signal is applied to the capacitor of a stage through a current source. A summing operation can be realized by simply connecting several current sources (one for each input) to the same capacitor. The only difficulty is to properly match all offset charges (accommodation of respective Q_{lead}'s), so that they would cancel each other out at the summing node.

The charge amplifier could be used to realize complex discrete-time analog filters, with multiple feed-back loops. Currently, such filters are usually built using switched-capacitor stages. An implementation with charge amplifier stages would be slightly different however. In a switched-capacitor filter, a combination of MOSFET switches and capacitors are normally used to emulate resistors between low-impedance nodes, thus turning an opamp with feedback capacitor into a discrete-time integrator. The charge amplifier implements a fixed transfer function of the form $A\ z^{-1}$ in the Z domain (delay element with gain). Blocks with such transfer function can be combined into higher-order filters, in a way similar to fully digital filters rather than traditional switched-capacitor filters. An approach in which a charge amplifier stage would be used as an integrator may be feasible as well (the capacitor inherently integrates current), but this possibility has not yet been fully investigated.

10.16. Input and Output Stages

The charge amplifier scheme has the disadvantage that the signal variable is a

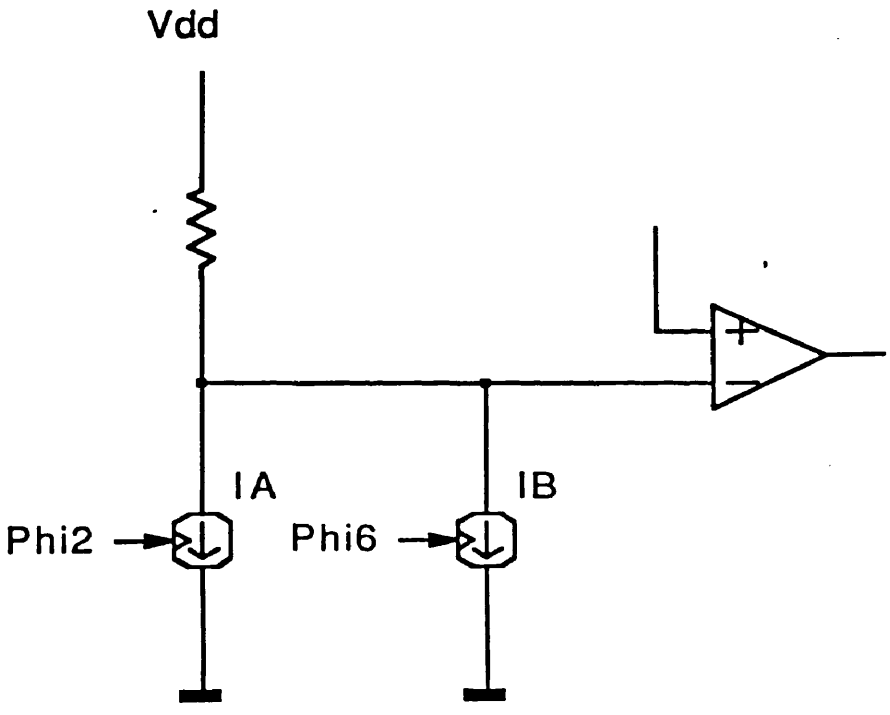


Fig. 119: Multiplexing of Comparator References (Bipolar)

charge (or a pulse width) rather than a voltage, like is the case in most conventional applications. While this is not a limitation *within* a particular system (like a complex filter or analog/digital converter), it creates a problem of compatibility with other systems.

Some applications use PWM coded signals (e.g. certain speech products), in which case the interface to a charge-based system is readily made. The PWM signal can be used to directly control a current source, like is being done inside each charge stage anyway. In a similar way the comparator output of the last stage of the charge-based system can be buffered and used as a PWM output.

Direct application of the charge amplification principle is also possible in systems that have a charge as input stimulus. One particular example is the sensing of photodiodes or photodiode arrays, like is being performed more and more in optical scanners for computers and FAX machines. In such systems, a very small electrical charge (a few pico-coulombs) is generated in a diode through illumination. Such charge can very efficiently be scanned and converted to a PWM signal through a comparator / current source combination.

Direct use of a charge amplifier could also be feasible to scan the output of CCD's (Charge Coupled Devices). This application could potentially be valuable for processing or digitizing video signals from cameras, which very often use CCD image sensors.

However, if the charge-based system has to be interfaced with traditional voltage-based systems, a special input and output stage are necessary. Such stages can easily be realized using conventional switched-capacitor or sample/hold techniques.

One simple input stage uses a unity gain follower and three switches, in a fashion identical to a stray-insensitive switched-capacitor stage [67, 68] (figure 120). The difference is that the main amplifier is replaced by a current source/comparator combination. This scheme is subject to non-linearity of the input follower, non-linearity of the capacitor and clock feed-through. Another scheme is a variation on a popular sample/hold configuration, with reduced clock feed-through effects [92, 93, 94] (figure 121).

An output stage can easily be realized by connecting a switch and a high-

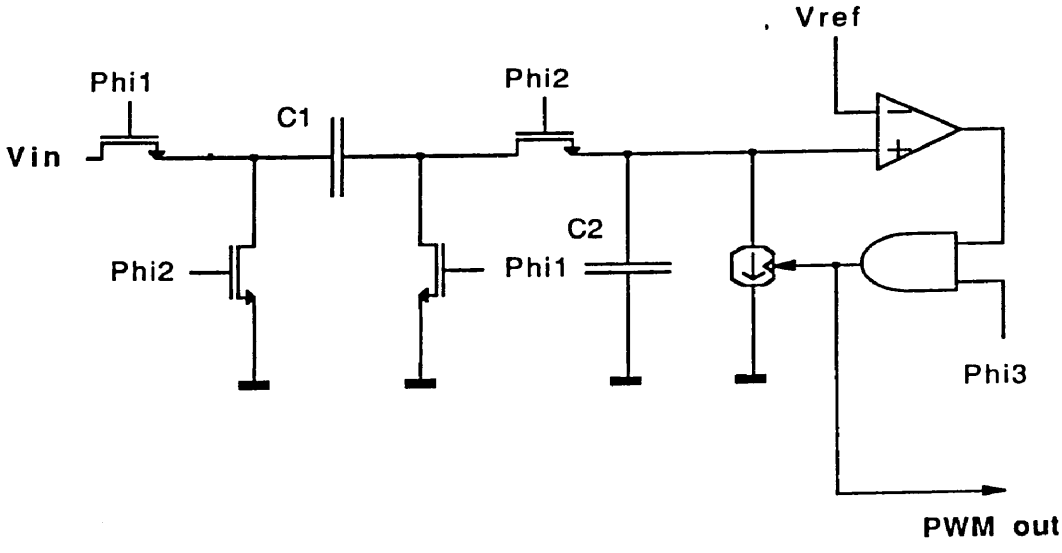


Fig. 120: Input Stage

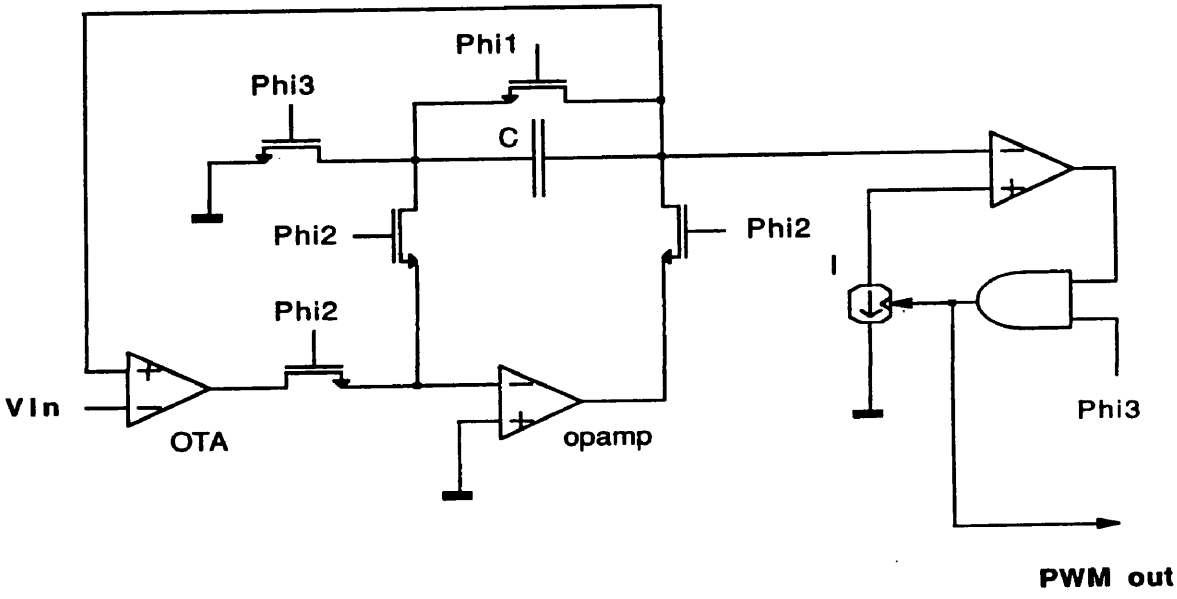


Fig. 121: Alternative Input Stage

quality opamp to the capacitor of a charge amplifier, again in a way similar to a switched-capacitor scheme (figure 122).

The fact that switched-capacitor techniques have to be applied for the input or output stage is not a major disadvantage, since these building blocks are normally only used once in the whole system. Minor gain errors can also be tolerated fairly easily at the input or output. As a contrast, a complex system (e.g. a high-order filter or a high-accuracy pipelined A/D converter), may require more than ten *internal* amplifiers, which do have to be very accurately matched.

10.17. Linearity

The linearity of the charge amplification process is limited by several effects, but *not* by the linearity of the capacitors. The first important limitation is the output impedance of the current sources, which has to be as high as possible, since a high output expresses the fact that the output current does not change significantly when the output voltage changes over a wide range.

$$Z_{out} = \frac{\delta V_{out}}{\delta I_{out}} \quad \text{or} \quad dI_{out} = \frac{dV_{out}}{Z_{out}} \approx 0 \tag{172}$$

As an example, if the useful signal range of a stage is 2V, the current I_2 is $20\mu A$ and the output impedance is $1G\Omega$, the maximum variation in current will be

$$dI_{out} = \frac{dV_{out}}{Z_{out}} = \frac{2V}{1G\Omega} = 2nA \quad \text{or} \quad 0.1\% \tag{173}$$

This current variation can be used to calculate the worst-case non-linearity due to finite output impedance. Depending on whether the configuration is inverting or non-inverting, the finite output impedance of current sources in the next stage may partially cancel, or worsen the effect of finite output impedance in the given stage. The inverting configuration is more favorable, since the currents always flow in the same direction.

So far, only DC output impedance has been considered. However, it is clear that under rapidly switching conditions (MHz rates), this output impedance cannot be considered constant, or even resistive anymore. A frequency domain simulation of the output impedance of the regulated cascode current source (to

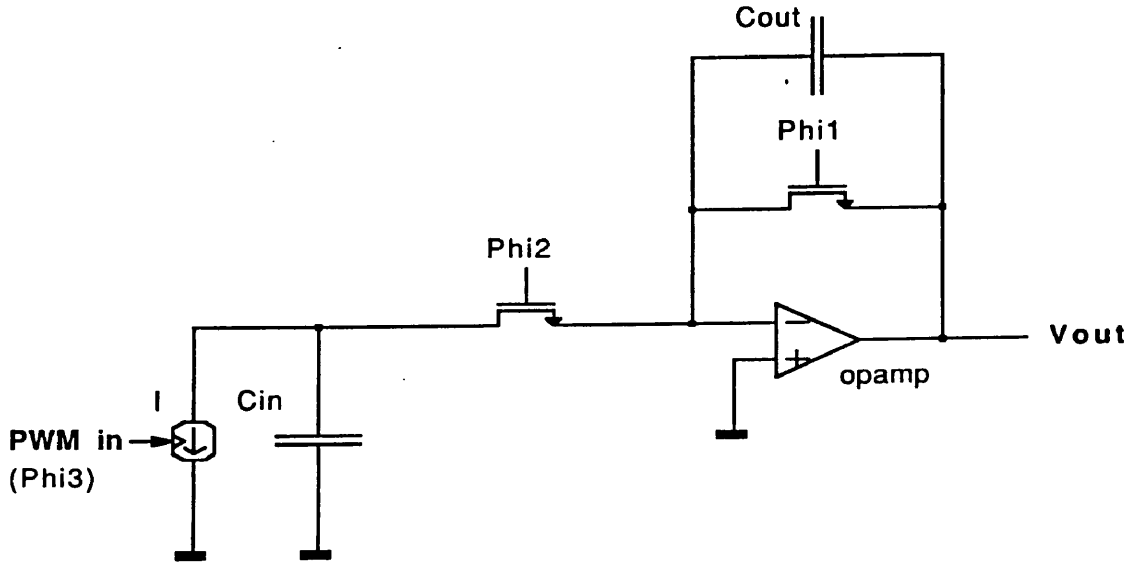


Fig. 122: Output Stage

be discussed below), reveals that the output impedance has a dominant pole at about $1kHz$, which at first seems to be a prohibitively low frequency for practical applications.

However, this pole is mainly created by the parasitic drain-gate capacitance of the output transistor. This capacitance, though small (a few fF), creates a large time constant when combined with the large output impedance of the system. In practice, its effect is very limited, since it is essentially in parallel with the main capacitor.

The main non-linearity contribution of this drain-gate capacitor occurs when the current source is switched off, since this implies that the gate of the output transistor is abruptly shorted to ground, like will be described below. Depending on what voltage was initially present at the gate, the amount of charge injection onto the main capacitor will be different, thus creating a potential gain error and/or non-linearity. However, for the regulated cascode current source (to be described), the signal-dependent voltage variation at the gate is at least 100 times smaller than the variation at the output. In addition, the drain-gate capacitance is 100 to 1000 times smaller than the main capacitance, making this effect really negligible up to the 14-16 bit level.

Severe non-linearity can also arise if the current sources of the system are not given enough time to either reach their steady-state current, or decay back to zero. This effect can be avoided altogether if proper care is taken to accommodate the lead and trail times as described above.

Another effect that could potentially contribute to non-linearity is parasitic source-bulk leakage current in the output transistor. This leakage is potentially signal-dependent. Fortunately, the effect is fairly limited (nano-amp range), especially when the output transistor operates in saturation. If a well-controlled IC technology is used, non-linearity due to this effect normally will not show up before the 13-15 bit level.

10.18. Implementation of the Current Sources

Precision current sources with extremely high output impedance can easily be implemented both in CMOS and in bipolar technologies, using just a few transistors. One CMOS alternative is the triple cascode configuration (figure

123). Another scheme, with similar output impedance but slightly better voltage range, is the regulated cascode, as described in [7] (figure 124). The regulated cascode also offers a lot of flexibility as far as switching is concerned.

The main requirement for the current sources, next to a high output impedance, is the possibility to be switched ON (fixed, steady-state current value) or OFF (zero current) in a very short time. This can be achieved very elegantly with the regulated cascode configuration.

The DC current of a regulated cascode is determined by the gate voltage of M_2 (figure 124). The current is approximately equal to the reference current, which is copied by current mirror M_1, M_2. The drain of M_2 is kept at almost constant potential (regulated) through the feedback loop formed by M_4, M_5 and M_3. As a result, the output current through M_2 and M_3 becomes highly constant. The output impedance of the source (at the drain of M_3) is of the order of g_m^2/g_o^3, typically in the $G\Omega$ range.

It is possible to switch the current ON or OFF using the gate of transistor M_2, by controlling the reference current or by switching the gate of M_2 between its normal value and ground. However, this creates some practical problems, like switching noise injection into the reference current node and slow ON/OFF transition, due to reduced transconductance in the system while the current levels decay.

A much better switching strategy is to use the gate of M_3, which is normally controlled by the negative feedback amplifier M_4, M_5. The whole current source can very effectively be switched OFF by shorting this gate to ground, using an additional N-channel MOSFET switch (M_7). When this switch is turned ON, it shorts M_4 and forces the gate of M_3 to ground almost immediately. As soon as this happens, the current at the output of the system disappears entirely, and the source of M_3 is quickly pulled to ground by M_2.

When M_7 is turned OFF again, the gate of M_3 is pulled up by M_6, which operates in saturation (constant current). Current through M_3 is re-established very fast, and this current pulls up the drain of M_2 to its previous equilibrium value (source follower). Once the drain potential of M_2 has sufficiently neared its final value, the whole situation approximately turns into a linear settling process.

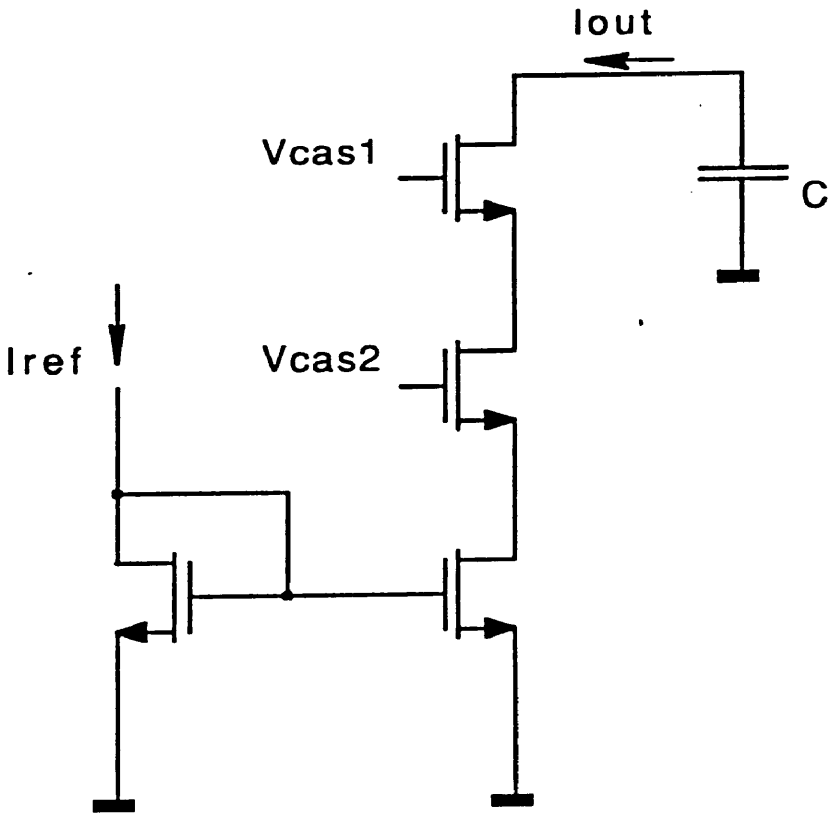
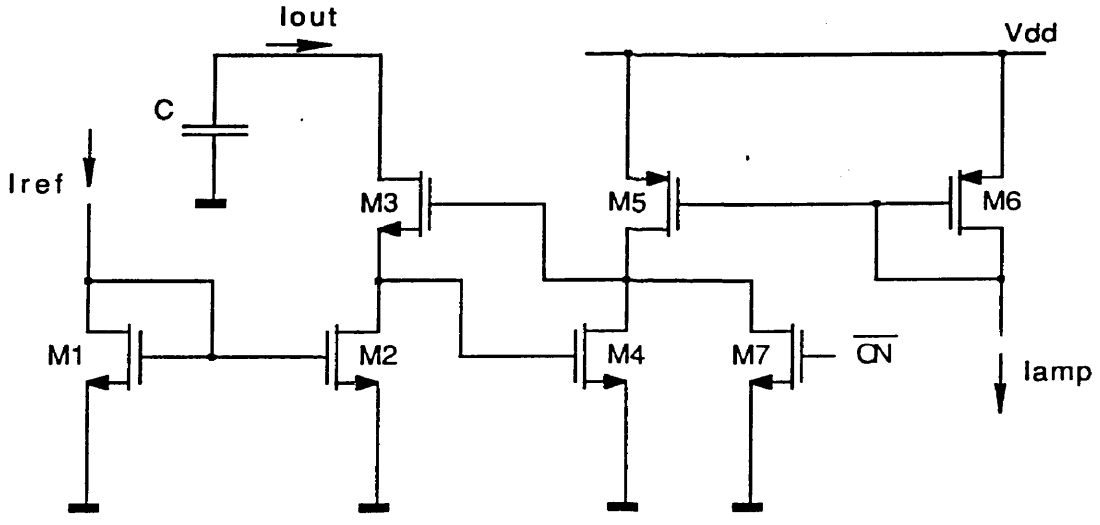


Fig. 123: Triple Cascode Current Source



(Note: All Transistors are 4 x 2 microns)

Fig. 124: Regulated Cascode Current Source

In order to get a good understanding of the speed of the system, the time constants of this settling must be estimated.

This kind of considerations cannot very well be simulated using a transient analysis on a computer (e.g. SPICE), since we are concerned about very high accuracy levels, and SPICE does not easily provide less than a few percent relative accuracy on this kind of simulations. An analytical expression for the transient behavior is not easy to obtain either, since this is a high-order system. However, a frequency domain analysis with SPICE, for the particular situation depicted in figure 124, does provide all necessary information to understand the switching behavior.

In practice, the current source is turned ON and OFF through the gate of M_3. When this gate is forced to ground, all transistors get cut off almost immediately and the system switches OFF extremely fast. The speed is determined by the time constant of the switch ($R_{on} \approx 1k\Omega$) and the internal parasitic capacitances ($10\ fF$ range), which is a fraction of a nanosecond. Clearly, shutting OFF the current will not be the limiting factor as far as speed is concerned.

However, when the gate of M_3 is released, the whole current source will re-activate itself through its internal currents (primarily the current provided by M_5, then gradually current coming out of the load capacitor). Once all transistors come out of their previous cut-off state again, the system will exponentially settle to its equilibrium value. This settling will determine the speed of the OFF-ON transition, and for this reason it is important to estimate all time constants (poles and zeros) of the system, around its normal operating point.

Doing a frequency domain analysis in which an input voltage is applied to the gate of M_3 would not make sense, since this would break the feed-back loop. Instead, one could apply an external *current* to the gate of M_3. However, this would only have a minor effect on the output current. A better understanding of the settling behavior (after switch M_7 is abruptly opened) can be obtained by analyzing how the system reacts to a variation in the reference current (current through M_1).

Although this does not reflect the *actual* switching operation, it accurately describes the settling behavior. When the switch is opened, M_2 is suddenly pulled

out of cut-off (by M_5 and M_3), which is functionally equivalent to what happens when a sudden change occurs in the reference current of current mirror M_1, M_2.

Figure 125 shows the frequency response of the output current (current from the capacitor) for an ac variation in reference current. This response resembles a second order band-pass characteristic. There is a double pole at a frequency of about 100 Hz and another one at about 300 MHz. The low-frequency pole reflects the fact that the current source does not have an infinite output impedance or output range. If the reference current is changed, the output current (from the capacitor) will follow the change at first, but as the capacitor voltage decreases (or in the ideal case becomes negative), the current will eventually decay.

In order for the charge amplifier to operate reliably (with constant current), the time constant with which the current decays must be very long compared to the time the current is ON. This is quite obviously the case, since for the applications we are looking at, current sources never stay ON for longer than about 1 μs, while the time constant with which the current decays is in the range of 1 ms.

The second requirement is that the current would quickly settle to its desired value. This is expressed by the second time constant, corresponding to the double high frequency pole in the transfer curve, at about $300MHz$. The corresponding time constant is about .5 ns, which predicts very accurate settling of the output current in just a few nanoseconds. Figure 126 sketches the transient response of the capacitor current for a step-like variation in reference current.

Figure 127 shows the frequency response of the capacitor voltage (drain of M_3), to an ac variation in reference current. figure 128 shows the associated time domain response.

Figure 129 represents the frequency response of the gate of M_3. This voltage is important, because there is a parasitic capacitance between this node and the current source output. The capacitance is caused by gate-drain overlap in M_3 and is normally fairly small compared to C (a few hundreds of times smaller). However, it could potentially contribute to some (very minor) non-linearity, if the voltage were to follow the output (capacitor) voltage.

Such a situation would mean that the gate voltage of M_3 would be signal-

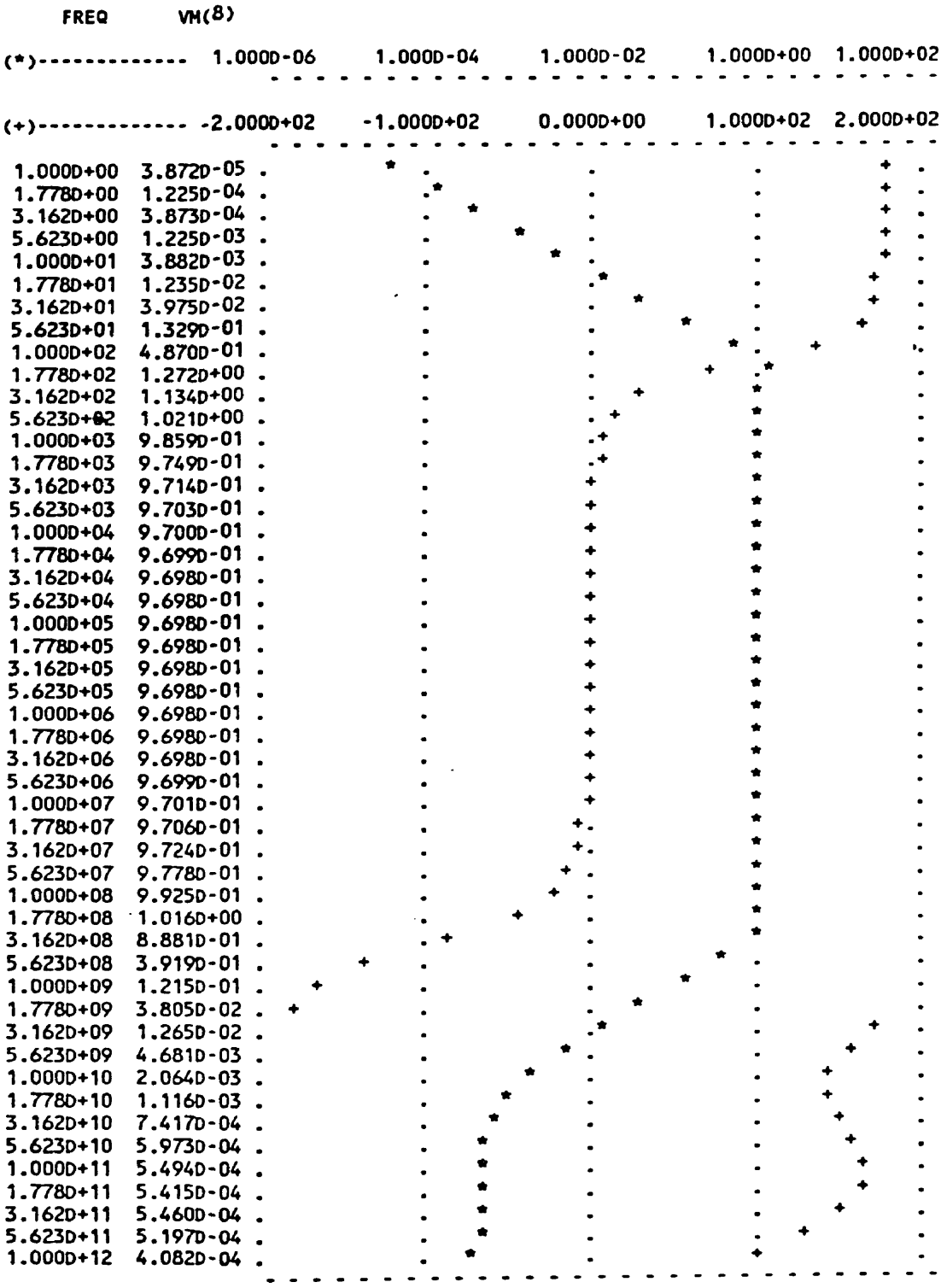


Fig. 125: Output Current Response (AC)

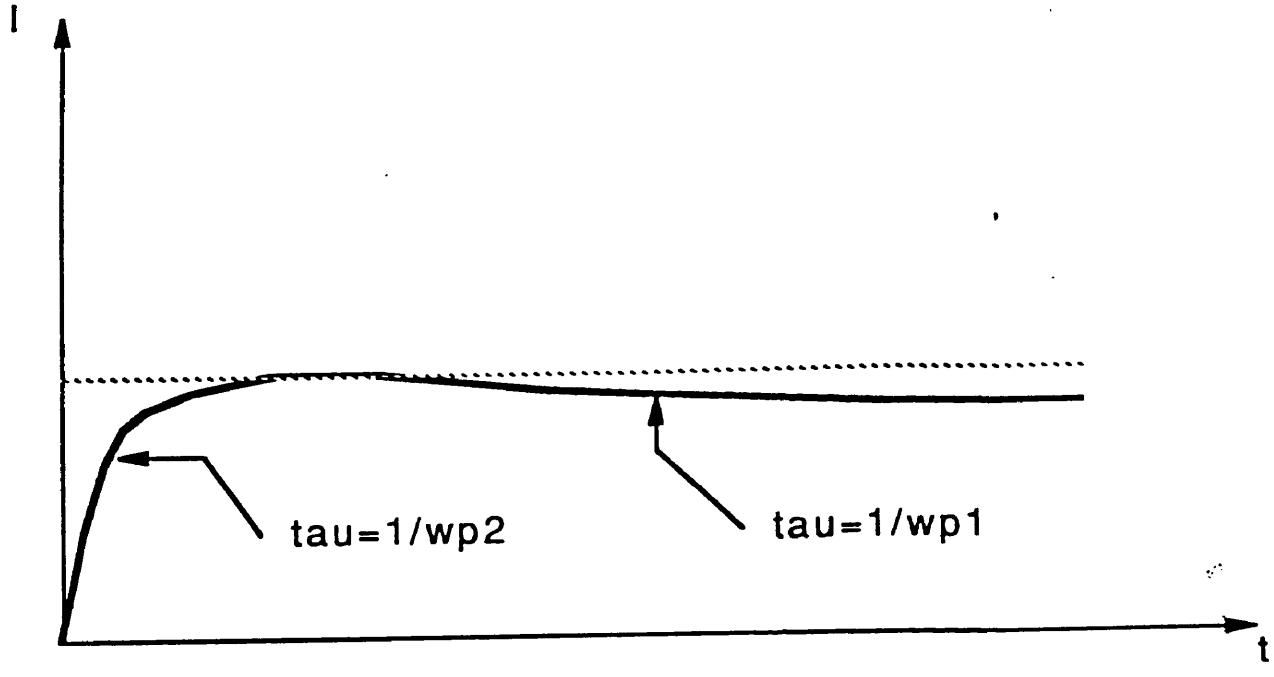


Fig. 126: Output Current Response (Step)

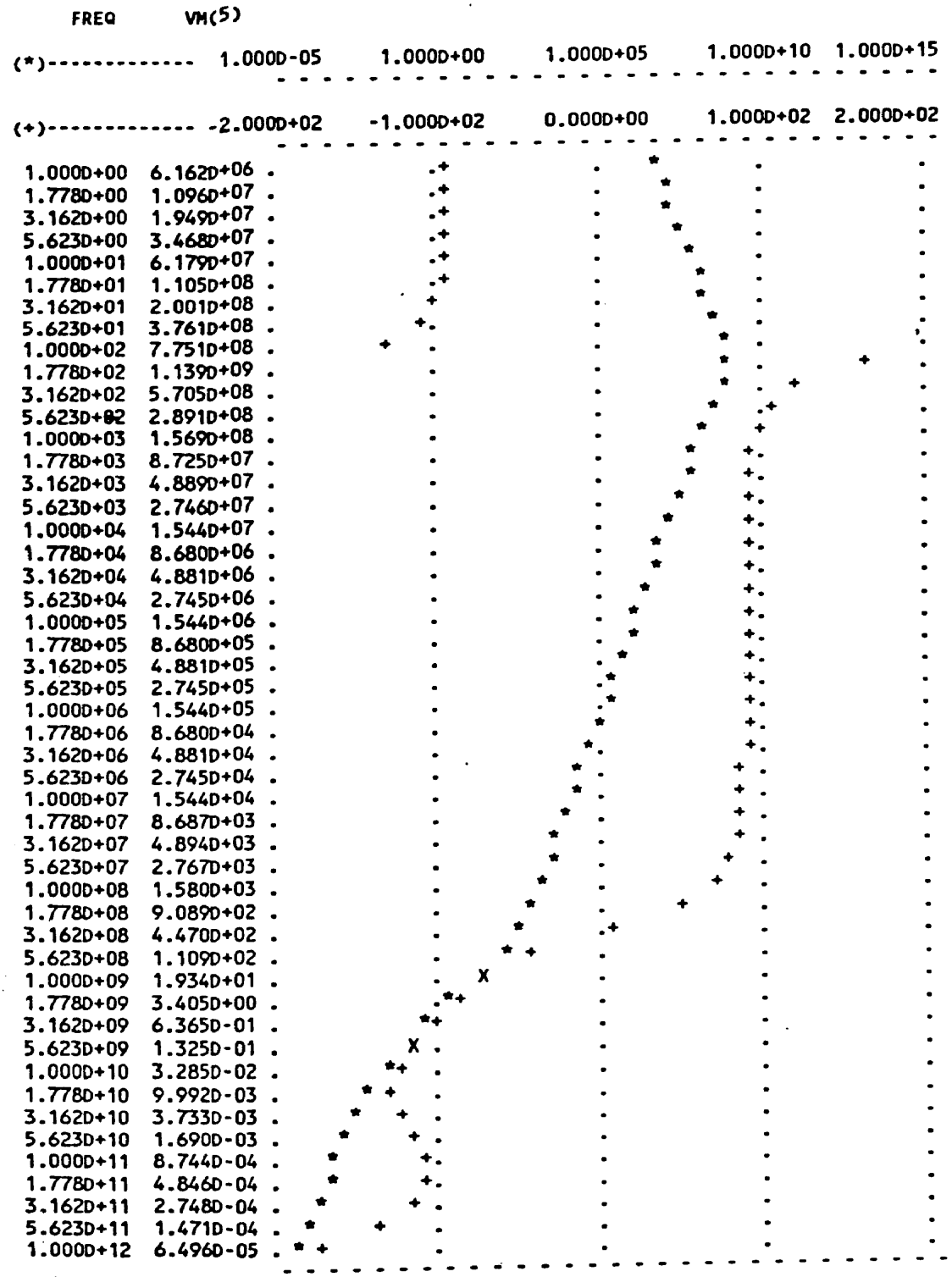


Fig. 127: Capacitor Voltage Response (AC)

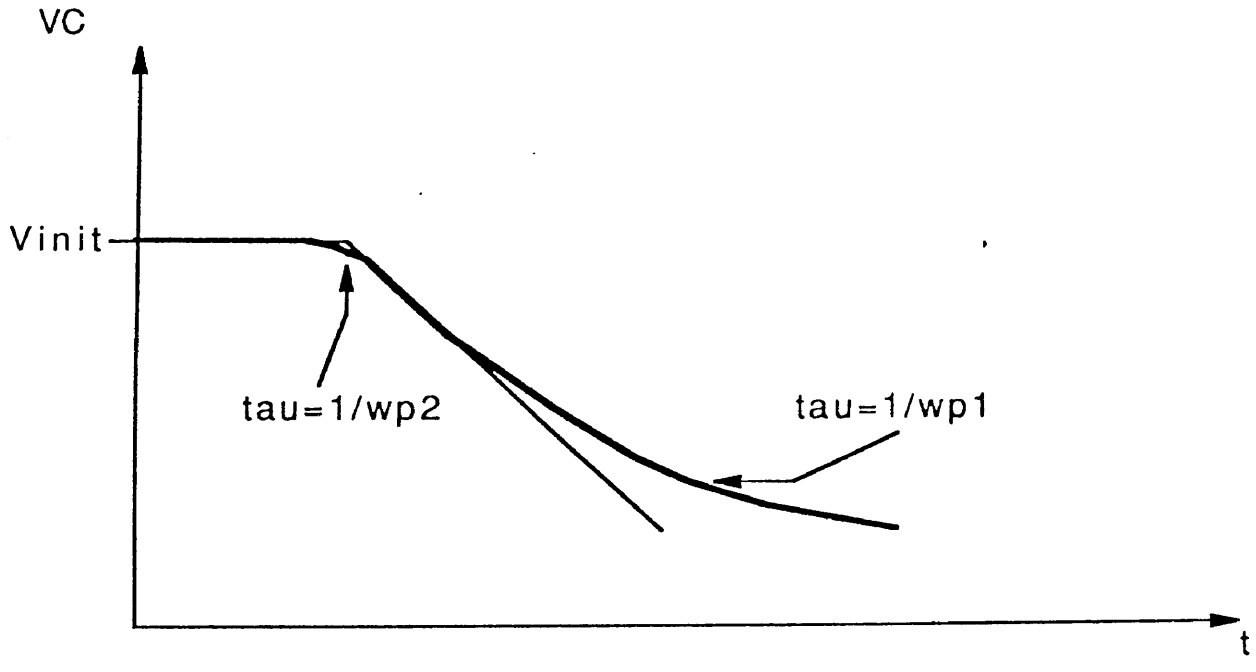


Fig. 128: Capacitor Voltage Response (Step)

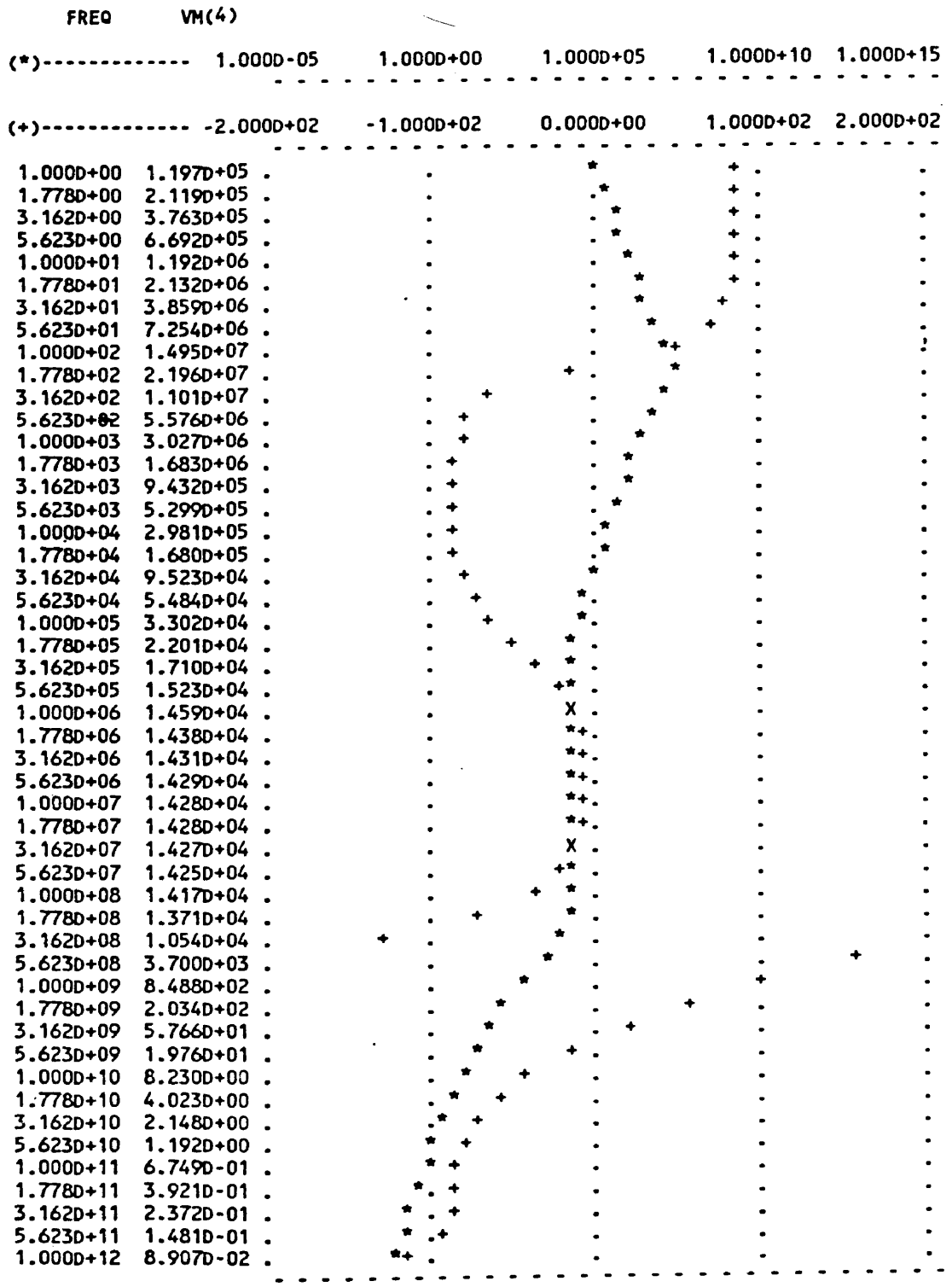


Fig. 129: Voltage Response, Gate of M_3 (AC)

dependent, and when M_3 is shorted in order to switch the current off, a signal-dependent charge would be injected into C. Rather fortunately, from the frequency domain analysis, the time domain response can be derived (figure 130), and this shows that the gate voltage settles to a steady value, mainly determined by the reference current. This means that the gate-drain overlap capacitance can be considered to be in parallel with the main capacitor, and no *signal-dependent* feed-through will occur when switch M_7 is closed again.

10.19. Channel Charge Redistribution

Since the main transistors of the current sources (M_2 and M_3) only operate in saturation, channel charge redistribution is not too much of a problem. When the current source is ON, a certain negative charge is present in order to form the (N-type) channel of these transistors. When the current source is switched OFF, this negative charge disappears into the signal capacitor of the charge stage, creating a slight drop in capacitor voltage. This drop is very small, since the channel charges are typically three orders of magnitude smaller than the maximum signal charge. In addition, any systematic component of this offset is effectively canceled by the correlated double sampling.

However, if the total channel charge were signal (output) dependent, it could possibly create some gain error or non-linearity around the 10-11 bit level. In practice, this will not be the case. The channel charge of M_2 is practically output-independent, since M_2 maintains a very constant drain potential. The drain of M_3 however, *is* the output, but since M_3 is always saturated (pinched off), the channel charge will still be very constant. The charge is approximately given by the following formula, which to the first order does not depend on the drain voltage

$$Q_{c,M3} \approx \frac{2}{3} C_{ox} W L V_{GS,M3} \tag{174}$$

Any variation in $V_{GS,M3}$ is typically at least 100 times smaller than the corresponding variation in drain voltage, due to the feedback.

10.20. Power Consumption

The charge amplifier scheme uses power extremely conservatively. The signal-carrying capacitors in the system are either charged by a single transistor, which

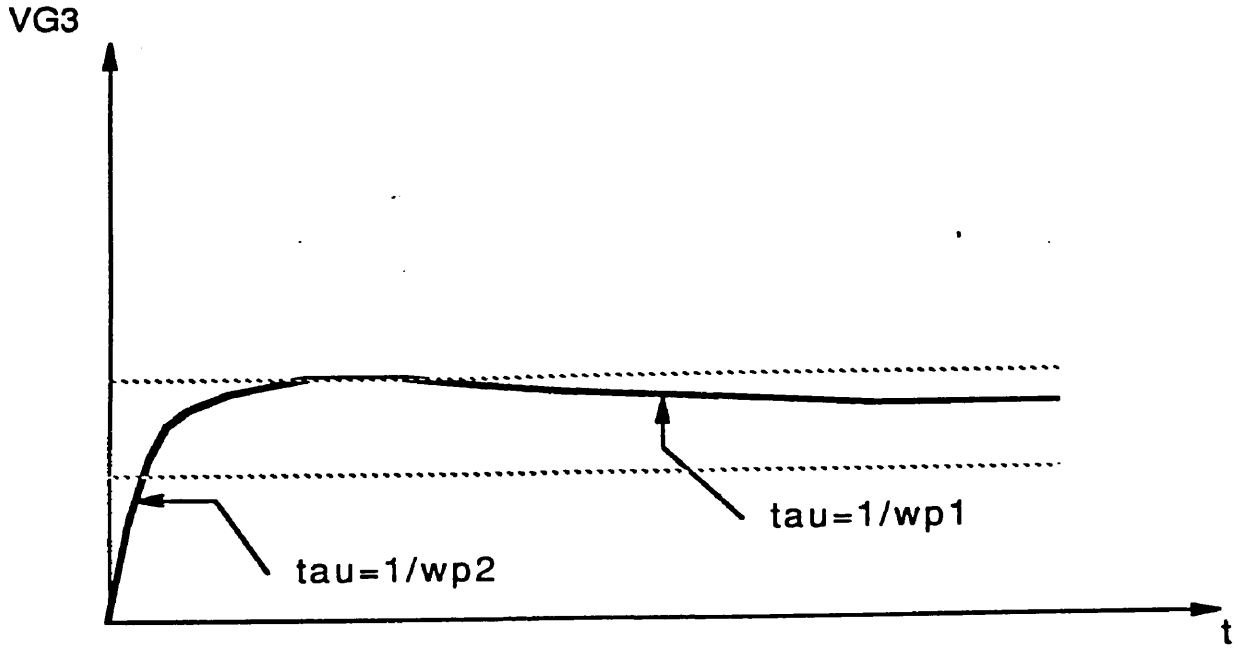


Fig. 130: Voltage Response, Gate of M_3 (Step)

is inherently the most efficient process, or they are discharged by fixed current sources. These current sources are ON for a short, signal-dependent time, during which they have almost 100 % efficiency (all the current coming from the power supply ends up on the capacitor). The rest of the time, they are OFF. The only significant DC power that is "wasted", is the bias current of the regulating amplifiers of the current sources, and the current in the current mirrors that set the DC current value.

The scheme also uses a number of comparators. These only have to be active during one or two clock phases each clock period. During that time, they consume a certain DC power, which can be increased in order to improve the switching speed of the system (decreased undershoot). However, since the comparators do not have to drive any strongly capacitive loads, adequate speed can usually be achieved with relatively low bias currents, of the same order as the currents in the current sources or even lower (micro-amps). During the clock phases when the comparators are not active, they could be shut off (power-saving mode), but this should rarely be necessary.

10.21. Speed Advantage Over an SC Stage

The charge amplifier can potentially be made extremely fast, without significantly increasing its power consumption. It is anticipated that this scheme will have a much better throughput rate (clock frequency) over power ratio than a corresponding switched-capacitor stage, all other specifications being the same.

The main advantage, as far as speed is concerned, is that the capacitor of the charge amplifier scheme is discharged in a linear (constant slew rate) way, rather than through an exponential settling process that may require several time constants before sufficiently accurate settling is obtained.

Another advantage is that the comparator that is used in the scheme does not have to drive the main capacitor. The only load it has to drive, is a logic gate. As a result, the comparator can be made very simple (small) and very fast using extremely low bias currents. There are no slew rate considerations and the required current driving capability is minimal.

Finally, a fundamental advantage of this kind of comparator is that stability is not an issue. The comparator in the charge amplifier essentially operates in

an open-loop fashion. As a result, compensation (reducing the gain-bandwidth product to a value well below the non-dominant poles) is not necessary, thus maximizing speed.

A simple comparator structure is shown in figure 131. This structure exhibits a switching time of a few nanoseconds when realized with minimum-sized transistors in 2μ CMOS technology, at a total current consumption of about $50\mu A$. The simplicity of this comparator stands in strong contrast with the high-bandwidth, high-gain, high slew rate, internally compensated opamp needed to drive the (big) feedback capacitor in a switched-capacitor stage.

10.22. Noise Performance

Due to its switching nature, a noise analysis of the charge amplifier is difficult. Noise is usually analyzed in the frequency domain, under the assumption that the circuit under consideration is biased at a constant operating point. Unfortunately, this assumption cannot be upheld for the charge amplifier, since the current sources in the system switch at a frequency of up to several MHz. In addition, the system includes correlated double sampling, which affects noise properties.

The noise analysis of the charge amplifier will require going back to the basics about noise, being the statistical properties of stochastic processes in the time domain. This analysis will be performed in chapters XI, XII, XIII. In chapter XIV, the conclusions from these three chapters will be used to derive approximative noise models for the charge amplifier.

10.23. Supply Noise Rejection

The charge amplifier is fairly insensitive to power supply variations. The extent to which such variations could potentially feed through to the actual signal, if at all significant, would be threefold.

First, there are the current sources. If these are implemented as regulated cascode current sources, the only path for the power supply to reach the output, would be through the active load (P channel MOSFET) of the auxiliary amplifier. However, this would mainly cause a variation on the output node of that amplifier, and the sensitivity of the output current to the voltage on that node is fairly small.

Second, there is the comparator. If its structure is asymmetric, the differential

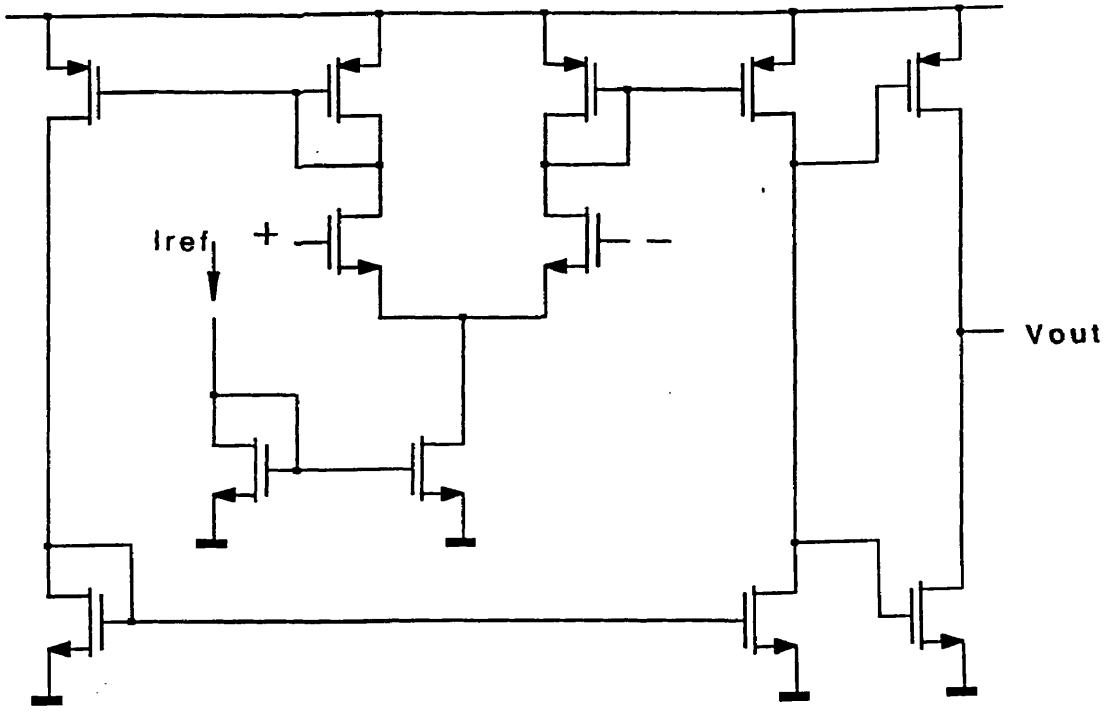


Fig. 131: Simple Comparator Scheme

offset voltage may be affected to a certain extent by variations in supply voltage.

Third, there is the effect of the AND gate. When the power supply changes, the change will be reflected in its threshold voltage. This effect, as well as the previous one, may affect the signal if the power supply variation occurs during the time that the comparator switches.

It is clear however that only random, high-frequency variations will affect the charge amplification process. Variations that are synchronous with the clock, or variations that are of sufficiently low frequency (e.g. 60 Hz), will be effectively suppressed by the correlated double sampling.

But one of the main reasons why this structure will not be affected much by feed-through of parasitic signals from the power supply (e.g. feed-through from another stage), is that the structure does not *generate* much power supply noise.

During Φ_1, the capacitor is precharged to a certain voltage. This precharge operation is synchronous and repeatable, since the capacitor is always charged starting from the same initial voltage. The resulting power supply glitch will be perfectly deterministic and will not introduce any distortion or random noise in other stages.

During subsequent clock phases, the capacitor is discharged a few times using current sources. This is an extremely favorable situation, since the capacitor and current source are essentially connected in a loop, and no current is injected into the power supply lines. As a result, the supply lines remain clean and feed-through to other stages is extremely limited. This situation is much better than in the switched-capacitor stage, where large, signal-dependent currents are drawn from the power supplies in order to charge or discharge the (large) capacitances of that system.

As a result of naturally good supply rejection characteristics, as well as good linearity, fully differential approaches should normally not be needed for a charge amplifier stage.

10.24. Conclusion

The charge amplifier uses a capacitor, high output impedance current sources and a simple comparator in order to achieve a certain charge gain from one stage

to the next one. One current source is used to discharge the capacitor in a linear way. The comparator monitors the capacitor voltage, detects when it reaches a pre-set reference level and switches the current OFF. The output of the same comparator also controls a second current source, with which the initial capacitor charge is copied onto the next stage.

The basic scheme is straightforward, yet the structure can achieve excellent performance in terms of speed and linearity. Because the gain is based on charge rather than voltage, the linearity is not limited by the quality of the capacitors. In addition, since the gain is set by a ratio of current sources, it can easily be externally adjusted.

Current efficiency is nearly 100%, which results in a strongly reduced power consumption, as compared to an SC amplifier stage with associated opamp. Inherent offset and low-frequency ($\frac{1}{f}$) noise due to the comparators can be strongly reduced using a simple cancellation technique, functionally equivalent to correlated double sampling. This peculiarity, as well as the fast switching of the current sources, makes the estimation of noise difficult. As a result, separate chapters will be devoted to the subject.

The structure is fairly insensitive to power supply noise, which should reduce the need for fully differential systems. It is also strongly immune to circuit parasitics and mismatches. It does not require MOS switches operating in the ohmic region, which limits charge redistribution problems like clock feedthrough and makes it ideal for implementation in bipolar technology. It is *not* subject to instability, which further eliminates a lot of design constraints and eases the requirements on the IC technology. Due to its inherent robustness, this structure should make a good candidate for applications that need to be radiation-hardened or able to function reliably in harsh environments.

It could also prove extremely valuable for applications that need direct analog signal processing or analog to digital conversion of signals that become available as a charge. This is typically the case in photodiode-based optical scanners or CCD-based video cameras.

CHAPTER XI

GENERAL NOISE THEORY

11.1. Introduction

This chapter reviews some fundamental concepts related to the statistical description of noise. The purpose of doing so, is to eventually develop the formulas needed to do a noise analysis of the charge amplifier. This analysis cannot easily be performed using conventional noise calculation techniques, since the circuit contains rapidly switching current sources and uses correlated double sampling. As a result, there is no point in describing the noise as a small variation around a fixed bias point.

Noise is usually analyzed in the frequency domain rather than in the time domain. The calculation techniques for frequency domain analysis are simple, elegant and well understood. The underlying fundamental reason why this kind of analysis can be performed so easily, is that noise, while being a random phenomenon in the time domain, has a deterministic power spectrum in the frequency domain.

The formulas given in this chapter are not new; they can be found in mathematical and electrical engineering textbooks. However, it was found useful to create a compilation of them within this work, because literature on the subject is either very mathematical, or else somewhat vague about how different concepts like autocorrelation, energy spectrum and power spectrum fit together, and how they relate to noise in actual electronic circuits. All those aspects had to be very well defined in order to derive formulas for noise in switched systems (chapter XII) and the influence of correlated double sampling on noise (chapter XIII).

11.2. Switched Versus Constant-Bias Systems

When an electronic circuit has a fixed bias point, the noise characteristics of all resistors and semiconductor devices remain relatively constant over time. As a result, the statistical properties corresponding to the noise voltages or currents can be determined. They are usually expressed in the form of a power spectrum, which depends on the specific bias conditions of each device. A small-signal

frequency domain analysis can be performed in order to combine the power spectra of individual devices into a total noise spectrum at the system output. The output spectrum can then be integrated over a certain "useful" frequency range in order to find the total expected or root mean square (rms) noise power. It can further be related to the transfer function of the system in order to calculate input-referred noise, dynamic range or noise figure.

Things are different in switched systems. In such circuits, a single bias point cannot be identified and the conventional noise analysis techniques cannot simply be applied. When the system (e.g. a current source) is ON, a bias point could be defined. But since the system is periodically switched OFF, all devices periodically go to a zero-current state, during which practically no noise is present. This may drastically influence the properties of the noise at the moment the system is switched back ON. In other words: the deterministic switching of the system may affect the statistics of the noise (random phenomenon) at the moment it is switched back ON.

In the practical situation at hand (charge amplifier), current sources are normally switched ON and OFF at a rate of several MHz. To complicate things, when they are ON, they are so for a length of time which may be dependent on the input signal (pulse width modulation).

An important question needs to be answered in order to model the noise in such a system: " *What happens to the statistical properties of the noise every time the system is turned back ON?*". Does the noise start at zero and gradually build up? Does it pick up at the value it had immediately before the system was turned OFF? Does it start at some random value, as if the system had continuously been ON? Or is the actual situation somewhere in between? Neither literature nor experts seem to provide a ready answer. Yet, in the particular system of the charge amplifier, the estimate for the total integrated noise on the output signal can be several orders of magnitude higher or lower, depending on what hypothesis is put forth.

To complicate things, the situation may be different depending on what kind of semiconductor device is used. A MOSFET for example, is characterized by white noise, due to the thermal agitation of electrons, and colored (1/f) noise, mainly due to random trapping of electrons in surface states created by the

interface between silicon and gate oxide. The assumptions that are to be made in order to come up with a comprehensive noise model, may be different for both kinds of noise.

It is intuitively clear that *if* it can be assumed that *the noise starts at zero* every time a current source is turned ON, the output current noise will not have the time to build up during the few hundreds of nanoseconds that the source is ON. This is the most favorable hypothesis, predicting remarkably good noise performance of the charge amplifier, especially as far as 1/f noise is concerned. The switching would then act as a sort of "reset" mechanism, which would periodically (at a rate of several MHz) zero out the noise signal. This resetting would cancel most low-frequency noise components, which are obviously dominant in 1/f noise.

The hypothesis may be overly optimistic. It is possible that at least some of the oxide traps that cause 1/f noise, will be preserved while the MOSFET is switched OFF, and will instantly contribute to the total noise when the device is switched ON again. If *all* trapped carriers remained in the same state while the device is currentless, the switching could be equivalent to momentarily "freezing" the noise, rather than zeroing it out. If carriers could be trapped (or untrapped) in the OFF state like they normally are in the ON state, switching might not even influence the 1/f noise at all.

Switching does not affect white noise as much. The white noise current component in a MOSFET is totally random in nature, due to its thermal origin. The instantaneous value of the white noise current at one time is in principle uncorrelated to the value at any other time, resulting in a practically infinite bandwidth in the frequency domain. This seems to insinuate that the white noise component *is not* influenced by the switching, since the fact that the noise was zero at some earlier time, does not influence its possible value at some later time, no matter how soon after.

Chapter XII will investigate the question in more detail. But Eventually, experimental data will have to determine which theoretical assumptions are correct. An unscientific indication that may point towards the fact that noise needs time to build up after being switched, is routine. Some analog designers (maybe not the most rigorous ones!) who design switched systems, neglect parts of the frequency spectrum below the sampling frequency (including possibly dominant

contributions of 1/f noise), although they usually are not able to prove the validity of such procedure.

In the meantime, even a worst-case (i.e. assuming that switching does *not* reset the noise) analysis of the charge amplifier indicates satisfactory noise performance. If needed, this could possibly be improved even further through the use of periodic recalibration (ratio setting) of critical current sources with respect to a single reference source. The latter procedure would reduce the effect of drift (low-frequency noise) in case the switching does not already take care of it.

If we were to assume that when a current source (or any other circuit) is turned ON, the noise starts building up from zero, we would need to quantify the phenomenon. In particular, we would need a means of estimating the root mean square noise voltage or current (usually referred to as the expected noise power) after a finite time T. This analysis is slightly unusual; traditional frequency-domain analysis does not directly provide an answer. However, using statistical concepts, both from the frequency domain and the time domain viewpoints, a reasonably simple theory can be developed which will allow us to calculate the expected noise power.

11.3. General Noise Concepts and Assumptions

Many of the statistical concepts that will be discussed, are borrowed from the textbook "Probability, Random Variables, and Stochastic Processes" by Athanasios Papoulis [95]. However, they will be placed in the practical context at hand, and sometimes extended for application in support of the our "switched noise" theory.

In any electronic system, signals can be represented as a sum of two components: a deterministic signal (the "expected value"), which can be predicted through the circuit equations, and a small, superimposed noise signal, which by definition cannot be predicted, but about which certain conclusions can be drawn using statistical techniques. From a circuit theory point of view, the noise signal is a "small-signal" excursion around a biasing point, which is deterministic in nature. This implies that linear circuit theory can be applied to the noise signal once the large-signal biasing conditions of the circuit are known.

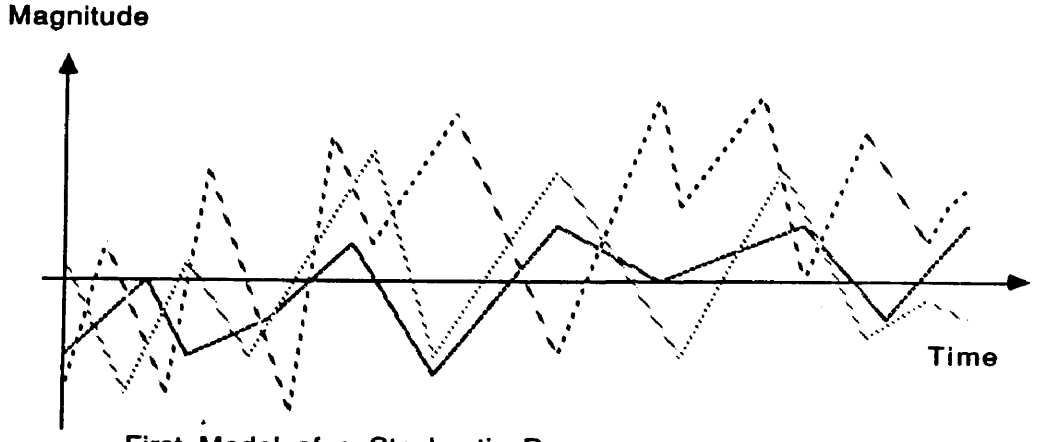
From a statistical point of view, noise is a continuous-time stochastic process.

A stochastic process is an infinite set of random variables, one of which is associated with every possible value of the time t. The different random variables have continuous probability density functions (e.g. over the possible values of current or voltage), which are not necessarily independent. There may be a certain correlation between the probability of a voltage occurring at one time and another voltage occurring at another time.

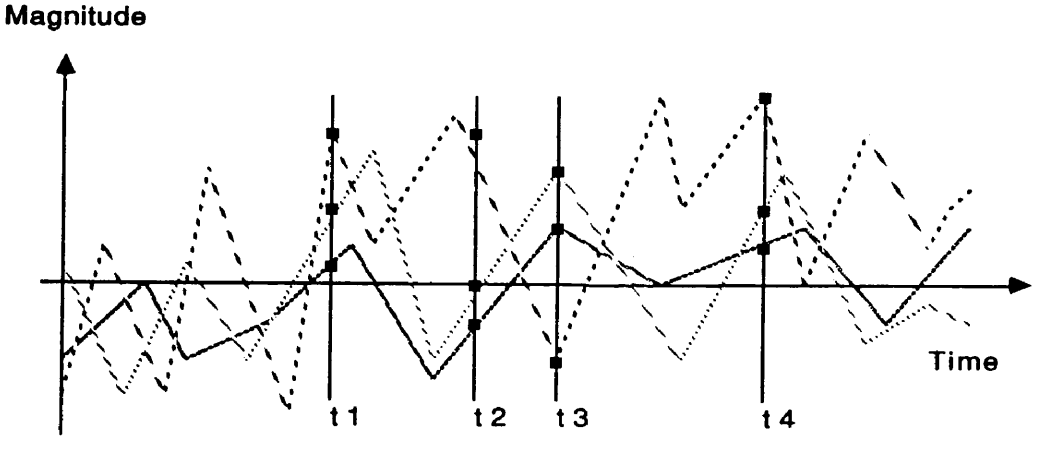
The *actual* value of the voltage (or current) cannot be predicted, but the *probability* of specific values at specific times can be determined. In particular, the stochastic process can be represented as an infinite set of "possible" signals, which are called "samples" of the process. Each sample is a different function of time. The two alternative ways to model a stochastic process are illustrated on figure 132. The *actual* noise signal corresponds to *one* sample out of all the possible ones, and its properties (probability, expected value, expected power etc.) can be found using statistical techniques. The statistical properties of the process are determined by the particular biasing conditions of the semiconductor devices involved in the circuit. In their simplest form, they are given as a power spectrum, of which the meaning will be elaborately described below.

Throughout the discussion, we will assume that the noise in electrical devices can be modeled as a *strict-sense stationary* (SSS) stochastic process. This kind of stationarity means that the statistical properties of the noise are insensitive to a shift in time origin. In other words: it does not matter at which specific time we observe the noise, its properties will be the same, be it yesterday, now, in a microsecond or tomorrow. Another important assumption we will make about the noise we wish to consider, is memorylessness. We will assume that the past of the stochastic process has no influence on its future if the present is specified. If we know the value of the noise at a time $t = 0$, we can determine the probability of a certain value at a time $t = T$, no matter what the values were for times $t < 0$. Such processes are also called Markoff processes.

A property of noise signals, closely related to strict-sense stationarity, is wide-sense stationarity (WSS). Every SSS noise process is WSS, but not vice-versa. Wide-sense stationarity requires that the noise signal would have zero expected value at all times, which follows directly from our definition of the noise signal. It also requires that the autocorrelation (to be defined) of the noise signal between



First Model of a Stochastic Process:
Each sample is one out of a number of possible time functions



Second Model of a Stochastic Process:
The value at each time is a stochastic variable

Fig. 132: Sample of a Stochastic Process

two particular times t_1 and t_2 be a function of the time difference $t_2 - t_1$ only. Many of the statistical properties we will discuss below apply to WSS stochastic processes (e.g. the existence of a power spectrum), and by extension to electrical noise (SSS process). However, some properties explicitly rely on the assumption of strict-sense stationarity, which is more stringent.

That assumption is certainly justified. As a matter of fact, high-quality measurement equipment like a spectrum analyzer implicitly rely on strict-sense stationarity to measure noise spectra of electrical systems, despite the fact that only *one* sample can be observed by the instrument at any given time. This is called ergodicity and will briefly be described below. Memorylessness is harder to demonstrate. However, it is intuitively clear that the electrical noise we are interested in, is caused by a multitude of microscopic processes, which are additive but mutually independent. There is no mechanism allowing for the noise to "remember" past values over extended periods of time, and hence this kind of noise must satisfy the Markoff definition.

Although we have considered strict-sense stationarity and memorylessness separately, one can actually demonstrate that the latter property follows from the former. If the statistical properties of a stochastic process are independent of time, they cannot be dependent upon the past history of the process, since a shift in time origin would obviously affect historic events. As a result, a strict-sense stationary process is also a Markoff process.

11.4. Fourier Transforms

Because of their importance throughout the discussion of noise, a few definitions and properties related to Fourier transforms will be summarized here. Most proofs will be omitted, since they can be found in almost any textbook about Fourier analysis.

The Fourier transform of a signal $x(t)$ is defined as

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\ \epsilon^{-j\omega t}\ dt \tag{175}$$

The integral converges to a finite value for *any* value of ω if the following two conditions are met:

* The signal (time function) $x(t)$ is absolutely integratable, or $\int_{-\infty}^{\infty} |x(t)|\ dt$ is finite
* The time function $x(t)$ has a finite number of discontinuities, each of finite magnitude

When its Fourier transform exists, the signal itself can be written as a Fourier integral (inverse Fourier transform):

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\ e^{j\omega t}\ d\omega \tag{176}$$

It can be shown that there is a one to one relationship between a time domain function and its Fourier transform. The time domain function uniquely defines the Fourier transform (frequency domain function) and vice-versa.

Even when the two convergence conditions are not met, Fourier transform and inverse Fourier transform can sometimes still be defined, using generalized functions like the Dirac impulse (delta function $\delta(t)$), with amongst others, the following properties

$$\int_{-\infty}^{\infty} \delta(\tau)\ d\tau = 1 \tag{177}$$

$$\int_{-\infty}^{\infty} f(t)\ \delta(t-\tau)\ d\tau = f(t) \tag{178}$$

$X(\omega)$ is the Fourier coefficient at frequency ω. If $x(t)$ represents a voltage, with units of volts (V), and the time is given in seconds (s), $X(\omega)$ will have units of Vs or V/Hz.

$X(\omega)$ is, in general, complex and can be written as $X(\omega) = R(\omega) + jI(\omega)$. If $x(t)$ is a real function of time, it can be shown that $R(-\omega) = R(\omega)$ and that $I(-\omega) = -I(\omega)$. In other words, the real part of the Fourier transform is an even function of frequency; the imaginary part is an odd function of frequency.

If $x(t)$ is a real, even function of time ($x(t) = x(-t)$), all Fourier coefficients will be real and $x(t)$ can be written as an infinite sum (integral) of cosine waves.

$$x(t) = \frac{1}{2\pi}[X(0) + 2\int_{0+}^{\infty} X(\omega)\ \cos(\omega t)\ d\omega] = \frac{1}{2\pi}[R(0) + 2\int_{0+}^{\infty} R(\omega)\ \cos(\omega t)\ d\omega] \tag{179}$$

If $x(t)$ is a real, odd function of time ($x(t) = -x(-t)$ and hence $x(0) = 0$), all Fourier coefficients will be purely imaginary and $x(t)$ can be written as an infinite sum of sine waves.

$$x(t) = \frac{2j}{2\pi} \int_{0+}^{\infty} X(\omega) \sin(-\omega t)\, d\omega = \frac{1}{\pi} \int_{0+}^{\infty} I(\omega) \sin(\omega t)\, d\omega \qquad (180)$$

A shift in time origin in the time domain corresponds to a pure phase shift in the frequency domain.

$$F(x(t - t_0)) = e^{-j\omega t_0}\, X(\omega) \qquad (181)$$

11.5. Parceval's Theorem and Energy Spectra

Another well-known property of the Fourier transform is that the transform of the convolution of two time-domain functions (here represented in the general, complex sense) is equal to the product of their Fourier transforms and vice-versa.

$$F[\int_{-\infty}^{\infty} x(t - \tau)\, y^*(\tau)\, d\tau] = X(\omega)\, Y^*(\omega) \qquad (182)$$

$$F^{-1}[\int_{-\infty}^{\infty} X(\omega - \lambda)\, Y^*(\lambda)\, d\lambda = 2\pi\, x(t)\, y^*(t) \qquad (183)$$

When the latter equation is evaluated for $y(t) = x(t)$ and $\omega = 0$, Parceval's Theorem follows

$$|x|^2(t) = \frac{1}{2\pi} F^{-1}[\int_{-\infty}^{\infty} X(\omega - \lambda)\, X^*(\lambda)\, d\lambda] \qquad (184)$$

$$F[|x|^2(t)] = \int_{-\infty}^{\infty} |x|^2(t)\, e^{-j\omega t}\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega - \lambda)\, X^*(\lambda)\, d\lambda \qquad (185)$$

Or for $\omega = 0$:

$$\int_{-\infty}^{\infty} |x|^2(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, X^*(-\omega)\, d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, X^*(\omega)\, d\omega \qquad (186)$$

This important theorem expresses a relationship between the total energy dissipated by the signal over time, and its squared Fourier coefficients (which will be identified to the energy spectrum). The total energy is defined as the time integral of the instantaneous signal power, which in turn is defined as the squared instantaneous value of the signal.

$$E = \int_{-\infty}^{\infty} P(t)\ dt = \int_{-\infty}^{\infty} |x|^2(t)\ dt \tag{187}$$

With the instantaneous power defined as the squared value of the signal (normally in V^2 or A^2).

Parceval's theorem makes it possible to calculate the total energy of a signal by integrating the squared Fourier coefficients over frequency. One can easily verify that the units of $X(\omega)X^*(\omega)$ are $(Vs)^2$, or V^2s/Hz. They have the dimension of an energy density function over frequency. For that reason, $|X(\omega)|^2$ is called the *energy spectrum* of $x(t)$.

Another theorem, called Wiener-Khintchine's Theorem, provides an additional relationship between the time domain and the frequency domain. It will be given here without proof. The theorem says that the energy spectrum of a signal can be found as the Fourier transform of the autoconvolution of the time domain signal.

We use the term autoconvolution to refer to what several authors call the autocorrelation function of the signal, because the term autocorrelation will be used later, with a related but distinct meaning. The autoconvolution $R_{xx}(t)$ of a signal $x(t)$ is defined as

$$R_{xx}(\tau) = x(\tau) * x(-\tau) = \int_{-\infty}^{\infty} x(t+\tau)\ x^*(t)\ dt \tag{188}$$

Clearly, $R_{xx}(\tau)$ is an even function of τ. The function expresses the degree of dependence (or correlation) of $x(t)$ upon the same signal, delayed by a time τ. Wiener-Khintchine's Theorem can then be written as

$$|X(\omega)|^2 = \int_{-\infty}^{\infty} R_{xx}(\tau)\ \epsilon^{-j\omega\tau}\ d\tau \tag{189}$$

The theorem is usually used in the other direction. Some signals are hard to measure in the time domain. Instead, a frequency domain measurement makes it possible to determine the energy spectrum. If the signal is random in nature, the *expected value* of the energy spectrum can be estimated. From the energy spectrum, the autoconvolution of the signal can be reconstructed through inverse Fourier transformation. Although this does not uniquely define the time-domain signal, it does provide useful statistical information.

It has to be noted that an energy spectrum is normally only defined for finite-energy (pulse-like) signals. If the signal is does not decay over time, like is the case for noise signals, the total energy is infinite and Parceval's theorem predicts an energy spectrum which is not integratable, and possibly not even finite. However, similar frequency domain concepts can be derived considering the *expected power* (energy per unit time) of the noise signal, rather than the total energy. These considerations are the basis for noise analysis in the frequency domain.

11.6. Autocorrelation Functions and Power Spectra

The autocorrelation function of a zero-mean stochastic process $x(t)$, between two times t_1 and t_2, is defined as the expected value of the product of the signal at those two times.

$$R(t_1, t_2) = E(x(t_1)\ x^*(t_2)) \tag{190}$$

Since a noise signal is a real function of time, the corresponding stochastic process is real, such that $x^*(t) = x(t)$.

In the case of a WSS signal, the autocorrelation (by definition) reduces to

$$R(t_1, t_2) = R(t_2 - t_1) = R(\tau) = E(x(t)\ x^*(t + \tau)) \tag{191}$$

Which is independent of t. It is an even function of τ, since $R(\tau) = R(-\tau)$.

This definition is very similar to the definition of autoconvolution (equation (188)). The autocorrelation is an expected value (average), whereas the autoconvolution is an integral over the entire time axis. The autocorrelation has the dimension of power (V^2), while the autoconvolution has the dimension of energy ($V^2 s$).

The noise power spectrum is defined as the Fourier transform of the autocorrela-

tion function of the corresponding stochastic process.

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau)\ e^{-j\omega\tau}\ d\tau \tag{192}$$

This spectrum is always real and even as a function of frequency, since the autocorrelation function is an even function of time. The dimension of the power spectrum is V^2/Hz. It is interesting to note the parallel between the definition of power spectrum with respect to the autocorrelation function, and the energy spectrum with respect to the autoconvolution, as expressed in Wiener-Khintchine's Theorem.

11.7. Expected Power

The expected power of a noise signal is defined as the covariance of the corresponding stochastic process.

$$E(P(t)) = \sigma^2(x(t)) = E(x(t)\ x^*(t)) \tag{193}$$

For a WSS stochastic process, $E(P(t))$ is by definition equal to $R(0)$ and independent of t. This means that the total expected power can also be considered the total average power over a long time (principle of ergodicity). $E(P(t))$ then corresponds to the average power of "*a specific sample*" of the stochastic process ("*a specific*" instance of the noise signal), determined during the (ideally infinite) time that the signal has been observed.

$$E(P(t)) = P = lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\ x^*(t)\ dt \tag{194}$$

From the definition of the autocorrelation function and the power spectrum, it follows that

$$P = E(x(t)\ x^*(t)) = R(0) = F^{-1}(S)|_{\tau=0} \tag{195}$$

Or

$$P = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)\ e^{j\omega\tau} d\omega|_{\tau=0} = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)\ d\omega = \frac{1}{\pi} \int_0^{\infty} S(\omega)\ d\omega \tag{196}$$

Or in terms of frequency:

$$P = 2 \int_0^\infty S(2\pi f)\, df \tag{197}$$

In other words, the total expected power P can be found by integrating the power spectrum over the entire frequency axis. (Some circuits are designed to only operate over a limited frequency range, in which case integration only needs to be performed over that range). This indicates again that the power spectrum has the dimension of a power spectral density (psd), expressed in V^2/Hz or A^2/Hz. Again, a strong parallel can be observed between the power spectrum and the expected power on the one hand, and the total energy and the energy spectrum on the other hand, as expressed in Parceval's Theorem. This duality will become even more obvious in some additional properties.

11.8. White and Colored Noise

White noise is defined as a stochastic process for which the values $x(t_1)$ and $x(t_2)$ at any two times t_1 and t_2, are independent stochastic variables. As a result, the autocorrelation function $R(\tau)$ is an impulse, equal to the expected value of $P(t)$ for $\tau = 0$ and equal to zero otherwise. If the white noise power were infinite, the autocorrelation would be a delta impulse, and the power spectrum would be finite and constant over frequency ("flat"). As a result, the power of perfect white noise (not band-limited) would be infinite. In practice, there is always a certain correlation between the white noise at two different times when the time difference becomes very small. The instantaneous power is limited as well, and an actual autocorrelation function looks more like a narrow pulse of finite height, centered around zero (figure 133). In the frequency domain, this translates into a power spectrum that is constant up to some finite frequency and then rolls off.

Other noise spectra are "colored". Their noise spectrum usually exhibits larger magnitudes at lower frequencies, due to a certain correlation between the particular value of the noise signal at a particular time, and the value a small amount of time before. This correlation obviously decreases for increasing correlation times. 1/f noise in MOSFETs has a colored spectrum, which has been found to increase with decreasing frequencies, with a slope of approximately 10 dB per decade (power spectrum multiplied by $\sqrt{2}$ each time the frequency is halved).

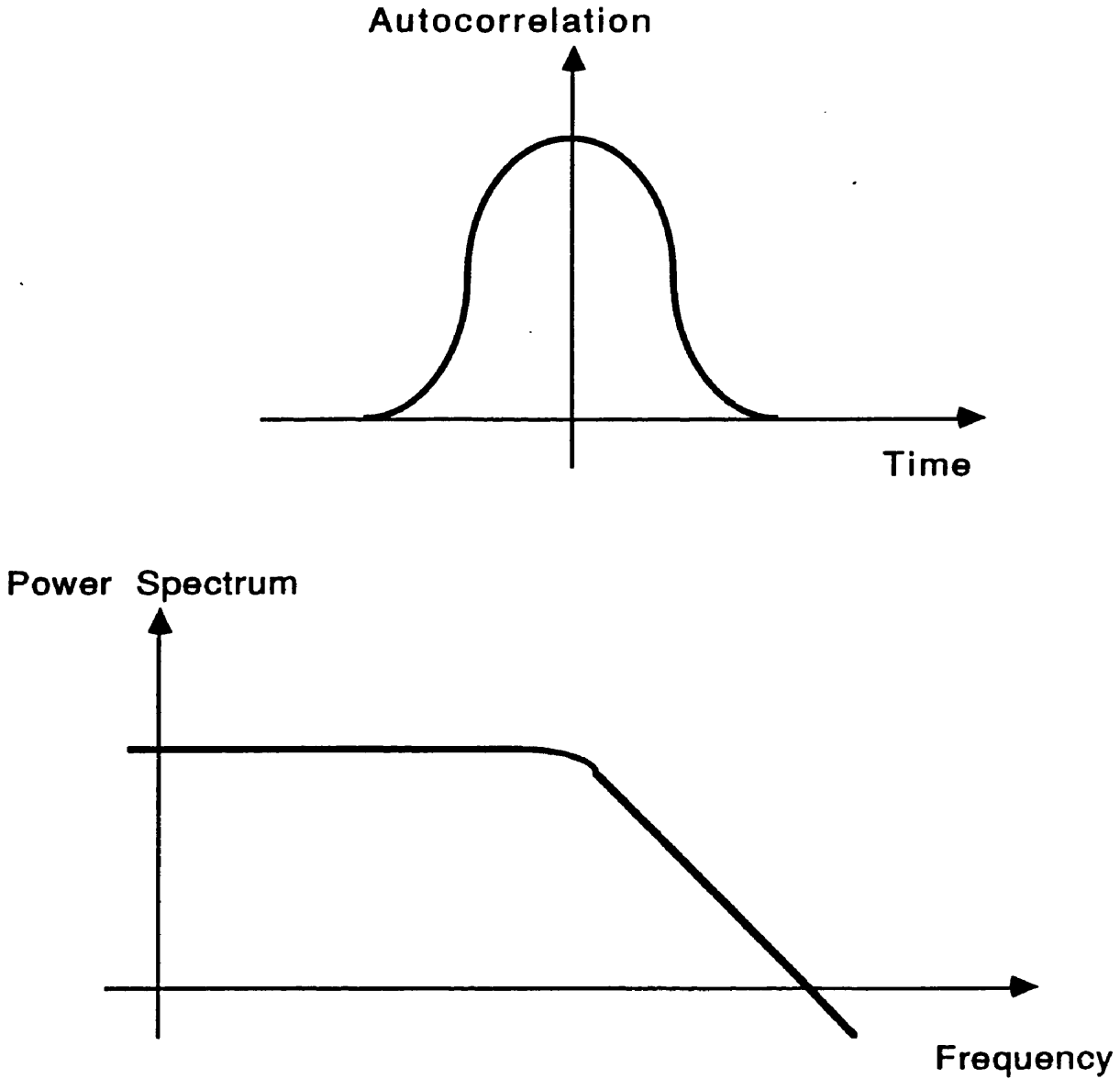


Fig. 133: Autocorrelation of Band-Limited Noise

11.9. Fourier Coefficients, Energy and Power Spectra

If $x(t)$ is a WSS stochastic process, its Fourier transform has an expected value of zero over the entire frequency range.

$$E(X(\omega)) = E(\int_{-\infty}^{\infty} x(t)\ e^{-j\omega t}\ dt) = \int_{-\infty}^{\infty} E(x(t))\ e^{-j\omega t}\ dt) = 0 \qquad (198)$$

This does not necessarily mean that the magnitudes of the Fourier coefficients have an expected value of zero. In practice, it does mean that their phases exhibit a random distribution between $-\pi$ and π, such that their expected *complex* value would be 0. This implies that a phase value of ϕ is exactly as likely to occur as a phase $\phi + \pi$.

If the stochastic process is SSS (like the noise we wish to consider in electronic circuits), the possible phase values exhibit a *uniform* probability distribution. An intuitive explanation is the following. Since the process is SSS, its statistical properties do not change over time. As a result, a signal $x(t)$ should have the same properties as the same signal, delayed by T, or $x(t-T)$. If the first signal has expected values $X(\omega)$ for its Fourier coefficients, the second one will have expected values equal to $e^{-j\omega T}\ X(\omega)$, which are identical, except for a difference in phase ωT. But because of the stationarity, the probability distribution of the phases for the second signal should be the same as the distribution for the first signal. Since this is true for *any* value of T, the probability distribution of the phases must be uniform (figure 134).

11.10. Autocorrelation of Fourier Coefficients

The autocorrelation between two Fourier coefficients can always be expressed as the two-dimensional Fourier transform of the autocorrelation function of $x(t)$.

$$E(X(\omega_1)\ X^*(\omega_2)) = E(\int_{-\infty}^{\infty} x(t)\ e^{-j\omega_1 t}\ dt \quad \int_{-\infty}^{\infty} x^*(t)\ e^{j\omega_2 t}\ dt) \qquad (199)$$

$$E(X(\omega_1)\ X^*(\omega_2)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E(x(t_1)\ x^*(t_2))\ e^{-j(\omega_1 t_1 - \omega_2 t_2)}\ dt_1\ dt_2 \qquad (200)$$

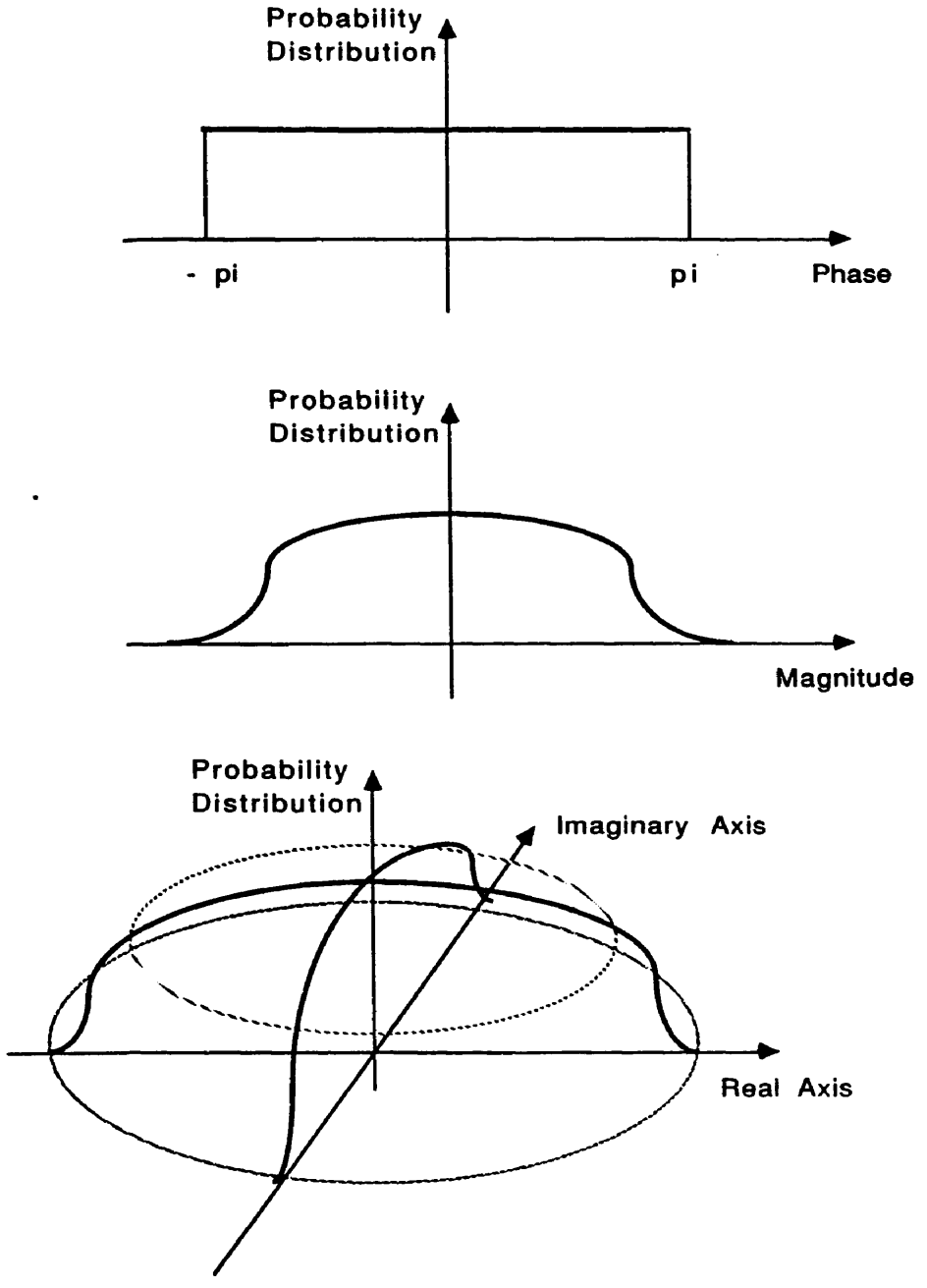


Fig. 134: Probability Distribution of Fourier Coefficients

$$E(X(\omega_1)\ X^*(\omega_2)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(t_1,t_2)\ e^{-j(\omega_1 t_1-\omega_2 t_2)}\ dt_1\ dt_2 \qquad (201)$$

For a WSS process, the previous expression can be rewritten as

$$E(X(\omega_1)\ X^*(\omega_2)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(t_1-t_2)\ e^{-j(\omega_1 t_1-\omega_2 t_2)}\ dt_1\ dt_2 \qquad (202)$$

Or after the substitution $t_1 = t_2 + \tau$:

$$E(X(\omega_1)\ X^*(\omega_2)) = \int_{-\infty}^{\infty} e^{-j(\omega_1-\omega_2)t_2} \quad \int_{-\infty}^{\infty} R(\tau)\ e^{-j\omega_1\tau}\ d\tau\ dt_2 \qquad (203)$$

$$E(X(\omega_1)\ X^*(\omega_2)) = S(\omega_1) \int_{-\infty}^{\infty} e^{-j(\omega_1-\omega_2)t_2}\ dt_2 \qquad (204)$$

Since the last integral equals $2\pi\ \delta(\omega_1-\omega_2)$, one can conclude that

$$E(X(\omega_1)\ X^*(\omega_2)) = 2\pi\ \delta(\omega_1-\omega_2)\ S(\omega_1) \qquad (205)$$

Two important conclusions can be derived from this equation. First, the value of the power spectrum at a particular frequency ω, is proportional to the expected value of the squared magnitude of the Fourier coefficient at that frequency. Since the squared magnitudes have been shown to represent the energy spectrum, one can conclude that the power spectrum is proportional to the energy spectrum. However, the proportionality factor contains a delta function. This follows from the fact that stationary noise signals have finite power, but infinite total energy, due to their infinite duration. As a result, the power spectrum is finite, while the energy spectrum is infinite.

Second, there is zero correlation between the Fourier coefficient of the noise signal at a particular frequency and the Fourier coefficient at a *different* frequency. The magnitudes may be correlated, but the phases of the Fourier coefficients of a WSS noise process are *independent* random variables. It has been demonstrated above that in the case of processes that are also strict-sense stationary, each of these phases has a uniform probability distribution.

11.11. Power Spectra and Expected Power

We have demonstrated that the expected power of a noise signal can be found through integration of its power spectrum over frequency. Another derivation leading to the same conclusion can be made relying on the inverse Fourier transform and the autocorrelation of Fourier coefficients. This derivation will later be generalized in order to apply to finite-time noise.

$$E(P(t)) = E(x(t)x^*(t)) = E(\frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} X(\omega_1)\ e^{j\omega_1 t} d\omega_1 \int_{-\infty}^{\infty} X^*(\omega_2)\ e^{-j\omega_2 t} d\omega_2) \tag{206}$$

$$E(P(t)) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(X(\omega_1)\ X^*(\omega_2))\ e^{j(\omega_1-\omega_2)t}\ d\omega_1\ d\omega_2 \tag{207}$$

Or since noise is a WSS process, $E(X(\omega_1)\ X^*(\omega_2)) = 2\pi\ S(\omega_2)\ \delta(\omega_1 - \omega_2)$, and the expression for the expected power reduces to

$$E(P(t)) = \frac{1}{(2\pi)^2}\ 2\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\omega_2)\ \delta(\omega_1 - \omega_2)\ e^{j(\omega_1-\omega_2)t}\ d\omega_1\ d\omega_2 \tag{208}$$

$$E(P(t)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)\ d\omega \tag{209}$$

Which is independent of t and confirms earlier results concerning the meaning of $S(\omega)$.

The equation allows us to calculate the expected (rms) noise power from the power spectrum of the noise. The assumption is that the noise process would be WSS. This requires that the noise be continuous over time. Physically, this would imply that the noise source should be ON for an infinite time. Obviously, this will not be the case for the charge amplifier, since all current sources and associated noise sources are switched OFF and back ON at a very high rate.

11.12. Conclusion

This chapter summarized a number of very important definitions and theorems related to noise in electronic systems. It relied heavily on concepts from

statistics, complex algebra and Fourier analysis. An overview of the difference between constant-bias and switched systems was given. Stochastic processes were introduced. Stationarity, wide-sense stationarity and memorylessness were defined and applied to noise in electronic systems. Fourier transforms and some of their properties were reviewed. Parceval's and Wiener-Khintchine's theorems were described, and their application to energy and power spectra demonstrated, through the concept of autoconvolution and autocorrelation.

Finally, the relation between the noise spectra (frequency domain) and expected energy or power (time domain) were derived, stressing all important underlying conditions and assumptions. The main goal of this chapter was to lay the mathematical framework for specific noise theories related to switched systems and correlated double sampling, which are used in the charge amplifier.

CHAPTER XII

NOISE IN SWITCHED SYSTEMS

12.1. Introduction

Chapter XI introduced a number of important concepts related to noise analysis in the time domain and in the frequency domain. However, the main assumption had always been that the noise process was at least wide-sense stationary (WSS). This assumption is justified for systems that have a fixed bias point, so that the noise can be regarded as a small-signal excursion around that bias point. In the charge amplifier, that assumption can no longer be justified, since the amplification process relies on switching current sources either completely ON or completely OFF. This chapter discusses the derivation of an important formula, which in some cases may make it possible to relate the noise of the (switched) current sources to the steady-state noise power spectrum (defined under constant-bias conditions).

12.2. Hypotheses and Assumptions

One could wonder about what would happen to the noise of a system that is first ON (steady-state), then switched OFF (all devices currentless, no noise) and finally switched back ON. Like mentioned in the beginning of chapter XI, several hypotheses could be put forth. Essentially, they boil down to two possible, major assumptions.

(1) The fact that a current source is switched OFF for some time, *does not* affect the noise characteristics when the current source comes back ON. This is equivalent to assuming that the noise process continues without interruption while the current source is OFF, although it is not visible at the output. The noise is merely "pulsed", or multiplied by a unity square wave. It is clear that if this were the case, the expected noise power at the output of the current source would always be equal to what it would have been if it had operated without interruption. The *expected* value of the noise power in the ON condition is independent of time (at least when the device is ON) and

can be found by integration of the steady-state power spectrum, as described in chapter XI.

Some people argue that before integration, the spectrum should be convolved with the spectrum of the multiplying square wave, to reflect the frequency contents of the actual, output noise signal. This seems unwarranted, since we assume that *when* the current source is ON, the noise exhibits steady-state characteristics. As a consequence the expected noise power should be time-independent and the same as if the noise had *not* been shut OFF.

(2) Switching OFF the current source for a short time "resets" the noise, since there *is* no noise in the OFF state. When the current source is switched back ON, the instantaneous value of the noise is zero, for continuity reasons. After this, the noise gradually builds up again, according to the stationary characteristics it normally has in the ON condition.

The theory we will now develop, refers to the *second* assumption. We will derive a formula which allows us to calculate the expected value of the noise after a certain, finite time T (e.g. the time the current source remains ON), under the assumption that at the time the current source is turned ON (for mathematical convenience, referred to as $t = 0$), the instantaneous value of the noise (and hence of the noise power) is zero.

It is our feeling that the second noise assumption has some physical foundations. The reason is the following. First of all, the distinction is only relevant as far as low frequency, finite bandwidth noise is concerned. Like has been pointed out earlier, white noise has an almost infinite bandwidth, which in the time domain translates to zero correlation between noise values at two different times. As a result, switching will not affect the behavior of white noise.

However, switching could drastically affect the time-domain behavior of $1/f$ noise in MOSFETs. $1/f$ noise physically originates from surface states (traps) at the silicon-oxide interface [96, 97]. The trapping of charge carriers is a dynamic, random phenomenon. Carriers are continuously trapped or released, but eventually a certain equilibrium condition is reached, in which the number of carriers being trapped *approximately* matches the number of trapped carriers being released again. This results in an approximately constant number of carriers in the

channel at any time. The small, random variations from this dynamic equilibrium condition result in $1/f$ noise.

According to this model, traps have a finite average lifetime. They do not become active all at once, nor do they release the trapped carriers all at once. The probability that a traps with a short lifetime would release carriers, is higher than the probability of traps with longer lifetimes. This causes $1/f$ noise, in which large bursts are less likely than small ones. In the frequency domain, this results in a power spectrum that rolls of with a slope of 10 dB per decade.

However, when the MOSFET is OFF, all trapped carriers are eventually released, since the channel is currentless. Carriers cannot remain in the traps, because if they did, an electric field would subsist with respect to the channel. As a result, we will assume that all traps are inactivated when the MOSFET is turned OFF. The $1/f$ noise will be "reset". When the MOSFET is turned back ON, an equilibrium current will start flowing almost instantaneously. As soon as this happens, the traps will resume their random action, but starting from the zero value.

A problem with this hypothesis, is that the time constants with which the traps release the carriers when the device is OFF, are not known. As a result, it is possible, if not likely, that the time during which the MOSFET is turned OFF (in our case possibly in the nanosecond range), will not be sufficient to release all carriers. As a result, turning the MOSFET OFF for such a short time may not completely reset the noise, but only partially. In the remainder of this chapter, we will ignore this dillemma. We will derive formulas based on the clear, simplified assumption that when a MOSFET is turned OFF, no matter how shortly, all noise will be completely reset.

In the following paragraphs, we will extensively refer to noise statistics, both from the time domain and from the frequency domain standpoint. We will eventually relate time domain noise behavior at the switch-ON time to the steady-state power spectrum (frequency domain) of the noise. The power spectrum is easy to measure and can be calculated using well-known formulas.

The main characteristic of the noise we are interested in, is the expected value of the instantaneous noise power (alternatively: the standard deviation of

the noise signal). It is well-known (this was also proved in chapter XI) that under steady-state conditions, this expected noise power is time-independent, because the noise is stationary (ergodic). However, we will try to demonstrate that in switched systems, this may not be the case. When the system is initially switched ON, low-frequency noise components will need time to build up, and the expected power will only gradually reach its steady-state, time-independent value.

The expected noise power will be related to standard deviations in the time domain and in the frequency domain. It is important to note that at no point, we will rely on a specific (e.g. Gaussian) magnitude distribution of the noise. This is physically correct, since it is well-known that $1/f$ noise does not have such a distribution. $1/f$ noise is usually characterized by short "bursts" of noise, followed by periods of relative calm. The length of the bursts and quiet periods, are random in nature. This is depicted in figure 135. It is clear that such noise is not Gaussian and not white. However, from the auto-correlation point of view, the noise process can still be considered stationary. (It seems warranted to consider 1/f noise to be ergodic, since 1/f spectra can reliably be measured using a spectrum analyzer, which implicitly relies on ergodicity.)

12.3. Noise Statistics in Switched Systems

In our mathematical switched-noise theory, we will explicitly apply the restriction that at $t = 0$ (switching instant), the noise is zero. Another assumption we will make implicitly, is that it takes negligible time to turn a MOSFET (current source) ON. In other words, for $t = 0^-$, the devices in the circuit are entirely OFF. For $t = 0$, the circuit operates under normal, steady-state bias conditions, without noticeable transient behavior.

If the continuous, steady-state noise is represented mathematically by a stationary stochastic process $x(t)$, it is clear that the restriction $x(0) = 0$ defines *one* subset of this process. The subset contains all "samples" of the original process for which the condition $x(0) = 0$ is met. The statistics of the restricted process are obviously different from the statistics of the unrestricted process, since all probabilities turn into conditional probabilities due to the condition $x(0) = 0$. In particular, the *restricted* process is not strictly stationary, or even WSS anymore. (The value at $t = 0$ is deterministic.)

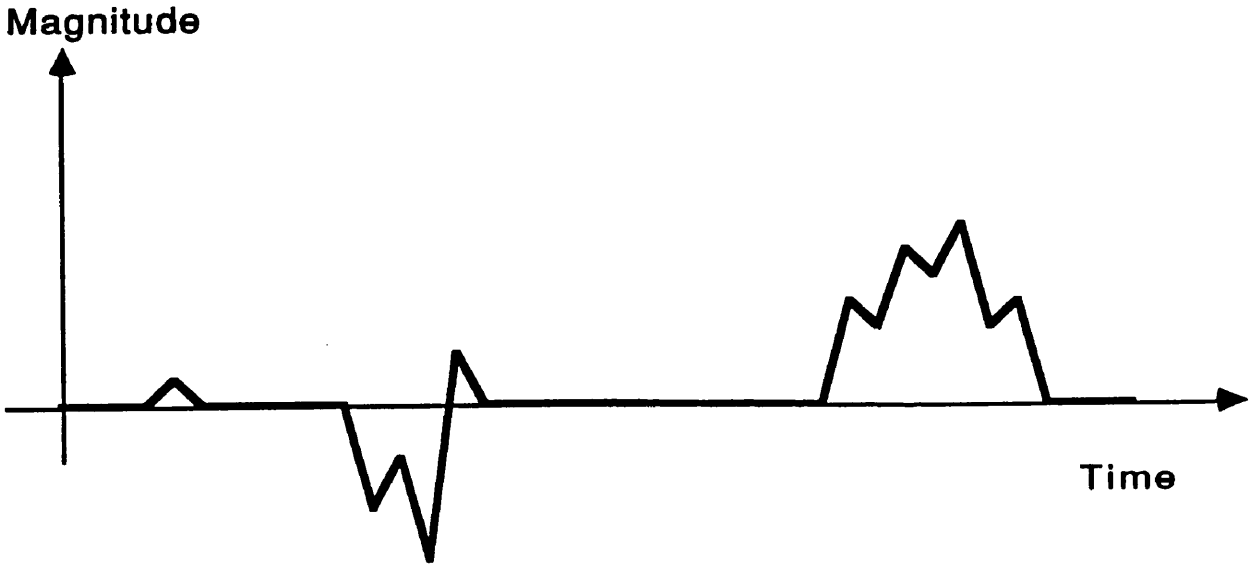


Fig. 135: Typical 1/f Noise Bursts

One could argue that we ought to consider a subset for which not only the condition $x(0) = 0$ is satisfied, but also the condition $x(t) = 0$ for $t < 0$. This distinction does not have to be made because of the initial assumption that the original process was a memoryless (Markoff) process. We are interested in the expected value of the power for time $T > 0$, or $E(|x_0|^2(T))$, under the condition that $x(0) = 0$. For a Markoff process, any *additional* condition referring to negative times would not influence the probability distribution of $x(t)$ for $t > 0$. As a result, the statistics of *any* process that is a subset of the original process in a way such that $x(0) = 0$, will have the same properties for any $t > 0$, and hence the same value for $E(P(T))$.

12.4. Expected Energy Spectrum of a Restricted Process

It will be shown that some properties of the Fourier coefficients associated with the restricted process, will remain similar to those of the unrestricted process. In particular, this is true for the expected value of the squared magnitude of the Fourier coefficients associated with a sample of the restricted process (the expected energy spectrum of the restricted process). This can be easily demonstrated in an intuitive way.

The restriction $x(0) = 0$ defines *one* subset of the original process, which we will refer to as $x_0(t)$. But an infinite number of other subsets could be defined in a similar fashion. If we considered all possible noise samples for which $x(t_1) = 0$, we would get another subset. All samples for which $x(t_2) = 0$ yet another one, etc. If we took the union of all such subsets (for which $x(t) = 0$), for all possible values of t), we would find the original WSS stochastic process, since the process is by definition zero-mean, and all samples must cross zero at one time or another. It should be noted that these subsets do not define a partition of the original process, since one sample could belong to several subsets at the same time, depending on the number of zero-crossings.

Since each (non-stationary!) subprocess is the same, except for a shift in time (corresponding to a phase shift in the frequency domain), each subprocess will have an identical expected energy spectrum (squared Fourier coefficients). Since the union of all the subsets yields the original, stationary process, it is also clear that the expected energy spectrum associated with each subset must be the

same as the expected energy spectrum of the unrestricted process. The latter has been shown to be proportional to the *power spectrum* of the unrestricted process (chapter XI). As a result, the following equation can be written for the restricted process $x_0(t)$:

$$E(X_0(\omega)X_0^*(\omega)) = \frac{1}{2\pi}\delta(0)\ S(\omega) \tag{210}$$

With $S(\omega)$ the power spectrum of the unrestricted, stationary process.

The proof can be visualized using a simplistic comparison. A school has a number of students, divided into several classes, such that the average age in each class would be the same. If it is known that the average age in the school is 16, what is the average age in each class?

12.5. Phase Considerations

It is more difficult to find the probabilities associated with the phase values that correspond to the Fourier coefficients of the restricted process $x_0(t)$. In an entirely general way, we could write a sample $x_0'(t)$ of the restricted process as a Fourier integral.

$$x_0'(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_0'(\omega)\ e^{j\omega t}\ d\omega \tag{211}$$

Or, if we separate X_0' into real and imaginary part:

$$x_0'(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} (R_0'(\omega) + jI_0'(\omega))\ e^{j\omega t}\ d\omega \tag{212}$$

From the condition $x_0'(0) = 0$, it is clear that the integral corresponding to the real parts of all the Fourier coefficients must be zero. (A similar condition does not exist for the imaginary parts, since all sine terms are automatically equal to zero for $t = 0$.)

$$\int_{-\infty}^{\infty} R_0'(\omega)\ d\omega = 0 \tag{213}$$

This condition is hard to deal with mathematically. Instead, we could wonder what would happen if we replaced it by a more stringent condition, in which we set all real parts $R_0(\omega)$ equal to zero, instead of their sum. The answer is obviously that this would limit us to all the samples that have odd symmetry

($x(t) = -x(-t)$). This would represent a further restriction of the original process, adding the condition of odd symmetry to the condition $x(0) = 0$.

However, this new restriction represents no loss of generality as far as positive times are concerned. Every possible sample can still be written as a Fourier integral with purely imaginary coefficients, although this kind of notation would impose odd symmetry. (Any time function with odd symmetry can always be written as an infinite sum of sine waves, but conversely, any function that can be written as an infinite sum of sine waves must be odd).

We will now attempt to demonstrate that instead of considering the restricted process for which $x(0) = 0$, we could consider *another* process, for which all Fourier coefficients are always purely imaginary, without changing the statistics of the noise for positive times. A strict mathematical proof will not be given (it is even possible that the odd process would have slightly different statistics, but even this would be hard to prove mathematically). What follows is an intuitive demonstration, relying on the assumption that the noise process is a Markoff (memoryless) process. This will considerably simplify the mathematical treatment, and the results will make appreciable physical sense.

12.6. Odd-Symmetry Random Process

What we would like to find, is an expression for the expected value of the instantaneous noise power (the variance of the noise) as a function of time, for a device that is switched ON. We have assumed that at $t = 0$, the expected noise power is zero (for continuity reasons). We have also assumed that the noise process, for positive times, is a subset of the original steady-state process, subject to the condition that $x(0) = 0$. The original process was assumed to be a Markoff process. In addition, we have already shown that the energy spectrum $|X_0|^2$ of the restricted process, is equal to the energy spectrum $|X|^2$ of the original, unrestricted process.

As a result of the original process being a Markoff process and the condition that $x_0(0) = 0$, the restricted process could be split into two statistically independent, half processes. The first process contains a number of samples which are time functions defined for negative times only ($f_-(t)$ defined for $t \leq 0$, with $f_-(0) = 0$). The second process contains a number of samples which are time

functions defined for positive times only ($f_+(t)$ defined for $t \geq 0$, with $f_+(0) = 0$). This is illustrated in figure 136.

The probabilities for each sample in the positive and negative half processes to occur, are obviously related to the statistics of the original, unrestricted process. The relationship is the following. If we were to consider all possible samples of the original process for which $x(0) = 0$, we could associate an (infinitely small) probability to each one of them, with total sum of 1. Each probability would be equal to the conditional probability (condition $x(0) = 0$) of the associated negative half sample to occur within the original process, multiplied by the conditional probability of the associated positive half sample, since the statistics of the process for negative and for positive times are independent (Markoff).

What we are interested in, are the statistics of the positive half sample only (whatever device we are looking at, is assumed to be OFF for $t < 0$). In particular, we will consider the expected value of the instantaneous power over time. Conceptually, there are two ways to calculate this power for positive times. Although the number of samples is actually infinite, we will illustrate the concept with an example shown in figure 137, containing three negative half samples and three positive half samples. The probabilities of each half sample are 1/4, 1/2 and 1/4 respectively. All samples are restricted by the condition $X(0) = 0$, and the probabilities should be interpreted as conditional probabilities with respect to the unrestricted process.

1. Every *full* sample $f(t)$ $(-\infty \leq t \leq \infty)$ is squared, multiplied by its probability and added. In total, there are nine possible samples (figure 138), corresponding to the nine combinations of three half samples. The probabilities are 1/16, 1/8 and 1/4 respectively, as shown on the figure. These probabilities are obtained by multiplying the respective probabilities of the respective half processes, since they are statistically independent (Markoff). The expected instantaneous power is also sketched. The positive half of the expected power function is the one we are interested in.
2. Since positive and negative half samples are statistically independent, one could obtain the same result by only considering the three possible positive half samples, with their respective probabilities (figure 139).

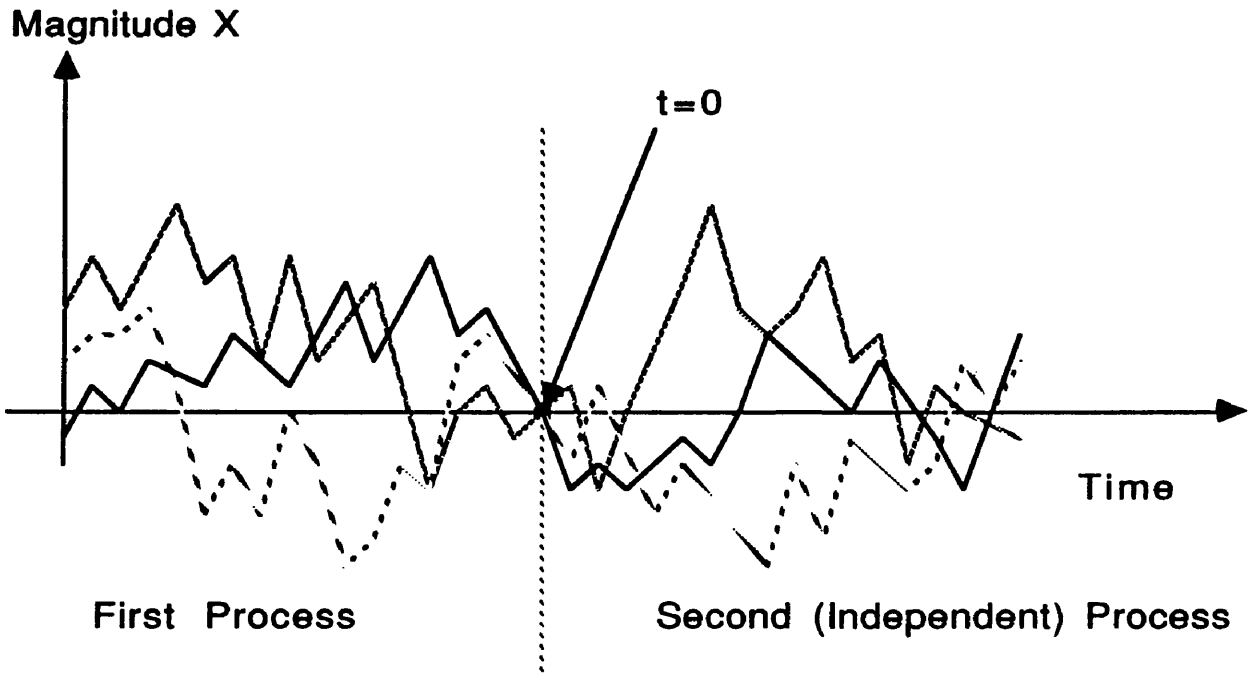


Fig. 136: Half Processes

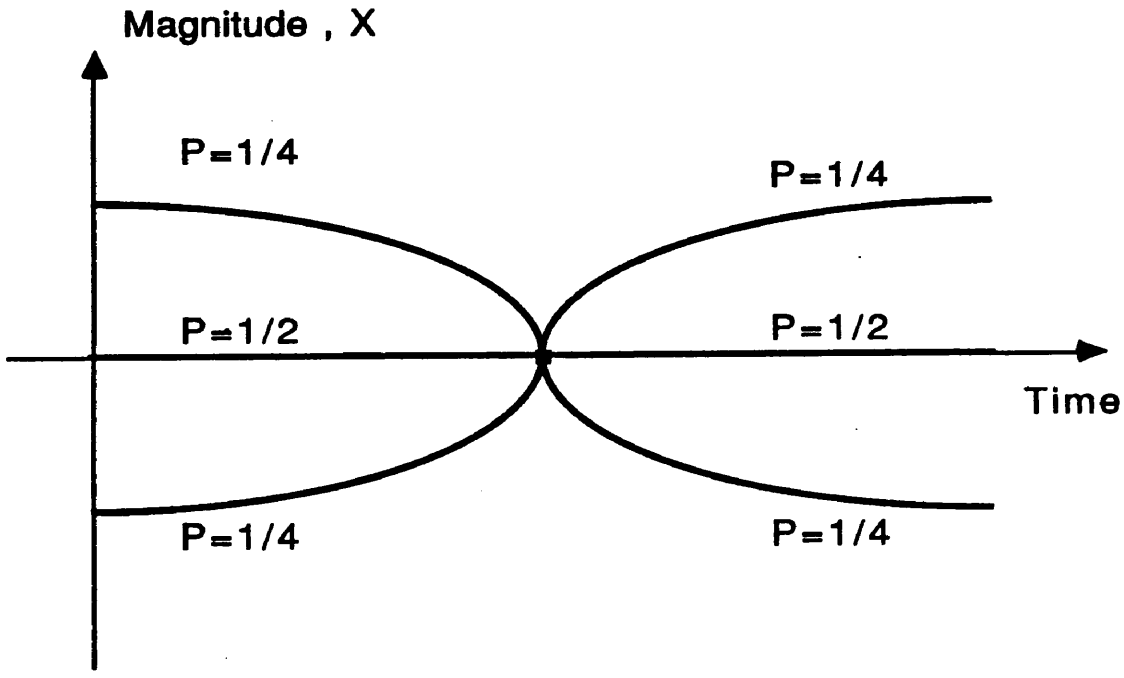


Fig. 137: Finite Process Example

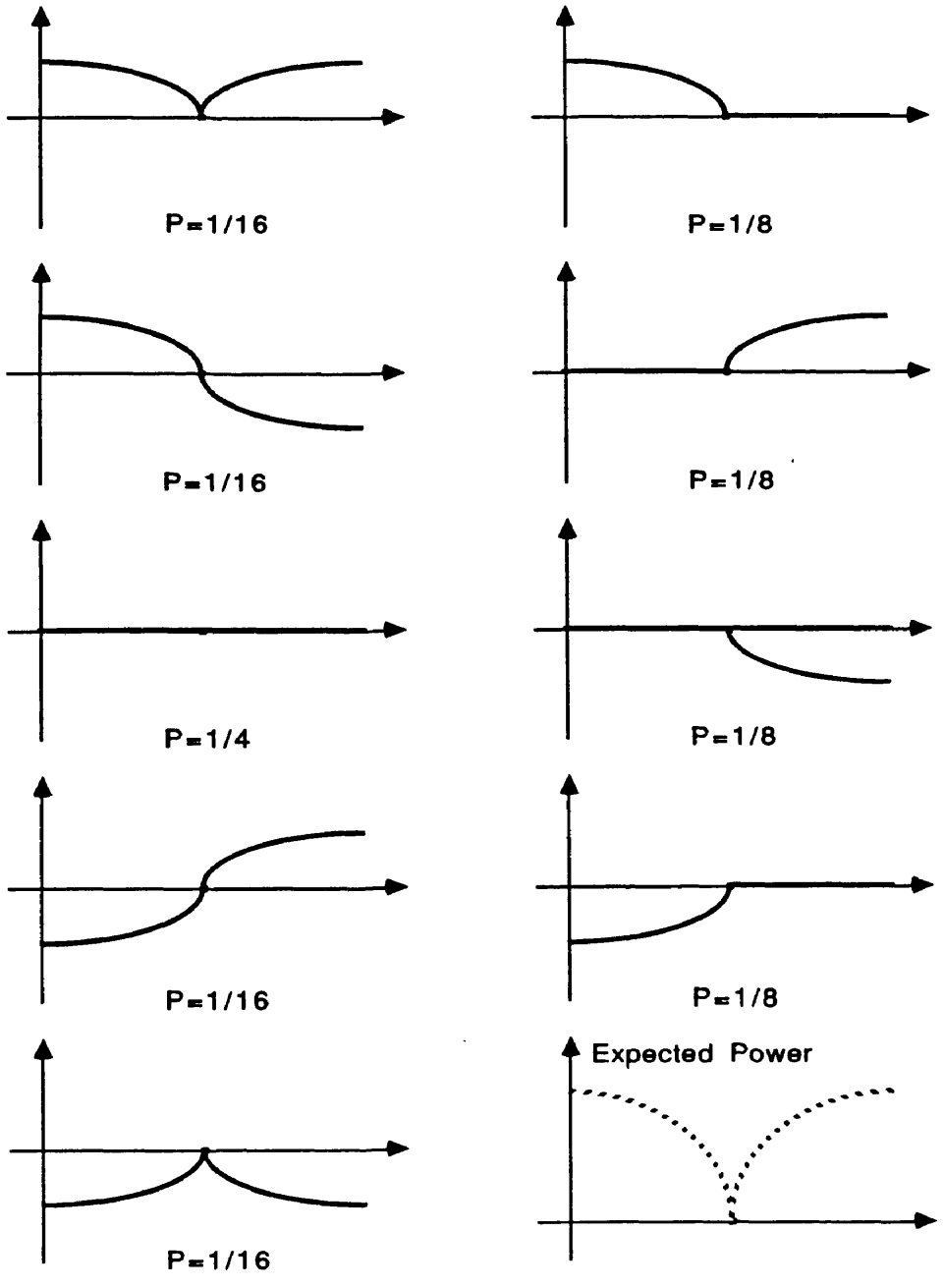


Fig. 138: Combinations of Samples in the Finite Process

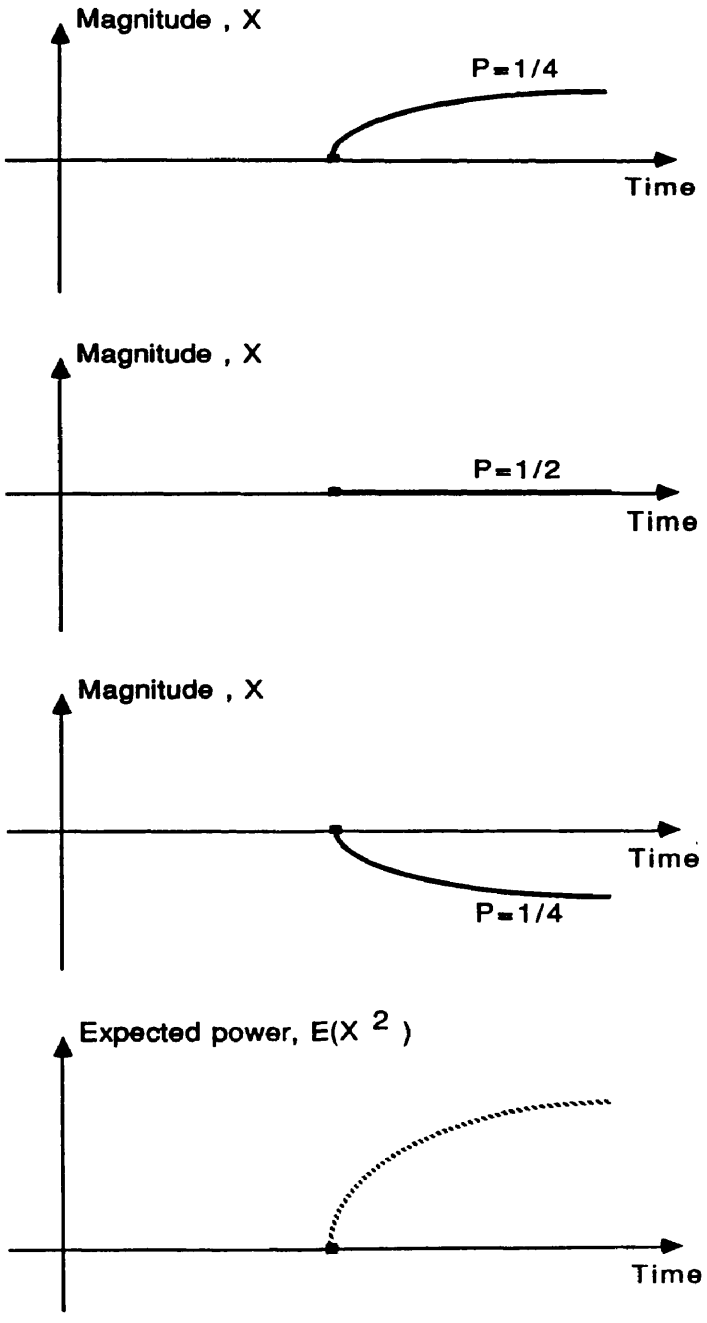


Fig. 139: Positive Half Samples

The second method can be generalized. We could consider a stochastic process which is only defined for positive times, and estimate the variance (expected noise power) as a function of time, over each sample. In the procedure that follows, every positive half sample will have exactly one associated negative half sample which is the symmetric extension (odd symmetry). However, this does not make any difference, since only the positive half of each sample is considered.

We will define the process as follows. We assume that the unrestricted noise process is known, and defined by its power spectrum. The magnitudes of the power spectrum (proportional to the expected magnitude of Fourier coefficients) over frequency are known (deterministic). All phase values of the Fourier coefficients are assumed to be random, uniformly distributed and statistically independent (no correlation between phase values at different frequencies). The probability distribution of each complex Fourier coefficient obviously exhibits circular symmetry in the complex plane, as shown on figure 140.

The restricted process we will consider is derived from the unrestricted process by taking all the samples that have a phase value of $\pm\pi/2$. This corresponds to taking a cross-section along the imaginary axis in the complex plane (figure 141).

As a result, this new process will only have imaginary Fourier coefficients ($I_0(\omega)$), and each sample can be written as an infinite sum of sine waves. Each sample will have odd symmetry in the time domain. The expected energy spectrum (squared Fourier coefficients) of this process will be equal to the expected energy spectrum of the original, unrestricted process (taking the cross-section does not alter the expected magnitude of the coefficients, due to the circular symmetry). The phase values at different frequencies are still uncorrelated ($\pm\pi/2$), and the expected complex value of each coefficient is still zero. A power spectrum cannot be defined, since the process is not WSS.

As a result of taking the cross-section, the following relations still hold.

$$E(X_0(\omega_1)X_0^*(\omega_2) = E(I_0(\omega_1)I_0(\omega_2)) = \frac{1}{2\pi}\delta(\omega_1 - \omega_2)\ S(\omega_2) \qquad (214)$$

$$E(|X_0(\omega)|^2) = E(I_0^2(\omega)) = \frac{1}{2\pi}\delta(0)\ S(\omega) \qquad (215)$$

This new process has the following properties:

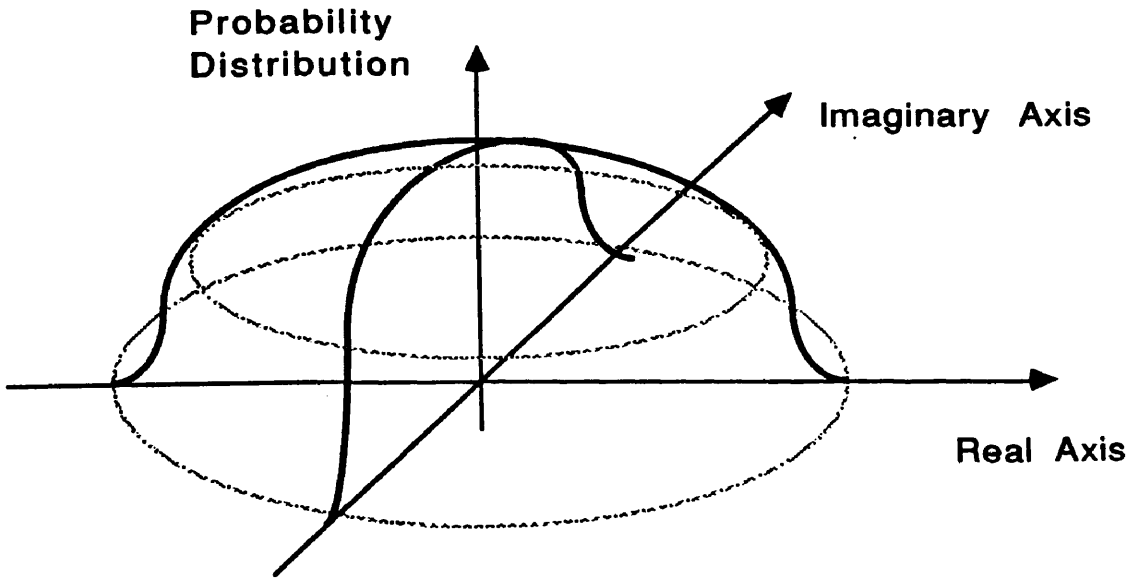


Fig. 140: Circular Symmetry of Fourier Coefficients

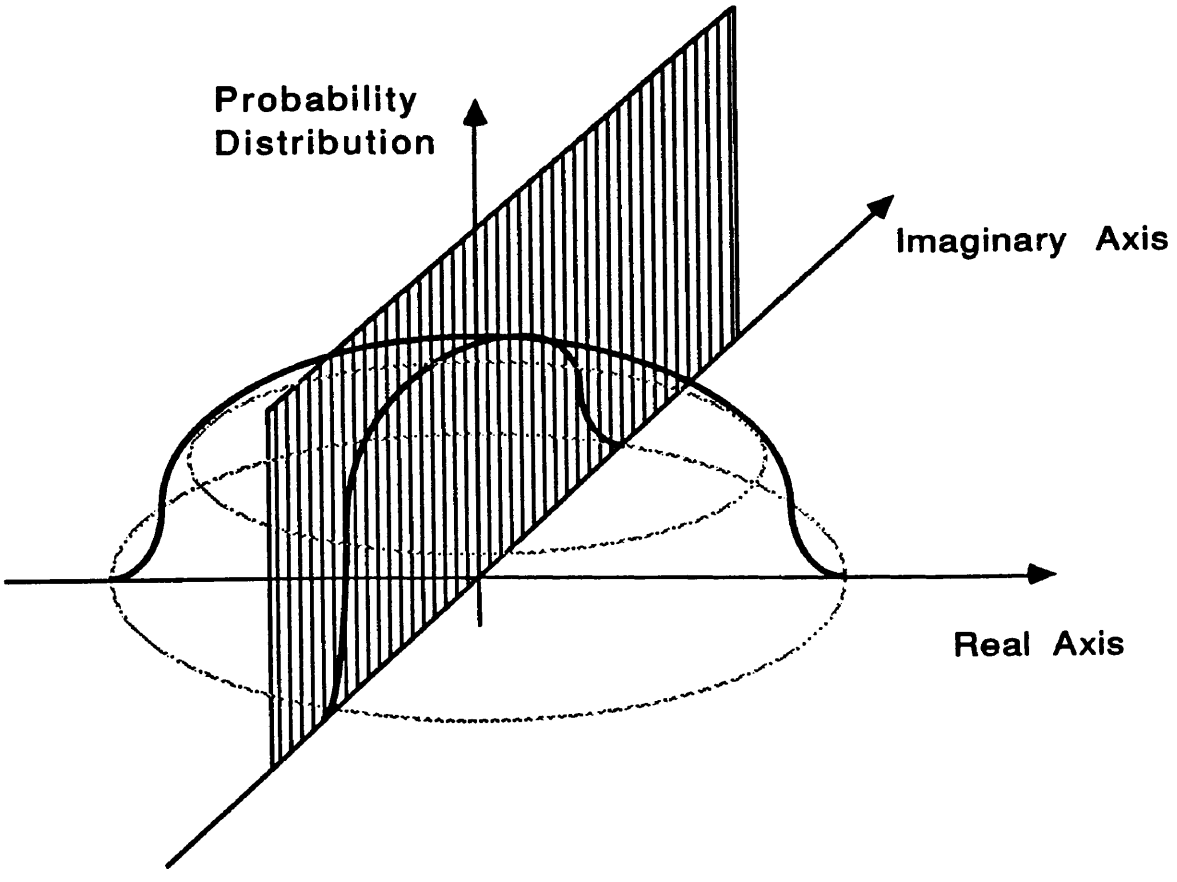


Fig. 141: Cross-Section of Fourier Coefficients

1. At a time of zero, the value of any sample is zero ($x_0(0) = 0$).
2. The energy spectrum is the same as the energy spectrum of the original process, and by extension also the same as the simple (without odd symmetry) restricted process.
3. Each sample of this odd-symmetry process corresponds to a whole class of samples of the original process, being all the samples with a given energy spectrum and either positive *or* negative phase ($0 \leq \phi < \pi$ or $-\pi \leq \phi < 0$). Each sample of the original process belonging to the same class, had the *same probability of occurrence*, since the phase distribution was uniform (the energy spectrum, or the squared magnitudes of the Fourier coefficients do not have a uniform distribution).

As a result of the third property, it seems plausible to conclude that for positive times, the statistics of the instantaneous noise power are the same for either the restricted process or the process containing only odd-symmetry samples (imaginary Fourier coefficients). From a strict mathematical point of view, this is not quite *exact*. However, deriving an exact formula to relate the statistics in both cases, would be extremely tedious and impractical.

For the remainder of this analysis, we will calculate the statistics of the odd-symmetry process, which have a much easier mathematical derivation. This is an approximation, whose validity will not be demonstrated here. It is based on strong intuition, and the fact that the results based on this assumption make great physical sense. In chapter XIII however, the results of this chapter will be related to exact mathematical formulas derived to describe the effect of correlated double sampling. At that time, the consistency between the two approaches will be demonstrated in detail. From that comparison, the implications of the approximation made here will become apparent.

12.7. Expected Noise Power in Switched Systems

Any sample $x_0(t)$ for which $x_0(0) = 0$ can be written in an entirely general way so that for $t \geq 0$

$$x'_0(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} jI'_0(\omega)\, e^{j\omega t}\, d\omega = \frac{1}{\pi} \int_0^{\infty} I'_0(\omega)\, \sin(\omega t)\, d\omega \qquad (216)$$

We can now calculate $E(P(T))$ in a manner similar to the calculation of $E(P(t))$ for the stationary process, in chapter XI.

$$E(P(T)) = E((\frac{1}{\pi}\int_0^\infty I_0'(\omega)\,\sin(\omega T)\,d\omega)^2) \tag{217}$$

$$E(P(T)) = \frac{1}{\pi^2}\int_0^\infty\int_0^\infty E(I_0'(\omega_1)I_0'(\omega_2))\,\sin(\omega_1 T)\sin(\omega_2 T)\,d\omega_1\,d\omega_2 \tag{218}$$

$$E(P(T)) = \frac{2}{\pi}\int_0^\infty\int_0^\infty S(\omega_2)\,\delta(\omega_1-\omega_2)\,\sin(\omega_1 T)\sin(\omega_2 T)\,d\omega_1\,d\omega_2 \tag{219}$$

$$E(P(T)) = \frac{2}{\pi}\int_0^\infty S(\omega)\,\sin^2(\omega T)\,d\omega \tag{220}$$

$$E(P(T)) = \frac{1}{\pi}\int_0^\infty S(\omega)\,(1-\cos(2\omega T))\,d\omega \tag{221}$$

This is the important conclusion of our discussion about finite-time noise. If we assume that when a current source, like the ones used in the charge amplifier, is switched ON, the noise gradually builds up form zero, the expected value of the noise power after a finite time T can be calculated by integrating the power spectrum $S(\omega)$ from 0 to ∞, after multiplication by the function $(1/\pi)(1-\cos(2\omega T))$.

This is not very different from the traditional procedure, which consists of integrating the spectrum itself from 0 to ∞, (after multiplication by $1/\pi$). Such straight integration yields the expected power under steady-state conditions (which is independent of time for a circuit with fixed biasing point). The new formula that was just derived is a generalization of the same concept, which allows to calculate the expected noise power, a finite time T after the noise source came on.

The multiplying function $(1/\pi)(1-\cos(2\omega T))$ becomes significantly smaller than unity for $\omega << (2\pi)/(4T)$ (figure 142). This agrees with the intuitive

perception that in systems that are ON for a limited time only, the effect of low-frequency noise components is strongly reduced. The smaller T becomes, the more pronounced the effect. In the limit, for $T \rightarrow 0$, all noise components of finite frequency will be rejected, since for finite ω, $\cos(2\omega T)$ approaches unity and the multiplying function approaches zero. This agrees with the fact that the noise needs a finite time to build up when it starts out at a value of zero.

In the limit, for $T \rightarrow \infty$, one would anticipate the expected power $E(P(T))$ to be equal to $E(x^2(t)) = R(0)$, the expected power of the stationary process that corresponds to the *steady-state* noise. This can easily be confirmed by rewriting the expression for $E(P(T))$.

$$E(P(T)) = \frac{1}{\pi}\int_0^\infty S(\omega)\, d\omega - \frac{1}{\pi}\int_0^\infty S(\omega)\cos(2\omega T))\, d\omega \qquad (222)$$

$$E(P(T)) = R(0) - \frac{2}{\pi}\int_0^\infty S(\omega/2)\cos(\omega T)\, d\omega \qquad (223)$$

The first term is the expected power of the steady-state process. The second term has the appearance of an inverse Fourier transform (since $S(\omega/2)$ is an even function of ω, the complex exponential terms reduce to cosines only; the sine terms cancel out). If $S(\omega)$ is a band-limited spectrum, the integral will tend to zero for large values of T. This can easily be proven using Parceval's Theorem. If $\int_{-\infty}^{\infty} S^2(\omega/2)\, d\omega$ is finite, then the square of $F^{-1}(S(\omega/2))$ is integratable over the time axis ($\int_{\infty}^{\infty} F^{-1}(S(\omega/2))\, dT is finite$), and hence the time function must tend to zero for large values of T).

Not only does equation (221) make physical sense, it also bears a significant resemblence with a formula that will be derived in chapter XIII, to quantify the effect of correlated double sampling on the expected noise power. The implications of this similarity will be high-lighted at the end of that chapter.

12.8. Conclusion

In this chapter, a closed mathematical formula was derived for the calculation of the expected noise power of a switched system. The assumption was that every time the system is switched OFF, the noise is reset to zero (since there is no current

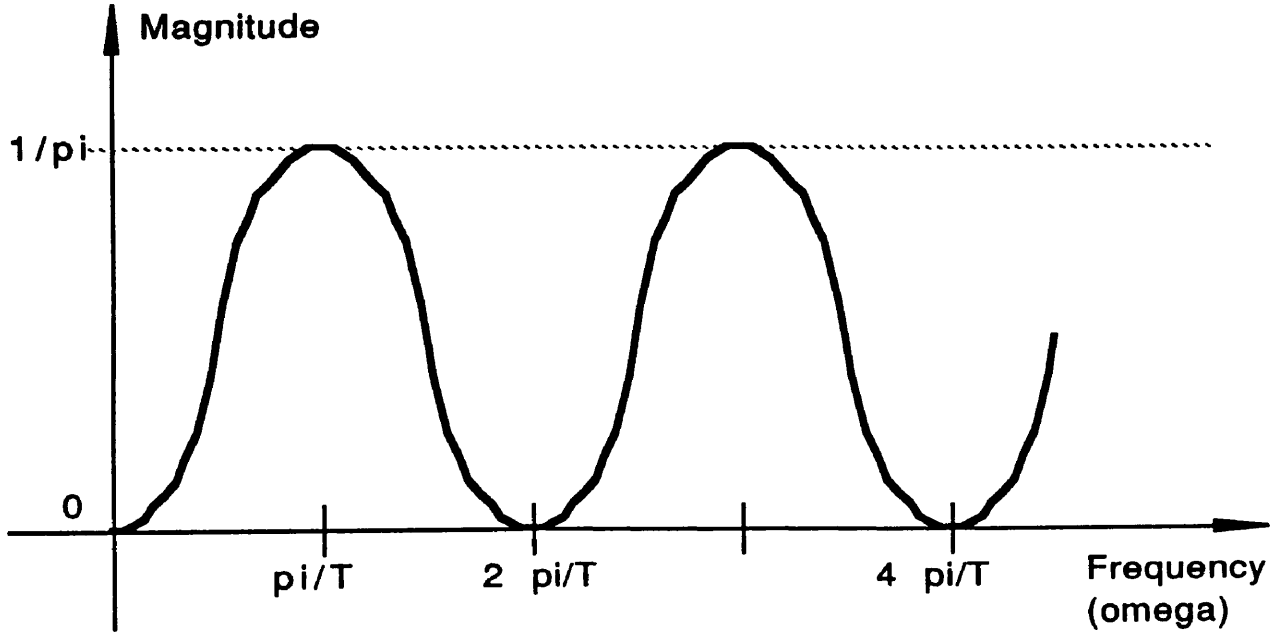


Fig. 142: Multiplying Function

in any of the devices). When the system is switched ON again, the noise starts out at zero, for continuity reasons, and then gradually builds up for a certain time T. The expected instantaneous noise power at that time T can be calculated by integrating the power spectrum of the noise (defined for steady-state conditions), after multiplication with a frequency-dependent correcting function.

CHAPTER XIII

CORRELATED DOUBLE SAMPLING AND NOISE

13.1. Introduction

In the previous section, we have seen how the influence of low-frequency noise components is reduced in switched systems, at least if it can be assumed that the switching periodically resets the noise. This section deals with a similar effect, although it relies on a different principle. What will be described here is how correlated double sampling (CDS) cancels low-frequency components in a signal by forming a difference between two samples of the signal over a finite time. The effect of CDS on the noise spectrum can easily be quantified, according to a formula similar to the one used to estimate finite-time noise.

As opposed to the analysis of finite-time noise, this analyis is mathematically exact. In addition, it adequately describes how CDS affects noise performance, and it is simpler than other analyses reported in literature [98, 99, 100].

13.2. Correlated Double Sampling in the Charge Amplifier

Correlated double sampling can be applied to sampled-data systems in general, in order to reduce the effect of systematic offsets or low-frequency noise that would otherwise be introduced by the sampling device. Usually, the sampling device is a sample/hold amplifier (e.g. designed in switched-capacitor technology), although in this case we will eventually apply the theory to the charge amplifier. In particular, the technique makes the gain stage largely immune to offsets and noise inherent to the comparator. Without CDS, the pulse width of the transmitted signal would be subject to time jitter, which in turn translates into a random (noise) charge on the holding capacitor of the next stage. This jitter becomes increasingly important at higher sampling rates (shorter pulse widths), and CDS should be considered an integral part of the operation of the charge amplifier.

13.3. Principles of Correlated Double Sampling

In order to quantify the effect of CDS, one could conceptually picture the sampled data system as shown in figure 143. A useful signal $u(t)$ is to be sampled

by the sampling device at regular time intervals t_i, with $t_i = i\ T$, and i an integer value between $-\infty$ and ∞. The sampling device has a certain gain A and a finite input-referred systematic offset os, both of which are considered time-independent. The sampling device also introduces a certain amount of noise, which we will assume to be stationary.

As described in earlier chapter XI, the assumption of stationarity implies that the statistical properties of the noise do not change over time, which in turn implies that the sampling device has a constant bias point and is not a switched system. If this condition is met, the noise can be described by a continuous power spectrum.

Although the noise is generated inside the sampling device, its effect can be modeled as the addition of an input-referred noise signal $x(t)$ to the input of the device. As a result, the sampling action can be described by the following equation (s_i is the sampled value, at the output of the device at time $t_i = i\ T$).

$$s_i = (u(t_i) + x(t_i) + os)\ A \tag{224}$$

The idea of CDS is to reduce the effect of the noise $x(t)$ by sampling twice instead of just once. At time $t_i - T'$, the sampling device is triggered with zero, (shorted) input. Obviously, the corresponding output would be

$$s_i' = (x(t_i - T') + os)\ A \tag{225}$$

At time t_i, the actual input signal $u(t)$ is sampled, resulting in the output value predicted by equation (224). CDS relies on forming the difference between s_i and s_i', and using that value as the final sampled value, instead of simply s_i. The timing is summarized in figure 144.

Of course, the mechanism under which the difference is formed, must be such that no additional noise is introduced, but this can be accomplished in practice. In the charge amplifier, the difference is formed implicitly by the combination of initial and final discharge phases.

The effective sampled value at time t_i, referred to as s_i", is the difference between the two samples.

$$s_i" = s_i - s_i' = (u(t_i) + (x(t_i) - x(t_i - T')))\ A \tag{226}$$

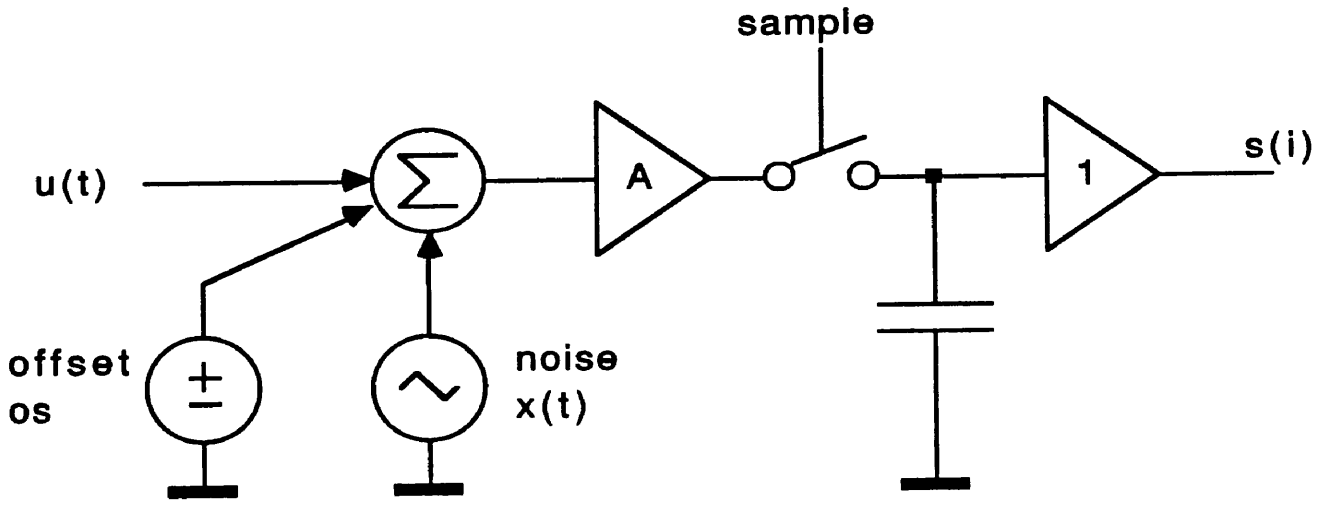


Fig. 143: Concept of Correlated Double Sampling

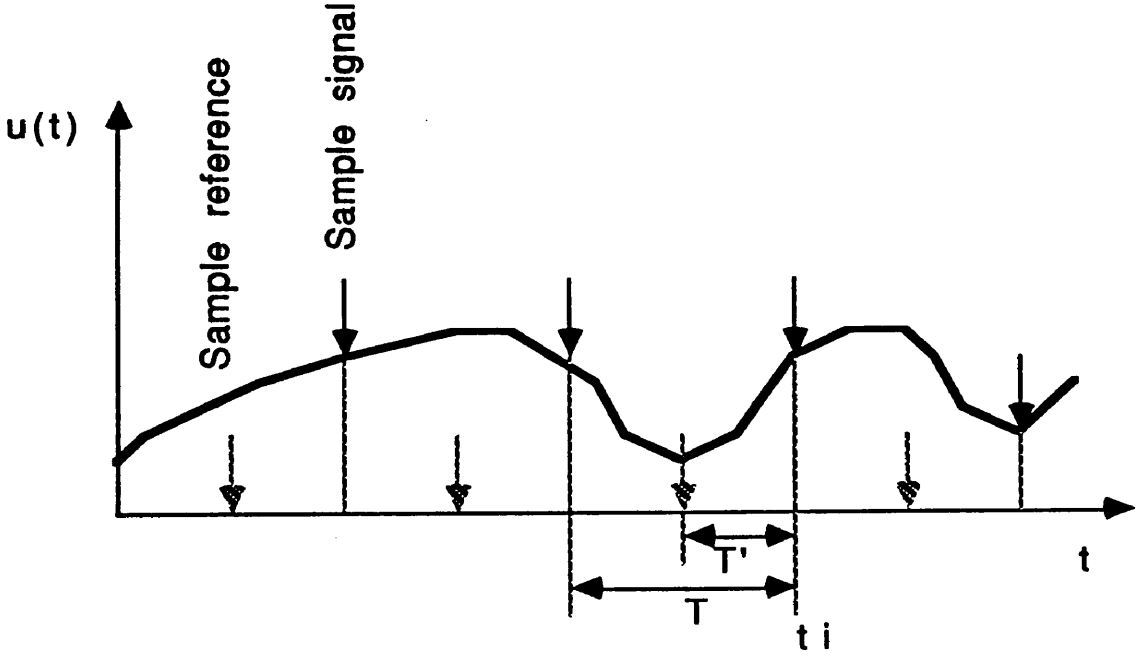


Fig. 144: Timing of the Samples

It is clear that the input-referred offset term *os* has vanished, which was desired. In addition, we will now show that low-frequency components in the noise signal $x(t)$, are strongly attenuated. Equation (226) contains the term $(x(t_i) - x(t_i - T'))$, which is the difference between the input-referred noise signal, sampled at $t = t_i$ and at $t = t_i - T'$.

Although we are looking at a sampled-data system, and the noise is only considered at discrete times, we have originally made the assumption that the noise was stationary, or time-continuous. This means that despite the fact that we only sample the noise at discrete times, the noise process occurs in an undisturbed way even when it is *not* sampled. For a switched-capacitor S/H amplifier, this could mean that the opamp remains ON between samples. For the charge amplifier, this would mean that the fast comparator (main contributor to the time jitter on the pulse width) remains ON, under approximately constant bias conditions. Under those conditions, the statistical concepts (including power spectrum and expected power) introduced earlier, can still be applied.

13.4. Noise Cancellation

The expected noise power at $t = t_i$ is $E(P(t_i)) = E(x^2(t_i))$, and obviously equal to the expected noise power at $t = t_i - T'$. The exact value can be found by integrating the power spectrum of $x(t)$ for ω between $-\infty$ and ∞. However, we are not interested in either value, but in the expected value of $(x(t_i) - x(t_i - T'))^2$. In order to calculate this, we will consider a fictitious, stationary noise process $x_{T'}(t)$, defined as follows:

$$x_{T'}(t) = x(t) - x(t - T') \tag{227}$$

Since $x(t)$ is stationary, obviously $x(t - T')$ and $(x(t) - x(t - T'))$ will be stationary processes, as well and as a result a power spectrum can be defined for the process $x_{T'}(t)$.

The autocorrelation function of $x_{T'}(t)$ is given by

$$R'(\tau) = E(x_{T'}(t)\, x_{T'}(t+\tau)) = E((x(t) - x(t-T'))\,(x(t+\tau) - x(t-T'-\tau))) \tag{228}$$

$$R'(\tau) = E((x(t) - x(t - T'))\,(x_(t + \tau) - x(t - T' - \tau))) =$$

$$E(x(t)\ x(t+\tau)) + E(x(t-T')\ x(t-T'+\tau)) -$$
$$E(x(t)\ x(t-T'+\tau)) - E(x(t-T')\ x(t+\tau)) \tag{229}$$

$$R'(\tau) = 2R(\tau) - R(\tau - T') - R(\tau + T') \tag{230}$$

With $R(\tau)$ the autocorrelation function of $x(t)$.

Since the power spectrum of $x_{T'}(t)$ is defined as the Fourier transform of $R'(\tau)$, it will be the sum of three components, each related to the original spectrum of $x(t)$ (a shift of the time axis by T' represents a phase shift of $j\omega$ in the Fourier domain). The new power spectrum $S'(\omega)$ can be written in terms of the original spectrum $S(\omega)$.

$$S'(\omega) = 2S(\omega) - e^{-j\omega T'}\ S(\omega) - \epsilon j\omega T'\ S(\omega) \tag{231}$$

$$S'(\omega) = S(\omega)\ (2 - \epsilon - j\omega T' - \epsilon j\omega T' = 2S(\omega)(1 - \cos(\omega T')) \tag{232}$$

Integration of the power spectrum of $S'(\omega)$, for ω ranging from $-\infty$ to ∞, yields the expected power of $x_{T'}(t)$. By definition, this expected power is equal to

$$E(P(t)) = E((x(t) - x(t-T'))^2) \tag{233}$$

This expression is independent of time. It is clear that it reflects the expected squared error due to the (double) sampling process, or the variance (σ^2) on the sample. (If the sampling device had non-unity gain, the output-referred variance would be $A^2\ \sigma^2$).

$$\sigma^2 = \frac{2}{\pi} \int_0^\infty S(\omega)(1 - \cos(\omega T'))\ d\omega \tag{234}$$

A common misconception in the analysis of sampled-data systems is that one would have to calculate the *spectrum* of the actual sampled signal. This would involve convolution of the noise spectrum of the sampling device with the spectrum of the sampling function (e.g. a pulse train). This procedure would be the correct one if one wants to use the sampled *analog* output of the system directly (e.g. to apply to a low-pass smoothing filter). In the case of ADC's however, the sample

is eventually converted to a single, discrete value. What is relevant from a signal processing point of view, is not the spectrum of the sample, but the expected error at the instant of sampling (σ^2).

Equation (234) is very similar to the formula derived in chapter XII, to estimate the expected power of switched noise, a finite time after the switching. Just like switching could reduce low-frequency noise components, it is clear from equation (234), that noise components at frequencies ω such that $\omega << \pi/T'$, will be strongly attenuated by CDS. Again, in the limit, one could demonstrate that for $T' \rightarrow 0$, σ^2 reduces to 0, which expresses that if the correlation time of the CDS becomes very small, more and more low frequency components of the noise are canceled. If $T' \rightarrow \infty$, σ^2 reduces to $2/\pi \int_0^\infty S(\omega)\, d\omega$, which expresses that for long correlation times (or large noise bandwidth), the two consecutive samples become uncorrelated and the expected value of the squared error actually doubles compared to the single-sampling case.

This means that CDS does not *always* result in noise reduction. It only works if a significant portion of the total noise power is made up of components at frequencies $\omega << \pi/T'$. This is typically the case of $1/f$ noise in MOSFETs that are part of S/H amplifiers. It certainly is the case of charge amplifiers switching at high frequencies (MHz rates).

13.5. Correlated Double Sampling and Switched Noise

Although derived in a different way, to describe a different phenomenon, equation (234) bears a remarkable resemblance with the expression derived in chapter XII to describe the effect of switching upon the noise. Under the (unproven) assumption that the noise is reset to zero each time by the switching, the expression for the expected noise power after a finite time T was found to be:

$$\sigma^2 = E(P(T)) = \frac{1}{\pi} \int_0^\infty S(\omega)(1 - \cos(2\omega T'))\, d\omega \qquad (235)$$

The differences between equation (234) and (235) are two factors 2. One occurs before the integral sign of equation (234) and expresses the fact that in a CDS scheme, two values are subtracted, resulting in a possible variance on the resulting value of the sample of up to twice the variance of a single sample (for

long correlation times). The second factor 2 is in the cosine argument of (235), and *not* in the cosine argument of (234). This peculiarity deserves some closer attention.

The cosine argument $2\omega T$ of (235), versus $\omega T'$ in (234), seems to imply that more low frequency noise components would be rejected by a CDS scheme with a correlation time t, than by a sampling device switched OFF and then switched back ON for an identical time t. This is actually consistent, as will be shown next. The variance between the noise at time $-t$ and the noise at time t can be written in an entirely general way as

$$E((x(t) - x(-t))^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, -t, x_2, t)\ (x_2 - x_1)^2\ dx_1\ dx_2 \qquad (236)$$

With $p(x_1, -t, x_2, t)$ the joint probability density of a value x_1 occurring at time $-t$ AND a value x_2 occurring at time t.

The same quantity can be expressed using (234) (formula for correlated double sampling), since for stationary (ergodic) noise, the expected value of the product is only dependent upon the time *difference* between the two sampling instants. Clearly, in this case the time difference is $2t$.

$$E((x(-t) - x(t))^2) = \sigma^2_{CDS,2t} = \frac{2}{\pi} \int_0^{\infty} S(\omega)(1 - \cos(\omega\ 2t))\ d\omega \qquad (237)$$

Using Poisson's rule, it can also be written generally in a way that involves the value of the noise at time 0.

$$E((x(-t) - x(t))^2) =$$

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1, -t, x_2, t)|(x_0, 0)\ (x_2 - x_1)^2\ dx_1\ dx_2 \right)\ p(x_0, 0)\ dx_0 \qquad (238)$$

With $p(x_1, -t, x_2, t)|(x_0, 0)$ the joint *conditional* probability density of a value x_1 occurring at time $-t$ AND a value x_2 occurring at time t, given that the value of the noise at time $t = 0$ is x_0. The probability density of a value x_0 occurring at time 0 is expressed by $p(x_0, 0)$.

Because the noise is assumed to be a Markoff process, for a *given* value of x at time 0, the conditional values of x at times t and $-t$ become independent statistical variables.

$$E((x(-t)-x(t))^2) =$$

$$\int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1,-t)|(x_0,0)\, p(x_2,t)|(x_0,0)\, (x_2-x_1)^2\, dx_1\, dx_2 \right) \quad p(x_0,0)\, dx_0 \tag{239}$$

The value of the noise at $t = 0$ can be introduced into this expression:

$$E((x(-t)-x(t))^2) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1,-t)|(x_0,0)\, p(x_2,t)|(x_0,0) \right.$$

$$\left. ((x_2-x_0)+(x_0-x_1))^2\, dx_1\, dx_2 \right) \quad p(x_0,0)\, dx_0 \tag{240}$$

$$E((x(-t)-x(t))^2) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x_1,-t)|(x_0,0)\, p(x_2,t)|(x_0,0) \right.$$

$$\left. ((x_2-x_0)^2+(x_0-x_1)^2+2\,(x_2-x_0)(x_0-x_1))\, dx_1\, dx_2 \right) \quad p(x_0,0)\, dx_0 \tag{241}$$

Since x_1 and x_2 are independent statistical variables, this can be simplified to:

$$E((x(-t)-x(t))^2) = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} p(x_1,-t)|(x_0,0)\, (x_0-x_1)^2\, dx_1 + \right.$$

$$\left. \int_{-\infty}^{\infty} p(x_2,t)|(x_0,0)\, dx_2 \; + \; 0 \right) \quad p(x_0,0)\, dx_0 \tag{242}$$

The second integral in the parentheses represents the expected variance of the noise at time t, given a known value x_0 at time 0. The first integral represents the variance of the noise at time $-t$ under the same conditions. Because of the symmetry of the autocorrelation function, as well as the Markoff property of the noise, it is clear that both integral should have the same value. Expression (242) then reduces to:

$$E((x(-t)-x(t))^2) = 2 \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} p(x_2,t)|(x_0,0)\, dx_2 \right) p(x_0,0)\, dx_0 \tag{243}$$

In many cases, the amplitude distribution of the noise process is such that $p(x_0, 0)$ is only significant for values of x_0 that are close to 0. Under those conditions, equation (243) can be approximated by:

$$E((x(-t) - x(t))^2) \approx 2 \int_{\infty}^{\infty} p(x_2, t)|(0,0)\ dx_2 \tag{244}$$

It is clear that this is an expression for twice the variance of the noise at time t, with the restriction that its value is 0 for time 0. This is precisely the condition we introduced in chapter XII, to analyze the expected power of switched (finite-time) noise. As a result, the variance between the noise at time $-t$ and the noise at time t can be written as

$$E((x(-t) - x(t))^2) \approx 2\ E(P(t)|(x(0) = 0)) \tag{245}$$

Or, using the expression for switched noise derived in chapter XII:

$$E((x(-t) - x(t))^2) \approx 2\ \frac{1}{\pi} \int_{0}^{\infty} S(\omega)(1 - \cos(2\omega t))\ d\omega \tag{246}$$

This expression is identical to expression (237), which was derived using the formula for CDS (equation (234)) directly. This conclusion demonstrates the consistency between the formulas for the expected power of switched noise and for CDS. It also confirms the assumptions made while deriving the formula for switched noise in chapter XII.

13.6. Conclusion

This chapter briefly discussed the use of correlated double sampling in order to limit low frequency noise (or alternatively: the time jitter on the pulse width) in the charge amplifier. It started with a general description of the technique. It was shown how a systematic offset in the sampling device is cancelled. An expression for the expected noise power of the sampled signal was calculated, taking into account the effect of correlated double sampling over a certain correlation time T'. It was shown how low frequency noise components are rejected, while high frequency components may actually get amplified by a factor of 2. Finally, it was pointed out how the correlated double sampling formula bears significant resemblance with the formula derived to calculate the effect of finite-time noise in switched systems (chapter XII).

CHAPTER XIV

NOISE IN THE CHARGE AMPLIFIER

14.1. Introduction

Chapter X contained a detailed analysis of the charge amplifier. It was pointed out that estimating the noise in the structure, is not straight-forward. The noise (especially the low-frequency components of the $1/f$ noise) is strongly influenced by the correlated double sampling, and possibly by the switching. The influence of these effects was investigated from an entirely general perspective in chapters XI through XIII. On top of the low-frequency noise, there also exist several mechanisms by which white noise can affect the accuracy of the charge amplifier.

This chapter will systematically analyze all the noise effects within the scheme. It should be stressed that the noise models and calculation techniques that are proposed in this chapter, do not aim at being perfectly accurate. It would actually be almost impossible to develop a comprehensive noise model for a complex, switched system like the charge amplifier. The considerations that are presented, are based on simplified assumptions, which have the purpose of developing a qualitative feeling for the important noise issues, as well as the means to come up with a rough but reasonable quantitative estimate of the noise performance.

14.2. Charge Amplifier Components

Figure 145 shows the main components in the charge amplifier scheme. The situation shown implements one stage, which takes a pulse width modulated (PWM) input signal and produces a PWM output. The gain of the system is -2, which is typical of the accuracy-bootstrapped stages to be used in a prototype, pipelined ADC. The situation of figure 145 is slightly simplified, in the sense that the flash and DAC sections of the stage are not shown, for the sake of simplicity.

One can easily verify that the flash section normally does not contribute to the total noise within a pipelined ADC, provided enough redundancy is built into the scheme. The DAC section however, can be an important contributor. In

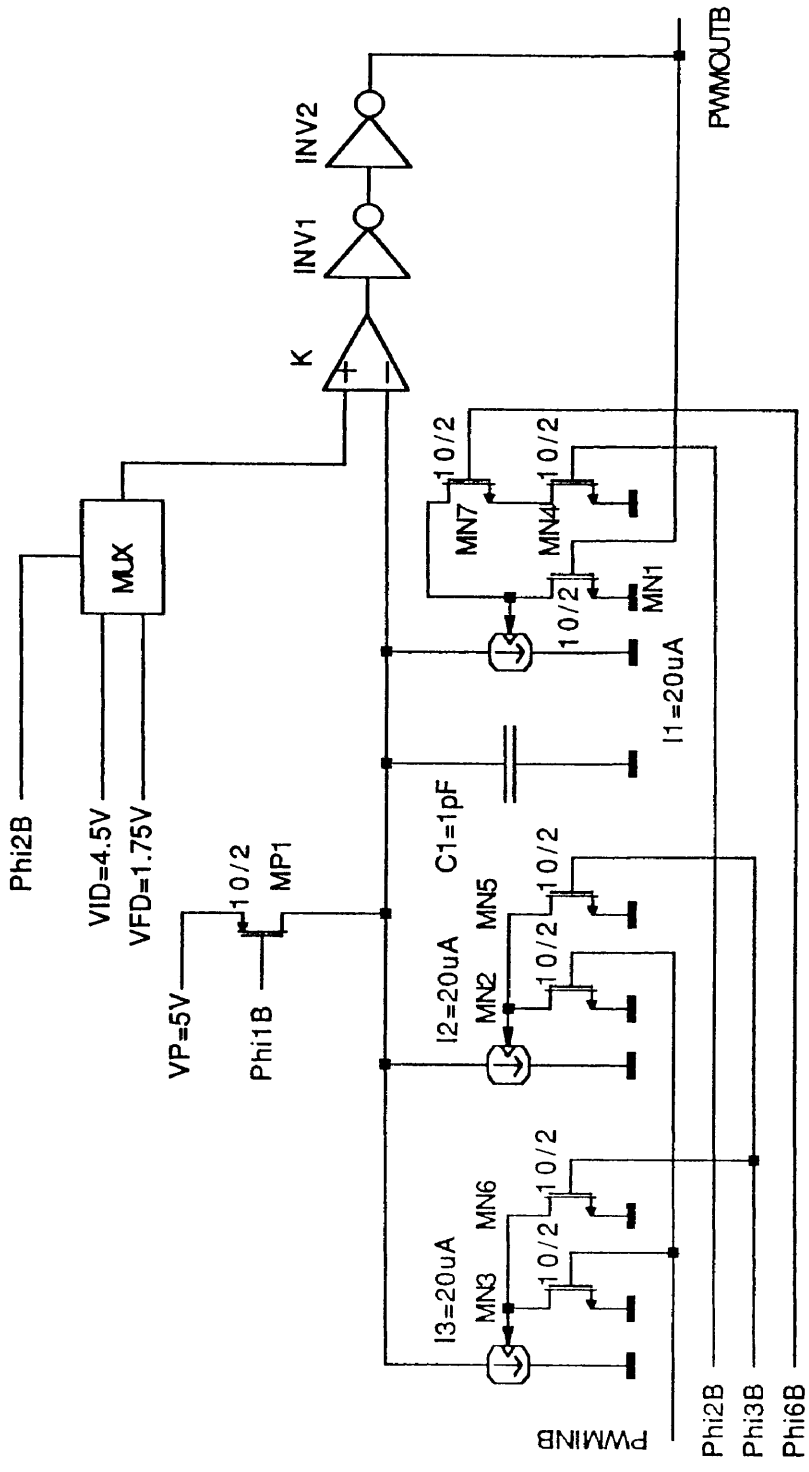


Fig. 145: Charge Amplifier Components

chapter XV, it will be seen how the DAC section can be implemented according to the charge amplification philosophy. It will become apparent then that the mechanisms by which the DAC section injects noise into the system are the same as the mechanisms by which the gain operation generates noise. As a result, it should become clear how this situation can be generalized to the more complex case of a complete ADC stage. It should also be clear how the analysis can be extended towards different cases, e.g. with a different gain.

Operation of the charge stage shown in the example is as follows. During the first clock phase, Φ_1, the holding capacitor, C_1, is precharged to a fixed voltage V_p. This is accomplished by P-channel MOSFET switch M_{P1} and the external voltage source V_p. The gate of M_{P1} is controlled by the active low clock signal $\bar{\Phi}_1$ ($PHI1B$ on the figure).

During Φ_2 (the initial discharge phase), the holding capacitor is discharged to a voltage V_i, using comparator K, analog multiplexer MUX, reference voltage source V_{ID} and current source I_1. The current source is enabled when Φ_2 is active, or in this case when $\bar{\Phi}_2$ ($PHI2B$) is low.

A detailed schematic of the current source is shown on figure 146. The configuration is a typical regulated cascode, in this case implemented in CMOS technology (The discussion presented below could however be generalized to bipolar technologies). The N-channel transistor M_{N3} of figure 146 is used to disable the current source. Its equivalent in figure 145 is the active pull-down network formed by transistors M_{N1}, M_{N4} and M_{N7}. It is clear that this pull-down combination enables current source I_1 when either Φ_2 OR Φ_6 are active, AND at the same time, the voltage on the holding capacitor exceeds the reference level set by MUX. The clock signals are active low in this example, which results in the active pull-down transistors being cut OFF and the current sources being enabled when these signals are activated.

During Φ_3, the input is enabled. The PWM coded input signal is used to control current sources I_2 and I_3. Each one of the current sources is nominally identical to I_1, and implemented according to figure 146. Since the current sources are in parallel, their total current is nominally equal to twice the current of I_1, which is required in order to realize the gain of -2.

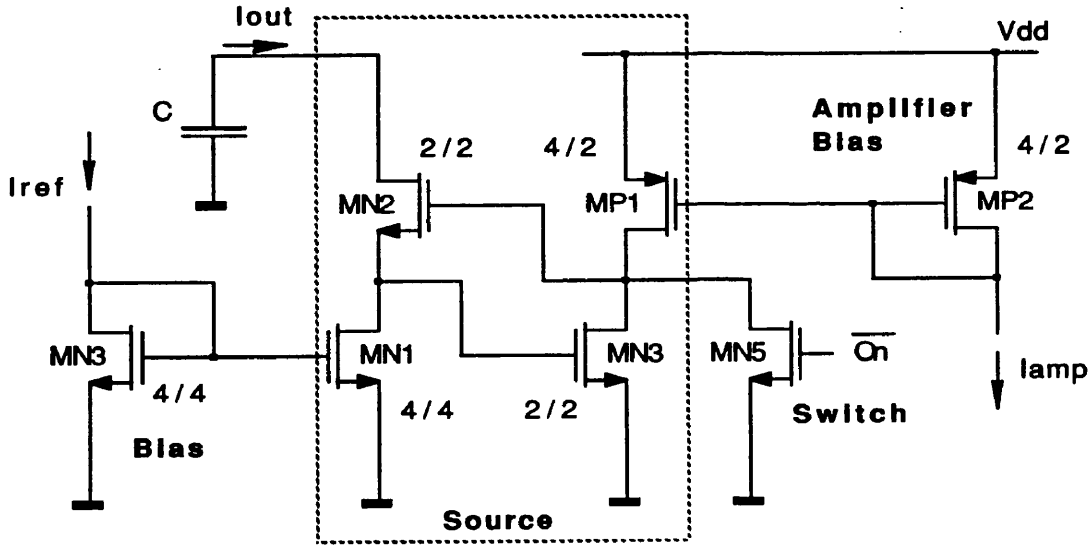


Fig. 146: Regulated Cascode

The active pull-downs of I_2 and I_3 are formed by transistors M_{N3} and M_{N6}, and M_{N2} and M_{N5}, respectively. It is clear that the current sources are enabled when Φ_3 is active AND the PWM-coded, active-low input signal $PWMINB$ is low. In this example, the PWM output signal as well as the input signal, are of the active low type. Such configuration makes it possible to use simpler logic (less inversions) to control the N-channel based current sources (actually current sinks), which in turn are faster and more practical than their P-channel counterparts.

According to earlier conventions, clock phases Φ_4 and Φ_5 are used for the flash (latching of comparator outputs) and DAC operations within an ADC stage. Accordingly, they will not be considered here.

Finally, during Φ_6, the holding capacitor is discharged down to the final discharge level V_f, again using comparator K, analog multiplexer MUX and current source I_1. This time, the multiplexer selects a reference voltage imposed by an external voltage source V_{FD}. A very simple CMOS implementation of the multiplexer is depicted in figure 147.

It is during this phase Φ_6 that the gain is realized. One can easily verify that the pulse width of the output (active low!) is nominally equal to a fixed amount (the offset, which is set by the combination of initial discharge and final discharge levels), minus two times the input pulse width. During the final discharge operation, current source I_1 is enabled while the voltage on the holding capacitor exceeds the reference level set by MUX. This can be verified by inspection of the active pull-down network formed by M_{N1}, M_{N4} and M_{M7} (figure 145).

14.3. Noise Sources

During this analysis, we will assume that the input PWM signal comes in noiseless. In other words, we will focus on the noise contribution of this amplifier stage only. This does not imply the assumption that the input signal would have infinitely sharp transitions (rise and fall times). It does however imply that the rise and fall portions of the waveform are perfectly repeatable (deterministic).

Noise is injected into the system from several sources. Essentially, every transistor in the system has the potential to be noisy. Since we are performing a noise analysis based on a MOSFET design, each transistor will contribute both white noise and $1/f$ noise. Fortunately for this scheme, all transistors are not ON

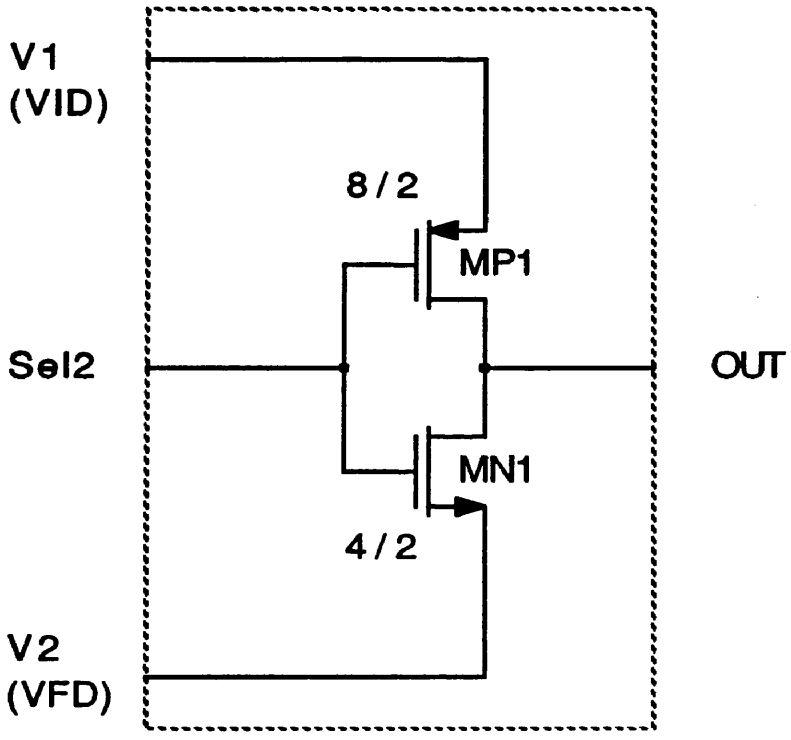


Fig. 147: Reference Voltage Multiplexer

all the time. We will assume throughout this discussion that when a transistor is OFF (currentless), its noise contribution is zero. As a result, different system components will contribute noise at different instants during the complete clock cycle.

The noise analysis that follows is based on the expected noise power concept, which has extensively been covered in earlier chapters. Essentially, each transistor (at least when ON) is characterized by the generation of a zero-mean, time-invariant (ergodic) noise current, which can be described by a deterministic power spectrum in the frequency domain. When a fixed bias point can be defined for a transistor, its current power spectrum can be calculated according to well-known formulas [7] for white ($S_{i,w}$) and $1/f$ noise ($S_{i,f}$).

$$S_{i,w} \approx \frac{8kT\ g_m}{3} \tag{247}$$

With k boltzmann's constant, T the absolute temperature and g_m the small-signal transconductance of the device.

$$S_{i,f} \approx K_f\ \frac{I_D^a}{f\ WL} \tag{248}$$

With I_D the DC drain current of the device, W the gate width, L the gate length, f the frequency and K_f a technology-dependent parameter. The exponent a is usually close to zero, and as a result the factor I_D^a is often omitted.

Under steady-state conditions, the noise currents of different transistors in the circuit are combined according to the frequency response of the circuit, based on the assumption that the total power spectrum (squared current per unit of frequency) is the sum of individual spectral contributions. From the total power spectrum, the expected instantaneous noise power can be derived by integration. The actual calculations involve sometimes complicated frequency responses, as well as the addition of many different noise contributions. As a result, no attempt will be made to derive closed algebraic formulas.

Instead, the analysis procedure will be demonstrated for the particular example at hand, using computerized circuit simulations (SPICE). The sizing of the devices, as well as the associated SPICE models that were used, can be

considered typical of a charge amplifier, optimized for integration in a typical, generic 2 micron

CMOS process. Such processes are offered at advantageous conditions to academic institutions by the MOSIS program, and are still widely used in industry for medium-speed, non-critical signal processing applications. The fact that the charge amplifier can successfully be used with such IC technologies, is a major advantage of the scheme.

The results from the frequency domain simulations will be related to theoretical considerations from the previous chapters, in order to account for the influence of the switching and the correlated double sampling upon the noise. However, approximations will have to be made, due to the complicated nature of the phenomena involved, in particular the absence of a fixed bias point.

14.4. Noise During the Precharge Phase

Initially, before Φ_1, the holding capacitor is assumed to be in a discharged state. The P-channel pull-up transistor M_{P1} is OFF. At the beginning of Φ_1 (falling edge of $\bar{\Phi}_1$), M_{P1} is turned ON. For a short time, the transistor will operate in saturation, until enough current reaches the capacitor in order to raise its voltage to a level where M_{P1} starts operating in the triode (ohmic) region. This situation is maintained as long as Φ_1 is active.

After a certain time, an equilibrium situation is reached, in which the capacitor voltage is approximately equal to the voltage of the reference voltage source. For practical purposes, M_{P1} acts as a resistor with value R_{eq}, determined by the aspect ratio of the transistor and its gate-source voltage.

$$R_{eq} = \frac{1}{\frac{K'W}{L}(V_{GS} - V_T)} \tag{249}$$

With K' a technological parameter, W the transistor width, L the transistor length, V_{GS} the gate-source voltage and V_T the threshold voltage.

In this state, M_{P1} contributes a (white) noise current with power spectral density $S_{i,w} = 4kT/R_{eq}$ (equation (247)). The corresponding voltage power spectral density of the capacitor node can be calculated by calculating the first-order transfer function of this system in the s domain (V_C is the capacitor voltage

and I_n the noise current of the transistor).

$$V_C = I_n \frac{R}{1 + s\ RC} \tag{250}$$

The voltage power spectrum S_v can be determined by multiplying the current power spectrum by the squared magnitude of this transfer function. The expected noise power (squared voltage) on the capacitor, $E(v_C^2)$, can then be determined by integration of this spectrum from $-\infty$ to ∞. The result of this calculation turns out to be independent of the equivalent resistance R_{eq}, since the DC voltage power spectrum is proportional to R_{eq} while the bandwidth of the system is inversely proportional to R_{eq} [81].

$$S_v = \frac{kT}{C} \tag{251}$$

At the end of Φ_1, $\bar{\Phi}_1$ goes high and M_p is gradually cut OFF. During the process, the effective resistance of the switch, R_{eq}, changes due to changing gate-source voltage. However, the transistor remains in the ohmic region, and equation (251) remains valid. Eventually, the transistor will be completely cut OFF and the capacitor will hold its last voltage. Obviously, the variance (expected noise power) on this voltage will still be given by equation (251) ($1/f$ noise is not considered).

14.5. Noise During the Initial Discharge Phase

For typical values of the holding capacitor (pF range), the noise introduced by the precharge phase will be very limited. However, its precise value does not even have to be known, since all noise present on the holding capacitor after Φ_1 will be cancelled by the initial discharge operation. Unfortunately, this operation will introduce new noise, determined mainly by the comparator, the reference level and the active pull-down network used to disable current source I_1. A limited noise contribution will also be made by current source I_1 itself, but its effect is limited.

At the beginning of Φ_2, analog multiplexer MUX is used to select the initial discharge level of C. Current source I_1 is enabled and the capacitor voltage starts decreasing in a linear way. Simulations of the capacitor voltage are shown in figure 148, for a complete clock cycle of the charge amplifier stage. At the end

of phase Φ_1, before the voltage starts decreasing during Φ_2, the characteristic exhibits a significant "bump". This is due to clock feed-through and channel charge injection from the PMOS switch, added onto the capacitor at the end of the precharge phase.

Shortly after Φ_2 becomes active, current source I_1 is turned ON. The current will exhibit a certain transient behavior, resulting in a smooth, rounded capacitor voltage behavior. Figure 149 shows the capacitor charge current for a complete clock cycle of the charge amplifier. The current is shown as being positive when the capacitor is discharged. During Φ_1, the capacitor is precharged very quickly, resulting in a large, negative current peak, which is not shown for the sake of clarity.

Figure 150 shows a detail of the current during phase Φ_2. In this case, Φ_2 is supposed to start at a time of 100 ns. The current only starts rising after about 10 ns. After about 20 ns, the current reaches its full nominal value. The current peak (overshoot of about 2 μA, decaying after about 10 ns) is a transient due to the regulated cascode circuit turning ON. The extra current is used to charge the channel of transistor M_{N1} in figure 146.

As far as the initial discharge phase is concerned, the clock feed-through, switch-ON transient, as well as possible noise associated with these, are unimportant, since only the final value of the capacitor voltage is "remembered". That final value is determined by the switch-OFF transient, which occurs when the capacitor voltage drops down to the comparator reference level, in this case V_{ID}.

The current source stays ON while the voltage on C_1 remains above the reference level (while the comparator output is low). When the capacitor voltage approaches the reference level (a possible comparator offset is not considered here), the comparator output starts switching high, and pull-down transistor M_3 starts conducting current. It should be noted that initially this transistor is in the OFF state, and that when it comes ON, it will be operating in saturation. When the current of M_3 exceeds the bias current of the regulating amplifier of the current source (M_5 in figure 146), the current source is turned OFF.

In this example, the comparator is a simple high-gain opamp operating in an open-loop configuration (no compensation). The precise schematic is shown

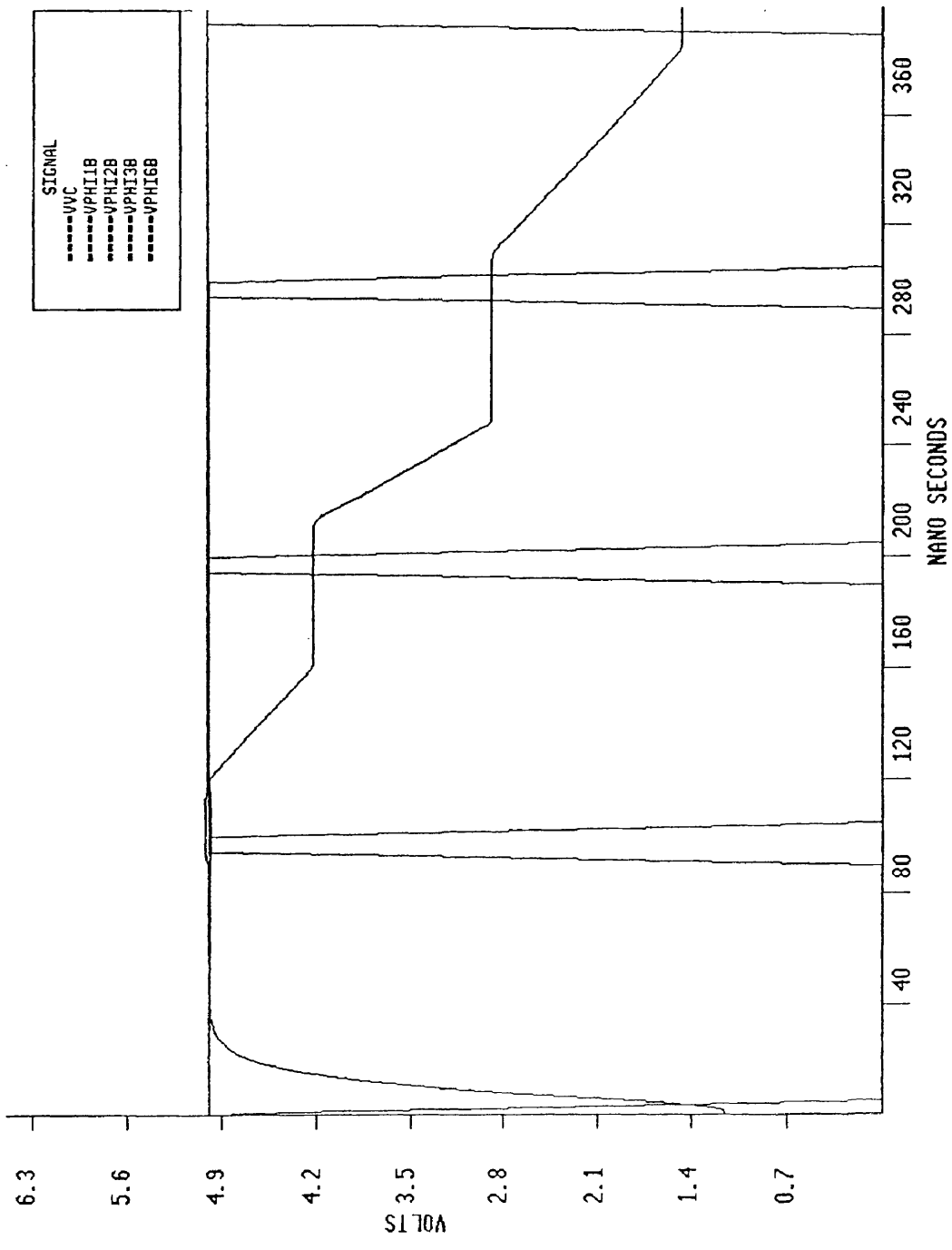


Fig. 148: Capacitor Voltage

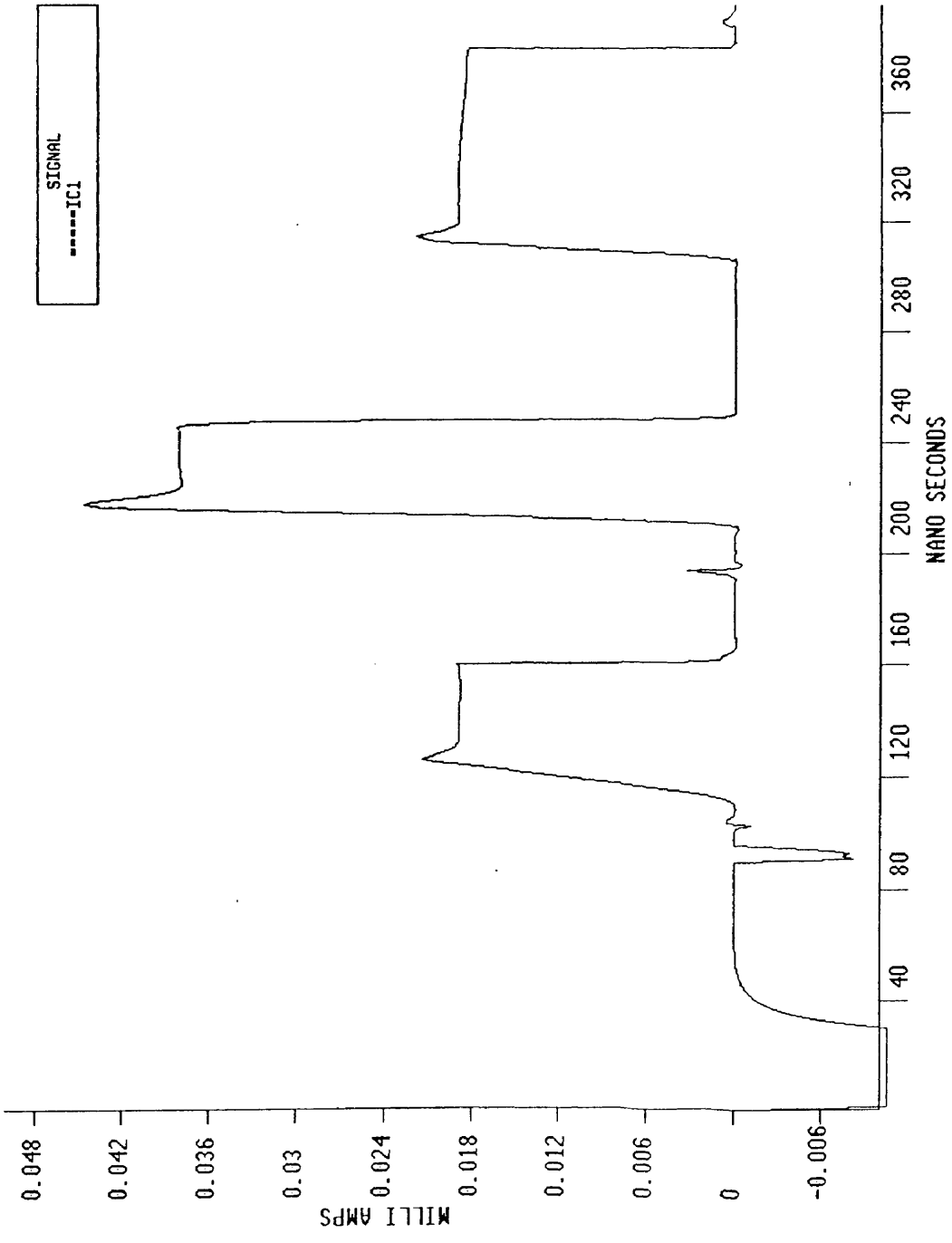


Fig. 149: Capacitor Charge Current

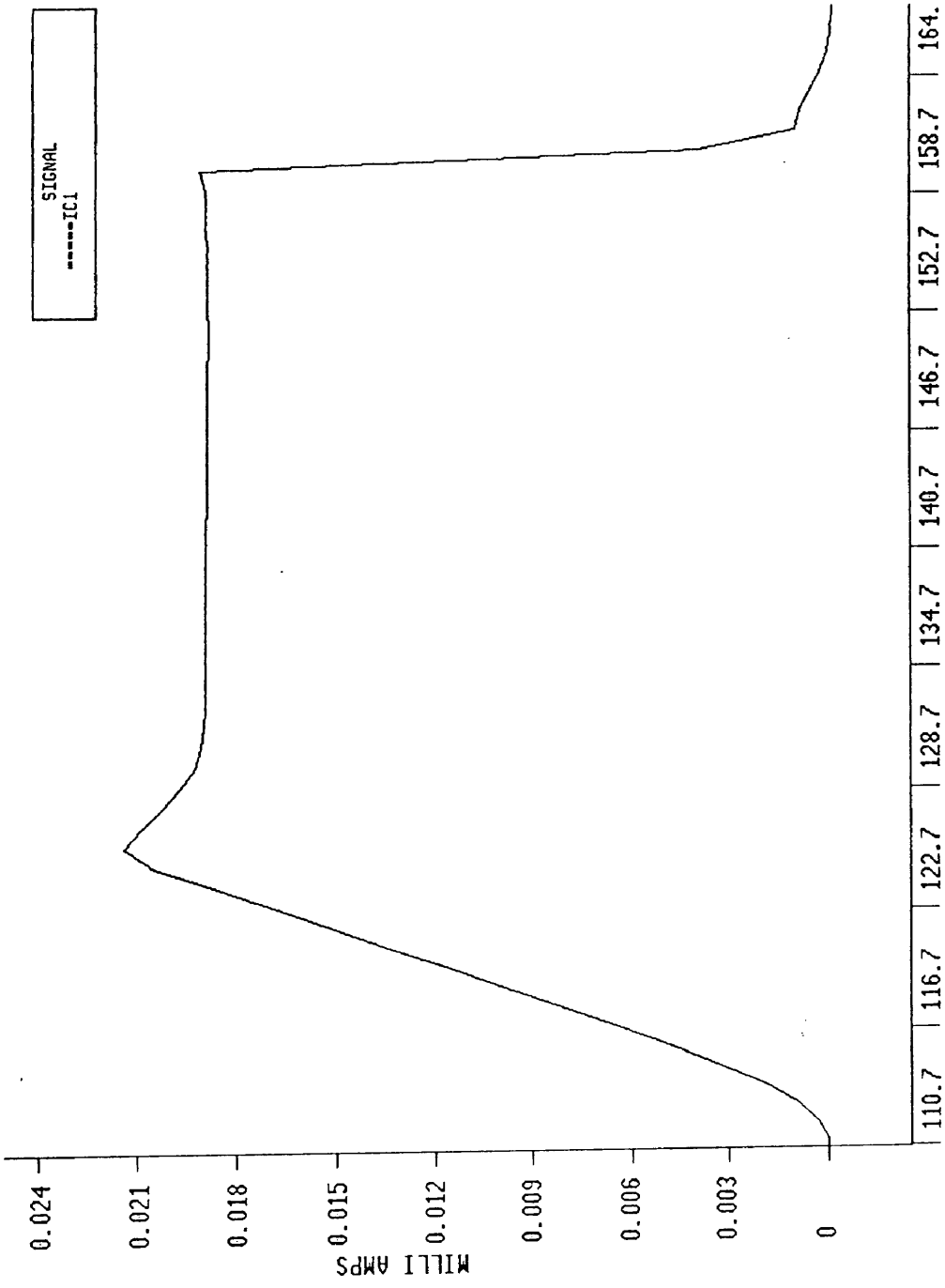


Fig. 150: Capacitor Charge Current in Phase 2

in figure 151. It was chosen for its fairly high gain-bandwidth product (speed) of about 500 Mhz, combined with limited area and power consumption (in this case $20\mu A$). Of course, other schemes are possible. It should be noted that in this example, the comparator is followed by two CMOS inverters. Their purpose is to sharpen the edges of the PWM signal. It will be shown that they do not introduce any significant additional delay, and that their noise contribution is limited.

The switching delay associated with the comparator, inverters, active pull-down and the current source itself, will determine the amount of undershoot on the capacitor voltage, below the reference level. This is a systematic effect, which results in an offset (which in turn will partially be cancelled by the final discharge operation). Unfortunately, the switching operation will also be subject to noise. This effect can be described from a magnitude point of view as "noise". An alternative way to describe it, is from a time point of view. The variability can then be seen as "time jitter" on the switching waveform. This duality is illustrated in figure 152. Although it is common to analyze noise from a magnitude (or expected power) point of view, we will see that the time jitter interpretation may be more useful when certain conditions are met.

In particular, we will consider two different switching noise models, based on two different assumptions. Which assumption is more justified, depends on the particular implementation of the charge amplifier. The first model assumes that the combination of comparator, current source (and active pull-down switch) form a stable negative feed-back system. This would rarely be the case in practice, but the situation has some theoretical merit. The second, more realistic, model assumes the opposite: that the system would be totally unstable. This is typically the case (like in this example) when the sequence of comparator (and inverters), switch and current source has a high gain and band-width. In that case, the switching will characterized by consecutive, rapid level changes of the different components, between the supply voltage and ground.

14.6. First Switching Model: Stable Feed-Back

Figure 153 sketches the switch-OFF transient behavior of I_1. Initially, while the current source is ON, a certain amount of current noise is generated. This noise is mainly determined by the noise of transistor M_{N1} (figure 146), since this

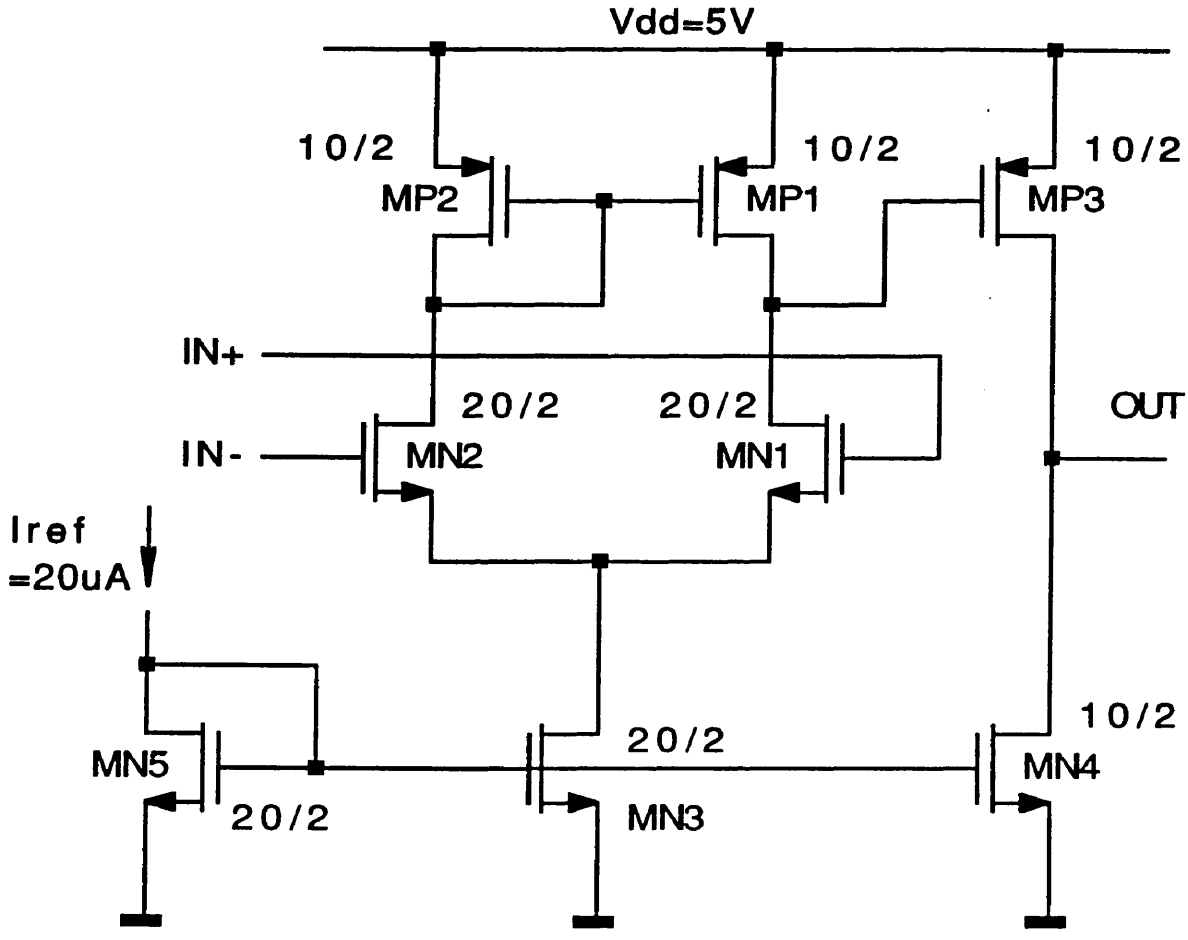


Fig. 151: Comparator Schematic

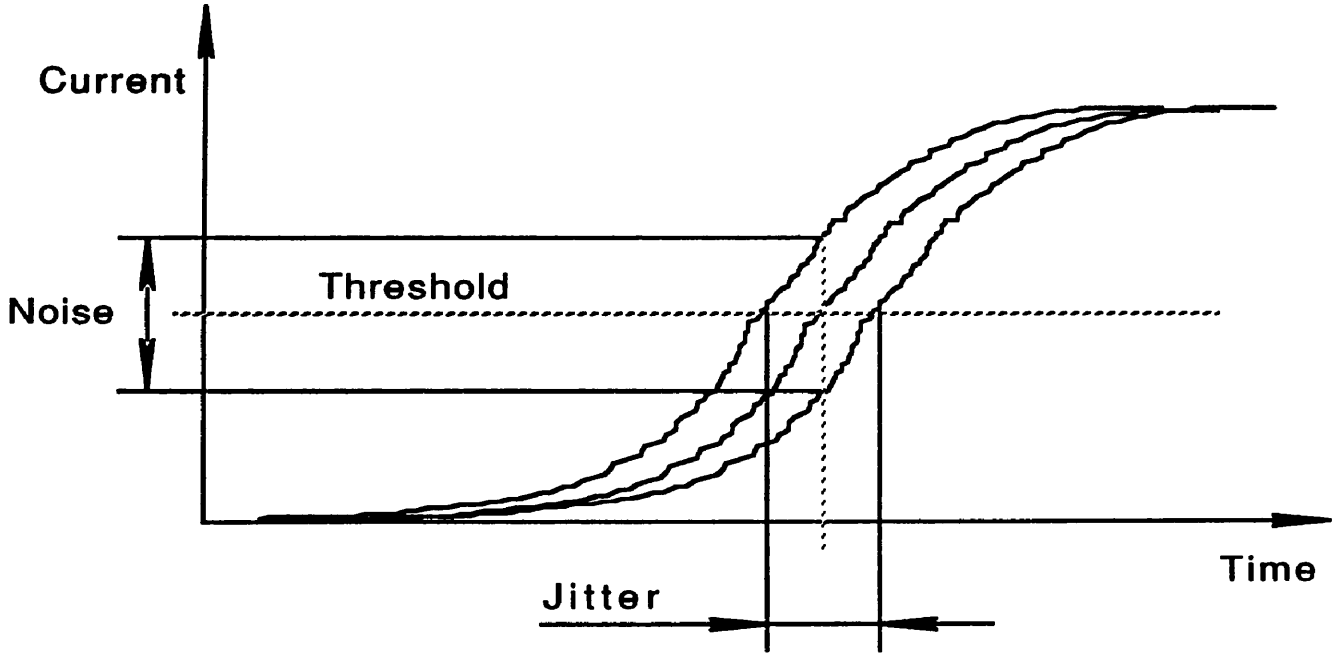


Fig. 152: Time Jitter

transistor forms the basic current source that sets the output current. The bias voltage to this transistor, $IBIAS$, is normally generated by an additional, diode-connected transistor, which is not shown on the figure. This other transistor obviously contributes to the current noise as well. All other transistors of figure 146 are part of a feed-back network that increases the output impedance of the system by regulating the drain voltage of M_{N1}. They do not strongly influence the output current, and hence contribute little noise.

At the beginning of the initial discharge operation, before the capacitor voltage reaches the reference level, noise on the output current of I_1 is essentially irrelevant. All this kind of noise could do, is slightly change the amount of time it takes the capacitor voltage to reach the reference level. However, once the comparator flips and the current source starts switching, noise becomes an issue.

According to the first switching model, the sequence formed by the comparator, the active pull-down transistor (M_{N1}) and the current source, form a high gain transconductance amplifier (a small voltage variation on the comparator inputs creates a large variation in the output current of I_1). This transconductance amplifier is configured in a negative feed-back mode, since the output node of the current source is also one of the input nodes of the comparator. When the capacitor voltage decreases, the output current decreases, thus counteracting the decrease in voltage.

However, this feed-back system has some peculiarities. First of all, it is only effective if the switching of comparator and current source are fairly slow, so that all devices in the loop would simultaneously operate at some more or less defined, small-signal operating point. It is clear that if either one of them is either fully ON or fully OFF, no small-signal transconductance can exist. This is due to the strongly non-linear, sigmoid shape of the I/V characteristic (figure 154). In addition, the feed-back loop contains an integrating element (the holding capacitor integrates the output current). Finally, the system is strictly unidirectional. The output current can only flow in one direction, when the capacitor voltage exceeds the reference level. If the capacitor voltage falls below the reference level, the current source is disabled. No negative current can flow from the current source back to the capacitor.

For slowly switching systems, an equivalent model of the feed-back network

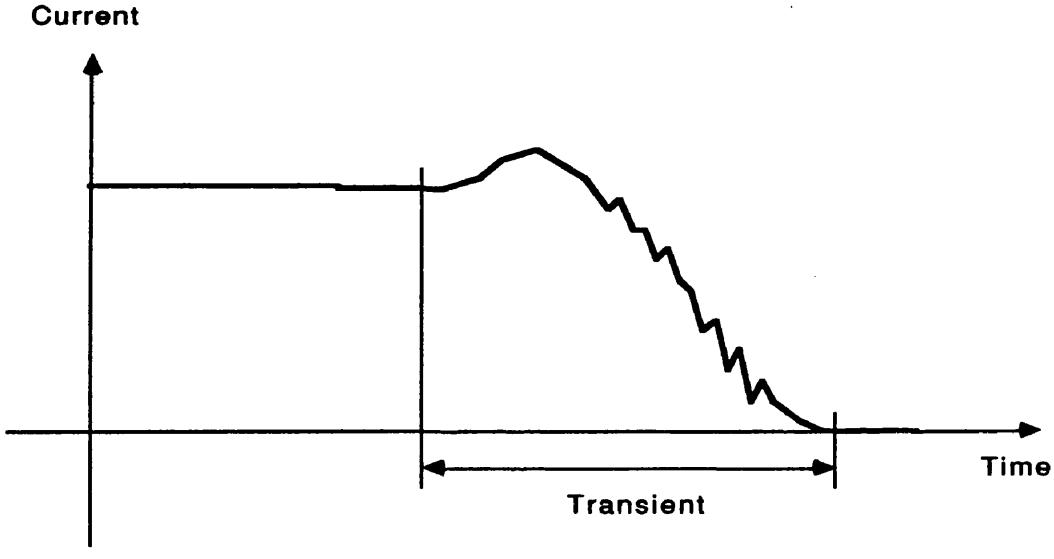


Fig. 153: Switch-OFF Current Transient

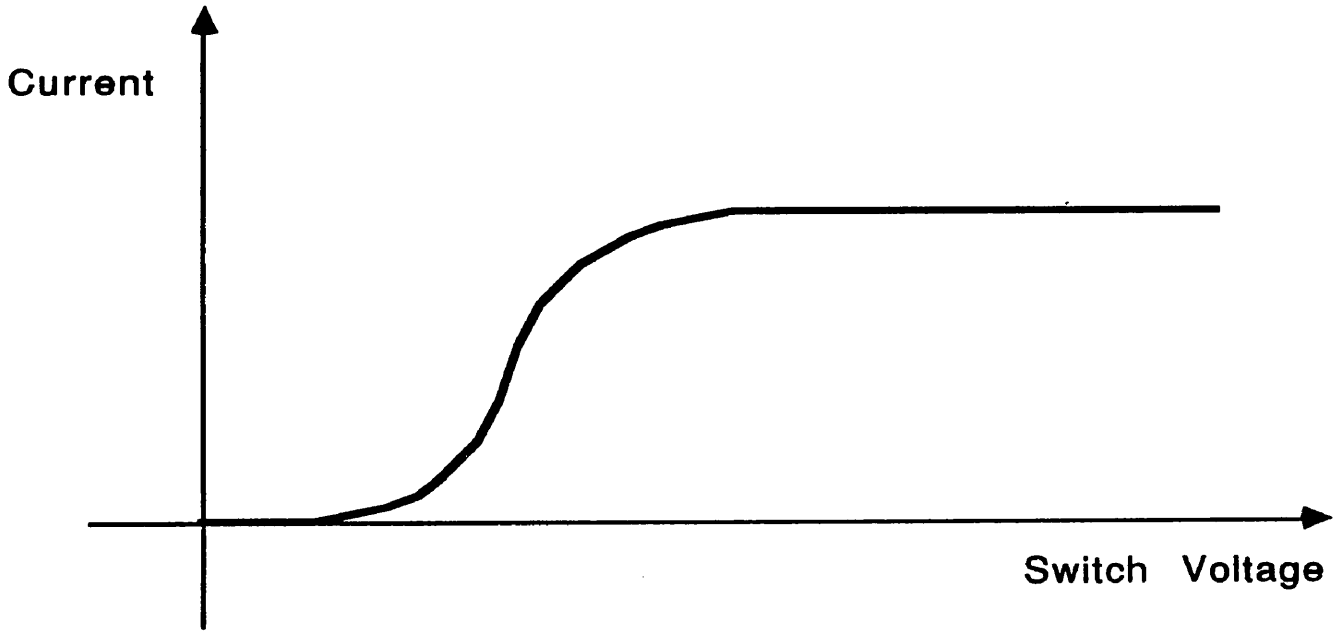


Fig. 154: I/V Characteristic of the Transconductor

can be constructed, as shown on figure 155. The model contains a differential transconductance amplifier (assumed linear and with infinite frequency response for the moment being), and the capacitor. The non-inverting input of the transconductance amplifier is formed by the reference level, which can be assumed to be applied in a step-like fashion. If the reference voltage is represented by the symbol V_{in}, the capacitor voltage by V_{out} and the transconductance gain by G, a linear, frequency domain relationship can be derived between the input and the output.

$$V_{out} = \frac{G}{sC} (V_{in} - V_{out}) \tag{252}$$

$$V_{out} = \frac{V_{in}}{1 + \frac{G}{sC}} \tag{253}$$

It is clear that this equation represents a first-order, exponential settling behavior. According to this model, the capacitor voltage will exponentially reach an equilibrium value equal to the reference voltage, without ever undershooting the reference level. This feed-back system is clearly stable. It should be stressed again that this is only because of the assumed, ideal frequency response of the transconductance element. In practice, stability is definitely not guaranteed, due to the presence of parasitic poles and zeros in the transconductor frequency response.

If such stability (no undershoot) can realistically be assumed, the noise analysis is fairly straight-forward. It is well-known that any amplifier (including transconductance amplifiers) can be modeled as an ideal, noiseless amplifier, with a noise source connected in series with the input. The expected power (variance) of this noise source can be calculated by first calculating the output-referred noise power (current in this case!) for the open-loop case with the capacitor as only load, and dividing it by the square of the gain. This value of the input-referred noise power (or rms noise voltage if the square root is taken) is also the output-referred noise power (voltage!) of the closed-loop system, because the feed-back system has unity voltage gain.

At a first glance, assuming that it would be physically possible to realize it, such stable feed-back system without undershoot may seem desirable. However,

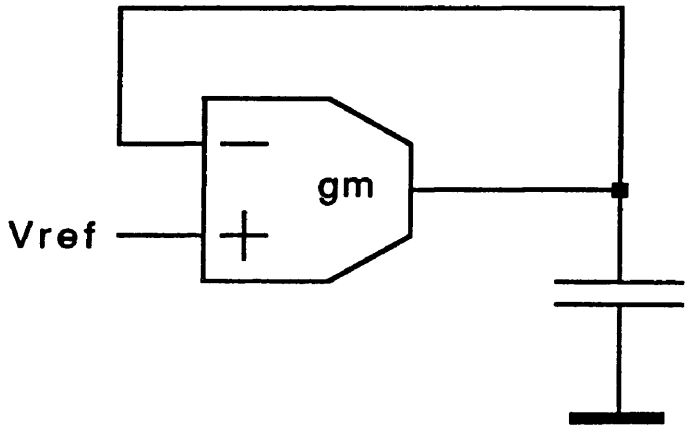


Fig. 155: Equivalent Model of the Feed-Back Network

in practice it is not. The exponential settling process can be very slow for normal values of the time constant G/C. In addition, this time constant becomes even longer (going to ∞) when the capacitor voltage approaches its final value, because of the strong sigmoidal non-linearity of the transconductance when the current becomes close to zero. This means that actual settling would be sub-exponential, and that the final value would never be reached entirely.

14.7. Second Switching Model: Sequential Switching

Although it results in an elegant way to estimate the noise, it is obvious that the stable feed-back model does not accurately represent the physical operation of the charge amplifier. A first discrepancy is illustrated by the fact that the capacitor voltage actually undershoots the reference voltage by a significant amount (almost 300 mV in the given example, as shown on figure 148).

Actually, there are many instances of systems that operate in a way analogous to the charge amplifier in every-day life. An extremely common example is that of a water tank, filled to a predetermined level using a mechanical shut-off valve controlled by a floater mechanism. The tank could be compared to the holding capacitor of the charge amplifier, the water supply pipe with a current source, the shut-off valve with the pull-down transistor (switch) and the floater mechanism with the comparator/inverter combination. It is clear from observation of such systems, that they do not operate in stable feed-back mode either. Once the preset level is reached, the water supply is shut off fairly abruptly, and eventually a situation is reached in which the final water level slightly exceeds the desired level, in such a way that the floater would keep the water supply solidly cut off.

Figure 156 shows the simulated switching transients in the different components of the charge amplifier, at the time the current source is switched OFF. The curves represent the comparator output voltage, the voltage at the output of the first inverter and the voltage at the output of the second inverter respectively.

It is clear that all transitions are very fast (nanosecond range). In addition, the figure shows that the three waveforms do not reach the middle of the voltage range (here 0 to 5V) at the same time. Instead, one waveform starts a transition at the time the previous one is about to finish its transition, or at least significantly after the DC logic threshold of the switching element has been reached. As a result,

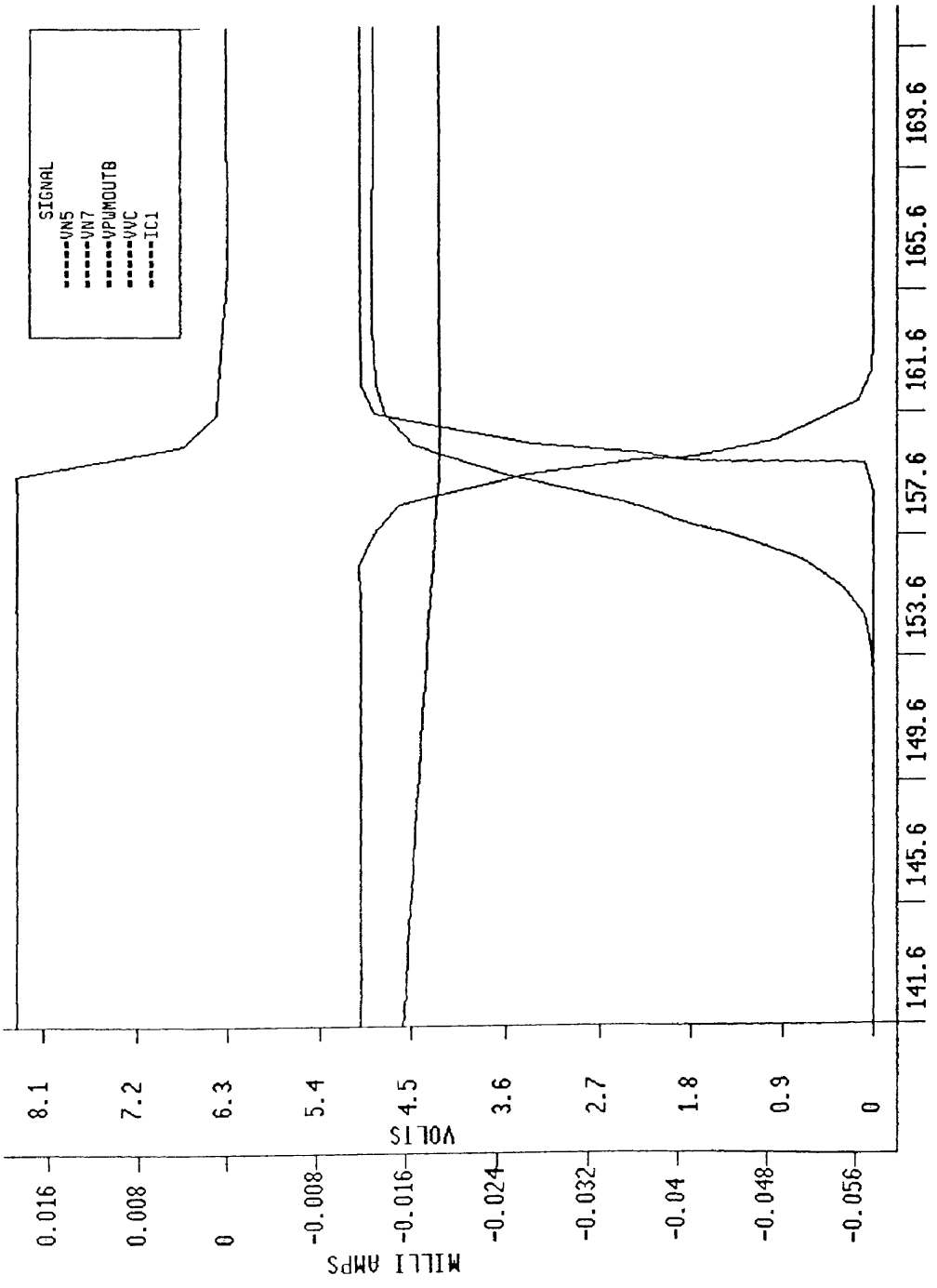


Fig. 156: Simulated Transients at Switch-Off

it cannot realistically be assumed that the comparator and the two inverters ever form a high-gain amplifier chain with a certain constant bias point. Although not shown on the same figure, the current waveform follows the same pattern. The current only starts falling *after* the output of the last inverter almost reaches the full supply voltage.

This situation can be modeled as if the switching elements (comparator and inverters) switched in sequence, one after the other. Each one of them could be characterized by a certain delay, like the equivalent gate delay in logic systems. if the delay were absolutely constant, no noise would be introduced. The capacitor voltage would still undershoot the reference level, but this is considered a systematic offset and not noise. In practice however, the switching delay will be subject to time jitter. This jitter is hard to predict, and can also be hard to measure on an actual circuit. However, figure 152 suggests a simple yet effective approximation.

First of all, it is clear that when an element (e.g. an inverter) is completely ON or completely OFF, its output-referred noise will be insignificant. Either the output is pulled to ground, or to the supply voltage. The worst-case situation from a noise point of view (at least for DC, but we will make the simplified assumption that this also holds during the switching) occurs when the input of the element is biased so that the output would be at about mid-supply. This is when the gain will be largest in the small-signal sense (steepest slope on the sigmoid). In addition, the mid-point is usually about where the DC switching threshold of the gate is located, and this threshold mainly determines the switching instant.

For mid-point bias conditions, the output-referred expected noise power (variance) can be calculated, using frequency-domain techniques or simulations (integration of the noise spectrum). An important point should be raised here. The switching element only remains at mid-point bias for an extremely short period of time. As a result, it is plausible (yet not proven at this point) that low-frequency noise components would be rejected. according to the switched-noise hypothesis that was set forth in chapter XII. If this is indeed the case, the spectrum should first be multiplied by the correcting function, which reflects the rejection of low frequencies (an approximative way to achieve the same effect is to only integrate the spectrum starting at a frequency of $1/T_s$, with T_s the "effective"

duration of the switching).

Even if this switched-noise assumption were not justified, a large portion of the low-frequency spectrum could still be ignored, due to the correlated double sampling mechanism that is built into the charge amplifier. This is due to the fact that any noise voltage (positive or negative) that is left on the capacitor after the switching, will be partially matched during the final discharge operation. However, the correlation time (time between the end of the initial discharge and final discharge operations) is significantly longer than the switching time, resulting in the rejection of less low-frequency components.

The square root of the variance, determined as described above, represents the standard deviation of the voltage, for mid-point bias conditions (σ_V). This standard deviation can be related to the standard deviation on the switching instant (defined as the time at which the output voltage of the switching element reaches a certain, predetermined level). Since the relationship between voltage and time can be determined by a standard transient simulation, the time derivative of the output voltage can easily be determined at the switching instant. It is clear that the standard deviation on the switching instant (σ_T) is a measure of the time jitter, or the variability of the switching delay of each element. This is illustrated in figure 157. The jitter can be calculated using the following formula:

$$\sigma_T = \frac{\sigma_V}{\frac{dV}{dt}} \tag{254}$$

A similar calculation can be performed for each switching element (comparator, inverters and current source). According to the sequential switching model, the different elements will switch one after another rather than simultaneously. This is obviously advantageous from a noise point of view, since noise contributions of earlier stages will not be multiplied by the high gain, formed by subsequent stages in the chain. The time jitter on the last switching stage (the variability of the time at which the current source is switched OFF), can be quantified by the *variance* (square of the standard deviation σ) on the switching instant. It is obvious that according to this model, this variance will be the sum of the time variances of individual stages. This is motivated by the fact that the total switching delay is the sum of individual switching delays, and that the random variations on each of these delays are independent stochastic variables.

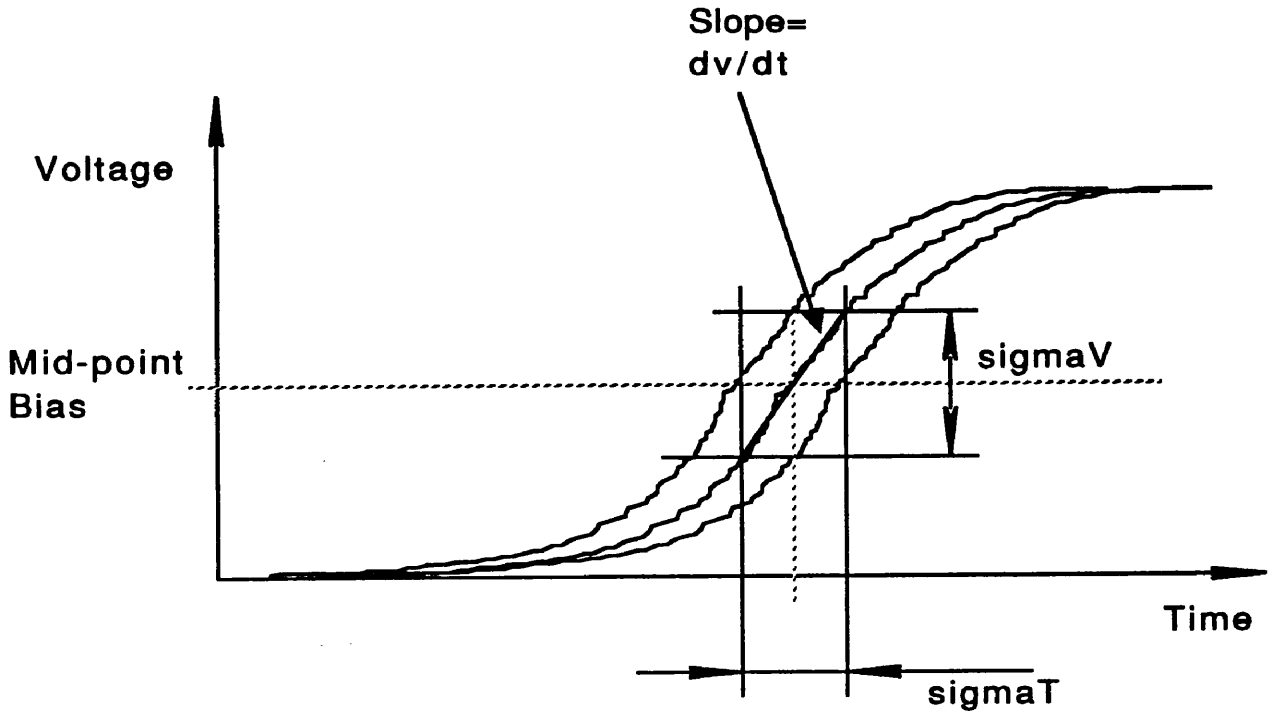


Fig. 157: Time Jitter Calculation

14.8. Ring Oscillator Analogy

Since the sequential switching model bears a lot of resemblance with the switching of digital gates, the theory could be verified using an actual, fully digital feed-back circuit, like a ring oscillator. It is well-known that a ring oscillator (figure 158) can be built using an odd number ($2N + 1$, with N integer) of inverters, so that the output of the last inverter would be fed back to the input of the first one. The oscillation frequency is given by the formula (the inverters are assumed equal):

$$f_{osc} = \frac{1}{(2N+1)\,T} \tag{255}$$

With T the elementary, equivalent delay of one inverter.

Due to noise in the semiconductor devices (e.g. CMOS type transistors), the oscillation frequency will be subject to random variations, which can be estimated from the "smearing" of the spectral peaks on a high-quality spectrum analyzer. (An actual measurement could be performed by first "beating" or digitally multiplying the output signals of two identical ring oscillator, since the beat signal will have a differential frequency component at low frequency). The variance of the frequency (σ_f^2) can be related to time jitter of individual inverters (variance σ_T^2). The individual gate delays are assumed to be independent and additive.

$$\sigma_T = \sigma_f \left|\frac{dT}{df_{osc}}\right| \tag{256}$$

$$\sigma_T = \frac{1}{2N+1}\,\frac{\sigma_f}{f_{osc}^2} \tag{257}$$

The standard deviation found this way should agree with the standard deviation predicted by noise magnitude simulations, and the slope of the switching waveforms, *if* the sequential switching model were accurate. The extent to which such measurements would differ from calculations, would be an indication of the accuracy of the sequential switching model.

14.9. Noise During the Load Phase

During the load phase, the input signal is applied to the charge amplifier in PWM-coded form. The duration of the input pulse determines the time during

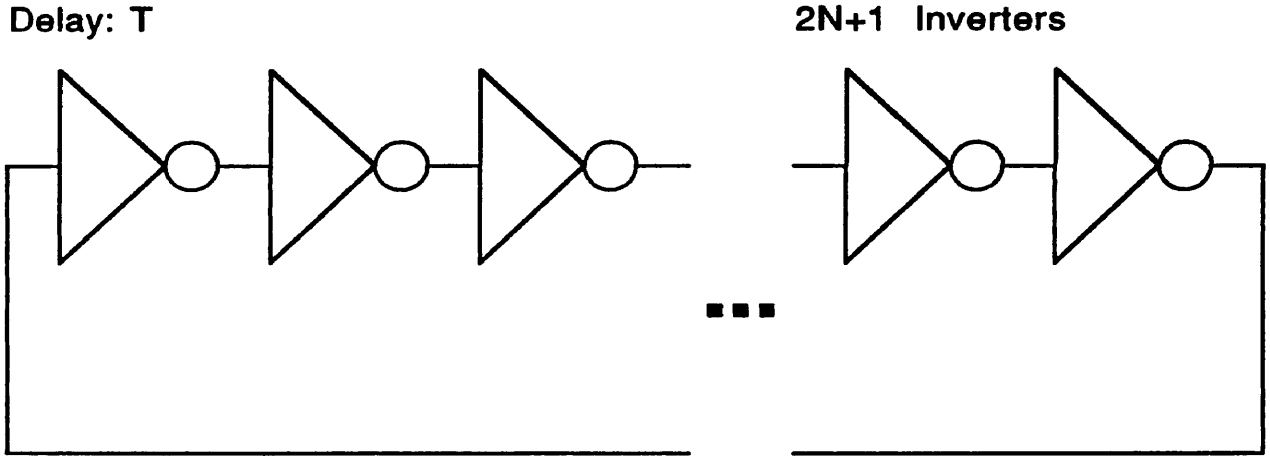


Fig. 158: Ring Oscillator

which current sources I_2 and I_3 are ON, and as a direct result, the amount of charge that is subtracted from the reference charge left on the holding capacitor after the initial discharge phase. (This configuration is inverting.) The total amount of charge that is displaced during this operation, is affected by noise in three different ways:

- Variability due to the switch-ON transient
- Variability due to the switch-OFF transient
- variability due to current noise during the steady-state operation of the current source.

Since these three effects are essentially independent, the total noise contribution can be determined by summing together the individual variances.

Depending on the particular implementation, the switch-ON transient of the current source can be determined by either the falling edge of $\bar{\Phi}_3$, or the actual leading edge of the PWM signal. The two signals have to be low at the same time before the current sources are enabled. This is realized by the active pull-down transistors M_{N3} and M_{N6}, and M_{N2} and M_{N5}, respectively. The distinction between the two approaches is the following. The PWM input signal can be assumed to come from a previous charge amplifier stage, where it is generated using a comparator (possibly followed by inverters), similar to comparator K in figure 145.

During the final discharge phase of the previous stage, the comparator reference voltage has to be equal to the desired final discharge voltage (possible undershoot is not considered at this point). This reference voltage can be applied either immediately after Φ_2 (initial discharge), or at the beginning of Φ_6 (final discharge). In the latter case, the comparator will need a finite time before its output goes low, resulting in additional delay. in the former case, the comparator output "waits" in a low state until the beginning of Φ_6, and as soon as the clock edge goes low as well, the current source I_1 is enabled. Since the comparator output signal is used as PWM-coded output signal of the stage, this signal will be low before the falling edge of the clock pulse. However, this is of no concern, since this signal will be combined with the $\bar{\Phi}_3$ signal of the next stage.

The first solution is not only faster, it is also more advantageous from a noise

point of view. The activation of I_2 and I_3 in the present stage is subject to clock jitter, which we will not consider, since it is not introduced by the charge amplifier, and it is normally matched by identical jitter in the other stages. It is also subject to noise introduced by the current sources and their pull-down devices, but *not* to comparator noise. This approach has been chosen in the simulated example.

At the beginning of Φ_3, on the falling edge of the clock, the gate voltage of M_{N6} and M_{N7} is brought to zero, which releases the current sources I_2 and I_3. The noise associated with this operation can best be described as uncertainty (jitter) on the switching instant, according to the procedure set forth in the sequential switching model. A numerical value can be associated with this jitter by calculating the total current noise power at the output of the current sources, under bias conditions that are such that the pull-down transistors M_{N6} and M_{N5} are partially ON. This noise power can be determined from frequency domain simulations, in which the current sources are biased so that their output current is about one half the nominal current. This noise power can be related to the time derivative of the current during normal switching conditions (determined from a transient simulation).

In a very similar way, the time jitter associated with switching OFF the current, can be estimated. The current noise will be the same from a magnitude point of view (same assumed mid-supply bias point), but the jitter may be different due to the different slope of the current transient. The two values of the time jitter are statistically independent, at least if only white noise is considered (if $1/f$ noise were included, a certain correlation would exist). The time skew associated with the current transient can be either positive or negative in both cases. The sum of the time variances can be multiplied by the square of the nominal current in order to determine the variance on the total capacitor charge due to the switching ($Q \approx IT$ and hence $\sigma_Q^2 \approx I^2\sigma_T^2$).

However, these are not the only effects that create variability on the capacitor charge. While the current sources are fully ON (steady-state condition), the output current will be subject to noise as well. This noise is integrated for the duration of the PWM pulse, and can be quite hard to predict accurately, since it is signal-dependent (the longer the integration time, the more noise). In addition, it is not immediately clear how the time domain behavior of the current can be

related to the frequency domain noise spectrum. If the capacitor were a perfect integrator, its voltage power spectrum could be derived from the current power spectrum, through multiplication by $1/(\omega C)^2$. Since this function approaches ∞ for small values of ω, it is clear that low frequency noise components will be dominant. However, the system is also switched, which could mean that low frequency noise components are rejected.

To avoid this kind of considerations, a simplified, worst-case calculation can be performed in order to estimate the total variance on the capacitor charge, right before the current source is disabled again. It will be assumed that when the current source comes on, the current I is subject to a certain amount of noise ΔI. This noise will change as a function of time, but for simplicity, we will assume that it does not. In other words, if the PWM input signal is active for a time T, we will assume that the displaced charge is

$$Q = (I + \Delta I)\ T \tag{258}$$

The precise value of ΔI is not known. However, the variance (expected squared noise current) can be calculated from the output noise spectrum of the current source. For the current source of figure 146, the noise is mainly determined by the noise currents of transistor M_{N1}, and of the bias transistor which is not shown. The effect of noise in the bias transistor can be reduced to almost zero if the same bias transistor is shared by the previous stage. In that case, any noise caused by the bias transistor in the current source used in the final discharge operation of the previous stage will be duplicated in the current sources of the present stage. As a result, the current ratio (gain) will be maintained despite the noise of the biasing transistor.

The effect of noise introduced by M_{N1} is more significant, since the noise contributions of the previous stage and the present one are independent, and hence additive form the point of view of their variance. However, it could be argued that since these transistors are periodically cut OFF, the low-frequency components of the noise spectrum will be attenuated, according to the switched-noise hypothesis we suggested in chapter XII. This has not been verified experimentally. However, if this were the case, a significant part of the noise spectrum (including the part that normally contains $1/f$ noise) would become insignificant (the current source

only stays ON for a very limited time T, being the maximum PWM pulse width). This would result in very favorable noise performance and low drift.

However, if this theory were not verified, the influence of drift in the current sources ($1/f$ noise) would have to be included. In particular, the drift may mean that current sources would have to be periodically recalibrated, using some analog self-calibration technique. Alternatively, if the charge amplifier were part of a digitally self-calibrated ADC (e.g. using accuracy bootstrapping), the converter should go through a calibration cycle every few milliseconds, in order to counteract the effects of $1/f$ noise components in the kilohertz range or below.

14.10. Noise During the Final Discharge Phase

The noise mechanisms affecting the final discharge phase are similar to the mechanisms affecting the initial discharge and the load phase. The situation is the following. After the initial discharge phase, a certain reference charge is present on the holding capacitor. This charge is subject to variations due to the switching of comparator, inverters and current source I_1. Next, during the load phase, an additional charge proportional to the pulse width of the input signal is subtracted. It has been shown that the amount of charge is subject to variability due to the noise during the OFF-ON transient, the ON-OFF transient and the steady-state operation of the current sources I_2 and I_3.

During the final discharge, the remaining charge on the holding capacitor is removed using current source I_1. comparator K is used to detect when the final discharge level is reached, in a way very similar to what happened during the initial discharge phase. It has been shown in chapter X that this results in a form of correlated double sampling, which effectively eliminates significant offset problems that could be introduced by the finite switching time of comparator, inverters and current source.

The correlated double sampling also eliminates a portion of the low-frequency switching noise, which according to the sequential switching model, shows up as time jitter affecting the switching instant. The idea is that if current source I_1 is switched OFF a little too late during the initial discharge phase, due to low-frequency noise, the same current source is likely to also be switched OFF a little too late during the final discharge phase, and the total charge that is displace

during the final discharge phase, will not be affected.

One word of caution should be used here. The theory about the effect of correlated double sampling on the noise spectrum, which was developed in chapter XIII, relied on some assumptions that are not entirely justified here. Nevertheless, the formula that was derived, can still be applied as a reasonable approximation. In particular, the noise model that was assumed in chapter XIII, was based a constant bias point of the sampling device. It is clear that the sampling device in this case, is the sequence of comparator, inverters and current source I_1. None of these operate at a fixed bias point, since all of them are switching. This may have two possible repercussions.

- The fact that those devices are OFF between the two operations, does *not* affect the noise characteristics when they come back ON. In that case, the assumptions on which the formula for the expected noise power in the presence of correlated double sampling is based, would still be valid. low-frequency noise components are rejected.
- The fact that the devices are OFF, does affect the noise, in the sense that the noise is reset every time, and starts building up from zero again. In that case, the assumption of correlated double sampling are not justified any longer. However, low-frequency noise components would still be rejected according to the switched-noise theory of chapter XII. The effect would be slightly different, but not much (in the two cases, a very similar formula was found for the calculation of the expected noise power).

As a result, one could say that no matter what the assumption is, low-frequency noise components (especially $1/f$ noise in MOSFETs) will be rejected. The two models may result in a slight numerical difference, but this is probably insignificant due to the approximative nature of the calculations.

In total, three major noise effects will affect the pulse width of the output signal:

1. Noise at the switch-ON moment. The current source is switched ON on the falling edge of $\bar{\Phi}_6$, when the gate voltage of active pull-down device M_{N7} (figure 145) is brought low. The time jitter on the switching instant, due to this effect, can be calculated by biasing M_{N7} so that the current source would be half ON. The expected noise power of this condition can be related to time

(jitter) through division by the time derivative of the current (determined from a transient simulation).

2. Noise on the steady-state current of current source I_1. This effect is similar to the effect described under the load phase. Actually, the noise current of the current source will mainly be affected by two elements: noise from transistor M_{N1} of figure 146 and noise due to the bias transistor. If the bias transistor is shared by two consecutive stages, its effect upon the charge gain (but *not* on the pulse width) will be strongly attenuated, since despite the noise, the current ratio will be maintained. The noise from transistor M_{N1} may be limited as well, if it can be assumed that the switching eliminates low-frequency noise components (switched-noise hypothesis).
3. Noise at the switch-OFF moment. This noise is introduced by the sequence of comparator, inverters and current source. Its quantitative value can be estimated according to the sequential switching model. However, it should be kept in mind that low-frequency noise components are rejected by the correlated double sampling (or the switching action, depending on which hypothesis is maintained).

14.11. Simulated Noise Performance

The noise calculation techniques based on the sequential switching model can be applied to the SPICE simulation results. As an example, we will roughly estimate the total time jitter associated with the switch-OFF current transient at the end of the final discharge phase. Due to the difficulty in determining how to handle low frequency noise, we will only consider the effect of white noise. Besides, it is plausible to assume that low frequency noise is adequately rejected, either inherently or by periodic calibration.

Also, we will ignore all effects related to switched noise and/or correlated double sampling. This is a pessimistic assumption: we will integrate noise spectra from 0 to ∞, although we have seen in chapters XII and XIII that the spectra may have to be multiplied first by a function that rejects a large part of the low frequency components. The idea is that in reality, noise performance should be *no worse* (and probably *much better*) than expressed by the numbers that are derived here.

All relevant waveforms (comparator output voltage, output voltage of the first inverter and output voltage of the second inverter) are shown on figure 156. The current waveform can be seen on figure 149.

Figure 159 shows the output-referred comparator noise, simulated for mid-supply bias conditions. The noise power spectral density is about $70\mu V/\sqrt{Hz}$, or 4.9 10^{-9} V^2/Hz at DC. The equivalent noise bandwidth is about roughly 1 MHz, resulting in an expected power of 0.49 V^2, or about 220 mV. The time derivative of the comparator output voltage during the switching is about 5 V in 4 ns, or 1.25 10^9 V/s. As a result, the time jitter would be about 0.22 $V/(1.25\ 10^9\ V/s) = 176\ ps$.

This number should be related to the possible pulse width range (maximum pulse width minus minimum pulse width) of the signal used in this scheme. As an example, we will assume that the pulse width range is 40 ns. As a result, the dynamic range of this amplifier could not be better than 40 $ns/176\ ps = 227$, or 47 dB. This number may seem extremely low. This is due to the fact that we have modeled the comparator as a one-stage amplifier as far as the noise was concerned.

That model is not quite realistic, since there are actually two stages (differential pair and output stage), which in practice switch almost sequentially (the output stage of the amplifier could be thought of as a PMOS inverter following the differential pair, and mainly introducing an additional delay). If the same calculation were repeated for the two amplifier stages separately, and the respective time jitters were added in the rms sense, a total time jitter of about 10 ps would be found, resulting in an upper limit for the dynamic range of about 4000, or 72 dB.

It is clear that this dynamic range can be extended by extending the pulse width range. For equal current, this can be accomplished by using a larger holding capacitor. It is also possible to reduce the nominal source current, but this may result in increased noise levels due to the current sources, as will be shown later. In any case, extending the pulse width range will result in lower throughput of the charge amplifier. There is an inherent trade-off between speed and noise.

To a certain extent, he same trade-off would become apparent if we attempted

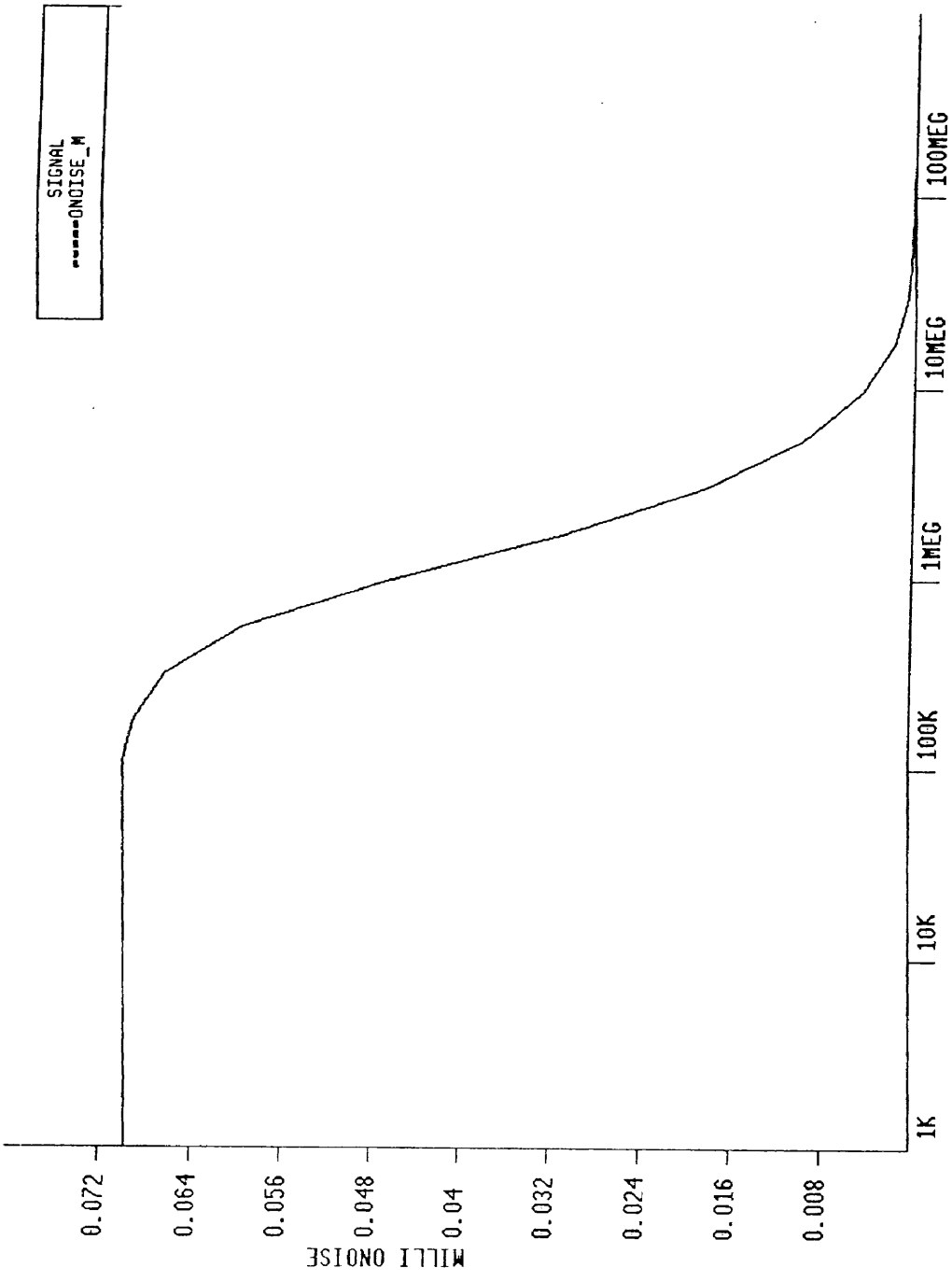


Fig. 159: Output-Referred Comparator Noise

to improve the noise performance by increasing comparator speed (higher dV/dt). Increasing the speed of an amplifier normally requires increasing the DC bias current, which normally also results in higher gain and higher output-referred noise. The inherent trade-off between speed and noise is also present in switched capacitor stages. The speed of a switched-capacitor stage can be increased by increasing the bandwidth of the opamps. Unless special care is taken to simultaneously increasing the gain (higher gain can result in lower *input*-referred noise, despite higher *output*-referred values), the increased bandwidth will result in an increase of the total expected noise power.

In a way similar to the calculation of the time jitter caused by the comparator, the time jitter due to the inverters can be calculated (white noise only). Figure 160 shows the simulated output-referred noise spectrum of one inverter, with the output biased at mid-supply. The output-referred voltage power spectral density is about 475 $nV/\sqrt{Hz}$, or 2.26 10^{-13} V^2/Hz. The equivalent noise bandwidth is about 10 MHz, resulting in a total expected noise power of about 2.3 10^{-6} V^2, or an effective noise voltage of 1.5 mV. The time derivative of the inverter output voltage during the switching is about 1 V/ns, resulting in an effective time jitter of about 1.5 $mV/(1\ V/ns) = 1.5\ ps$. It is clear that this value is small compared to the jitter introduced by the comparator. The same conclusion would hold for the jitter caused by the second inverter. This shows that the inclusion of the two inverters behind the comparator sharpens the transitions of the PWM signal without adding significant extra noise.

The last element that can introduce time jitter during switching is the combination of the current source itself, and the active pull-down network used to disable it. Figure 161 shows the simulated output-referred noise spectrum of a current source, of which the pull-down transistor is biased so as to maintain half the nominal current (10 μA instead of 20 μA). The current power spectral density is roughly 15 $pA/\sqrt{Hz}$, or 2.3 10^{-22} A^2/Hz. The effective noise bandwidth is about 10 MHz, resulting in a total expected current noise power of about 2.3 10^{-15} A^2, or an effective noise current of 48 nA. The time derivative of the current during switching is about 20 $\mu A/5\ ns = 4000\ A/s$. As a result, the total effective time jitter added by the current source is about 48 $nA/(4000\ A/s) =$ 12 ps.

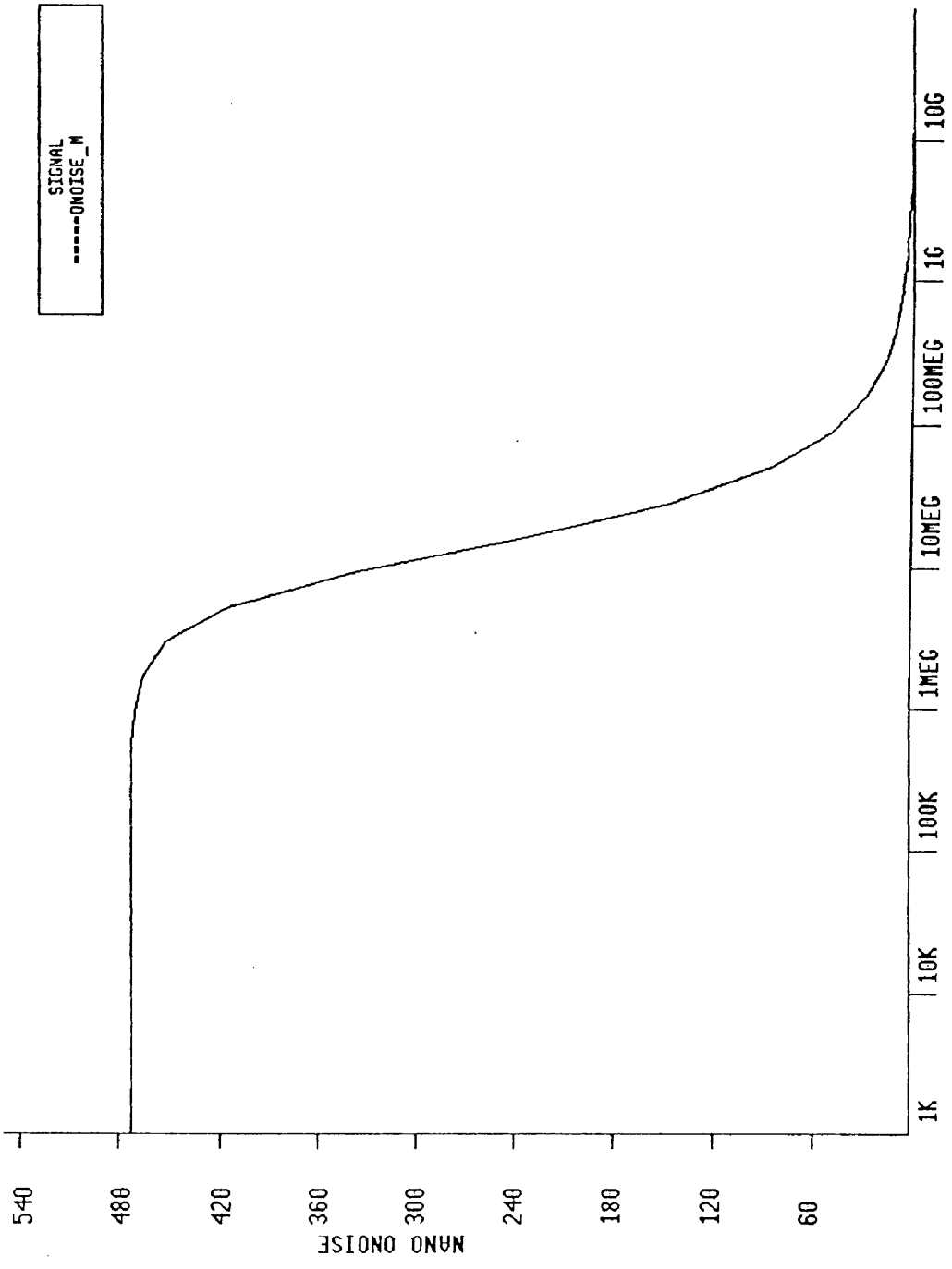


Fig. 160: Output-Referred Inverter Noise

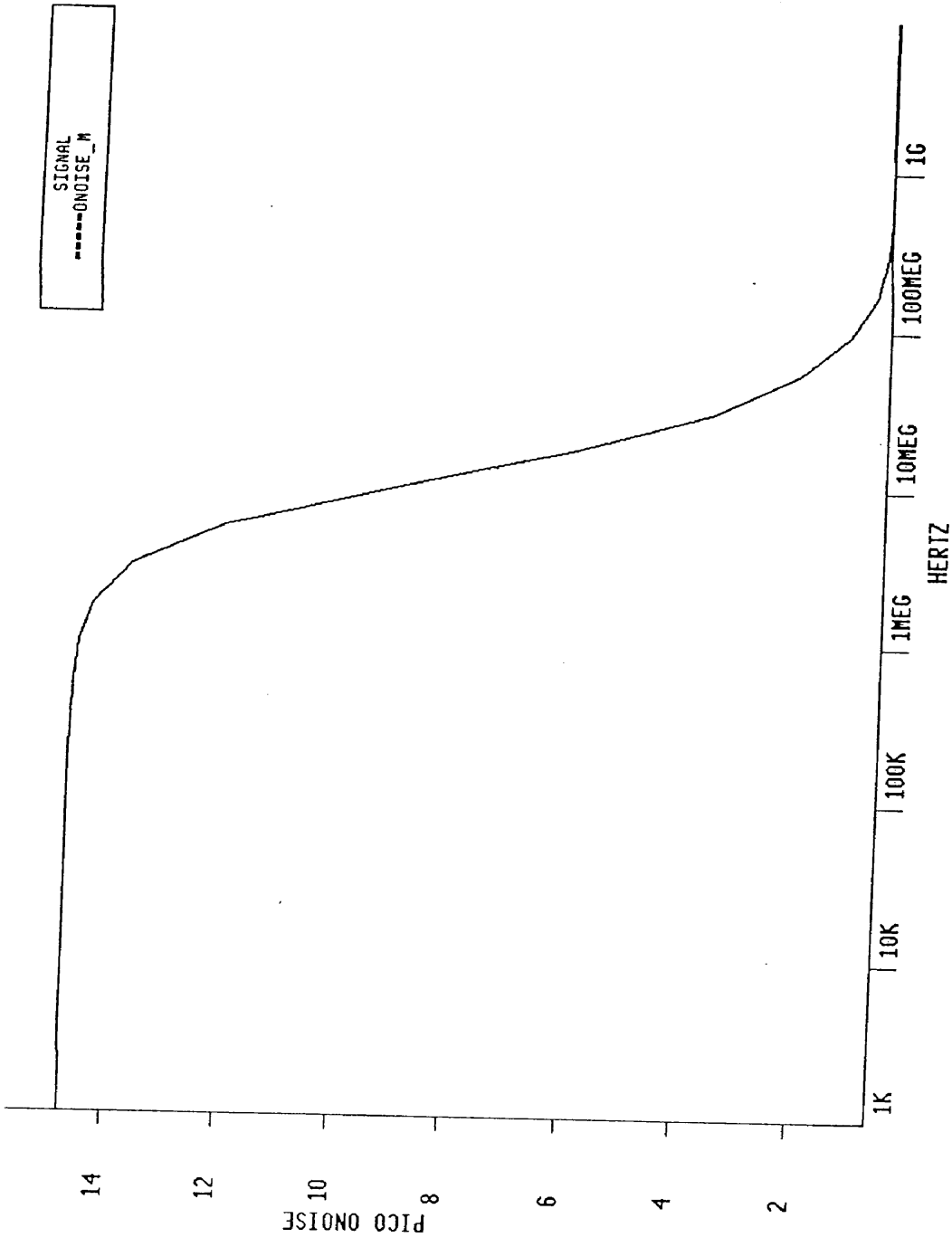


Fig. 161: Output-Referred Current Noise

This jitter seems comparable to the jitter introduced by the comparator. Actually, the situation may be slightly more favorable, since we assumed that current source and active pull-down network were active at the same time. In reality, sequential switching may occur, resulting in less time jitter. In any case, the conclusion of these calculation is that the comparator and current sources are the main contributors of time jitter in the whole scheme. Improvement of the noise performance should focus mainly on optimal design of these sub-circuits.

In addition to noise due to time jitter, the charge amplifier is also subject to steady-state current noise. Its effect can be calculated by estimating the total expected current noise power of a current source, and relating this number to the steady-state current value. Figure 162 shows the output-referred current power spectral density of a current source, under steady-state conditions (current source fully ON). At DC, the spectral density is about 1 $pA/\sqrt{Hz}$, or 10^{-24} A^2/Hz. The effective noise bandwidth is about 100 MHz, resulting in a total expected current noise power of 10^{-15} A^2, or a current variance of 32 nA. As a consequence, for this particular configuration, the dynamic range will be limited to 20 $\mu A/32$ $nA = 625$, or 56 dB.

It should be noted that on figure 162, the noise spectrum stabilizes to a value of about 0.6 $pA/\sqrt{Hz}$ at a frequency of about 1 GHz. This is due to the fact that within SPICE, the output MOSFET is modeled with an ideal (infinite bandwidth) noise current source between source and drain. This is merely a mathematical model; it is clear that in the physical world, the bandwidth of the noise current noise within the MOSFET must be limited (e.g. by some kind of distributed parasitic capacitance within the channel). Consequently, the flattening of the spectrum is not considered in this calculation.

The value of 56 dB may not seem great. However, it should be kept in mind that actual performance can be much better, due to the following:

- High-frequency switching may actually reduce the noise to a value well below the steady-state value.
- This particular set-up was not optimized for noise performance. A better current source design (transistor sizing) could yield much lower values.
- The simulation of the steady-state current noise used a biasing transistor

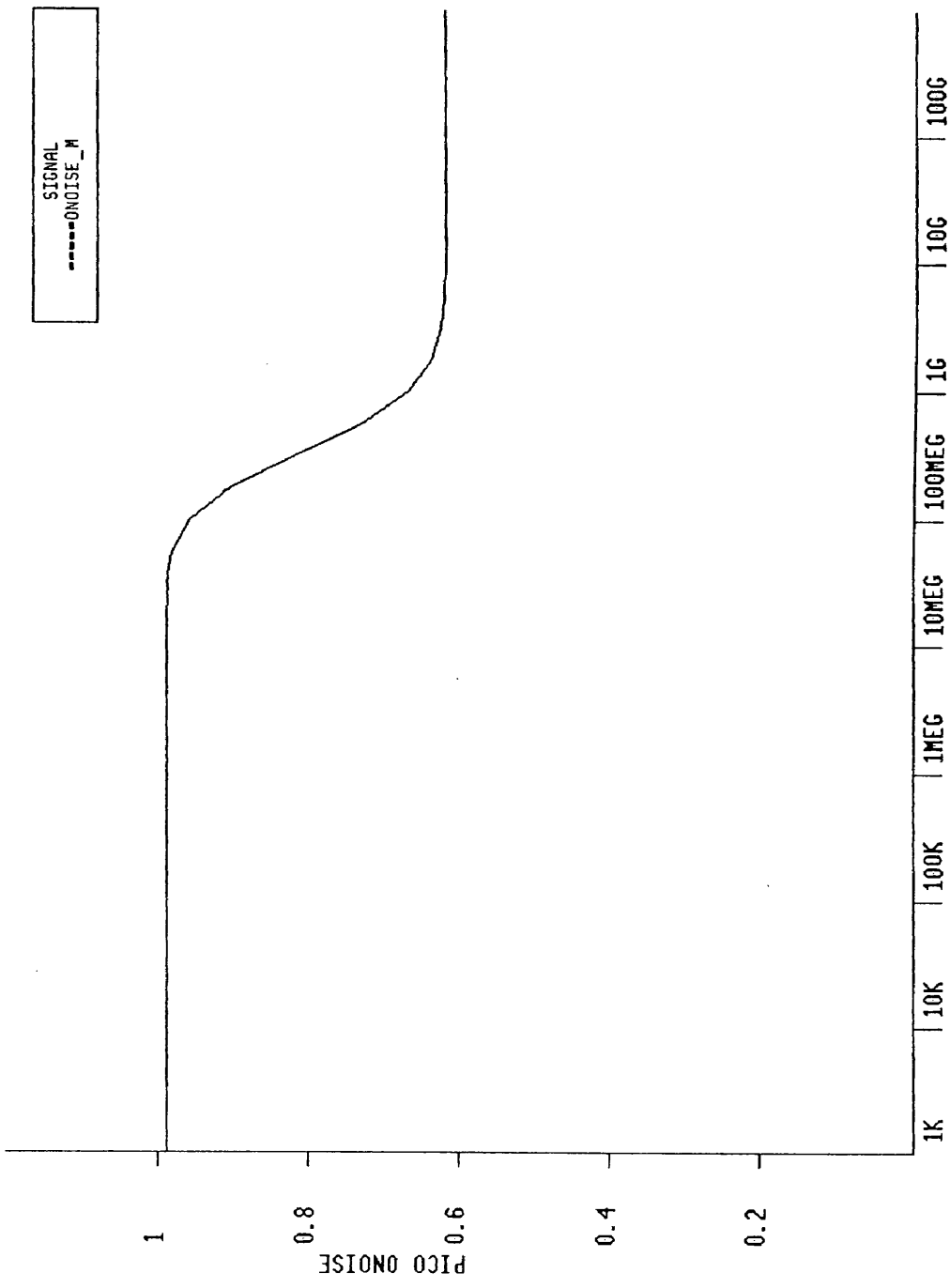


Fig. 162: Output-Referred Steady-State Current Noise

(current mirror), which by itself was responsible for half the noise power. In a charge amplifier configuration, the current sources that realize a current ratio share the same biasing line. Noise on the biasing line does not alter the current ratio. As a result, this effect should have been discarded, and an extra 6 dB of dynamic range would have been seen.

14.12. Supply Noise Rejection

A common source of "noise" in practical systems, is interference that reaches the analog signal path from the power supply. Voltage fluctuations may be present on the supply lines due to the switching of digital gates within the circuit, or due to the normal operation of other stages or sections of the same system (cross-talk).

If supply noise reaches the holding capacitor, it is mainly through a supply-dependent variation in comparator offset and logic thresholds. The effect is normally limited, especially since low-frequency components are rejected by the correlated double sampling. Another favorable aspect of the charge amplifier, is that very little supply noise is *generated* by each stage. As a result, it is less likely that this noise would interfere with other parts of the same system.

During the precharge phase, supply current is drawn, but the pattern is perfectly repetitive and signal-independent, and hence does not cause any cross-talk to other stages. This is especially true since the precharge phase of one stage normally never coincides with a critical phase (initial or final discharge, or load phase) of other stages. In a typical pipelined analog to digital converter configuration, there is a three-phase (half clock period) shift between successive stages (figure 163). Phase Φ_1 of the second stage corresponds to stage Φ_4 of the first stage and vice-versa (the load phase Φ_3 of the second stage corresponds to the final discharge phase Φ_6 of the first stage). As a result, when one half of the stages in the system pre-charges, the other half performs a flash operation, which has been shown to be very insensitive to errors.

During the other clock phases, the holding capacitors in the system are successively discharged using one or more of the current sources. During such operation, current sources and capacitors are connected in a loop, and no current variations other than those from the switching devices reach the supplies.

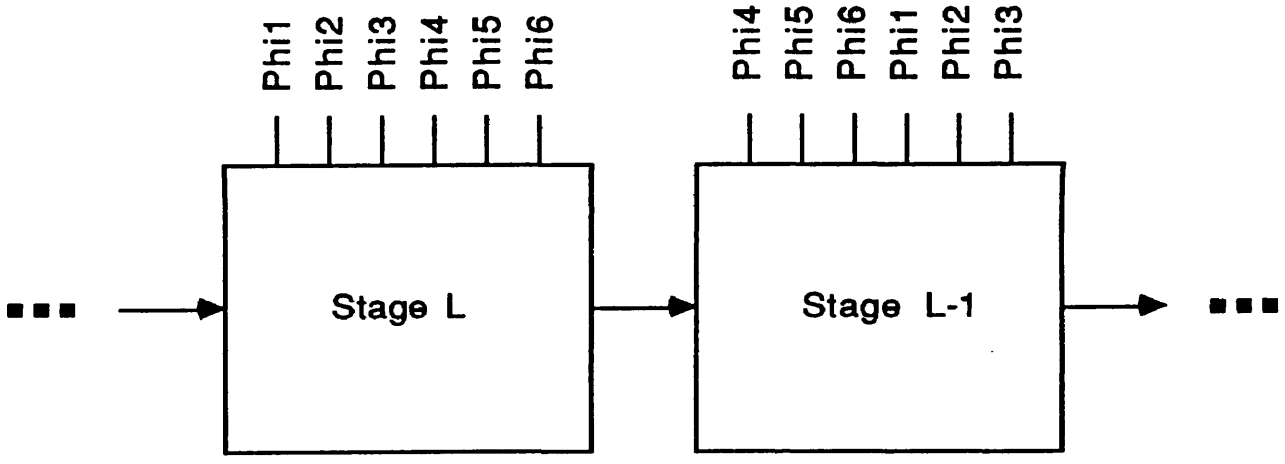


Fig. 163: Typical Pipeline Configuration

14.13. Conclusion

In this chapter, a systematic analysis of all the noise effects in the charge amplifier was given. To keep the discussion concrete, a particular configuration with a gain of -2 was described as an example. Circuit simulations (SPICE) for a typical 2μ technology were presented as illustrations. It was shown that due to the switching nature of this amplifier, a precise, mathematical noise analysis is almost impossible, as opposed to systems that operate at a fixed bias point. Instead, a number of significant, but justifiable approximations are made. In particular, two switching models were proposed: the stable feed-back model and the sequential switching model. The latter model seems more realistic, and bears a lot of resemblance with the effects that take place during the operation of a ring oscillator. Based on this model, the noise effects during the different phases of operation of the charge amplifier were discussed. Some numerical examples derived from simulations were given, based on white noise only. From these examples, it was implied that a dynamic range in excess of 60 dB should be easily realizable. Finally, a short discussion was given about supply noise rejection.

CHAPTER XV

IMPLEMENTATION OF THE PROTOTYPE CONVERTER

15.1. Introduction

This chapter describes a complete pipelined A/D converter prototype, based on charge amplifier stages. The self-calibration is accomplished using accuracy bootstrapping. The complete system consists of 16 nominally identical stages, each with a nominal gain of -2. This configuration realizes an overall resolution of 16 bits (one bit per stage). If all stages had a perfectly linear, reproducible gain, up to 16 bits of integral non-linearity could be realized. In practice however, the scheme is believed to be linear only to the 13-14 bit level, due to inherent linearity limitations of the charge amplifier stages and due to noise that could interfere with the self-calibration.

An actual prototype similar to the one described here has been integrated by Texas Instruments Inc. of Dallas, Tx, using a proprietary 5V, 1 μ CMOS technology. The total silicon area is less than 8 mm^2. It is believed that the conversion rate of the prototype could reach 2.5 MHz with a total current consumption below 10 mA. The conversion rate could be pushed as high as 5 MHz, using a master clock frequency of 40 MHz. However, the current consumption would have to be increased more than proportionally in order to achieve adequate switching speeds in comparators and current sources. At 5 MHz, the linearity could slightly deteriorate, since switching transients do not get as much time to die out.

Unfortunately, at the time of this publication, no experimental results were available yet to confirm the linearity after calibration. A lay-out error caused the device to remain in the self-calibration mode, instead of switching to continuous conversion. So far, only the operation of the clocking scheme, the calibration cycle and the current consumption have been verified.

All the building blocks of the prototype are described from a top-level point of view. When relevant, more detailed circuit schematics are discussed. Since the purpose of the prototype was only to demonstrate the capabilities of accuracy bootstrapping, a lot of the required arithmetic functions are performed off-chip

rather than real-time on-chip. However, possible on-chip implementation will be briefly discussed.

In addition to the pipelined approach, a few short sections are included about recycling and time-interleaved versions of the same hardware. The former achieves the same accuracy in a much smaller area, at the price of conversion rate. The latter approach realizes a much higher throughput, at the price of proportionally increased chip area.

It will be shown that time interleaving and accuracy bootstrapping could make it feasible to realize a massive, continuously self-calibrating, 12 bit, 40 MHz converter. This level of performance would be significantly beyond anything ever achieved in any technology at the time of this writing. Such performance could prove critical for certain high resolution digital image processing applications, like could be associated with high definition television (HDTV) systems. It should be noted that in TV and video applications, periodic recalibration of the chip can be carried out during the inter-frame image blanking time. This makes it possible to virtually eliminate any effect of drift, without sacrificing any useful conversion time.

15.2. Clock Generation

This converter requires a six-phase non-overlapping clock, of which two phases (Φ_3 and Φ_6) have double the length of the other phases (figure 164). The clocking scheme is derived from an externally supplied master clock, running at eight times the converter sampling rate. The generation of the clock signals occurs in two steps.

The first step is the generation of a two-phase non-overlapping clock. This is easily accomplished using the circuit of figure 165. The purpose of the two-phase clock is to drive a synchronous modulo-8 Gray counter (second step in the clock generation). The Gray counter is implemented as a standard state-machine with three data flip-flops and combinational feed-back circuitry (figure 166).

The logic is designed in order for the state variables (flip-flop outputs) to cycle through the combinations 000, 001, 011, 010, 110, 111, 101 and 100. Obviously, the advantage of this approach is that only one bit changes at a time. This will guarantee the absence of glitches during the decoding of individual clock phases.

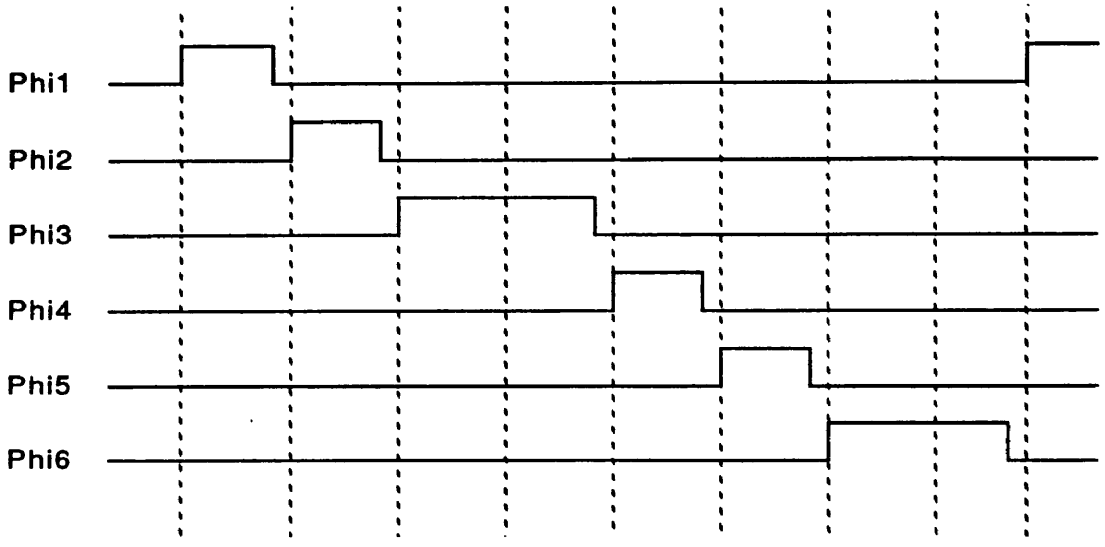


Fig. 164: Clock Timing

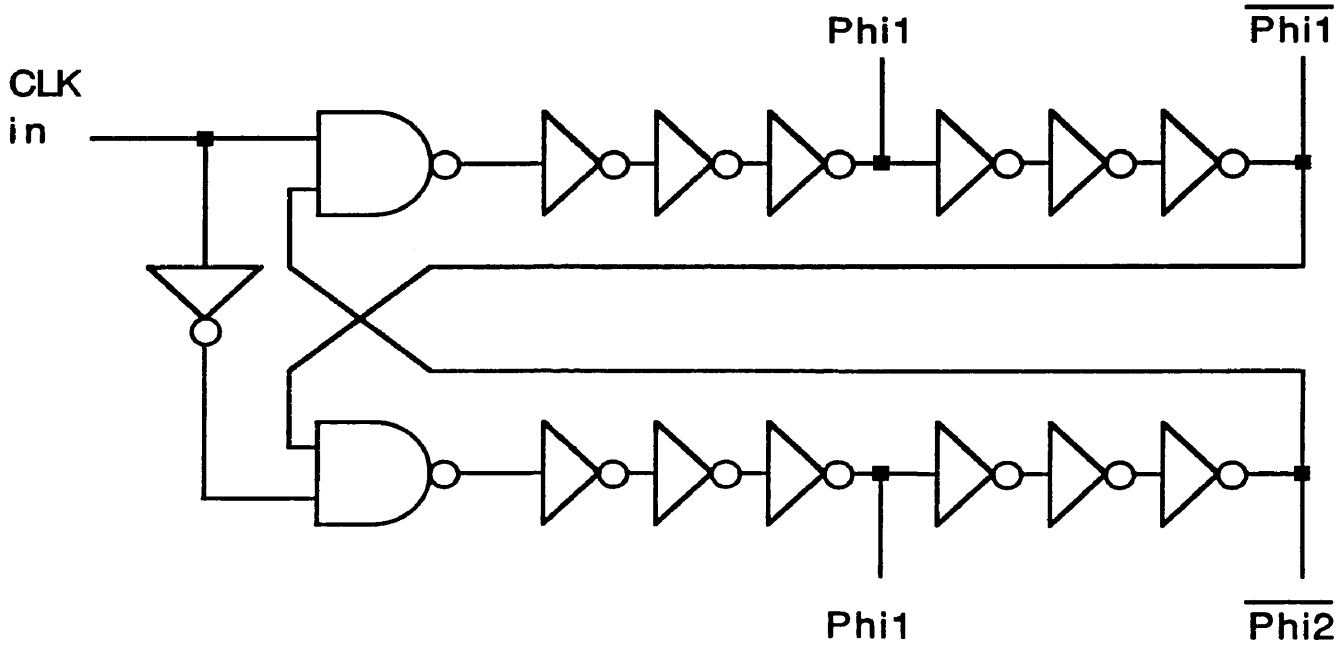


Fig. 165: Two-Phase Clock Generator

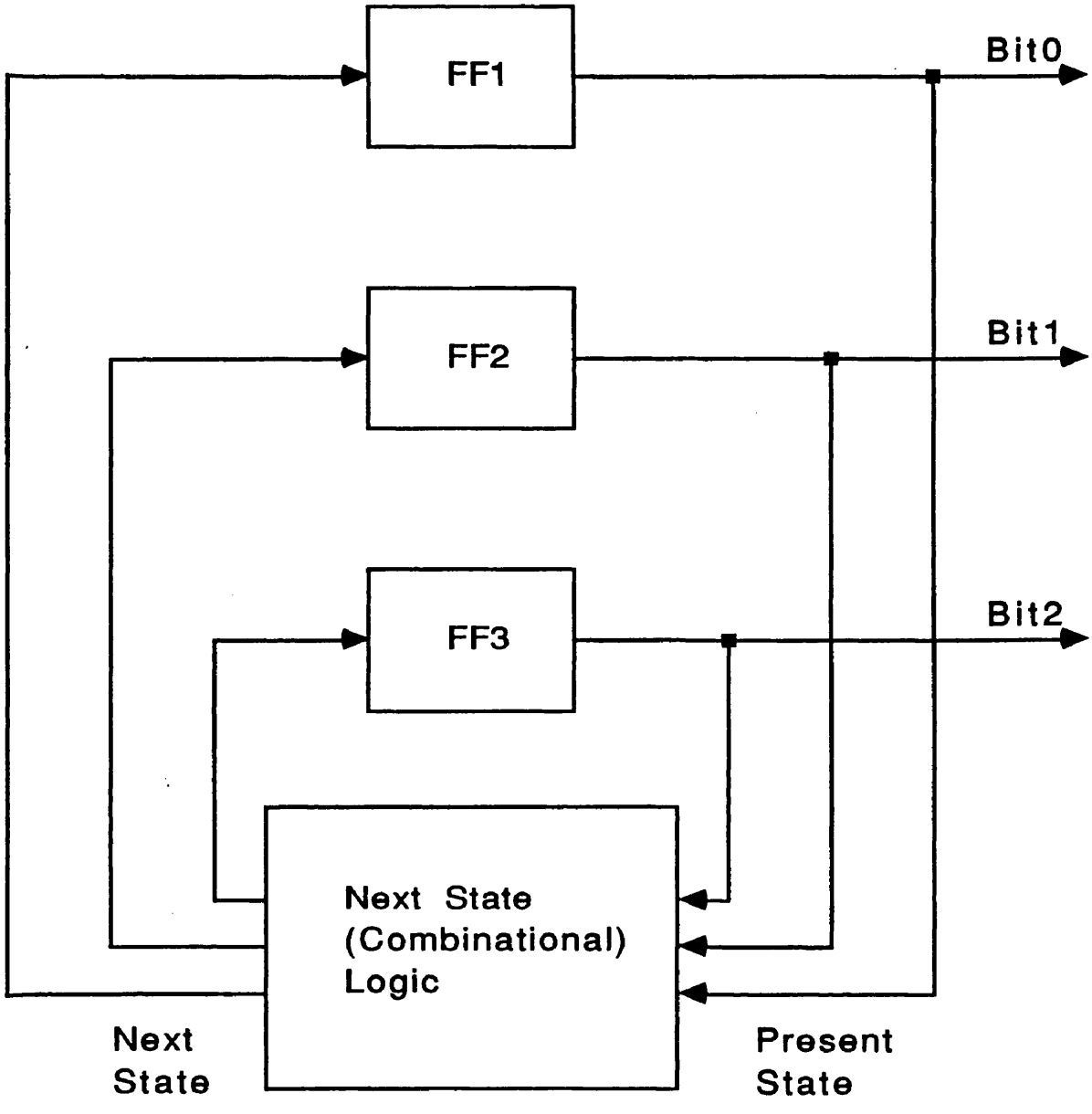


Fig. 166: State Machine

The clock phases are derived from the state variables using an array of AND gates. Φ_1 is obtained by decoding 000, Φ_2 from 001, Φ_4 from 110 and Φ_5 from 111. The two long phases (Φ_3 and Φ_6) are derived by decoding the two most significant bits only, for 01 and 10 respectively.

The six resulting phases are guaranteed to be glitch-free. However, they are not guaranteed to be non-overlapping. Separation of the phases is obtained with the circuit of figure 167, which is an generalization of the two-phase circuitry shown on figure 165.

These six clock outputs will be referred to as the six *main* clock phases. This is to be distinguished from the six phases used to clock individual charge amplifier stages, which can correspond to different main clock phases due to the pipelining.

15.3. Basic Pipeline Stage

The functional block diagram of the converter stage is identical to the one discussed in chapter VIII and is repeated in figure 168. The stage has a nominal internal range of 1 V, with $\pm 25\%$ overrange capability. However, the practical implementation is now based on charge amplification and PWM-coded signals. The basic gain stage is an extension of the charge amplifier stage discussed in chapter XIV and is shown on figure 169, as well as the associated flash section.

The holding capacitor has a nominal value of 1 pF; the current sources are implemented as regulated cascodes that sink a nominal current of 20 μA, resulting in a dV/dt of 20 $V/\mu s$ (40 $V/\mu s$ during the load phase). This is sufficient for operation at 2.5 MHz, since each clock phase has a nominal length of 50 ns. The holding capacitor is precharged to 5 V (nominal supply voltage), using a P-channel MOSFET with source connected to 5 V (power supply). The initial discharge level is 4.5 V, which is set through a resistive voltage divider and an analog multiplexer. The nominal time required to discharge the capacitor from 5 V to 4.5 V at 20 μA, is $0.5V * 1pF/20\mu A = 25ns$, which leaves a comfortable safety margin for switching transients.

The charge that corresponds to the initial discharge level of 4.5V, will be the initial charge of the system for the subsequent load phase. The highest nominal signal level on the capacitor, *after* the load phase, is 3.5 V has been chosen to be 1V below the initial discharge level. The associated capacitor charge corresponds

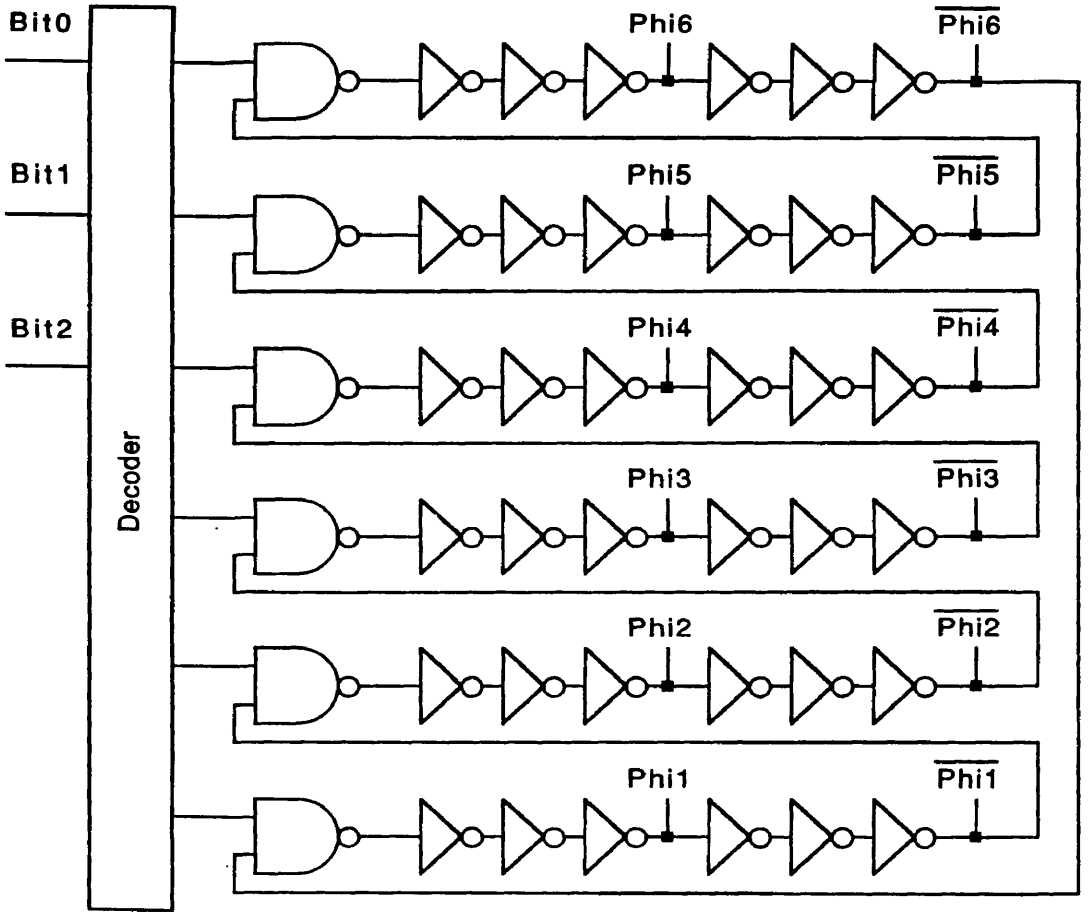


Fig. 167: Six-Phase Non-Overlapping Clock Generation

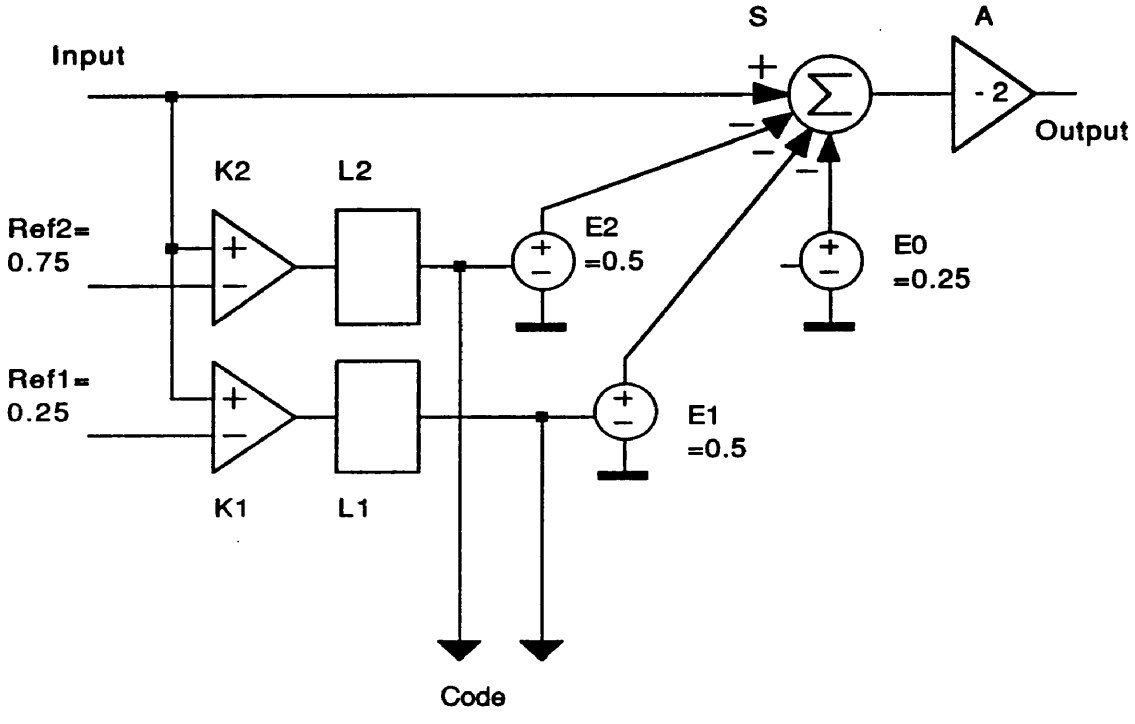


Fig. 168: Block Diagram of a Converter Stage

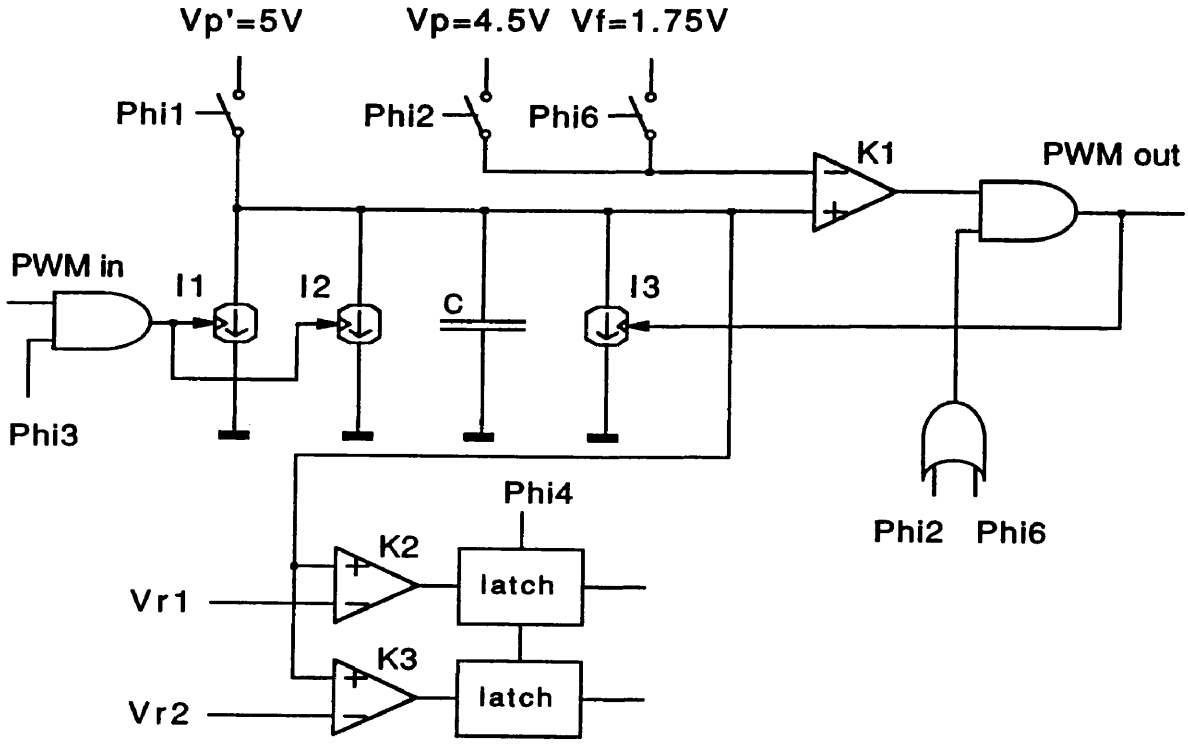


Fig. 169: Basic Gain Stage with Flash Section

to the high end of the nominal signal input range (a "1" level on figure 168) and is reached when two current sources (40 μA total) are turned ON for 25 ns, after the initial discharge. This implies that at 20 μA per current source, the minimum input pulse width should be 25 ns. This systematic time offset is necessary to accommodate finite PWM signal edges and switching transients when the current sources are switched ON.

The lowest nominal signal level on the capacitor, after the load phase, is 2.5 V. This corresponds to the "0" level of figure 168. The difference in pulse width required to generate a 1 and a 0 is 25 ns, which is the nominal PWM range. Due to errors in the components of previous stages, this PWM range may be slightly exceeded in both directions (shorter or longer pulse). However, up to ± 5 ns extra pulse width (or associated $\pm 25\%$ variation in capacitor charge, or ± 0.25 V) can be tolerated, due to the $\pm 25\%$ overrange capability of the converter scheme.

Including over-and underrange, the capacitor voltage after the load phase could be as high as 3.75 V and as low as 2.25 V. This voltage is compared against two reference levels (2.75 V and 3.25 V respectively) during the flash phase, and generates a two-bit thermometer code. The comparison is performed using one comparator for each level, followed by a transparent latch. During the load phase, the latch operates in transparent mode. During the flash phase, the latch is closed and the data allowed to stabilize to a logic 1 or 0 level. Depending on the resulting code, a certain charge is subtracted during the DAC phase. The charge increments of the DAC are such that they would generate a voltage drop that is a multiple of 0.5 V ($0V$, $0.5V$ or $1V$). The result of this operation is that after the DAC phase, the capacitor voltage will be between 2.25 V and 2.75 V. The DAC section is implemented by two switchable current sources and will be discussed below.

During the final discharge phase, the capacitor is discharged using one 20 μA current source, down to 1.75 V. The time required to do this, is obviously between 25 ns and 50 ns. This precisely meets the pulse width requirement to the next stage, and makes the whole scheme consistent. It should also be clear why the load phase and the final discharge phase are double the length of the other phases. Since only 50 ns are allowed for the single phases, the total pulse width (including switching transients and delays) may not fit within such a phase. doubling the

duration of the relevant phases solves this problem without having to sacrifice too much of the total clock period (and conversion rate).

It should be stressed that in this scheme, neither the precise values of the holding capacitors nor those of the current sources are critical, as long as reasonable matching is maintained. A systematic error in the values of all capacitors and/or current sources of the system would only result in a pulse width variation, and not in a gain error. However, is a *mismatch* were to exist between nominally identical components, a gain and/or offset error would result. As long as this error remains within a range of a few percent, the accuracy bootstrapping algorithm will be able to compensate for it.

15.4. Current Sources and DAC Section

The complete stage is implemented with only three main current sources. One current source is used for the initial and final discharge phases. Two current sources are used for the load phase. These two sources normally share the same bias line as the discharge source of the previous stage, in order to improve matching (nominal gain of -2) and noise immunity (if the bias line of simultaneously operating discharge and load sources is subject to the same variation, the current ratio is maintained and no significant error occurs).

In addition to those, up to three auxiliary current sources may be required to implement the DAC action. This is shown on figure 170. The three auxiliary current sources normally share the same bias line, also for noise immunity. The two current sources that implement the DAC action can be ON either individually or together, depending on the local conversion (thermometer) code. Depending on which case occurs, the time derivative of the capacitor voltage may be either 20 $V/\mu s$ or 40 $V/\mu s$. In this scheme, it has been found that the different slopes do not significantly affect the steady-state or transient currents of the current sources. Had this been the case, a separate clock phase could have been allocated to each source, in order to allow successive rather than simultaneous switching. This would have resulted in a slightly reduced overall conversion rate and additional circuit complexity.

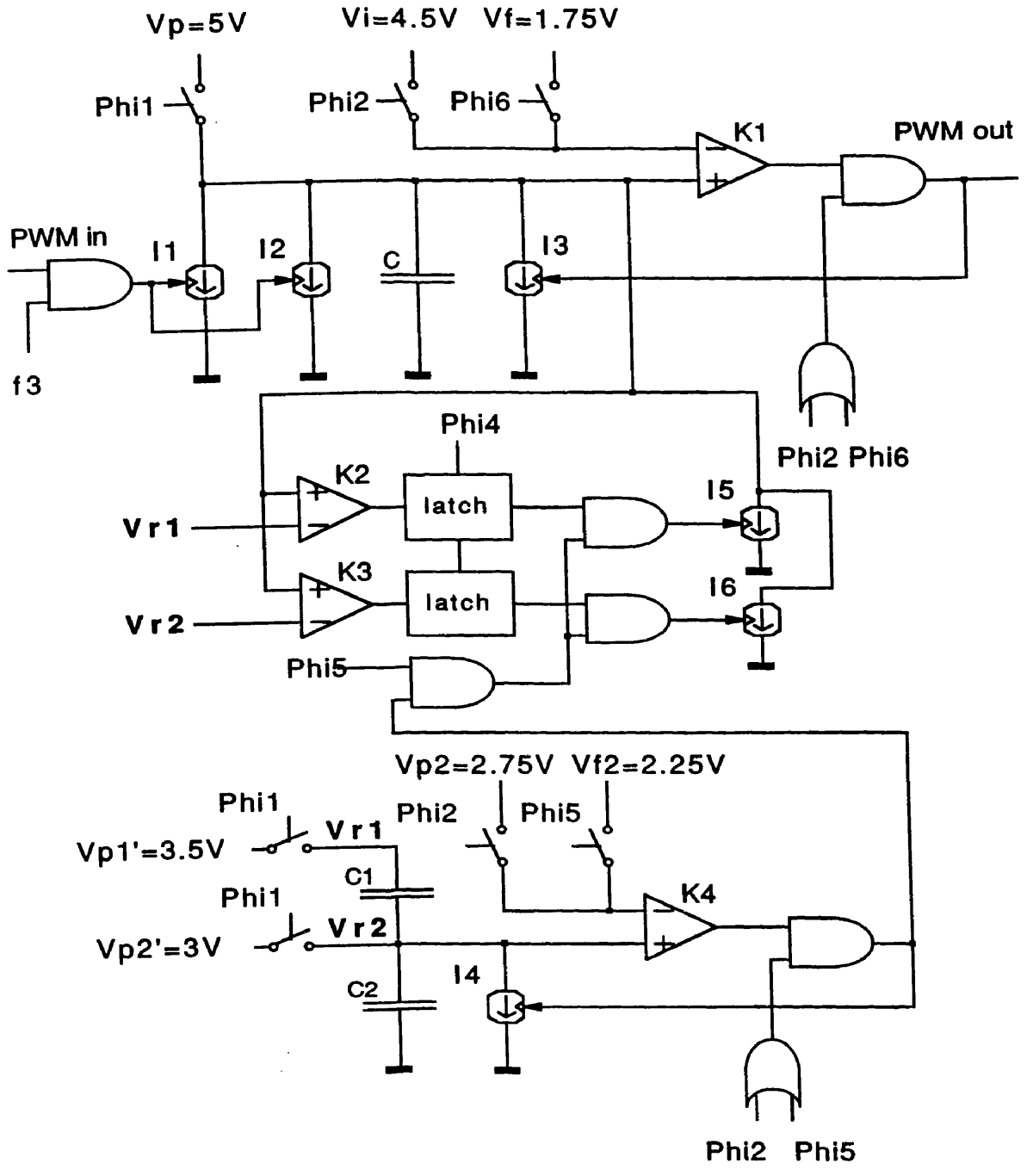


Fig. 170: Converter Stage with DAC Section

15.5. Offset Compensation in the Flash Section

Since the charge amplifiers are operated at extremely short pulse widths (25 ns difference represents the full signal range), cancellation of the offset that is introduced by finite comparator and current source switching delay becomes crucial. This delay is of the same order of magnitude as the PWM signal pulse width itself. Intrinsically, accuracy bootstrapping will only absorb offset errors graciously, but only as long as they are significantly below 25% of the signal range (over-range capability of these particular stages).

The effect of these finite time delays is reduced well below this critical level by the correlated double sampling scheme (initial and final discharge phases), at least as far as pulse width relationship between input and output of a stage is concerned. However, the scheme causes the voltage range on the capacitor (nominally 2.5 V to 3.5 V) to be subject to a significant offset (typically 100 to 300 mV in this technology, on a signal range of $1V$). Such offsets would corrupt the calibration if they were not compensated for by corresponding shifts on the comparator reference levels.

The compensation is obtained by shifting the comparator reference levels by the same amount. This can be achieved with a capacitive divider and an extra comparator, which matches the main comparator. Since a physically different comparator is used, the matching may not be perfect, but as long as the total offset is reduced to a few percent of the total signal range, the calibration algorithm will function correctly. One capacitor of the capacitive divider, and the comparator are shared with the DAC circuitry in order to reduce total silicon area and power consumption. This is shown on figure 170.

15.6. Data Latches

As mentioned before, the flash section of each stage is realized by two comparators and associated transparent latches. These can be standard digital building blocks, based on transmission gates (figure 171). During the load phase, the comparator output signals are passed through the input transmission gates.

At the end of the load phase (Φ_3), the comparator input signals (capacitor voltage) stabilize, and a little under 25 ns are available for the comparator output voltages to stabilize as well. At the beginning of the flash phase (Φ_4), the latches switch from the transparent mode to the hold mode. The input transmission gates

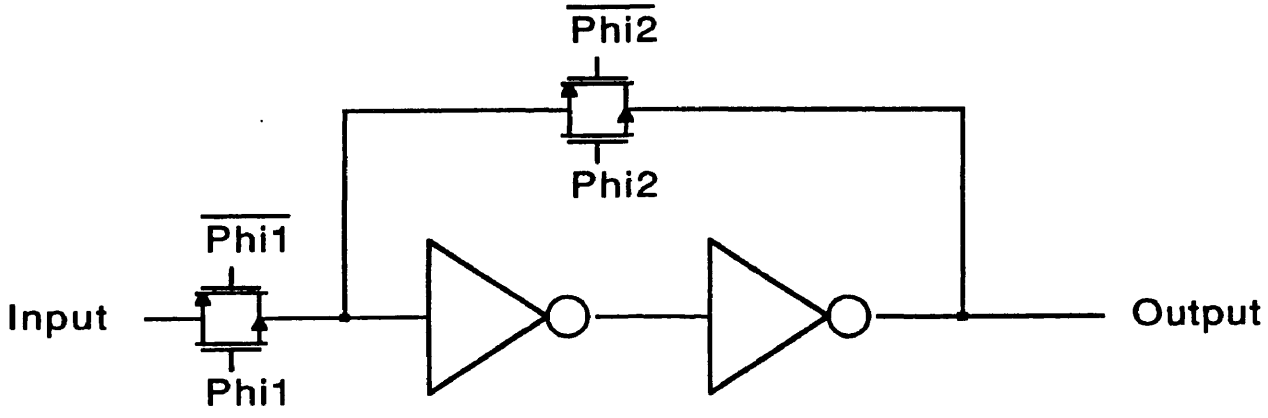


Fig. 171: Schematic of a Data Latch

are opened, and the feedback gates closed. Again, 25 ns (one clock period) are available for the latches to evolve to a stable equilibrium situation (0 or 1).

Of course, there is always a finite probability that such stable situation would not be reached within the allocated time (when the capacitor voltage is very close to the comparator reference level). This probability was considered insignificant in this design, due to the high loop gain and speed of the latches. Similarly, it is possible that between two conversions, the last state of a latch would influence its next state (hysteresis). Since in this scheme, relatively large comparator errors can be tolerated without affecting overall ADC linearity, this effect was also neglected.

15.7. Input Stage

The drawback of designing a data converter based on charge amplifier schemes, is that the input signal should come in PWM form. Since most practical applications provide signals that are coded as a voltage, a "conversion" from voltage to charge, and then from charge to pulse width must be accomplished at the input. The conversion from voltage to charge can be accomplished by sampling the input voltage onto a capacitor with the help of MOSFET switches, in a stray-insensitive fashion (like in the switched-capacitor gain stage of chapter IX). At the same time, the input switch acts as a sample/hold device, which freezes the (varying) input signal at a specific instant in time. The conversion from charge to pulse width is accomplished using a fast comparator and a switchable current source, like shown on figure 172.

The input section seems to be the weakest link in this particular system. It must be preceded by a low output impedance, unity gain buffer. This normally requires a high speed, high linearity, low noise operational amplifier. Because of the switches, clock feed-through (and associated non-linearity) is of major concern. At the same time, the switches must be fast enough to sample a rapidly changing signal without noticeable dynamic distortion. In addition, this input stage is *not* insensitive to capacitor non-linearity and offset errors like the stages within the pipeline are.

In the prototype, a folded cascode amplifier with output stage biased at about 50 μA was used (figure 173) for the unity gain buffer. This amplifier turned out to perform in a marginal to poor way as far as maximum sampling speed without

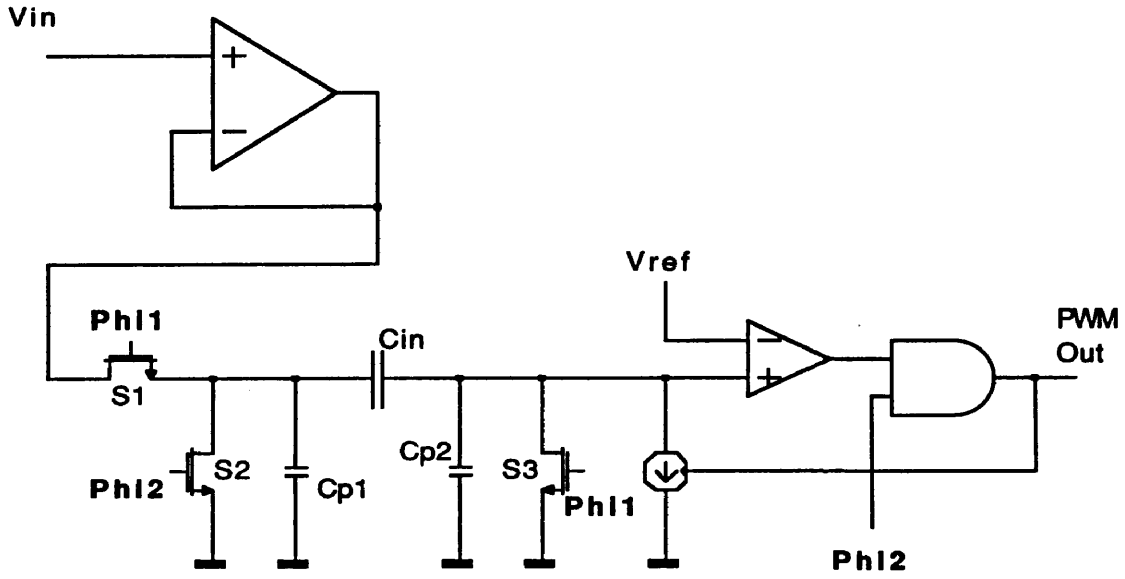


Fig. 172: Input Section of the Converter

linearity degradation is concerned. However, it would accurately sample the input signal for medium-frequency operation and characterization of the converter. For actual high-speed, high-linearity operation, a high slew rate Castello type amplifier would probably have been a better choice. Further research may also be useful, in order to develop a more direct conversion method from voltage to pulse width.

15.8. In-Stage Calibration Circuitry

In order to apply the accuracy bootstrapping calibration algorithm, a certain "block operation control" capability must exist within each converter stage. In this case, it must be possible to apply a certain reference level, and then that same reference level, minus the difference between two successive DAC levels (chapter VIII).

Within the charge amplifier scheme, the subtraction of a differential DAC level is accomplished by switching ON one of the DAC current sources for a finite time, set by the reference source and the comparator of figure 170. The amount of charge displaced by this operation is normally extremely independent of the signal level on the holding capacitor. This means that during calibration, the DAC operation will behave exactly as it would during normal conversion, especially if the calibration is performed at actual speed (same clock rate as normal conversion).

The reference level must be chosen so that it can adequately be converted by the next stage in the pipeline, and so that the signal obtained after subtraction of one differential DAC level could also be converted. One can easily verify that the reference level should correspond to a capacitor voltage of 2.75 V. Subtraction of one DAC increment ($0.5V$) would result in a final voltage of 2.25 V. These levels would cause a nominal output pulse width of either 25 ns or 50 ns, using a current of 20 μA and a 1 pF capacitor.

As a result of this arrangement, the initial discharge level during calibration should be 2.75 V rather than 4.5 V (excluding undershoot), since no additional charge is subtracted during the load phase. This voltage can be applied to the comparator during the initial discharge phase using an additional multiplexer, as shown in figure 174. In order to be able to reach the initial voltage level within 50 ns (one clock phase), the precharge level should now be 3.25 V rather than 5 V. This can be achieved using another switch (P-channel MOSFET), connected to a

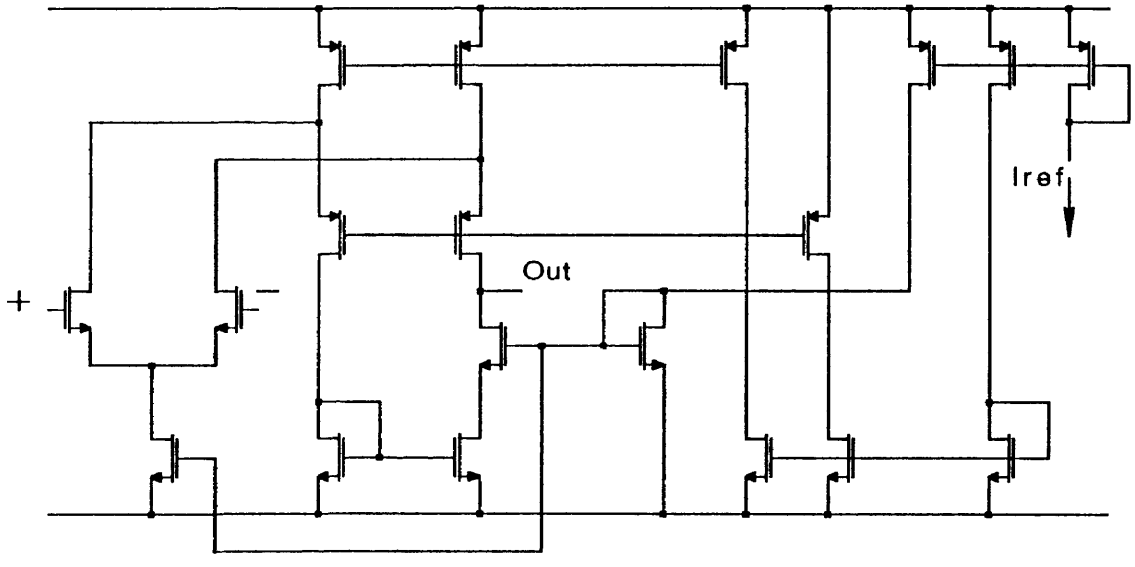


Fig. 173: Folded Cascode Amplifier

different reference voltage for the precharge operation of the calibration mode.

A digital signal is used to distinguish between calibration and normal operation of the stage. The pull-down logic used to disable the two current sources of the load phase is such that when the calibration signal is high, the current sources would be disabled. This effectively inhibits the input signal during the load phase. For the sake of simplicity, this is not shown on the figures During the complete calibration cycle of the converter, all stages will be connected in a loop. The stage that is currently being calibrated "breaks" the loop, since its analog input is not propagated any further.

In the calibration mode, the two DAC current sources are controlled by external calibration inputs rather than by the comparator outputs (thermometer code). This is achieved by using two digital multiplexers, controlled by the calibration signal.

15.9. Calibration Logic

The calibration algorithm for this type of pipelined converter was discussed in chapter VIII. The calibration starts from the last stage. The reference level is applied to that stage and the converter consisting of the first stage up to the last one is used to measure this level. The conversion result, N_1, is stored. The same thing is done for the reference level minus the first and the second DAC increment, resulting in the measurements N_2 and N_3, respectively. These three numbers are used to update the digital coefficients of the last stage (stage 0). Next, the stage before the last one (stage 1) is calibrated in a similar way (all stages are connected in a loop), and the process is repeated until the first stage (stage 15) is calibrated.

If the system were noise-free, one such cycle would be sufficient to calibrate the converter to the 15 bit level or beyond. However, since at the time the prototype was designed, no accurate noise estimate of the charge amplifiers existed, it was decided to average 16 calibration measurements every time. As a result, the calibration consists of three nested cycles, which are implemented by an on-chip sequencer.

The first cycle consists of applying a certain input value (reference level, or reference minus DAC increment) to one stage (the stage under calibration), and using the other stages to measure this quantity. This operation must be performed

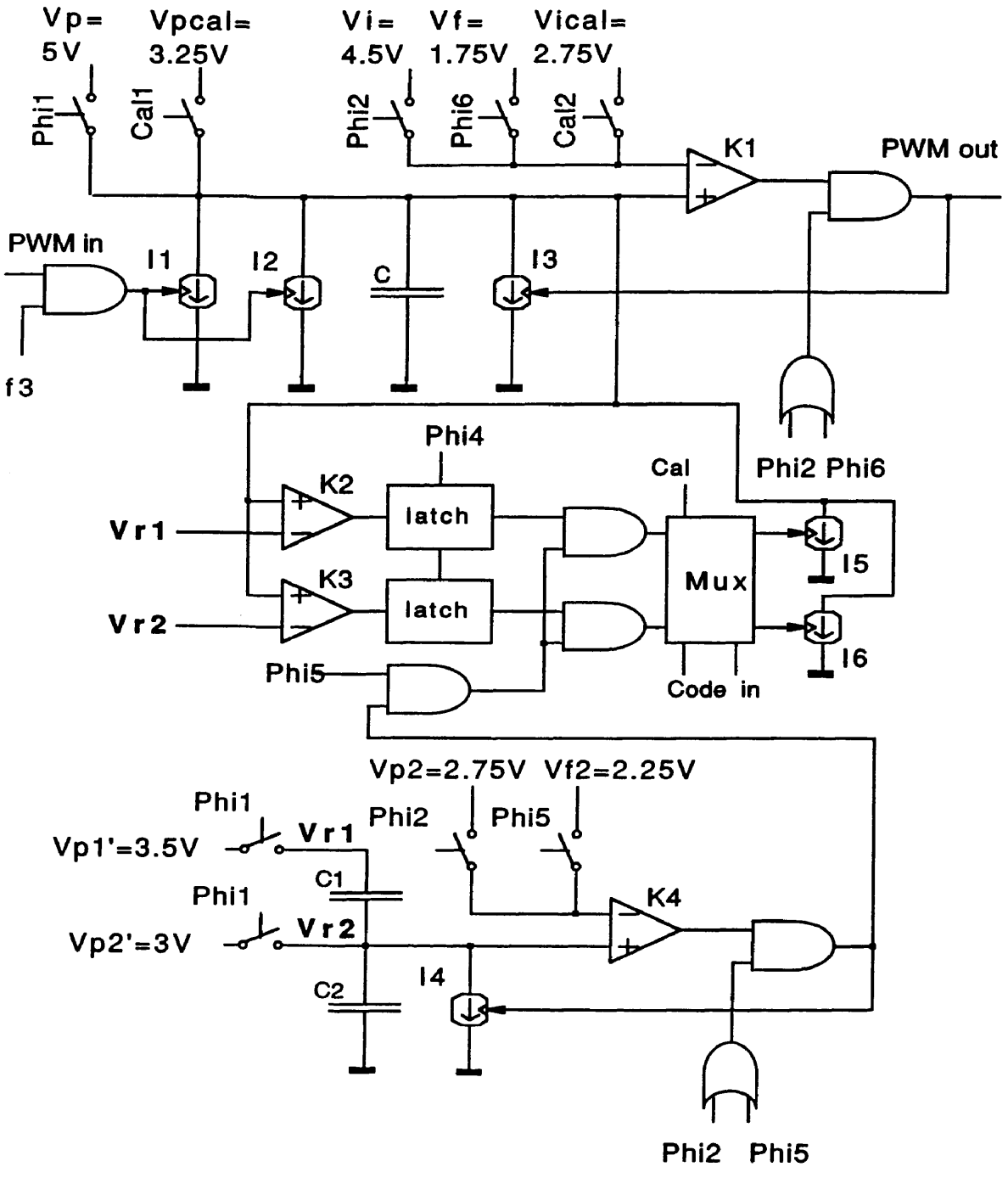


Fig. 174: In-Stage Calibration Circuitry

for each of the 16 stages in the system. Due to the pipelining, precise timing of this inner calibration cycle requires special attention.

First of all, it should be realized that although the accuracy bootstrapping algorithm requires the measurements of stage 0 to be *used* first (to update the coefficients of stage 1), it is not required that this measurement be *made* first. In fact, a completely difference measurement sequence can be applied, as long as the uncorrected measurement results (raw comparator output bits) of each stage are stored. Since the actual calibration of the prototype is to be performed off-chip, this condition can easily be maintained. As a result, there is no objection to starting the calibration from the first stage (stage 15) and proceeding to the last one (stage 0). This approach will simplify the sequencing hardware.

Upon power-up, a general calibration signal ("flag") is enabled, as shown in figure 175. This signal disables the converter input and connects all stages in a loop. This is achieved by multiplexing the input of the first stage (stage 15) between the system input (the PWM signal coming from the input section) and the PWM-coded residue from the last stage (stage 0).

An individual calibration signal (which puts one stage in calibration mode and disables its PWM input) is applied to stage 15 starting on phase Φ_1 of the first clock cycle. For reference, this first clock cycle will be called clock cycle 0, and phase ϕ_1 of this first cycle will be referred to as Φ_1^0. During phase Φ_1^0 and Φ_2^0, the appropriate reference level is introduced (2.75 V, corrected for undershoot due to comparator delay through an initial discharge). During phase Φ_5^0, one (or none) of the DAC charge increments is subtracted. Which increment is subtracted, is controlled by the two calibration input bits of the stage. During phase Φ_6^0, a final discharge operation is performed, which generates the PWM pulse that goes to the next stage.

Because of the pipelining, the final discharge phase Φ_6 of stage 15 corresponds to the load phase Φ_3 of stage 14 (there is one half clock period of skew between stages). Stage 14 will convert the signal during *its* Φ_4, which is Φ_1^1 of the main clock. It will pass on a residue to stage 13 during *its* Φ_6 (Φ_3^1 of the main clock). This means that stage 14 can only be put in calibration mode during the next clock phase (*its* next Φ_1, or Φ_4^1 of the main clock).

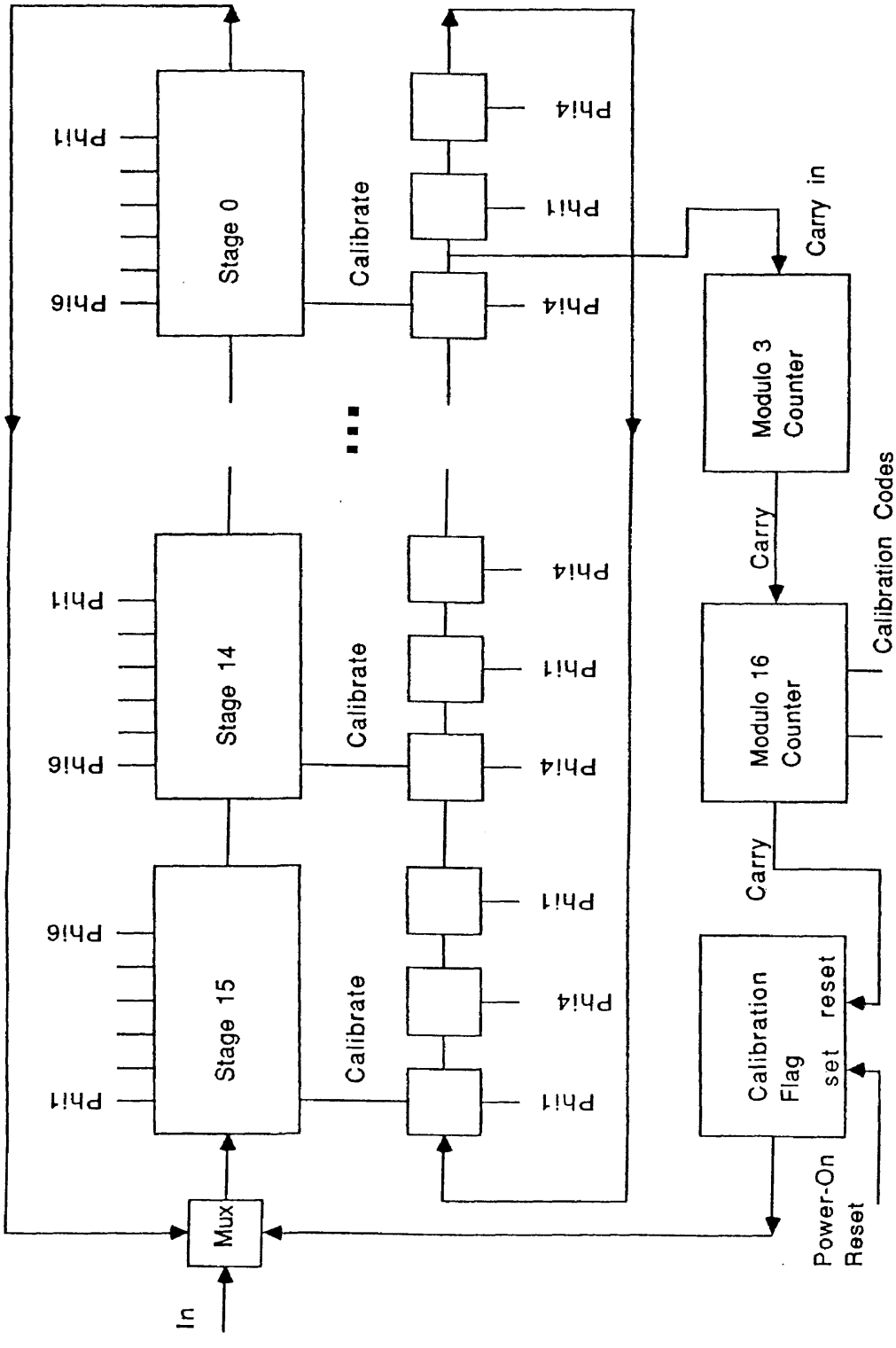


Fig. 175: Calibration Logic

From this discussion, it should be clear that there should be one and one half clock period skew between the times calibration signals are applied to successive stages. Since calibration signals are only applied to one stage at a time, they can conveniently be generated by a running one in a cyclic shift register (token ring), as shown in figure 175. The shift register consists of a sequence of transparent latches, clocked by complementary clock phases (master/slave configuration). In this case, the two phases are Φ_1 and Φ_4 of the main clock. To generate the correct timing, three latches are required for each stage of the ADC, bringing the total to 24.

The first latch of each stage will have a logic high output for exactly one clock cycle (the output of the latch changes state on the rising edge of clock phase Φ_1 or Φ_4, depending on whether the number of the stage is odd or even). As a result, this signal can be used as a calibration signal for that particular stage.

The token ring controls the inner cycle of the calibration sequence (successive measurement of one level in each of the 16 stages). However, there are three different levels to be measured for each stage. Alternation between these three levels forms the second loop of the calibration sequence. This cycle is implemented by a synchronous modulo 3 counter with carry input and output. The counter is implemented as a standard state machine, with two state variables (data flip-flops), one input (carry in) and three outputs (bit 0, bit 1 and carry out). The carry output goes high whenever the counter is in state 10, AND the carry input is high.

The two bit inputs are used to control the DAC current sources of the ADC stage being calibrated. The logic of the modulo 3 counter is designed to generate the sequence 00, 01 and 10 on the bit outputs. A transition is made in the clock cycle (Φ_1 and Φ_4) following a logic high on the carry input. The output of the first latch (clocked on Φ_4) of the last ADC stage is used as carry input. This ensures that the modulo 3 counter would change state every time the "token" one reaches the first latch of the first ADC stage (stage 15). The logic is shown on figure 175.

Once the three combinations have been run through the 16 stages, all relevant calibration information should be available. However, in order to reduce the effect of random, zero-mean noise during the calibration, 16 successive measurements are made, which can later be averaged. This is accomplished by a third loop in

the calibration sequence. The complete sequence is depicted in figure 176.

The hardware for this outer loop consists of a 4 bit (modulo 16) synchronous binary counter with carry input and output. The carry input of the modulo 16 counter is connected to the carry output of the modulo 3 counter. Whenever the token one reaches the second latch of the last ADC stage, a carry is generated into the modulo 3 counter. This carry is propagated whenever the counter state is 10. When the token one reaches the first latch of the first ADC stage, the modulo 3 counter returns to state 00, and the modulo 16 counter is incremented.

After 16 cycles, the modulo 16 counter returns to state 0000 and the calibration is complete. This is indicated by the carry output of that counter, which goes high during the last clock cycle of the complete calibration procedure. This carry output can be used to reset the "flag" flip-flop, which keeps the whole converter in its cyclic calibration configuration. Whenever the flag is disabled, the input of stage 15 is connected to the input section of the system (voltage to PWM translation), and continuous conversion starts.

It should be noted that due to the one and a half clock cycle delay for the token one to travel between ADC stages, and the one half cycle delay for the analog signal to travel between stages, the total number of codes (two thermometer-coded comparator outputs) available from each measurement will be more than 16. Indeed, one can verify that since both structures are cyclic, and since the ADC has 16 complementary-clocked stages while the shift register has 48 complementary-clocked latches, it will take 24 full clock cycles for the signal to catch up with the token one ($24 * 2 = 48 * 2 / 3$). As a result, each measurement of a stage will yield 23 bits (one bit per stage the signal travels through, minus one due to the fact that when a reference level is applied to a stage, no valid code is available).

This means that every time a measurement is made, the ADC pipeline is partially recycled. The first 8 (minus 1) stages are reused, yielding 7 extra bits. These bits provide additional resolution to the calibration. However, it has been shown in chapter VIII that calibration to the 15 bit level or beyond can be achieved without recycling (using a 16 bit measurement). As a result, within the device that computes the actual memory coefficients from the measurements (external digital computer in this case), these 7 bits could be discarded. However, if these bits are used, the compensation coefficients can be determined more accurately,

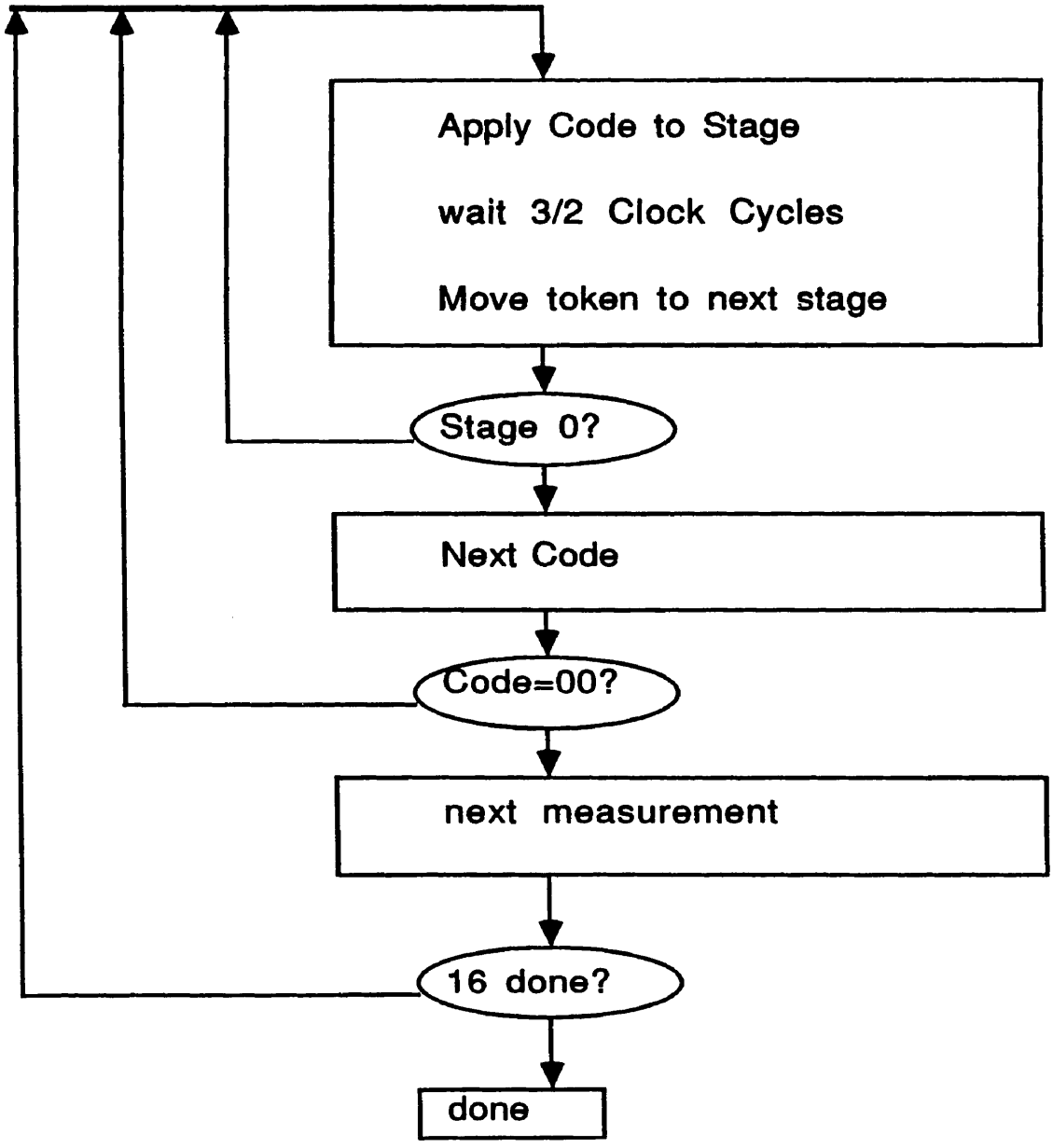


Fig. 176: Calibration Sequence

and the final, residual linearity error of the ADC will be beyond the 16 bit level. However, the resolution will be limited by the quantization error, which will always be exactly at the 16 bit level.

15.10. Reset Circuitry

The ADC prototype has two modes of operation: calibration and continuous conversion. The logic has been set up so that the device would start calibrating upon reset (16 complete calibration cycles), after which it would automatically start converting. The continuous conversion mode starts by applying an external input signal to the first stage of the ADC (stage 15), while calibration data are still coming out of the other stages. The timing of the calibration ensures that no conflict would occur.

It has been discussed how the calibration is stopped. The carry output of the modulo 16 counter is used to reset the calibration flag (flip-flop), which in turn connects the external input to the first stage, rather than the residue from the last stage (linear versus cyclic configuration). Upon reset, the opposite occurs. While the reset signal (either a true power-on reset or an external input to the chip) is active, all counters in the system are reset. This means that all latches in the cyclic shift register are set to 0, except for the first one, which is set to 1. The modulo 3 and modulo 16 counters are set to all zeros. The calibration flag flip-flop is set to one.

As soon as the reset line is released (this reset line should be synchronized for reliable operation), the one starts running and the device starts outputing calibration data. During the first clock cycle after reset no valid code is present (a reference level is applied to stage 15). One half clock cycle later, stage 14 outputs a valid code. One clock cycle after that, a reference level is applied to stage 14. After 8 complete clock cycles, the analog signal has traveled around the pipeline and all 16 stages output valid codes. The total calibration takes 24 * 3 * 16 complete clock cycles, after which stage 15 converts an input signal. 8 complete clock cycles after the end of calibration. the converter is fully in convert mode. Before that time, some stages are still outputing calibration data.

15.11. Output Multiplexing

A drawback of the redundancy built into each converter stage for error correcting purposes, is that the number of output bits (comparators) is doubled. There are two output bits per stage, which would require 32 output pins on the chip, for data alone. This number can be reduced back to 16 by a simple multiplexing scheme. Instead of outputing all data at once, for a complete clock period, all data from the odd-numbered stages (of which the first clock phase is Φ_1 of the main clock) are output during the first half of a main clock cycle. This represents 2 * 8 = 16 bits. During the second half of a main clock cycle, all data from the even-numbered stages (clocked on Φ_4) are output, through the same output pins. A schematic of the multiplexer is shown on figure 177.

15.12. Data Capture and Microcomputer Support

In order to measure the analog component errors of the converter as accurately as possible (reflecting actual operation, including as many dynamic errors as possible), the calibration has to be performed at actual speed. Since in this case, it is performed off-line by a separate computer, the calibration data must be temporarily captured and stored, at full speed. This is accomplished by an externally buffer memory module. As part of earlier research, such a module has been built, featuring 128k x 16 bits, operating up to 35 MHz. This is more speed than necessary, since the maximum output data rate of the converter is only 10 MHz (at 40 MHz input clock rate and 5 $Msample/s$ input data throughput).

The memory module communicates with a computer through an IEEE 488 (GPIB) interface, making it possible to download the captured data to a workstation or personal computer. On the computer, a certain amount of data processing must be performed:

- Demultiplexing the captured data from two 16 bit data words to 16 two bit codes.
- Shifting the appropriate codes back so as to compensate for the fact that all the codes of a same measurement become available during different clock phases (due to the pipelining).
- Changing the order of the measurements (the first stage was measured first; the calibration algorithm required the last one to be measure first).
- Converting the 16 two bit codes to an actual number, using the look-up

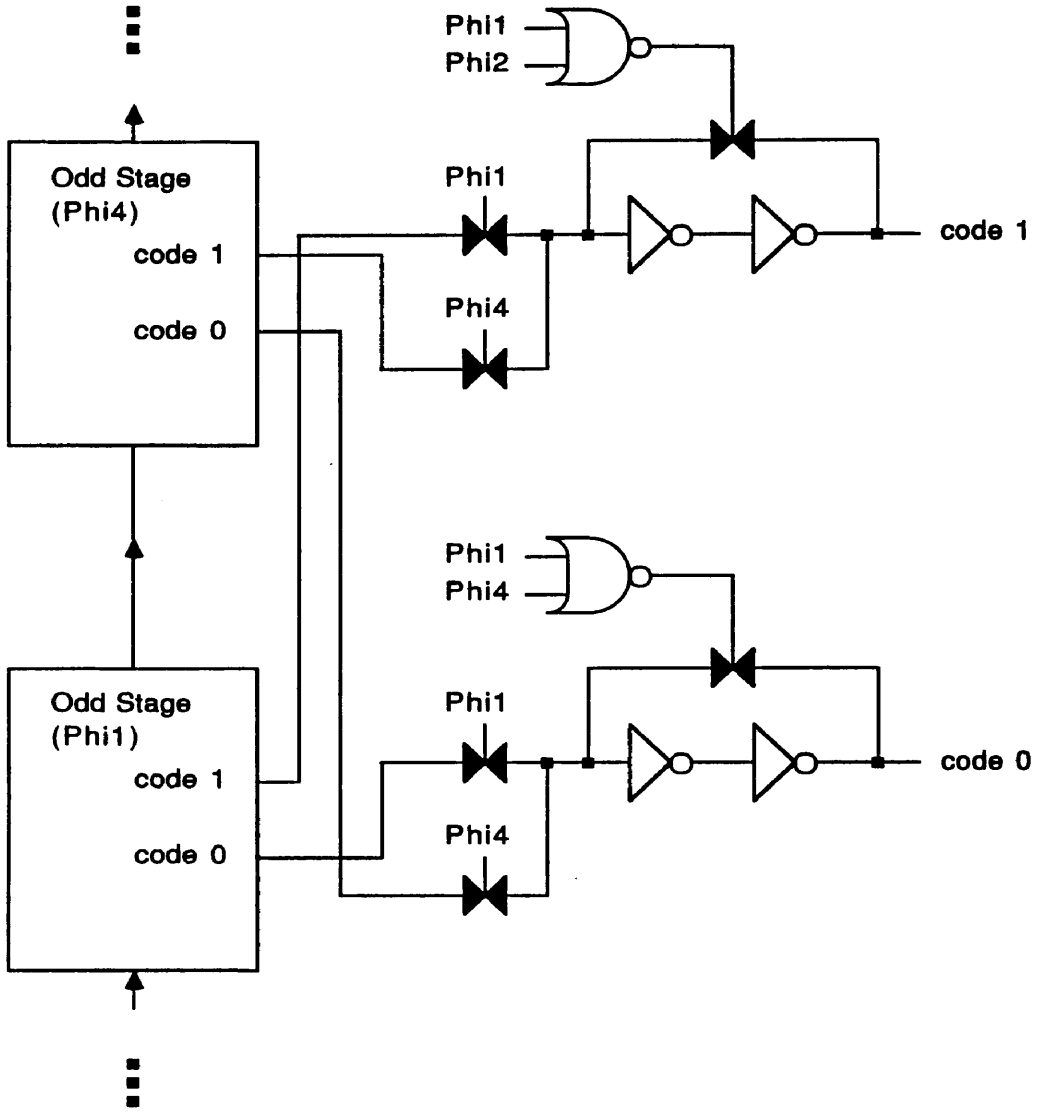


Fig. 177: Output Multiplexer

table (coefficients) associated with each ADC stage. The look-up tables are represented by memory locations which first contain nominal values, and are gradually updated according to the accuracy bootstrapping calibration algorithm.

Once the converter finishes calibration, it immediately switches to the continuous conversion mode. The digitized data (e.g. from an input ramp or sine wave, for test purposes) is downloaded by the same buffer memory board, and can be unscrambled by the same software that was used to unscramble the calibration data. At that point, a histogram or FFT plot can be constructed for evaluation.

15.13. On-Chip Digital Error Correction

The actual prototype does not contain the hardware for the calibration algorithm, nor the look-up tables and arithmetic blocks necessary for full-speed, continuous conversion. Both calibration and subsequent error correction are to be performed off-chip, as described in the previous section. This made it possible to reduce chip area and complexity, while still providing the means to test the calibration algorithm.

In some practical applications, a similar approach could be taken. This would be the case if the converter were to digitize relatively short burst of data, which are later to be processed by a micro-processor based system (e.g. in digital waveform analyzers). In other applications however, one or more of these functions may have to be placed on-chip. A practical way to perform the actual error correction, during continuous conversion, was discussed in chapter VIII. Figure 178 summarizes the approach. However, this architecture still requires that the correct digital be computed and placed in the appropriate memory locations.

This operation is more complex, and could still be kept off-chip. It could be accomplished using either a dedicated, or general purpose micro-processor. The complete self-calibrated ADC would then be realized as a chip-set, rather than one unit. If a monolithic approach is preferred, a dedicated arithmetic module would have to be integrated with the pipeline. Part of the arithmetic circuitry used during continuous conversion (figure 178) could be used to compute the coefficients. However, more research is needed in order to determine an optimal architecture, which would minimize additional circuitry and required data bus

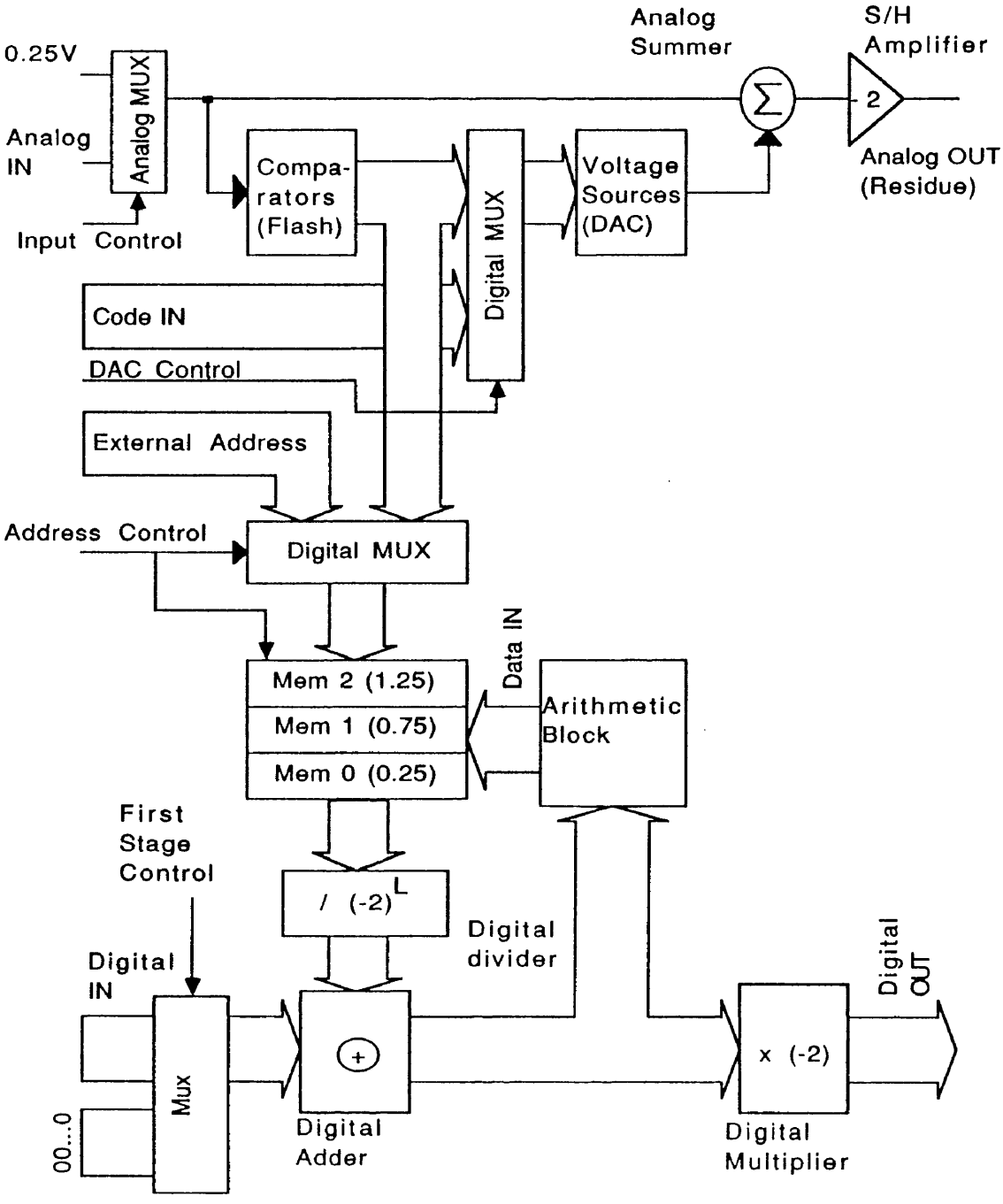


Fig. 178: On-Chip Calibration Hardware

area.

As a reference, figure 179 shows the lay-out of the present prototype converter. The total die size is 4 mm x 2 mm, which includes bond-pads, multi-phase clock generator, input S/H, calibration control logic, biasing circuitry and an independent 2-stage recycling converter with associated logic (not described here). The actual 16-stage pipeline uses about 35 % (3 mm^2) of the available area. The look-up tables and associated arithmetic blocks (mainly adders) were not included on this chip.

15.14. Recycling Converters

The fact that the data conversion and calibration sections of this architecture are so clearly separated, as well as the fact that all stages can be made nominally identical, are distinct advantages. In particular, the precise number of stages required to perform conversion to a specified number of bits, can be varied almost at will. Instead of a pipelined approach, a recycling approach can be taken, in which the analog signal is applied to one or two stages, and then recycled back to the input as many times as required to achieve the desired resolution.

Such architecture would be extremely useful to realize very compact, low power, (but slower) converters to be used as part of a single-chip micro-controller system. Since such system contains processor hardware anyway, the calibration algorithm can be executed without any area penalty. All that would be required, is a section of on-chip ROM memory containing the calibration program, in the form of firmware. The user of the micro-controller could then calibrate the converter and convert external analog signals, by calling the appropriate sub-routine.

15.15. Time-Interleaving

Time-interleaving of several pipelined converters has been discussed in chapter IV. It was shown how in a time-interleaved scheme, the conversion rate and the pipeline area are multiplied by the number of pipelines. This may make it seem like almost any conversion rate could be achieved, by just putting more pipelines in parallel. In practice however, some fundamental limitations exist. The main one is mismatch between pipelines.

In principle, it is possible to interleave a number of identical pipelines, which

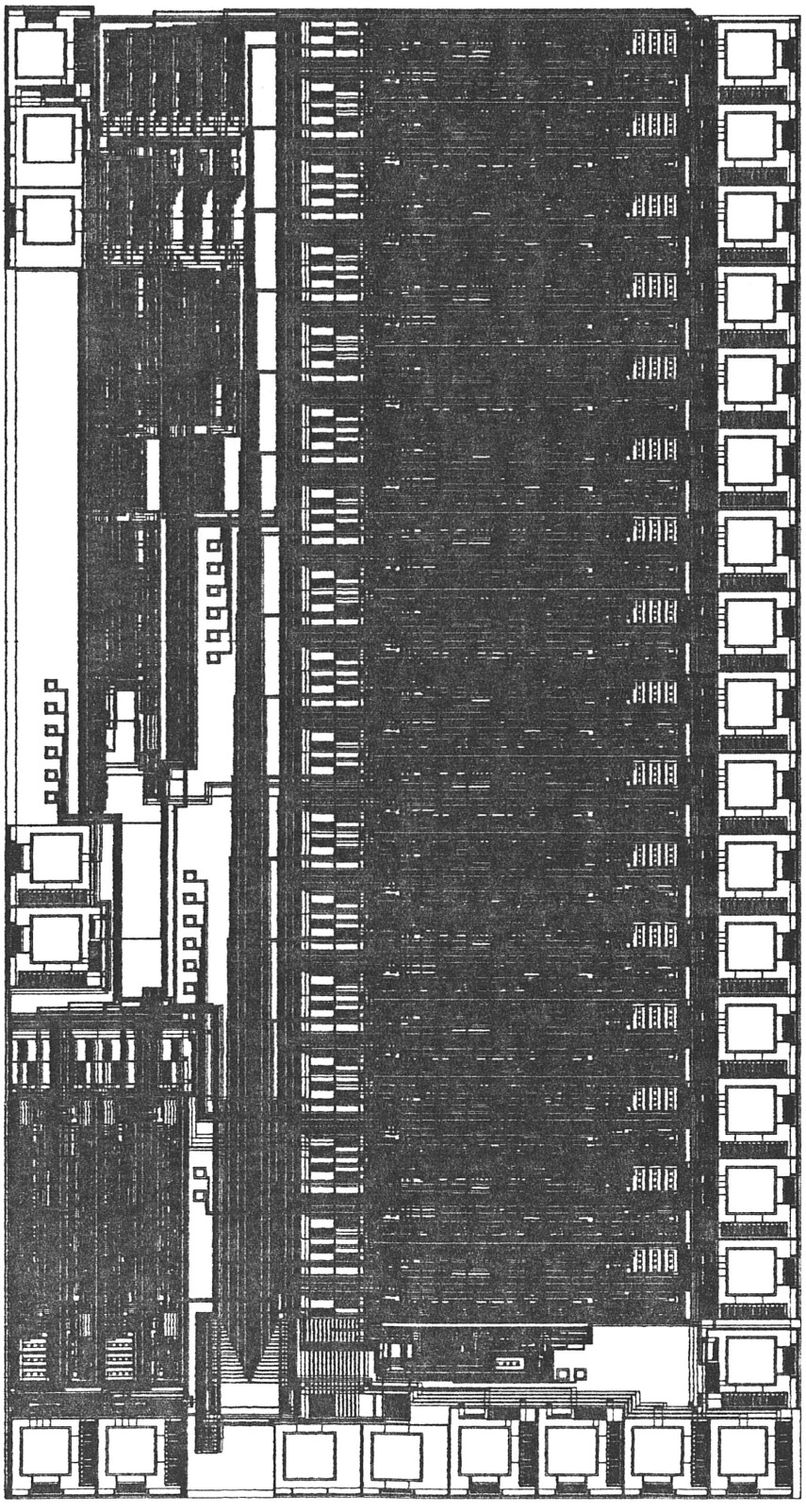


Fig. 179: Prototype Converter

are all linear up to a specified number of bits. However, linearity of a converter has been defined in a relative rather than in an absolute way (chapter I). After calibration through accuracy bootstrapping, the converter gain (code difference per voltage increment) of each pipeline, as well as the offset (output code for zero voltage) may be different. When the global output of the interleaved converter is cyclically connected to the different pipeline outputs, this may result in significant fixed-pattern (periodic) noise. This noise must be reduced to a level well below one *lsb* through careful matching. This is difficult to achieve for converters that use analog self-calibration techniques.

The situation is much more favorable for converters with full digital error correction, like the one described here. Since the conversion result is obtained as the sum of a number of terms (one per stage), gain and offset errors can be eliminated by proper scaling of each of these terms. The scaling term and factors can easily be obtained from a simple two-point measurement of a common reference, e.g. the measurement of a voltage corresponding to the lower end of the nominal input range, and a voltage corresponding to the full-scale input. The scaling itself can be achieved in the digital domain, using common arithmetic circuitry.

Another problem associated with high-resolution converters, is drift due to temperature effects, aging and low-frequency noise. Accuracy bootstrapping makes it possible to compensate for these effects through continuous calibration. For instance, one could conceive building a 12 bit, 40 MHz time-interleaved converter based on pipelines like described above. Since each pipeline could have a 5 MHz conversion rate, eight pipelines would have to be connected in parallel. This can still be achieved using a 40 MHz master clock, since the master clock frequency is divided by eight anyway, in order to generate the correct clock phases.

It is feasible to make a continuously calibrating system by adding one extra pipeline, bringing the total up to nine. That way, one pipeline can always go through a complete calibration cycle (which would take a few milliseconds, using on-chip calibration hardware), while the other eight are converting an input signal. An extra advantage of the (digital) PWM signals, is that they can be switched or multiplexed very easily. Each calibration cycle would have to be terminated by a two-point reference measurement in order to eliminate gain and offset mismatch.

Such calibration circuitry would be complex. It would probably necessitate the use of an elementary on-chip micro-processor, especially in order to implement the multiplication and division necessary to normalize the gain and offsets of each pipeline. In addition, each active pipeline would still require coefficient memory and data buses for continuous data conversion. These take up a significant amount of chip real-estate. Despite these draw-backs, it seems feasible to integrate a complete 40 MHz, 12 bit, eight-level interleaved system on less than 1 cm^2 of silicon, at a total power consumption of less than 400 mA.

15.16. Conclusion

This chapter described the different sections of a prototype 16 stage pipelined analog to digital converter using charge amplifier stages and accuracy bootstrapping for self-calibration. First, the generation of the six-phase, non-overlapping clock from one master clock input was described. A detailed analysis of each stage was given, including the operation of current sources, DAC section and data latches. A separate section was devoted to the input stage of the converter, which "translates" the external voltage input into a PWM signal. A complete description was given of the calibration sequencer used in the prototype. The reset and output multiplexing logic was described, as well as the requirements of the external data capture system, which is used to perform the off-chip calibration and error correction. Additional sections described possible further ramifications, such as the use of on-chip error correction, and the application of this architecture in recycling or time-interleaved converters.

CHAPTER XVI

CONCLUSION

16.1. Scope

The goal of dissertation was twofold. First, it was intended as a reference work about analog to digital converters in general. As such, it provides a fairly detailed overview of commonly used terminology, qualifiers and measures of performance. It also provides a broad review of the major architectures that are currently being used in practical circuits, as well as detailed theoretical and practical (design) considerations concerning high-speed pipelined (multi-step) architectures. These architectures were the main focus of the actual research that was performed for this work.

The second objective was the discussion of a number of newly developed techniques, both in testing and evaluation and in the design of high-performance analog to digital converters. These innovations were the result of practical research in the field of high-speed converters. They are usually not discussed separately, but included as an integral part of the text, as an extension of the discussion on existing concepts.

16.2. Summary

Chapters I and II are a discussion of general concepts in analog to digital converter characterization and testing, with emphasis on the DC transfer characteristic and static histogram testing. Chapter III presents an extension of the histogram concept towards possible use for dynamic evaluation. Chapter IV gives an overview of converter architectures, with particular focus on multi-step and pipelined architectures and their variations (recycling and time-interleaved approaches).

In chapter V, a detailed analysis is given of possible error mechanisms in multi-step structures. At the end, the concepts of over-ranging (redundancy) and digital error correction is first introduced. The same concepts are further worked

out and generalized in chapter VI, and presented as a method to virtually perfectly linearize the transfer characteristic of complex pipelined converters.

The drawback of full digital error correction is that the actual values of all components in the system must be known accurately. However, this can be accomplished using a new system identification and calibration technique, called accuracy bootstrapping. The technique is introduced from a general perspective in chapter VII, and applied to pipelined converters in chapter VIII. The validity of the technique for the calibration of pipelined converters is demonstrated by a combination of theoretical considerations and high-level simulations. The program that was used for most of these simulations is listed in the appendix.

The remainder of this dissertation focuses on a number of traditional and novel design techniques that are crucial for the realization of high-performance converters. Chapter IX contains speed and accuracy considerations of switched-capacitor gain stages. Chapter X introduces a new kind of sample/hold amplifier, called the charge amplifier. This configuration is believed to have several significant advantages over traditional switched-capacitor stages. However, one aspect of the charge amplifier turned out to be difficult to model : noise performance.

Its analysis required going back to the fundamental concepts of noise theory, which is done in chapter XI. In chapter XII, a formula is derived to describe the variance on the useful signal within a switched system like the charge amplifier. Chapter XIII analyzes the effect of correlated double sampling on noise, since this kind of sampling is implicitly built into the charge amplifier. Chapter XIV discusses simulated numerical examples of noise in a charge amplifier scheme, in an attempt to demonstrate the practicality of the approach. Finally, chapter XV discusses a prototype pipelined converter, which should be able to reach a linearity of 13 to 14 bits at 2 MHz or beyond.

16.3. Major Contributions

As mentioned, one contribution of this work was to clearly define a number of concepts that are commonly used for the characterization of analog to digital converters. Even the exact definition and inter-relation of common qualifiers like static transfer curve, differential and integral non-linearity are hard to find in previous literature.

In addition, several new concepts have been included into this work. In chapter II, it was shown how performing a static histogram test using a single low-frequency sweep rather than a random or pseudo-random input signal, can result in significant savings in measurement resources, as compared to traditional approaches.

In chapter III, the histogram concept is generalized in order to include the effect of the signal slope, rather than the magnitude alone. It was shown how such an approach could significantly enhance the measurement and understanding of dynamic errors within analog to digital converters, which has been recognized as a challenging area of research.

In chapter V, a complete, detailed analysis of static errors in multi-step converters is performed. The effect of each error upon the overall transfer curve of the converter is analyzed. In addition, it is shown how each individual static error can be recovered from the information obtained from a simple, external histogram test.

At the end of the chapter, the mechanism of redundancy and digital error correction in multi-step structures is introduced, from the perspective of histogram testing and probability density functions. This new analysis is thought to be useful, since it sheds a new light on concepts that are widely used (almost essential) in practical schemes, but very sparsely documented in existing literature.

Chapter VI presents a complete mathematical description of digital error correction in multi-step and pipelined converters. The formulas derived in this chapter precisely define the digital hardware required to perform error correction. The results are crucial to the development of efficient, digitally corrected structures that correct for more classes of errors simultaneously than previously reported schemes.

Chapters VII and VIII discuss accuracy bootstrapping. This novel identification and self-calibration technique seems to promise dramatic improvements in the achievable accuracy of complex (multi-stage) systems, designed to implement a very well-controlled overall transfer function. The technique has been observed to work amazingly well in digitally corrected, pipelined converters, of which the overall transfer curve must be made extremely linear. But calibration of other

systems may be possible as well.

The algorithm is iterative in nature. In particular systems, successive passes of the algorithm can yield increased accuracy. In that case, the fundamental limit to the achievable overall system accuracy is set only by the resolution of the system itself, noise and component variations over time.

Chapter IX contains a detailed discussion of the speed and accuracy of switched-capacitor gain stages, which have been essential building blocks of many existing analog to digital converters. The concepts that are presented are not new, but they are usually scattered as part of many publications. In addition, this chapter makes it easier to compare these traditional stages to a new structure, first presented in chapter X: the charge amplifier.

The charge amplifier uses an entirely new philosophy. A charge present on a capacitor is removed using a highly accurate current source, and simultaneously copied (and amplified) onto another capacitor. During the process, a digital, pulse-width coded representation of the analog charge is generated. It is believed that this amplification scheme can be made extremely fast and accurate, especially through the addition of a correlated double sampling scheme. Power consumption is very low, since current only flows directly onto or from capacitors.

It is expected that the performance of the charge amplifier could equal or surpass that of switched-capacitor stages, especially from the perspective of speed and power consumption. An additional advantage of the scheme is that no MOSFET switches are required, which eliminates some problems with clock feed-through, makes the scheme more robust and makes it possible to integrate it in purely bipolar technologies.

In addition, the scheme seems ideal to realize analog signal processing functions (filters) or direct analog to digital conversion on systems that have a charge input. This is the case in many opto-electronic devices based on photo-diodes (e.g. sensors for FAX machines) or charge-coupled devices (CCD's), which are often used in video cameras and auto-focus systems. Other potential applications could include high-power, low-accuracy control systems and servo-loops involving electric motors, heating elements or lights. The intermediate pulse width coded signal could be used advantageously for direct duty cycle power

control.

In chapters XI, XII, XIII and XIV, a fairly detailed analysis of noise is given, in view of analyzing the performance of the noise amplifier. The fact that the charge amplifier is a switched system, makes it hard to analyze using existing noise calculation techniques, especially when it comes to the effect of low-frequency (1/f) noise components.

From a theoretical point of view, some hypotheses are proposed. One of them is that when a MOSFET device is switched off, the noise is momentarily reset to zero, after which it takes low-frequency noise components a certain time to build up again. Whether this hypothesis proves wholly ore partially accurate, is better left in the middle for the moment being. However, assuming that it holds, a mathematical formula is derived to correlate statistical behavior in the time domain with frequency domain information (power spectrum).

The formula shows how and to what extent switching would reduce the influence of low-frequency noise components. Strangely enough, a very similar effect is obtained using correlated double sampling, which is also analyzed mathematically. The relationship between the two formulas is investigated. The hypothesis about switched noise remains to be verified by experimental results, but in the meantime, it results in the only mathematical tools available for an approximative analysis of noise behavior in systems like the charge amplifier.

Finally, in chapter XV, a prototype pipelined converter is discussed. This chapter shows how the charge amplifier can be used as the major component in a high-speed, high-accuracy, accuracy-bootstrapped converter. The proposed converter should achieve impressive performance (13 to 14 bits overall INL at 2 MHz). However, this is not the limit. The same basic converter core could be used in a massively parallel time-interleaved converter, which could achieve the same accuracy at sampling rates several times higher. Further speed improvement could also be gained using more power and/or even faster (bipolar?) IC technologies.

16.4. Suggested Further Research

It is obvious that this dissertation could not exhaust all potential research topics that were encountered. On the contrary, in many instances this work contains theoretical considerations and assumptions that should be verified by

practical measurements. The focus of the dissertation was the improvement of high-speed pipelined converters. In the process, new measurement (e.g. dynamic histogram), system identification (accuracy bootstrapping) and circuit design (charge amplifier) techniques were developed. In order to model the noise performance of the charge amplifier, hypotheses were put forth, which were later converted into useable mathematical expressions.

Each of these techniques and hypotheses could easily be a dissertation topic by itself. In the present case, they were not investigated any further than strictly needed to convince ourselves that they would allow the given application to work. Had more work been done on any of these topics (which would have necessitated much more time and resources), the initial focus could easily have been lost, and the publication of this text would undoubtedly have been significantly delayed.

However, it is our intention to continue the research, and to develop even more practical applications in silicon. It is also our hope that this dissertation, as well as related publications which should follow, will inspire others to experiment, and investigate questions we left unanswered.

For instance, more practical measurements should be made in order to demonstrate the usefulness of dynamic histogram testing as a diagnosis tool for high-speed converters. A tighter, mathematical proof should be developed to confirm that (given a limited amount of component errors), accuracy bootstrapping will *never* fail to calibrate the converter of chapter VIII. More simulations and experiments should be run in order to verify that accuracy bootstrapping could possibly also be used to compensate for interstage amplifier non-linearity.

A complex discrete-time filter should be built using charge amplifier instead of switched-capacitor stages. Precision measurements should be performed to confirm that high-frequency switching can reduce low-frequency noise components in semiconductor devices, and hence improve the noise performance of the charge amplifier. An optimal trade-off between noise, area, speed and power consumption should be determined for the configuration.

Finally, even faster, bigger and smarter pipelined converter prototypes should be integrated into silicon. One of the major conclusions of this work, is that the field of high-speed pipelined converters is amazingly vast, and still extremely fertile.

REFERENCES

[1] "Understanding High-Speed A/D Converter Specifications", *Application Notes*, Analog Devices Computer Labs, Wilmington, MA, 1974

[2] "Sinewave Curve Fitting in Dynamic Performance Testing of A to D Converters", *Product Note 5180A-2*, Hewlett Packard, Palo Alto, CA, 1985

[3] Stuart K. Tewksbury, Frederick C. Meyer, David C. Rollenhagen, H. K. Schoenwetter, Thomas M. Souders, "Terminology Related to the Performance of S/H, A/D, and D/A Circuits", *IEEE Transactions on Circuits and Systems*, Vol. CAS-25, No. 7, pp. 419-426, July 1978

[4] Joey Doernberg, Hae-Seung Lee, David A. Hodges, "Full-Speed Testing of A/D Converters", *IEEE Journal of Solid-State Circuits*, Vol. SC-19, No. 6, pp. 820-827, December 1984

[5] Solomon Max, "Fast Accurate and Complete ADC Testing", *Proceedings of 1989 International Test Conference*, pp. 111-117, 1989

[6] The Waveform Measurement and Analysis Committee of the IEEE Instrumentation and Measurement Society, "*Trial Use Standard for Digitizing Waveform Recorders*", IEEE Standard 1057, July 1989

[7] Randall L. Geiger, Phillip E. Allen, Noel R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, McGraw-Hill, New York, 1990

[8] Eric G. Soenen, Peter M. VamPeteghem, Humphrey H. Liu, James Cummings, Sriram Narayan, "Bridging the Gap Between Design and Testing of Analog Integrated Circuits", *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2716-2719, May 1990

[9] Fredric J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform", *Proceedings of the IEEE*, Vol. 66, No. 1, pp. 51-83, January 1978

[10] Christopher W. Mangelsdorf, "A 400-MHz Input Flash Converter with Error Correction", *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 1, pp. 184-191, February 1990

[11] Palaksha A. Setty, Peter M. Vanpeteghem, Randall L. Geiger, "An Algorithm to Compensate for Source Spectral Impurities in Dynamic Range Measurements", *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2236-2239, May 1990

[12] Rob E.J. Van De Grift, Ivo W.J.M. Rutten, Martien Van Der Veen, "An 8-Bit Video ADC Incorporating Folding and Interpolation Techniques", *IEEE Journal of Solid-State Circuits*, Vol 22, No. 12, pp. 944-953, December 1987

[13] Rudy Van De Plassche, Peter Baltus, "An 8b, 100-MHz A/D Converter with 500 MHz Resolution Bandwidth", *1988 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 222-223, February 1988

[14] Detlef Daniel, Ulrich Langmann, Berthold G. Bosch, "A Silicon Bipolar 4-bit 1-Gsample/s Full Nyquist A/D Converter", *IEEE Journal of Solid-State Circuits*, Vol. 23, No. 3, pp. 742-749, June 1988

[15] Benjamin J. McCarrol, Charles G. Sodini, Hae-Seung Lee, "A High-Speed CMOS Comparator for Use in an ADC", *IEEE Journal of Solid-State Circuits*, Vol. 23, No. 1, pp. 159-165, February 1988

[16] Toshio Kumamoto, Masao Nakaya, Shigeru Kusunoki, Tadashi Nishimura, Yasutaka Horiba, "An SOI Structure for Flash A/D Converter", *IEEE Journal of Solid-State Circuits*, Vol. 23, No. 1, pp. 198-200, February 1988

[17] Valdis E. Garuts, Yeou-Chong Simon Yu, Einar O. Traa, Tadanori Yamaguchi, "A Dual 4-bit 2-Gs/s Full Nyquist Analog-to-Digital Converter Using a 70-ps Silicon Bipolar Technology with Borosenic-Poly Process and Coupling-Base Implant", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 2, pp. 216-222, April 1989

[18] Francois Thomas, Francis Debrie, Maurice Gloanec, Michele Le Paih, Philippe Martin, Thao Nguyen, Stephanie Ruggeri, Jean-Marie Uro, "1-GHz GaAs ADC Building Blocks", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 2, pp. 223-228, April 1989

[19] Shiro Hosotani, Takahiro Maeda, Nobuharu Yazawa, "An 8-bit 20-MS/s CMOS A/D Converter with 50-mW Power Consumption", *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 1, pp. 167-172, February 1990

[20] J. Fernandes, S. Lewis, M. Mallinson, G. Miller, "A 14b, 10us Subranging A/D Converter with S/H", *1988 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 230-231, February 1988

[21] Tatsuji Matsuura, Toshiro Tsukada, Shimya Ahba, "An 8b, 20 MHz CMOS Half-Flash A/D Converter", *1988 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 220-221, February 1988

[22] Michael K. Mayes, Sing W. Chin, "A Multi-Step A/D Converter Family with Efficient Architecture", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 6, pp. 1492-1497, December 1989

[23] Joey Doernberg, Paul R. Gray, David A. Hodges, "A 10-bit 5-Msample/s CMOS Two-Step Flash ADC", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 2, pp. 241-249, April 1989

[24] Toshihiko Shimizu, Masao Hotta, Kenji Maio, Seiichi Ueda, "A 10-bit 20-MHz Two-Step Parallel A/D Converter with Internal S/H", *IEEE Journal of*

Solid-State Circuits, Vol. 24, No. 1, pp. 13-20, February 1989

[25] Donald A. Kerth, Navdeep S. Sooch, Eric J. Swanson, "A 12-bit 1-MHz Two-Step Flash ADC", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 2, pp. 250-255, April 1989

[26] Mdhav P. V. Kolluri, "A 12-bit 500-ns Subranging ADC", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 6, pp. 1498-1506, December 1989

[27] Stephen Lewis, Paul R. Gray, "Circuit Techniques for Monolithic CMOS Video A/D Conversion", *1987 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 185-186, February 1987

[28] Stephen Lewis, Paul R. Gray, "A Pipelined 5MHz 9b ADC", *1987 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 191-192, February 1987

[29] Stephen Lewis, Paul R. Gray, "A CMOS Pipelined 5 Msample/s 9b ADC with Digital Correction", *IEEE Journal of Solid-State Circuits*, Vol. 22, No. 12, pp. 954-961, December 1987

[30] Stephen Henry Lewis, "Video-Rate Analog-to-Digital Conversion Using Pipelined Architectures", Ph.D. Dissertation, Electronics Research Laboratory, University of California, Berkeley, CA 94720, Memorandum No. UCB/ERL M87/90, 18 November 1987

[31] B. Song, M. Tompsett, "A 12b, 1-MHz Capacitor Error Averaging Pipelined A/D Converter", *1988 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 226-227, February 1988

[32] B.-S. Song, M.F. Thompsett, K.R. Lakshimukar, "A 12-bit 1 Msample/s Capacitor Error-Averaging Pipelined A/D Converter", *IEEE Journal of Solid-State Circuits*, Vol. SC-23, No. 12, pp. 1324-1333, December 1988

[33] S. Sutarja, Paul R. Gray, "A 250Ks/s, 13b Pipelined A/D Converter", *1988 IEEE International Solid-State Circuit Conference*, Digest of Technical Papers, February 1988

[34] Joey Doernberg, "High-Speed Analog-to-Digital Conversion Using 2-Step Flash Architectures", Ph.D. Dissertation, Electronics Research Laboratory, University of California, Berkeley, CA 94720, Memorandum No. UCB/ERL M88/73, 21 November 1988

[35] Yasuhiro Sugimoto, Satoshi Mizoguchi, "An Experimental BiCMOS Video 10-Bit ADC", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 4, pp. 997-999, August 1989

[36] Kenzo Watanabe, Gabor C. Temes, Tomohisa Tagumi, "A New Algorithm for Cyclic and Pipeline Data Conversion", *IEEE Transactions on Circuits*

and Systems, Vol. 37, No. 2, pp. 249-252, February 1990

[37] David G. Nairn, C. Andre T. Salama, "A Ratio-Independent Algorithmic Analog-to-Digital Converter Combining Current Mode and Dynamic Techniques", *IEEE Transactions on Circuits and Systems*, Vol. 37, No. 3, pp. 319-325, March 1990

[38] B. Ginetti, P. Jespers, "A 1.5 Ms/s 8-Bit Pipelined RSD A/D Converter", *ESSCIRC Digest of Technical Papers*, pp. 137-140, September 1990

[39] David G. Nairn, C. Andre T. Salama, "Current-Mode Algorithmic Analog-to-Digital Converters", *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 4, pp. 997-1003, August 1990

[40] Yuh-Min Lin, Beomsup Kim, Paul R. Gray, "A 13-b 2.5-MHz Self-Calibrated Pipelined A/D Converter in 3-μm CMOS", *IEEE Journal of Solid-State Circuits*, Vol. SC-26, No. 4, pp. 628-635, April 1991

[41] Peter Real, David H. Robertson, Christopher W. Mangelsdorf, Theodore L. Tewksbury, "A Wide-band 10-b 20-Ms/s Pipelined ADC Using Current-Mode Signals", *IEEE Journal of Solid-State Circuits*, Vol. SC-26, No. 8, pp. 1103-1109, August 1991

[42] G.C. Temes, "High-Accuracy Pipeline A/D Converter Configuration", *IEE Electronics Letters*, Vol. 21, pp. 761-762, August 1985

[43] Kh. Hadidi, G.C. Temes, K.W. Martin, "Error Analysis and Digital Correction Algorithms for Pipelined A/D Converters", *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 1709-1712, May 1990

[44] James L. McCreary, Paul R. Gray, "All-MOS Charge Redistribution Analog-to-digital Conversion Techniques - part I", *IEEE Journal of Solid-State Circuits*, Vol. SC-10, No. 6, pp. 371-379, December 1975

[45] Hae-Seung Lee, David A. Hodges, Paul R. Gray, "A Self-Calibrating 15 Bit CMOS A/D Converter", *IEEE Journal of Solid-State Circuits*, Vol. SC 19, No. 6, pp. 813-819, December 1984

[46] Hae-Seung Lee, David A. Hodges, "Self-Calibrating ADC's; Accuracy Considerations", *IEEE Transactions on Circuits and Systems*, Vol. CAS-32, pp. 590-597, June 1985

[47] Kh. Hadidi, Vincent S. Tso, Gabor C. Temes, "An 8-b 1.3 MHz, Successive-Approximation A/D Converter", *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 3, pp. 880-885, June 1990

[48] Hidetoshi Onodera, Tetsuo Tateishi, Keikichi Tamaru, "A Cyclic A/D Converter That Does Not Require Ratio-Matched Components" *IEEE Journal of Solid-State Circuits*, Vol. 23, No. 1, pp. 152-158, February 1988

[49] J. C. Candy, O. J. Benjamin, "The Structure of Quantization Noise from Sigma-Delta Modulation", *IEEE Transactions on Communications*, Vol. COM-29, pp. 1316-1323, September 1981

[50] J. C. Candy, "A Use of Double Integration in Sigma-Delta Modulation", *IEEE Transactions on Communications*, Vol. COM-33, pp. 249-258, March 1985

[51] J. C. Candy, "Decimation for Sigma-Delta Modulation", *IEEE Transactions on Communications*, Vol. COM-34, pp. 72-76, January 1986

[52] R. M. Gray, "Oversampled Sigma-Delta Modulation", *IEEE Transactions on Communications*, Vol. COM-35, pp. 481-489, May 1987

[53] L. E. Larson, T. Cataltepe, Gabor C. Temes, "Multibit Oversampled Sigma-Delta A/D Converter with Digital Error Correction", *Electronic Letters*, Vol. 24, pp 1051-1052, August 1988

[54] Jacques Robert, Philippe Deval, "A Second-Order High-Resolution Incremental A/D Converter with Offset and Charge Injection Compensation", *IEEE Journal of Solid-State Circuits*, Vol. 23, No. 3, pp. 736-739, June 1988

[55] Steven R. Norsworthy, Irving G. Post, H. Scott Fetterman, "A 14-bit 80-kHz Sigma-Delta A/D Converter: Modeling, Design, and Performance Evaluation", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 2, pp. 256-266, April 1989

[56] Mike Rebeschini, Nicholas R. Van Bavel, Patrick Rakers, Robert Greene, James Caldwell, John R. Haug, "A 16-b 160-kHz CMOS A/D Converter Using Sigma-Delta Modulation", *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 2, pp. 431-436, April 1990

[57] Vittorio Comino, Michel S.J. Steyaert, Gabor C. Temes, "A First-order Current-steering Sigma-delta Modulator", *IEEE Journal of Solid-State Circuits*, Vol. SC-26, No. 3, pp. 176-183, March 1991

[58] Hans A. leopold, Gunter Winkler, Paul O'Leary, Karl Ijzer, Juergen Jernej, "A Monolithic CMOS 20-b Analog-to-digital Converter", *IEEE Journal of Solid-State Circuits*, Vol. SC-26, No. 7, pp. 910-916, July 1991

[59] Brian P. Brandt, Bruce A. Wooley, "A CMOS Oversampling A/D Converter with 12b Resolution and Conversion Rates Above 1MHz", *1989 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 64-65, February 1991

[60] Yasuyuki Matsuya, Kuniharu Uchimura, Atsushi Iwata, Tsutomu Kobayashi, Masayuki Ishikawa, Takeshi Yoshitome, "A 16-bit Oversampling A-to-D Converter Using Triple-Integration Noise Shaping", *IEEE Journal of Solid-State Circuits*, Vol. SC-22, No. 12, pp. 921-929, December 1987

[61] Frank Op't Eynde, Guang Ming Yin, Willy Sansen, "A CMOS Fourth-Order 14b 500k-Sample/s Sigma-Delta ADC Converter", *1989 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 62-63, February 1991

[62] David B. Ribner, Richard D. Baertsch, Steven L. Gaverick, Donald T. McGrath, Joseph E. Krisciunas, Toshiaki Fujii, "A Third-Order Multi-Stage Sigma-Delta Modulator with Reduced Sensitivity to Nonidealities", *IEEE Journal of Solid-State Circuits*, Vol. SC-26, No. 12, pp. 1764-1774, December 1991

[63] Yiannos Manoli, "A Self-Calibration Method for Fast High-Resolution A/D and D/A Converters", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 3, pp. 603-608, June 1989

[64] Wouter Groeneveld, Hans Schouwenaars, Henk Termeer, "A Self Calibration Technique for Monolithic High-Resolution D/A Converters", *1989 IEEE International Solid-State Circuit Conference, Digest of Technical Papers*, pp. 22-24, February 1989

[65] K. Hwang, *Computer Arithmetic - Principles, Architecture and Design*, John Wiley & Sons, New York, 1979

[66] J.T. Caves, M.A. Copeland, C.F. Rahim, S.D. Rosenbaum, "Sampled Analog Filtering Using Switched Capacitors as Resistor equivalents", *IEEE Journal of Solid-State Circuits*, Vol. SC-12, pp. 592- 599, December 1977

[67] K. Martin, "Improved Circuits for the Realization of Switched-capacitor Filters", *IEEE Transactions on Circuits and Systems*, Vol. CAS-27, pp. 237-244, April 1980; Also published as BNR Technical report TRIE 81-78-06, March 1978

[68] K. Martin, A.S. Sedra, "Strays-insensitive Switched-capacitor Filters Based on the Bilinear Z-Transform", *Electronic Letters*, Vol. 19, pp. 365-366, June 1979

[69] David J. Allstot, "MOS Switched Capacitor Ladder Filters", Ph.D. Dissertation, University of California, Berkeley, 1979

[70] P.R. Gray, D.A. Hodges, R.W. Brodersen, *Analog MOS Integrated Circuits*, IEEE Press Book, IEEE Press, New York, 1980

[71] Gabor C. Temes, "Finite Amplifier Gain and Bandwidth Effects in Switched-Capacitor Filters", *IEEE Journal of Solid-State Circuits*, Vol. SC-15, No. 3, pp. 358-361, June 1980

[72] A.S. Sedra, K.C. Smith, *Micro-Electronic Circuits*, CBS College Publishing, New York, 1982

[73] K. Martin, "New Clock Feedthrough Cancellation Technique for Analogue

MOS Switched-Capacitor Circuits", *Electronic Letters*, Vol. 18, No. 1, pp. 39-40, January 1982

[74] David J. Allstot, William C. Black, Jr., "Technological Design Considerations for Monolithic MOS Switched-capacitor Filtering Systems", *Proceedings of the IEEE*, Vol. 71, No. 8, pp. 967-985, August 1983

[75] Roubik Gregorian, Kenneth W. Martin, Gabor C. Temes, "Switched-Capacitor Circuit Design", *Proceedings of the IEEE*, Vol. 71, No. 8, pp. 941-966, August 1983

[76] Rinaldo Castello, Paul R. Gray, "A High-Performance Micropower Switched-Capacitor Filter", *IEEE Journal of Solid-State Circuits*, Vol. SC-20, No. 6, pp. 1122-1132, December 1985

[77] Roubik Gregorian, Gabor C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, John Wiley & Sons Inc., New York, 1986

[78] Peter M. Van Peteghem, "Accuracy and Resolution of Switched Capacitor Circuits in M.O.S. Technology", Doctoral Dissertation, Katholieke Universiteit Leuven, Belgium, June 1986

[79] J.A. Fisher, "Noise Sources and Calculation Techniques for Switched Capacitor Tpelgromechniques", *IEEE Journal of Solid-State Circuits*, Vol. SC-17, No. 4, pp. 742-752, April 1982

[80] C.A. Gobet, A. Knob, "Noise Analysis of Switched-Capacitor Networks", *IEEE Transactions on Circuits and Systems*, Vol. CAS-30, No. 1, pp. 37-43, January 1983

[81] K.-C. Hsieh, "Noise Limitations in Switched-capacitor Filters", Ph.D. Dissertation, Electronics Research Laboratory, University of California, Berkeley, CA 94720, May 1982

[82] H.W. Klein, W.L Engl, "Minimization of Charge Transfer Errors in Switched Capacitor Stages", *IEEE Journal of Solid-State Circuits*, Vol. SC-18, No. 6, pp. 764-766, June 1983

[83] D. McQuigg, "Residual Charge on a Switched Capacitor", *IEEE Journal of Solid-State Circuits*, Vol. SC-18, No. 6, pp. 811-813, June 1983

[84] W.B. Wilson, H.Z. Massoud, E.J. Swanson, R.T. George, Jr., R.B. Fair, "Measurement and Modeling of Charge Feed-Through in N-Channel MOS Analog switches", *IEEE Journal of Solid-State Circuits*, Vol. SC-20, No. 6, pp. 1206-1213, December 1985

[85] Bing J. Sheu, Je-Hurn Shieh, Mahesh Patil, "Modeling Charge Injection in MOS Analog Switches", *IEEE Transactions on Circuits and Systems*, Vol. CAS-34, No. 2, pp. 214-216, February 1987

[86] J. Shieh, M. Patil, B.L. Sheu "Measurement and Analysis of Charge Injection in MOS Analog Switches", *IEEE Journal of Solid-State Circuits*, Vol. SC-22, No. 4, pp. 277-281, April 1987

[87] E.G. Soenen, R.L. Geiger, "The Charge Amplifier, an Alternative for Switched-Capacitor Stages", *1991 Government Micro-Circuit Applications Conference, Orlando, FL, Digest of technical papers*, pp. 281-284, November 1991

[88] Marcel J. M. Pelgrom, Twan Jochijms, "Matching Properties of MOS Transistors", *1988 European Solid-State Circuit Conference, Digest of Technical Papers*, pp. 327-330, September 1988

[89] David Vallancourt, Yannis P. Tsividis, "A Fully Programmable Sampled-Data Analog CMOS Filter with Transfer-Function Coefficients Determined by Timing", *IEEE Journal of Solid-State Circuits*, Vol. SC-22, No. 6, pp. 1022-1030, December 1987

[90] Rudy J. Van De Plasche, "Dynamic Element Matching for High-accuracy Monolithic D/A Converters", *IEEE Journal of Solid-State Circuits*, Vol. SC-11, pp. 795-800, December 1976

[91] D. Wouter Groeneveld, Hans J. Schouwenaars, Henk A. H. Termeer, Cornelis A. A. Bastiaansen, "A Self-Calibration Technique for Monolithic High-Resolution D/A Converters", *IEEE Journal of Solid-State Circuits*, Vol. 24, No. 6, pp. 1517-1522, December 1989

[92] M. Nayebi, B.A. Wooley, "A 10-Bit Video BiCMOS Track-and-Hold Amplifier", *IEEE Journal of Solid-State Circuits*, Vol. SC-24, No. 12, pp. 1507-1516, December 1989

[93] Peter J. Lim, Bruce A. Wooley, "A High-Speed Sample-and-Hold Technique Using a Miller Hold Capacitance", *IEEE Journal of Solid-State Circuits*, Vol. SC-26, No. 4, pp. 643-651, April 1991

[94] U. Gatti, F. Maloberti, G. Palmisano, "An Accurate CMOS Sample-and-Hold Circuit", *IEEE Journal of Solid-State Circuits*, Vol. SC-27, No. 1, pp. 120-122, January 1992

[95] Athanasios Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1984

[96] Kwok K. Hung, Ping K. Ko, Chenming Hu, Yiu C. Cheng, "A Unified Model for the Flicker Noise in Metal-Oxide- Semiconductor Field-Effect Transistors", *IEEE Transactions on Electron Devices*, Vol 37, No. 3, pp. 654-665, March 1990

[97] D.M. Fleetwood, "Evidence That Similar Point Defects Cause 1/f Noise and Radiation-Induced-Hole Trapping in Metal-Oxide-Semiconductor Tran-

sistors", *Physical review Letters*, Volume 64, No. 5, pp. 579-582, January 1990

[98] R.J. Kansy, "Response of a Correlated Double Sampling Circuit to 1/f Noise", *IEEE Journal of Solid-State Circuits*, Vol. SC-15, No. 3, pp. 373-375, March 1980

[99] F. Maloberti, F. Montecchi, "Low-Frequency Noise Reduction in S.C. Ladder Filters", *IEE Electronics Letters*, Vol. 18(1), pp. 39-40, January 1982

[100] H.M. Wey, W. Guggenbuhl, "Noise Transfer Characteristics of a Correlated Double Sampling Circuit", *IEEE Transactions on Circuits and Systems*, Vol. CAS-33, No. 10, pp. 1028-1030, October 1986

APPENDIX: ACCURACY BOOTSTRAPPING PROGRAM

```
/*****************************************************************/
/* ACCURACY BOOTSTRAPPING TEST PROGRAM (C)                       */
/* This listing contains a few of the key routines used          */
/* to simulate and evaluate the performance of the algorithm     */
/*****************************************************************/

#include <stdio.h>
#include <stdlib.h>

double A= -2.; /* Nominal value of interstage gain (integer) */
int L=16;      /* Number of stages in the pipeline */

int absA;      /* Absolute value of the interstage gain */
double totA;

/* Stage parameters: flash and DAC levels, gains, weights */
double ADC[16][16],DAC[17][16],As[16],W[17][16];


void nominal() /* Calculate nominal parameters for configuration */

  {
  int l,i;

  for (l=0;l<L;l++)
    {
    for (i=0;i<absA;i++)
      ADC[i][l]=1./(2.*absA)+i*1./absA;
    for (i=0;i<=absA;i++)
      DAC[i][l]= -1./(2.*A)+i*1./absA;

    As[l]=A;
    }
  }


void prt_stg() /* Print stage parameters */

  {
  int l,i;

  printf("ADC:\n");
  for (l=0;l<L;l++)
    {
    printf("%2d:",l);
    for (i=0;i<absA;i++)
     printf(" %5.3f",ADC[i][l]);
    printf("\n");
    }

  printf("\nDAC:\n");
  for (l=0;l<L;l++)
    {
    printf("%2d:",l);
    for (i=0;i<=absA;i++)
     printf(" %5.3f",DAC[i][l]);
  printf("\n");
    }
```

```
  printf("\n\n");
  }


void init_W() /* Initialize digital weights (look-up tables) */
              /* (initially, they are all nominal */
  {
  int l,i;

  for (l=0;l<L;l++)
    for (i=0;i<=absA;i++)
      W[i][l]= -1./(2.*A)+i*1./absA;
  }


void dostage(l,vin,cd,vres) /* simulate behavior of one stage */
                            /* l: stage number, cd: output code, */
                            /* vin: input signal, vres: residue */

  int l,*cd;
  double vin,*vres;

  {
  int i;

  i=0;
  while ((vin*As[l]>=ADC[i][l])&&(i<absA))
    i++;

  *vres=vin*As[l]-DAC[i][l];

  *cd=i;
  }


double convert(vin,l) /* Convert a signal vin */
                      /* with l as first stage in pipeline */

  double vin;
  int l;

  {
  double x,y,sx,sy;
  int k,st,c;

  sx=0.0;
  st=l;
  x=vin;

  for (k=0;k<16;k++)
    {
    if (st<0)          /* This simulates the cyclic arrangement */
      st=L-1;

    dostage(st,x,&c,&y);
    x=y;

    sy=A*sx+W[c][l]/totA; /* This simulates the operation of the */
                          /* look-up tables, multiplication and */
                          /* addition */
    sx=sy;
```

```
    st--;
    }

  return(sy);
  }


void calibrate() /* Calibrate whole pipeline stage by stage */
                 /* Using accuracy bootstrapping procedure */
  {
  int l,c;
  double inputref,N[17];

  inputref=1./(2.*absA);
  for (l=0;l<L;l++)
    {
    N[0]=convert(inputref,l-1);
    for (c=0;c<absA;c++)
      N[c+1]=convert(inputref-(DAC[c+1][l]-DAC[c][l]),l-1);

    for (c=1;c<=absA;c++)
      W[c][l]=W[c-1][l]+(N[c]-N[0]);
    }
  }


double absol(x) /* Absolute value */
  double x;

  {
  double a;

  if (x>=0)
    a=x;
  else
    a= -x;

  return(a);
  }


double maxdif(l) /* Maximum difference from nominal weights, */
                 /* for linearity criterion */
  int l;

  {
  int i,j;
  double diff,max;

  max=0.;

  for (i=0;i<absA;i++)
    for (j=i+1;j<=absA;j++)
      {
      diff=absol((W[j][l]-W[i][l])-(DAC[j][l]-DAC[i][l]))
      if (diff>max)
        max=diff;
      }
```

```
   return(max);
   }


double maxconv(l) /* Calculate maximum possible conversion */
                  /* for pipeline with stage l as first stage */
   int l;

   {
   int st,i;
   double sum,denom;

   denom=1.;
   st=l;
   for (i=0;i<L;i++)
     {
     if (st<0)
       st=L-1;

     denom*=A;
     sum+=W[absA][st]/denom;
     st++;
     }

   return(sum);
   }


double minconv(l) /* Calculate minimum possible conversion */
                  /* for pipeline with stage l as first stage */

   int l;

   {
   int st,i;
   double sum,denom;

   denom=1.;
   st=l;
   for (i=0;i<L;i++)
     {
     if (st<0)
       st=L-1;

     denom*=A;
     sum+=W[0][st]/denom;
     st++;
     }

   return(sum);
   }


void init() /* Calculate overall gain of pipeline (A^L) */

   {
   int i;

   if (A>=0)
     absA=(int)A;
```

```
  else
    absA= -(int)A;

  totA=1.0;
  for (i=1;i<=L;i++)
    totA*=A;
  }


void testnom() /* Test a few pipelines with different positive */
               /* and negative gain values */
               /* Each time, apply input and look at output code */

  {
  double v;

  for (A= -2.;A>= -5.;A--)
    {
    printf("\nA=%f\n\n",A);

    init();
    nominal();
    prt_stg();
    init_W();
    calibrate();

    for (v= -1./absA;v<=0;v+=1/(30.*absA))
      printf("v=%f, s=%f\n",v,convert(v,L-1));
    }

  for (A=2.;A<=5.;A++)
    {
    printf("\nA=%f\n\n",A);

    init();
    nominal();
    prt_stg();
    init_W();

    for (v=0;v<=1./absA;v+=1/(30.*absA))
      printf("v=%f, s=%f\n",v,convert(v,5));
    }
  }


main()

  {
  testnom();
  }
```