Marcel J.M. Pelgrom

Analog-to-Digital Conversion

Springer

Analog-to-Digital Conversion

Marcel J.M. Pelgrom

Analog-to-Digital Conversion

Marcel J.M. Pelgrom
NXP Semiconductors
HTC-32
Eindhoven 5656AE
The Netherlands
marcel.pelgrom@nxp.com

Preface

A book is like a window that allows you to look into the world. The window is shaped by the author and that makes that every window presents a unique view of the world. This is certainly true for this book. It is shaped by the topics and the projects throughout my career. Even more so, this book reflects my own style of working and thinking.

That starts already in Chap. 2. When I joined Philips Research in 1979, many of my colleagues used little paper notebooks to keep track of the most used equations and other practical things. This notebook was the beginning for Chap. 2: a collection of topics that form the basis for much of the other chapters. Chapter 2 is not intended to explain these topics, but to refresh your knowledge and help you when you need some basics to solve more complex issues.

In the chapters discussing the fundamental processes of conversion, you will recognize my preoccupation with mathematics. I really enjoy finding an equation that properly describes the underlying mechanism. Nevertheless mathematics is not a goal on its own: the equations help to understand the way the variables are connected to the result. Real insight comes from understanding the physics and electronics. In the chapters on circuit design I have tried to reduce the circuit diagrams to the simplest form, but not simpler... I do have private opinions on what works and what should not be applied. Most poor solutions have simply been left out, sometimes you might read a warning in the text on a certain aspect of an interesting circuit.

Another of my favorites is the search for accuracy. In Chap. 11 you will find a detailed description, but also in the earlier chapters, there is a lot of material referring to accuracy.

Circuit design and analog-to-digital circuit design is about bridging the gap between technology and systems. Both aspects have been treated less than they deserve. Still I hope it will be sufficient to create an interest to probe further.

This book is based on my lectures for graduate students who are novice in analog-to-digital design. In the classes my aim is to bring the students to a level where they can read and interpret the literature (such as IEEE Journal of Solid-State Circuits) and judge the reported results on their merits. Still that leaves a knowledge gap with the designer of analog-to-digital converters. For those designers this book may serve as a reference of principles and background.

Inevitably this book has some strong points but also weak points. There are still so many wonderful ideas, that are not addressed here but certainly would deserve some space, but simply did not fit in this volume. Still I hope this book will let you experience the same thrill that all analog-to-digital designers feel, when they talk about their passion. Because that is the goal of this book: to encourage you to proceed on the route towards even better analog-to-digital converters.

Acknowledgments

University for their collaboration, prof. Gabor Temes and dr. Vlado Valence for inviting me to lecture in the MEAD and EPFL courses.

A special word of thanks goes to all the students for their questions, remarks and stimulating discussions.

Contents

List of Symbols

A	General variable	
A	Area	[cm^2]
C	Capacitance	[F]
C_{ox}	Oxide capacitance	[F/cm^2]
D_n	Diffusion coefficient of electrons	[cm^2/s]
d_{ox}	Oxide thickness	[cm $= 10^8$ Å]
E	Electric field	[V/cm]
E_{Fn}	Fermi energy level of electrons	[eV]
E_{Fp}	Fermi energy level of holes	[eV]
E_G	Band gap energy	[1.205 eV]
E_i	Energy level of an intrinsic semiconductor	[eV]
f	Frequency	[Hz]
f_c	Clock frequency	[Hz]
f_i	Frequency of input signal	[Hz]
f_s	Sample rate	[Hz]
H	Transfer function	[1]
I	Large signal or DC current	[A]
i	Small signal current	[A]
J	Current density	[A/cm^2]
J_n	Electron current density	[A/cm^2]
J_p	Hole current density	[A/cm^2]
K	Substrate voltage influence on the threshold voltage	[$\sqrt{V}$]
k	Boltzmann's constant	[1.38×10^{-23} J/K]
L	Length of transistor gate	[cm]
L_w	Inductance of a wire	[H]
M	Multiplex factor	[1]
N	Resolution	[1]
N_a	Substrate doping concentration	[cm^{-3}]
n	Volume density of electrons	[cm^{-3}]
n_i	Intrinsic charge volume density	[1.4×10^{10} cm^{-3} (300 K)]
p	Volume density of holes	[cm^{-3}]

p_{p0}	Volume density of holes in a p-substrate in equilibrium	[cm^{-3}]
R	Resistance	[Ω]
Q	Charge	[C]
q	Electron charge	[1.6×10^{-19} C]
T	Time period	[s]
T	Absolute temperature	[K or $(T - 273.15)$ °Celsius]
T_0	Reference temperature	[K]
T_s	Sample period	[s]
t	Time as a running variable	[s]
V	Bias or DC potential	[V]
v	Small signal voltage	[V]
V_{DD}	Positive power supply	
V_{DS}	Drain potential with respect to the source potential	[V]
V_{FB}	Flat-band voltage	[V]
V_G	Gate potential with respect to ground	[V]
V_{GS}	Gate potential with respect to the source potential	[V]
V_T	MOS transistor threshold voltage	[V]
W	Width of transistor channel	[cm]
X, Y	General input and output variable	[1]
x	Dimension perpendicular to the wafer surface	[cm]
y	Dimension parallel to the transistor current	[cm]
Z	Complex impedance	[Ω]
β	Current gain factor of MOS transistor: $W\beta_\square/L$	[A/V^2]
$\beta_\square$	Current gain factor of a square MOS transistor	[A/V^2]
ϵ	Permittivity in vacuum	[8.854×10^{-14} F/cm]
$\epsilon_{ox}\epsilon$	Permittivity in silicon dioxide	[3.45×10^{-13} F/cm]
$\epsilon_s\epsilon$	Permittivity in silicon	[10.5×10^{-13} F/cm]
ϕ_F	Potential difference between intrinsic and hole Fermi level	[V]
μ_0	Magnetic permeability in vacuum	[$4\pi \times 10^{-7}$ H/m or N/A^{-2}]
μ_n, μ_p	Mobility of electrons and holes	[cm^2/Vs]
π	Angular constant	[3.14159]
ψ	Electrostatic potential	[V]
ψ_B	Electrostatic potential at which strong inversion starts	[V]
ψ_s	Electrostatic potential at the interface	[V]
$\sigma_{\Delta P}$	Standard deviation of ΔP	
σ_n	Electron capture cross-section	[5×10^{-15} cm^2]
τ	Time-constant	[s]
$\omega = 2\pi f$	Angular or radian frequency	[rad/s]

Reference Tables and Figures

Chapter 1
Introduction

Analog-to-digital conversion is everywhere around us. In all forms of electronic equipment the analog-to-digital converter links our physical world to digital computing machines. This development has enabled all the marvelous functionality that has been introduced over the last thirty years, from mobile phone to internet, from medical imaging machines to hand-held television.

Pure analog electronics circuits can do a lot of signal processing in a cheap and well-established way. Many signal processing functions are so simple that analog processing serves the needs (audio amplification, filtering, radio). In more complex situations, analog processing however lacks required functionality. There digital signal processing offers crucial extensions of this functionality. The most important advantages of digital processing over analog processing are a perfect storage of digitized signals, unlimited signal-to-noise ratio, the option to carry out complex calculations, and the possibility to adapt the algorithm of the calculation to changing circumstances. If an application wants to use these advantages, analog signals have to be converted with high quality into a digital format in an early stage of the processing chain. And at the end of the digital processing the conversion has to be carried out in the reverse direction. The digital-to-analog translates the outcome of the signal processing into signals that can be rendered as a picture or sound. This makes analog-to-digital conversion a crucial element in the chain between our world of physical quantities and the rapidly increasing power of digital signal processing. Figure 1.1 shows the analog-to-digital converter (abbreviated A/D-converter or ADC) as the crucial element in a system with combined analog and digital functionality.

The analog-to-digital converters and digital-to-analog converters discussed in this book convert high resolution and high speed signals to and from the digital domain. The basics of the conversion process is shown in Fig. 1.2. In the analog domain a ratio exists between the actual signal and a reference quantity. This ratio is reflected in the digital domain, where the digital code is a fraction of the available word width. The analog-to-digital converter tries to find an optimum match between these ratios at any moment in time. However, an essential rounding error must be accepted.
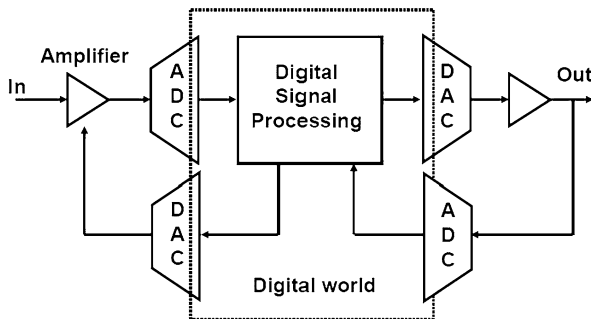
Fig. 1.1 The analog-to-digital and digital-to-analog converters are the ears and eyes of a digital system

Fig. 1.2 In analog-to-digital conversion a connection is made between the analog world of physical quantities and the digital world of numbers and bits
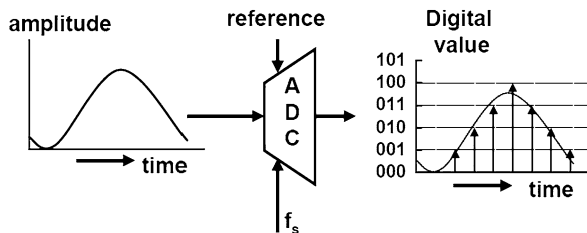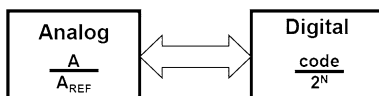
Fig. 1.3 Functions of the analog-to-digital converter: sampling, quantizing and linking to a reference

Signals in the digital domain differ from analog signals, which exist in the physical world, because digital signal are sampled and quantized, Fig. 1.3. Sampled signals only have meaning at their sample moments as given by the sample frequency. Moreover digital signals are arithmetic quantities, which are only meaningful in the physical world while there is somewhere an assignment that relates the digital number range to a physical reference value. These three main functions characterize the analog-to-digital converter, see Table 1.1. These functions will be visible in each stage of the discussion of analog-to-digital conversion and are reflected in the set-up of this book.

Table 1.1 Key functions in analog-to-digital conversion

Analog-to-digital	Digital-to-analog
Time discretization	Holding the signal
Amplitude discretization	Amplitude restoration
Reference to	Reference from
A conversion unit	A conversion unit

1.1 About this Book

An analog-to-digital converter and a digital-to-analog converter are electronics circuits that are designed and fabricated in silicon IC technology. The main focus in this book is on CMOS realizations. Chapter 2 summarizes the main physics and mathematics for understanding the operation of analog-to-digital converters. This chapter is meant to refresh existing knowledge.

In Chaps. 3 to 6 the three basic functions for conversion are analyzed. Chapter 3 describes the sampling process and give guidelines for the different choices around sampling in analog-to-digital conversion design. The design challenges around the sampling process are discussed in the design of sample-and-hold circuits in Chap. 4. Both sampling and quantization operation are non-linear, resulting in undesired behavior. The combination of sampling and quantization results in a fundamental error: the quantization error. The attainable performance of every analog-to-digital conversion is fundamentally limited by this error as is described in Chap. 5. Chapter 6 deals with the generation and handling of reference voltages.

The main task of a designer is to construct circuits. Chapter 7 describes the basics of digital-to-analog converter circuit design and some implementations of digital-to-analog converters. The design of analog-to-digital converters is detailed in Chap. 8. Oversampling and sigma delta conversion are a special class of conversion techniques and are discussed in Chap. 9.

Next to theory and circuit design, the proper operation of a converter relies on some additional aspects. The measurement methods for analog-to-digital converters and specification points are the subject of Chap. 10. Chapter 11 deals with some of the boundary conditions in conversion due to technological and physical limitations. Finally Chap. 12 deals with system aspects of the application of analog-to-digital conversion like sample frequency choices and the various forms of input handling. This section also introduces a Figure-of-Merit for conversion and compares the various implementation forms. In this way an optimal converter for a given system situation can be chosen.

Several books have been published in the field of analog-to-digital conversion. One of the first books was published by Seitzer [1] in 1983 describing the basic principles. Van der Plassche [2] in 1994 and 2003 and Razavi [3] in 1994 discuss extensively bipolar and CMOS realizations. Jespers [4] and Maloberti [5] address the theme on a graduate level. These text books review the essential aspects and focus on general principles, circuit realizations, and their merits.

Chapter 2
Components and Definitions

An electronic engineer combines in his/her work elements of mathematics, physics, network theory and other scientific disciplines. The creative combination of the elements allows the engineer to bridge fundamental theoretical insights with practical realizations. Often these theoretical disciplines are phrased in mathematical descriptions. Therefore it is relevant to start this book with a summary of these disciplines.

2.1 Mathematical Tools

Events and processes in semiconductor devices span a large range of numbers. Abbreviations for these numbers are shown in Table 2.1. The words "billion" and "trillion" must be avoided as they refer to different quantities in Europe and the USA.[1]

Mathematical expressions are built from functions of variables: $y = f(x)$. Table 2.2 lists some elementary mathematical functions and equations. Extensive lists of mathematical functions for engineers are found in [6–8]. When calculations in two or three dimensions are too complicated, it can help to map the problem on a circle or a sphere. Especially time-repetitive signals and electromagnetic field calculations use cyclic and spherical functions to simplify the analysis. Table 2.3 lists some mathematical properties of spheres and circles. Also the use of complex notation can help to visualize rotation, see Table 2.4.

Many events in nature have a cyclic and repetitive character. Sinusoidal wave forms describe these events and Table 2.5 gives a number of regular goniometrical expressions.

Derivatives of functions represent the way a function changes from one set of values of its variables to a next set. In electronics the derivative helps to understand the behavior of complex functions that are used in a small range of the main variables.

---

[1]This confusion is related to the use of the "long scale" numbering system in continental Europe and the "short scale" numbering system in the USA and UK.

Table 2.1 Multiplier abbreviations

Name	Abbreviation	Multiplier
Googol		10^{100}
Exa	E	10^{18}
Peta	P	10^{15}
Tera	T	10^{12}
Giga	G	10^{9}
Mega	M	10^{6}
kilo	k	10^{3}
hecto	h	10^{2}
deca	da	10
unity		1
deci	d	10^{-1}
centi	c	10^{-2}
milli	m	10^{-3}
micro	μ	10^{-6}
nano	n	10^{-9}
pico	p	10^{-12}
femto	f	10^{-15}
atto	a	10^{-18}

Table 2.2 Elementary algebraic functions [6–8]

$$n! = 1 \times 2 \times 3 \times \cdots \times n \qquad\qquad 0! = 1, \quad 1! = 1$$

$$\binom{n}{m} = \frac{n!}{m!(n-m)!} \qquad\qquad \binom{n}{n} = \binom{n}{0} = 1$$

$$(a+b)^n = a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 \qquad (a+b)^2 = a^2 + 2ab + b^2$$

$$+ \cdots + \binom{n}{n-1}ab^{n-1} + b^n \qquad (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b \qquad\qquad a^2 - b^2 = (a-b)(a+b)$$

$$+ \cdots + ab^{n-2} + b^{n-1})$$

$$ax^2 + bx + c = 0, \quad x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$$

$$\sum_{i=0}^{n} r^i = \frac{1 - r^{n+1}}{1 - r} \qquad\qquad \sum_{i=0}^{\infty} r^i = \frac{1}{1-r}, \quad |r| < 1$$

$$a^b = e^{b \ln a} \qquad\qquad a^0 = e^0 = 1, \quad e = 2.71828$$

$$\ln(b) = \ln(a) \times {}^a\log(b) \qquad\qquad \ln(10) = 2.303, \quad {}^{10}\log(10^n) = n$$

$$\qquad\qquad {}^{10}\log(2) = 0.301, \quad {}^{10}\log(3) = 0.477$$

Table 2.3 Circle and sphere functions [6–8]

Perimeter of a circle	$2\pi r$
Area of a circle	πr^2
Surface of a sphere	$4\pi r^2$
Volume of a sphere	$\dfrac{4}{3}\pi r^3$

Table 2.4 Complex notation [6–8]

$j^2 = -1$	$|a + jb|^2 = a^2 + b^2$
$z = a + jb$ conjugate(z): $z^* = a - jb$	$\mathrm{Re}(z) = \mathrm{Re}(z^*) = a,$ $\mathrm{Im}(z) = -\mathrm{Im}(z^*) = b$
$e^{\pm j\alpha} = \cos(\alpha) \pm j\sin(\alpha)$	$\sin(\alpha) = \dfrac{e^{j\alpha} - e^{-j\alpha}}{2j},$ $\cos(\alpha) = \dfrac{e^{j\alpha} + e^{-j\alpha}}{2}$

Table 2.5 Goniometrical relations used in this book. The argument of goniometric formulas is expressed in radians $(0\ldots 2\pi)$ not in degrees [6–8]

$\sin(-\alpha) = -\sin(\alpha)$	$\cos(-\alpha) = \cos(\alpha)$
$\sin(\alpha) = \cos(\frac{\pi}{2} - \alpha)$	$\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}$
$\tan(\alpha) = \dfrac{\sin(\alpha)}{\cos(\alpha)} = a$	$\arctan(a) = \alpha$
$\sin^2(\alpha) + \cos^2(\alpha) = 1$	
$\sin(2\alpha) = 2\sin(\alpha)\cos(\alpha)$	$\cos(2\alpha) = \cos^2(\alpha) - \sin^2(\alpha) = 2\cos^2(\alpha) - 1$
$\sin(3\alpha) = -4\sin^3(\alpha) + 3\sin(\alpha)$	$\cos(3\alpha) = 4\cos^3(\alpha) - 3\cos(\alpha)$
$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta)$	$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$
$\sin(\alpha - \beta) = \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta)$	$\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)$
$2\sin(\alpha)\sin(\beta) = -\cos(\alpha + \beta) + \cos(\alpha - \beta)$	$2\cos(\alpha)\cos(\beta) = \cos(\alpha + \beta) + \cos(\alpha - \beta)$
$2\sin(\alpha)\cos(\beta) = \sin(\alpha + \beta) + \sin(\alpha - \beta)$	$2\cos(\alpha)\sin(\beta) = \sin(\alpha + \beta) - \sin(\alpha - \beta)$
$\sinh(\alpha) = \dfrac{e^\alpha - e^{-\alpha}}{2}$	$\cosh(\alpha) = \dfrac{e^\alpha + e^{-\alpha}}{2}$
$\tanh(\alpha) = \dfrac{\sinh(\alpha)}{\cosh(\alpha)} = \dfrac{e^\alpha - e^{-\alpha}}{e^\alpha + e^{-\alpha}},$	$\tanh(\alpha) = -j\tan(j\alpha)$

Integration of functions is the main mathematical method to form summations over time, area, etc. Elementary manipulation of derivatives and integrals of functions are given in Table 2.6.

In many applications the function $f(x)$ is known, however the behavior of this (perhaps) complicated function is required over only a small fraction of the entire range of x. If the function's behavior close to point $x = a$ is needed, the derivative of the function gives the direction in which the function will change. By adding

Table 2.6 Standard manipulations for derivatives and integrals of functions [6–8]

$$\frac{du(v(t))}{dt} = \frac{du(v)}{dv}\frac{dv(t)}{dt} \qquad\qquad\qquad \frac{du(at)}{dt} = a\frac{du(t)}{dt}$$

$$\frac{d}{dt}(u(t)v(t)) = u(t)\frac{dv(t)}{dt} + v(t)\frac{du(t)}{dt} \qquad\qquad \frac{d}{dt}u^n(t) = u^{n-1}(t)\frac{u(t)}{dt}$$

$$\frac{d}{dt}\left(\frac{u(t)}{v(t)}\right) = \frac{v(t)\dfrac{du(t)}{dt} - u(t)\dfrac{dv(t)}{dt}}{v^2(t)} \qquad\qquad \frac{d}{dt}\ln(at) = \frac{a}{t}$$

$$\frac{d}{dt}\sin(at) = a\cos(at) \qquad\qquad\qquad \frac{d}{dt}\cos(at) = -a\sin(at)$$

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} \quad n \neq 1 \qquad\qquad \int \frac{1}{x}\,dx = \ln(x)$$

$$\int \frac{1}{a^2 + x^2}\,dx = \frac{1}{a}\arctan(x/a) \qquad\qquad \int_{x=0}^{\infty} \frac{1}{a^2 + x^2}\,dx = \frac{\pi}{2a}$$

$$\int f(x)\delta(x - x_0)\,dx = f(x_0)$$

$$\int_{x=0}^{x=L} \sin(2\pi nx/L)\sin(2\pi mx/L)\,dx = 0, \qquad \int_{x=0}^{x=L} \sin^2(2\pi nx/L)\,dx = 0.5 \quad n, m : \text{integer}$$
$$\text{if } m \neq n$$

Table 2.7 Taylor series expansions [6–8]

$$(1 + a)^n \approx 1 + na + \frac{n(n-1)}{2!}a^2 + \cdots, \quad |a| \ll 1 \qquad \sqrt{1 + a} \approx 1 + 0.5a, \quad |a| \ll 1$$

$$e^{1+a} \approx 1 + a + +\frac{a^2}{2!} + \frac{a^3}{3!}, \quad |a| \ll 1 \qquad \ln(1 + a) \approx a - \frac{a^2}{2} + \frac{a^3}{3}, \quad |a| \ll 1$$

$$\sin(\alpha) \approx \alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!}, \quad \alpha \text{ in radians}, |\alpha| \ll 1 \qquad \cos(\alpha) \approx 1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!}, \quad |\alpha| \ll 1$$

higher order derivatives a series expansion is formed, that is useful to represent a complicated function. A function as defined by its Taylor series, looks like:

$$f(x) = f(a) + \frac{(x-a)}{1!}\frac{df(x)}{dx}\bigg|_{x=a} + \frac{(x-a)^2}{2!}\frac{d^2 f(x)}{dx^2}\bigg|_{x=a}$$
$$+ \frac{(x-a)^3}{3!}\frac{d^3 f(x)}{dx^3}\bigg|_{x=a} + \cdots. \tag{2.1}$$

Table 2.7 gives some Taylor series expansions for regular functions.

The Taylor series expands a function at one moment in time. For static signals this representation is the basis for non-linear analysis. The Volterra series is a method for including time-dependent effects.

$$f(x,t) = \sum_{i=1}^{\infty} \frac{1}{i!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g_i(\tau_1, \ldots, \tau_i) x(t - \tau_1) \times \cdots \times x(t - \tau_i) \, d\tau_1 \cdots d\tau_i$$

$$= g_0 + \frac{1}{1!} \int_{-\infty}^{\infty} g_1(\tau_1) x(t - \tau_1) \, d\tau_1$$

$$+ \frac{1}{2!} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(\tau_1) g_2(\tau_1, \tau_2) x(t - \tau_1) x(t - \tau_2) \, d\tau_1 \, d\tau_2$$

$$+ \frac{1}{3!} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(\tau_1) g_2(\tau_1, \tau_2) g_3(\tau_1, \tau_2, \tau_3) x(t - \tau_1) x(t - \tau_2)$$

$$\times x(t - \tau_3) \, d\tau_1 \, d\tau_2 \, \tau_3 \ldots$$

The first term in the Volterra series is the convolution function as in (2.144). Laplace and Fourier analogies of the Volterra series are useful to evaluate time-dependent distortion. Various techniques exist to estimate the coefficients g_i.

2.1.1 The Fourier Transform

The Fourier transform is used to analyze the behavior of time repetitive signals $h(t)$. The Fourier transform and its inverse transform are defined as:

$$H(\omega) = \int_{t=-\infty}^{\infty} h(t) e^{-j\omega t} \, dt = \int_{t=-\infty}^{\infty} h(t) e^{-j2\pi f t} \, dt$$

$$h(t) = \frac{1}{2\pi} \int_{\omega=-\infty}^{\infty} H(\omega) e^{j\omega t} \, d\omega = \int_{f=-\infty}^{\infty} H(f) e^{j2\pi f t} \, df \qquad (2.2)$$

In case a sinusoidal current is applied, the steady-state voltage and current relations for coils, capacitors and resistors are described in the Fourier domain, resulting in:

$$v(\omega) = j\omega L i(\omega) = j2\pi f L i(\omega)$$

$$v(\omega) = R i(\omega) = i(\omega)/g$$

$$v(\omega) = \frac{i(\omega)}{j\omega C} = \frac{i(\omega)}{j2\pi f C}$$

where the complex notation "$j = \sqrt{-1}$" is used to indicate that a 90° phase shift exists between the current and the terminal voltage.

Some special relations exist between the time-domain and the Fourier domain. A physical signal is a real signal and its imaginary part[2] equals 0. The Fourier transform converts the real time-continuous function $A(t)$ in a complex function $A(\omega)$

---

[2]In modern communication theory, single side-band modulation coins its two signal components with a 90° phase relation (I/Q) as the "real part" and an "imaginary part". This is a form of notation on a higher level than straight-forward Fourier analysis of a physical signal.
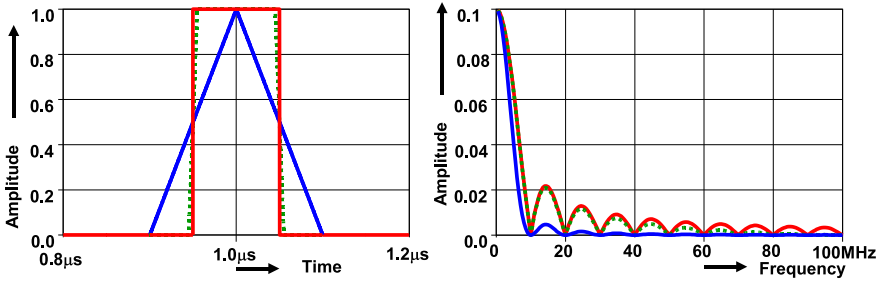
Fig. 2.1 The transform from a box function in the time domain leads to a $\sin(x)/x$ shape in the frequency domain. Three different rise and fall times show the transition from a block to a triangle shaped pulse

with $\omega = 2\pi f$, describing the signal in the frequency domain. After transformation a real signal results in a Hermitian function with the following properties:

$$A(\omega) = A^*(-\omega) \quad \text{or equivalently}$$
$$\text{Re}(A(\omega)) = \text{Re}(A(-\omega)) \quad \text{and}$$
$$\text{Im}(A(\omega)) = -\text{Im}(A(-\omega))$$

where "Re" and "Im" define the real and imaginary parts of a function.

Parseval's energy conservation theorem states that if two time functions $x(t)$ and $y(t)$ exist in the frequency domain as $X(\omega)$ and $Y(\omega)$ then:

$$\int_{t=-\infty}^{\infty} x(t)y^*(t)\,dt = \frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} X(\omega)Y^*(\omega)\,d\omega \tag{2.3}$$

The substitution of $x(t) = y(t)$ results in the "Energy theorem":

$$\int_{t=-\infty}^{\infty} |x(t)|^2\,dt = \frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} |X(\omega)|^2\,d\omega = \int_{f=-\infty}^{\infty} |X(f)|^2\,df \tag{2.4}$$

The energy of a signal over infinite time equals the energy over infinite frequency range.

The Fourier transform links the time domain to the frequency domain, Fig. 2.1. A box function in the time domain leads to a $\sin(x)/x$ function in the frequency domain. If the slope of the time signal is made less steep, this its Fourier transform moves to a $(\sin(x)/x)^2$ function in case of a triangle shape. As the Fourier transform is symmetrical to and from each domain, a box function in the frequency domain will result also in a $\sin(x)/x$ function in the time domain. This is e.g. visible in the definition of time-discrete filter coefficients, Sect. 3.2.1.

2.1.2 Fourier Analysis

The sine wave describes in time the position of an object following a repetitive circular motion. The mathematical technique to split up any form of repetitive signal

in sine waves is the Fourier series expansion. The expansion for a function $f(x)$ that is repetitive over a distance L is:

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi n x/L) + \sum_{n=1}^{\infty} b_n \sin(2\pi n x/L) \qquad (2.5)$$

with

$$a_0 = \frac{2}{L} \int_{x=-L/2}^{L/2} f(x)\,dx$$

$$a_n = \frac{2}{L} \int_{x=-L/2}^{L/2} f(x)\cos(2\pi n x/L)\,dx$$

$$b_n = \frac{2}{L} \int_{x=-L/2}^{L/2} f(x)\sin(2\pi n x/L)\,dx$$

As an example consider a square wave in time $f(t)$, defined as:

$$-T/2 < t < 0 \quad f(t) = -1$$
$$0 < t < T/2 \quad\quad f(t) = 1$$

which repeats over a period T. Evaluation shows that: $a_0 = 0$, and $a_n = 0$ for all n. Only the b_n coefficients are unequal to zero:

$$b_n = \frac{2}{T}\left[\int_{t=-T/2}^{0} (-1)\sin(2\pi n t/T)\,dt + \int_{t=0}^{T/2} (+1)\sin(2\pi n t/T)\,dt \right]$$

$$b_n = \frac{-1}{n\pi}\left[(-1)(\cos(0) - \cos(n\pi)) + (+1)(\cos(n\pi) - \cos(0)) \right]$$

$$= \frac{2(1 - \cos(n\pi))}{n\pi}$$

The resulting terms b_n equal zero for even n and $b_n = 4/n\pi$ for odd n. Consequently the Fourier expansion for a square wave is a sum of sine waves with frequencies that are odd multiples of the fundamental $n = 1$ frequency, Fig. 2.2:

$$f(t) = \sum_{n=1,3,5,\ldots}^{\infty} \frac{4}{n\pi} \sin(2\pi n t/T)$$

The result of a Fourier expansion may seemingly differ if another starting point on the curve is chosen. In this case the same square wave $f(t)$ could also be defined symmetrically around $t = 0$:

$$-T/2 < t < -T/4 \quad f(t) = -1$$
$$-T/4 < t < T/4 \quad\quad f(t) = 1$$
$$T/4 < t < T/2 \quad\quad f(t) = -1$$

resulting in $a_0 = 0$, and $b_n = 0$ for all n. Only the odd a_n coefficients are now unequal to zero:

$$f(t) = \sum_{n=1,3,5,\ldots}^{\infty} \frac{4\sin(n\pi/2)}{n\pi} \cos(2\pi n t/T)$$

Fig. 2.2 A square wave
signal is decomposed in sine
waves with frequencies that
are odd integers of the
fundamental frequency. The
first five components are
added into an approximation
of a square-wave
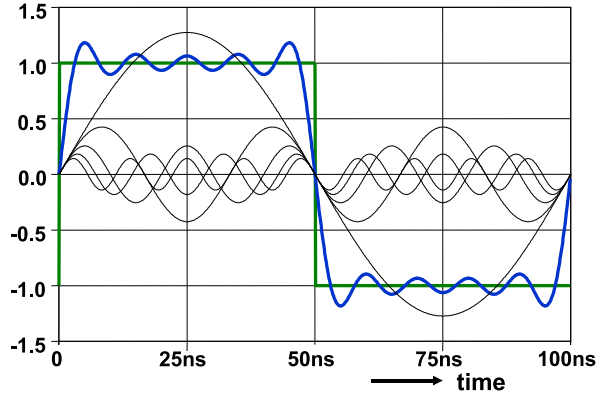
Table 2.8 Fourier series expansions of signals repeating at a period T [6–8]

Square wave (transition at $t = 0$)

$$f(t) = \begin{cases} -1 & -T/2 < t < 0 \\ +1 & 0 < t < T/2 \end{cases} \qquad\qquad f(t) = \sum_{n=1,3,5,...}^{\infty} \frac{4}{n\pi} \sin(2\pi nt/T)$$

Square wave symmetrical around $t = 0$

$$f(t) = \begin{cases} -1 & -T/2 < t < -T/4 \\ +1 & -T/4 < t < T/4 \\ -1 & T/4 < t < T/2 \end{cases} \qquad f(t) = \sum_{n=1,3,5,...}^{\infty} \frac{4\sin(n\pi/2)}{n\pi} \cos(2\pi nt/T)$$

Square wave with T_c high period

$$f(t) = \begin{cases} -1 & -T/2 < t < -T_c/2 \\ +1 & -T_c/2 < t < T_c/2 \\ -1 & T_c/2 < t < T/2 \end{cases} \quad f(t) = \sum_{n=1}^{\infty} \frac{4(-1)^{n+1}\sin(\pi nT_c/T)}{n\pi} \cos(2\pi nt/T)$$

Triangle

$$f(t) = \begin{cases} -\dfrac{t}{T} & -T/2 < t < 0 \\ +\dfrac{t}{T} & 0 < t < T/2 \end{cases} \qquad f(t) = \frac{1}{2} - \sum_{n=1,3,5,...}^{\infty} \frac{4}{n^2\pi^2} \cos(2\pi nt/T)$$

Saw tooth

$$f(t) = \frac{t}{T} \quad 0 < t < T \qquad\qquad f(t) = \frac{1}{2} - \sum_{n=1}^{\infty} \frac{1}{n\pi} \sin(2\pi nt/T)$$

The even coefficients are zero. Just as for the square wave the resulting Fourier
series experiences a 90° phase shift with respect to the first analysis.

Some more Fourier expansions are given in Table 2.8 [8, p. 405].

Fourier series expansions are used for determining distortion components. In this
example the square wave is composed of a fundamental sine wave and odd harmon-
ics. In order to find the power ratio between the fundamental and the harmonics, the
power content of both must be established.

$$
\begin{aligned}
P &= \frac{1}{T} \int_{t=-T/2}^{T/2} (f(t))^2 \, dt \\
&= \frac{1}{T} \int_{t=-T/2}^{T/2} \sum_{n=1,3,5,\ldots}^{\infty} \left(\frac{4}{n\pi} \sin(2\pi nt/T) \right)^2 dt \\
&= \frac{1}{T} \int_{t=-T/2}^{T/2} \sum_{n=1,3,5,\ldots}^{\infty} \left(\frac{4}{n\pi} \right)^2 \sin^2(2\pi nt/T) \, dt = \sum_{n=1,3,5,\ldots}^{\infty} \frac{8}{n^2\pi^2}
\end{aligned}
$$

The last formula conversion uses the fact that an integral over a period T of any product $\sin(2\pi nt/T)\sin(2\pi mt/T)$ with $n \neq m$ is equal to zero and only the DC terms of the squared sinusoidal terms result in non-zero contributions. On the other hand the power of the square wave can be easily calculated: the amplitude is either -1 or $+1$, which both result in a power of 1. Consequently:

$$
\sum_{n=1,3,5,\ldots}^{\infty} \frac{8}{n^2\pi^2} = \frac{8}{\pi^2} + \sum_{n=3,5,\ldots}^{\infty} \frac{8}{n^2\pi^2} = 0.81 + 0.19 = 1 \tag{2.6}
$$

from which the ratio between the first harmonic and the sum of the remaining components in a square wave can be derived:

$$
10^{10}\log\left(\frac{1}{\frac{\pi^2}{8} - 1} \right) = 6.31 \text{ dB}
$$

2.1.3 Distortion

In signal processing ratios between various quantities (signals, noise, distortion) are mostly specified as power ratios, e.g. the total harmonic distortion (THD):

$$
\text{THD} = \frac{P_{\text{distortion}}}{P_{\text{fundamental}}}
$$

In audio engineering the THD is expressed in %. As these ratios can amount many orders of magnitude, a logarithmic notation often replaces the exponential notation[3]:

$$
\text{THD} = 10^{\,10}\log\left(\frac{P_{\text{distortion}}}{P_{\text{fundamental}}} \right)
$$

The unit of ratio is called "decibel" or "dB", indicating the "deci" or one-tenth fraction of the unit "Bell".[4] In many cases a relation to the signal in the voltage or current domain is required:

$$
\text{THD} = 10^{10}\log\left(\frac{V_{\text{distortion}}^2/R}{V_{\text{fundamental}}^2/R} \right) = 20^{10}\log\left(\frac{V_{\text{distortion}}}{V_{\text{fundamental}}} \right) \tag{2.7}
$$

---

[3] 43.8 dB is a short hand for 4.167×10^{-5} power ratio. Use the exponential notation in complex calculations.

[4] The signal-to-noise ratio is defined in the opposite way: signal power divided by noise power. Therefore the minus sign that normally precedes the THD number is sometimes omitted.

The popular $20^{10}\log$(voltage ratio) is in fact a derived power ratio. In case of doubt always use power ratios.

Example Assume a transfer function of the form:

$$y = x + ax^2$$

where $a \ll 1$ is the generating term for second order distortion. This type of distortion is called "soft distortion" in contrast to "hard distortion" where a discontinuous jump in the signal or transfer is present. With an input signal $x(t) = V \sin(\omega t)$ and using some goniometric equivalences from Table 2.5, the distortion is calculated[5]:

$$y(t) = \frac{1}{2}aV^2 + V \sin(\omega t) - \frac{aV^2}{2} \cos(2\omega t)$$

Consequently the second order distortion component relative to the first order component is:

$$\text{HD2} = \frac{aV^2/2}{V} = \frac{aV}{2}$$

The second order component goes up quadratically if the signal amplitude rises linearly. Similar for a third order distortion:

$$y = x + bx^3$$
$$y = V \sin(\omega t) + \frac{3bV^3}{4} \sin(\omega t) - \frac{bV^3}{4} \sin(3\omega t)$$

The third order distortion relative to the first order component is:

$$\text{HD3} = \frac{bV^2}{4 + 3bV^2}$$

The third order component goes up with a third power when the input amplitude increases linearly. In the RF field this observation has led to a somewhat deviating formulation of the third order distortion: the third-order intercept point IP3, Fig. 2.3. In this point the extrapolated first order amplitude equals the extrapolated third order distortion amplitude. Using the previous analysis: $V_{\text{IP3}} = 2/\sqrt{b}$. The third-order intercept point can be related to the input axis (IIP3) or the output level (OIP3). Values of IIP3 exceeding 1 V are normally considered rather good.

In the previous analysis the stimulus consisted out of a single tone at a fundamental frequency. Its distortion products are situated at multiples of that frequency. In systems with filters, the transfer function for the distortion products can be different from the processing of the fundamental frequency. In that case an intermodulation method is used to determine the third order distortion. The input is chosen as the

---

[5]Using goniometric equivalences is an engineering short cut. The Fourier series expansion gives the same results.

Fig. 2.3 The third order intercept point is found by extrapolating the fundamental component and the third order distortion component
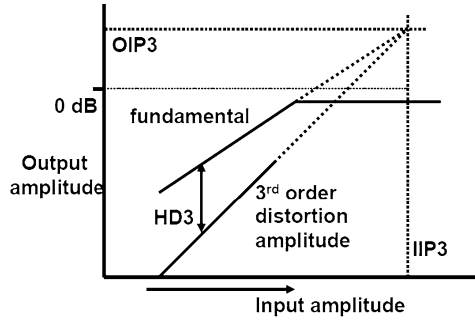
sum of two closely spaced carriers: $x = \frac{1}{2}V(\sin(\omega_1 t) + \sin(\omega_2 t))$ with $\omega_1 \approx \omega_2$. After passing a distorting stage this input results in:

$$y = \frac{1}{2}V(\sin(\omega_1 t) + \sin(\omega_2 t)) + \frac{bV^3}{8}(\sin^3(\omega_1 t) + 3\sin^2(\omega_1 t)\sin(\omega_2 t)$$
$$+ 3\sin(\omega_1 t)\sin^2(\omega_2 t) + \sin^3(\omega_2 t))$$

Using Table 2.5, the signal is now written as a sum of its spectral components:

$$y = \left(\frac{1}{2}V + \frac{9b}{32}V^3\right)(\sin(\omega_1 t) + \sin(\omega_2 t))$$
$$+ \frac{bV^3}{32}\{3(\sin(2\omega_1 t - \omega_2 t) + \sin(2\omega_2 t - \omega_1 t))$$
$$- 3(\sin(2\omega_1 t + \omega_2 t) + \sin(2\omega_2 t + \omega_1 t)) - (\sin(3\omega_1 t) + \sin(3\omega_2 t))\}$$

There are additional intermodulation distortion components at frequencies $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ with a relative amplitude:

$$IM3 = \frac{3bV^2}{16 + 9bV^2}$$

These intermodulation distortion products appear next to the pair of input frequencies at a spacing equal to the difference between the input frequencies. Their amplitude is a factor of 3 higher than the regular third order distortion. A proper choice of the input frequencies allows to process the intermodulation and the input components in the same way, thereby allowing a correct measurement of the distortion. The IM3 value, see Fig. 2.4 is roughly $\frac{3}{4}$ (-2.5 dB) of the magnitude of the third order distortion HD3.

2.1.4 Laplace Transform

Currents and voltages over resistors and transistors are preferably described with time-invariant models for the components. Coils and capacitors use first order

Fig. 2.4 A signal is distorted with $a = b = 0.1$ and gives second and third order distortion. A similar distortion is applied to two sine waves at a Δf frequency spacing but at half of the amplitude. This will result in IM3 products at Δf from each carrier
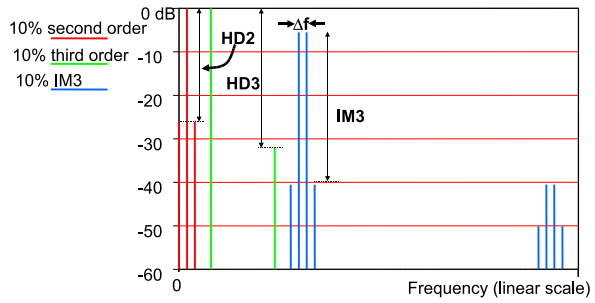
derivatives to relate their terminal voltages to the currents. A network with a number of these elements requires solving a higher order linear differential equation. In order to facilitate linear analysis the one-sided or unilateral Laplace transform allows to reformulate complex differential equations into simple polynomial calculus.

The Laplace transform of a function in time $f(t)$ is defined:

$$f(s) = \int_{t=0}^{\infty} e^{-st} f(t)\, dt = \int_{t=0}^{\infty} e^{-\alpha t} e^{-j\omega t} f(t)\, dt \qquad (2.8)$$

The Laplace transform exists if the function $f(t)$ is (piecewise) continuous in the positive time domain and if constants A, B, T exists such that $|f(t)| < Ae^{Bt}$ for $t > T$. The functions $f(t)$ and $f(s)$ have different mathematical formulations, as can be easily seen from Table 2.9. This is reflected in formal text books in different type settings of the function names, e.g. $f(t)$ and $F(s)$ or $\mathbf{f}(s)$. Even more type settings are needed after the introduction of Fourier and z-transforms. Yet all these mathematical descriptions refer to the same physical or electrical process. Therefore only in cases were the reader might get confused a difference in type setting is applied in this text.

The Laplace independent variable s consists of a real and imaginary part: $s = \alpha + j\omega$. The complex plane defines the exponential decay α on the horizontal axis and $j\omega$ as the radial frequency on the vertical axis. This plane is a plot of the variable "s", where for each s the function $f(s)$ has a complex value.[6] A curve in the s plane can separate e.g. the portions of the function $f(s)$ that have an amplitude larger or smaller than 1. Circles and crosses indicate where the function is zero or infinite (called a pole). If $f(t)$ is a real-valued function, the imaginary parts of $f(s)$ for $s = \alpha + j\omega$ and for $s = \alpha - j\omega$ are equal. A real function in the time domain results in a symmetrical function in the Laplace domain around the α axis.

The one-sided Laplace transform starts at $t = 0$ and allows the transformation of exponential, sinusoidal and polynomial functions of an independent variable, which in network theory is the time t.

The differential equations for e.g. capacitors and coils translate in the Laplace domain to first order polynomial expressions in s ($t > 0$), Table 2.9.

---

[6]One could imagine the complex value of the function plotted in the third and fourth dimensions.

Table 2.9 Laplace transforms. The notation $u(t)$ is the step function at $t = 0$ [6–8]

$v(t) = Ri(t)$	$v(s) = Ri(s)$
$v(t) = L\dfrac{di(t)}{dt}$	$v(s) = sLi(s) - Li(t=0)$
$i(t) = C\dfrac{dv(t)}{dt}$	$i(s) = sCv(s) - Cv(t=0)$
$v(t) = C\displaystyle\int_{\tau=-\infty}^{\tau=t} i(\tau)\,d\tau$	$v(s) = \dfrac{Ci(s)}{s}$
$v(t) = \displaystyle\int_{\tau=0}^{\tau=\infty} h(\tau)x(t-\tau)\,d\tau$	$v(s) = H(s)x(s)$
$v(t) = u(t)$	$v(s) = \dfrac{1}{s}$
$v(t) = u(t-T)$	$v(s) = \dfrac{e^{-sT}}{s}$
$v(t) = Ae^{-\alpha t}$	$v(s) = \dfrac{A}{s+\alpha}$
$v(t) = A\dfrac{t^n}{n!}e^{-\alpha t}$	$v(s) = \dfrac{A}{(s+\alpha)^{n+1}}$
$v(t) = A(1-e^{-\alpha t})$	$v(s) = A\dfrac{\alpha}{s(s+\alpha)} = A\left(\dfrac{1}{s} - \dfrac{1}{(s+\alpha)}\right)$
$v(t) = A\left[1 - e^{-\alpha t}\displaystyle\sum_{i=1}^{n}\dfrac{(\alpha t)^{i-1}}{(i-1)!}\right]$	$v(s) = A\dfrac{\alpha^n}{s(s+\alpha)^n} = A\left(\dfrac{\alpha^{n-1}}{s(s+\alpha)^{n-1}} - \dfrac{\alpha^{n-1}}{(s+\alpha)^n}\right)$
$v(t) = e^{-\alpha t}\sin(\omega t)$	$v(s) = \dfrac{\omega}{(s+\alpha)^2 + \omega^2}$
$v(t) = e^{-\alpha t}\cos(\omega t)$	$v(s) = \dfrac{s+\alpha}{(s+\alpha)^2 + \omega^2}$
$v(t) = e^{-\alpha t}\left[\cos(\omega t) + \left(\dfrac{\beta-\alpha}{\omega}\right)\sin(\omega t)\right]$	$v(s) = \dfrac{s+\beta}{(s+\alpha)^2 + \omega^2}$
$v(t=0+) = \lim\limits_{s\to\infty} sv(s)$	$v(t=\infty) = \lim\limits_{s\to 0} sv(s)$

The Laplace transform is linear and transforms a differentiation operation into a multiplication by s. An integration results in a division by s. When an analysis in the frequency domain is carried out, the real part of s is set to zero leaving the radial frequency $j\omega$ as the running variable. This is a quick route to come to a Bode analysis, see Fig. 2.50.

In electronic design the most common form of a signal or a transfer function in the Laplace domain is a fraction formed by two polynomials in s, e.g.:

$$\frac{N(s)}{D(s)} = \frac{s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}$$

$$= \frac{(s + s_{zm})(s + s_{z(m-1)})((s + s_{z2})^2 + s_{z1}^2)}{(s + s_{pn})(s + s_{p(n-1)})((s + s_{p2})^2 + s_{p1}^2)}$$

$$= \frac{N_n(s)}{(s + s_{pn})} + \frac{N_{n-1}(s)}{(s + s_{p(n-1)})} + \frac{N_2(s)}{((s + s_{p2})^2 + s_{p1}^2)}$$

The roots of the numerator polynomial ($s = -s_{zm}, s = -s_{z2} \pm js_{z1}$) are called "zeros" and the roots of the denominator polynomial ($s = -s_{pn}, s = -s_{p2} \pm js_{p1}$) are the "poles". A physical voltage-to-voltage transfer function goes to 0 for $s \to \infty$, which implies that in that case the numerator polynomial is of a lower order than the denominator polynomial. Returning to the time domain implies to factor the denominator and split the original formula in transformable parts. During the factoring there are a few possibilities for the poles:

- $s_n = 0$ which results in a constant term starting at $t = 0$.
- A real valued root s_{pn}. In the time domain this factor will result in an exponential term $e^{-s_{pn}t}$.
- A real valued root $s = -s_{pn} < 0$. In the time domain this factor will result in an exponentially decaying function.
- A real valued root $s = -s_{pn} > 0$. A design example of this exponentially growing function is in the analysis of a latch.
- A second order polynomial $(s + s_{p2})^2 + s_{p1}^2$ with roots $s = -s_{p2} \pm js_{p1}$. In the time domain pair of roots results in an exponentially decaying sinusoidal function if the real part of the root $-s_{p2} \leq 0$.

The axis of a complex plane in Fig. 2.5 show the real (horizontal) and imaginary (vertical) portions for the complex variable s and the complex variable z. In this plane crosses and circles indicate one real-valued pole, a pair of complex poles and a pair of imaginary zeros. Any pole in the Right-Half plane corresponds to an exponentially growing function. A zero in the Right-Half plane (RHZ) will show the phase behavior of a left plane pole and can therefore introduce undesired stability problems in feed-back.

The frequency transfer function as in the Bode-plot is found by moving over the $j\omega$ axis and measuring the absolute values of the distances from the frequency on the vertical axis to the poles and zeros. The distances to the zeros are multiplied and divided by the distances to the poles.

Fig. 2.5 The complex plane for the Laplace transform (s-plane) and the time-discrete plane (z-plane). A real pole, a pair of imaginary poles and a pair imaginary zeros are depicted
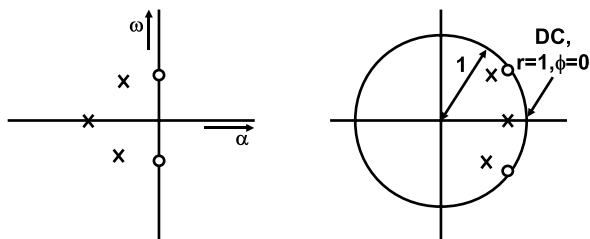
2.1.5 The z-transform

In a Laplace or Fourier analysis the functions and variables are assumed to be continuous with respect to the independent variable (mostly the time). In time-discrete signal processing another type of variable occurs: a sequence of values represents the signal. If $f(n) = f(nT_s), n = 0 \ldots \infty$ is a sequence of values corresponding to the value of $f(t)$ at points in time $t = nT_s$, this sequence can be described in the z-domain as:

$$f(z) = \sum_{n=0}^{n=\infty} f(n)z^{-n}$$

where z is a complex number in polar representation $z = re^{j\omega_z}$, which resembles the Laplace parameter $s = \alpha + j\omega$ with $r \leftrightarrow e^{\alpha}$, $\omega \leftrightarrow \omega_z$. The important difference is that the z-domain describes a sampled system where the maximum frequency is limited to half of the sample rate. While ω is expressed in rad/sec, $\omega_z \leftrightarrow \omega T_s$ is expressed in radians. The s-plane and the z-plane can be mapped on each other. Due to the polar description the $j\omega$ axis in the s-domain becomes a unity circle in the z-domain, with the DC point at $z = 1e^{j0} = 1$. Poles and zeros in left-side of the s-plane resulting in stable decaying exponential functions in the time domain move to the inner part of the unity circle in the z-domain, Fig. 2.5 (right).

The unilateral z-transform definition is used for causal system. Causal systems react only after the excitation is applied and not before. In the z-domain a delay of mT_s corresponds with a multiplication of z^{-m}.

The z-transform uses extensively series expansions:

$$\sum_{n=0}^{n=\infty} (az^{-1})^n = \frac{1}{1 - az^{-1}}$$

The summation is only bounded if the term in brackets is smaller than unity. If a transfer function therefore shows a term as in the right hand side of the equation, stability is only possible if $|z| > |a|$.

An integration in the time-discrete domain is carried out by adding the previous output result $Y((n-1)T_s)$ weighted with a factor a to the input:

$$Y(nT_s) = X(nT_s) + aY((n-1)T_s) \rightarrow H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - az^{-1}}$$

The pole in the z-domain is at $z = a$ and is stable as long as $|a| \leq 1$ and inside the unity circle.

The z-transform can be mapped on both the Laplace domain and the Fourier domain. Every mapping is applicable in a limited range. The most simple mapping uses the substitutions:

$$\begin{aligned} z &\leftrightarrow e^{sT_s} \\ z &\leftrightarrow e^{j\omega T_s} \end{aligned} \qquad (2.9)$$

The exponential function is not a design friendly function. In order to transform time continuous functions into the z-domain and vice versa, a linear or a bilinear transform can be used. These approximations of the exponential function are however only valid for the narrow frequency band $f \ll 1/T_s$:

$$z \leftrightarrow 1 + sT_s \quad s \leftrightarrow \frac{1}{T_s}(z-1)$$

$$z \leftrightarrow \frac{2 + sT_s}{2 - sT_s} \quad s \leftrightarrow \frac{2}{T_s}\frac{(z-1)}{(z+1)} \tag{2.10}$$

2.1.6 Statistics

Many phenomena in nature are so complex that an exact treatment of the problem is not possible. The outcome of throwing a die is fully calculable if the exact dimensions, position, forces, etc are known. However this is a typical example where statistics are more useful. Problems involving many similar events can be modeled effectively with the help of statistics.

Figure 2.6 shows the probability density function of throwing a single die: each outcome $1, 2, \ldots, 6$ has a chance or probability of $p = 1/6$. The plot of its probability is called a probability density function, whose sum or integral equals by definition 1. With the probabilities for all outcomes being equal the probability density function of a die is a "uniform" distribution function. If n dice are used the probability of having exactly k occurrences (e.g. $k = 3$ occurrences of the value $x = 4$) is described by a binomial probability density function.

$$p(k) = \frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \tag{2.11}$$

Example What is the probability of having exactly one die showing "6" if six dice are tossed. Now $n = 6$, $p = 1/6$ and $p(k = 1) = (5/6)^5$.

The binomial distribution combines the probability of a single event to happen into the probability of a number of these events to occur. In Fig. 2.6 (right) the prob-
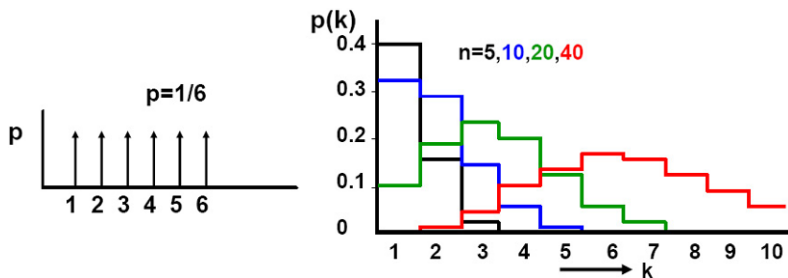
Fig. 2.6 The distribution function of one die, and the binomial distribution for multiple dice with $n = 5, 10, 20, 40$

Fig. 2.7 The Poisson
distribution with $\lambda = 1, 2, 5,$
10 and 20. The distribution
shifts from Poisson to
Gaussian for increasing
values of the parameter λ
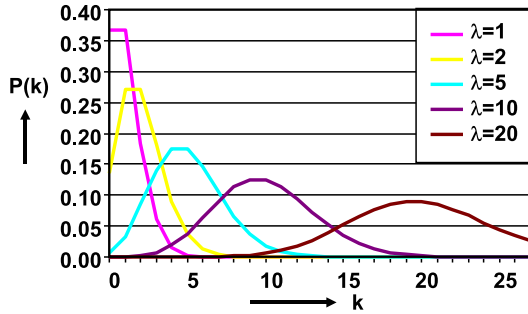
ability density for $n = 5, 10, 20$ and 40 is shown. This type of experiment uses a
discrete stochastic variable or random variable k and the probability $p(k)$ is given
by the probability density function. If n is large and p is small the binomial function
will converge to a Poisson distribution with $\lambda = n \times p$.   This probability function
describes a process where a number of events can happen during an observation pe-
riod. If the average number of expected events is λ than the probability that exactly
$k = 0, 1, 2, \ldots$ events will occur is a Poisson distribution:

$$p(k, \lambda) = \frac{e^{-\lambda} \lambda^k}{k!} \tag{2.12}$$

Example If a typist makes on average two errors per 10 minutes than the probability
that (s)he makes one error in that time period is $p(1, 2) = 0.27$.

A Poisson distribution is used if an average probability is known and a binomial
distribution if the probability of a single event is given.

For large λ the discrete Poisson distribution is again approximated by a Gaussian
distribution with $\mu = \sigma^2 = \lambda$. The shift from a Poisson to a Gaussian distribution is
shown in Fig. 2.7.

The normal probability density function or Gaussian distribution has a probabil-
ity density function, see Fig. 2.8:

$$p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.13}$$

The discrete variable k has moved into the continuous variable x. A normal distri-
bution is often denoted as $N(\mu, \sigma)$ and a standard normal distribution as $N(0, 1)$.
In the transition from a binomial distribution to a normal distribution, the following
parameter equality applies: $\mu = np$ and $\sigma = np(1 - p)$.

x is a continuous stochastic variable and normally the question $p(x = 3)$ has no
meaning or equals zero. The probabilities are now defined between two limits. The
probability that an event occurs between $x = x_1$ and $x = x_2$ is found by integration:

$$\text{Probability}(x_1 < x < x_2) = \frac{1}{\sigma \sqrt{2\pi}} \int_{x=x_1}^{x=x_2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx \tag{2.14}$$

With $x_1 = -\infty$ this integral represents the probability that an event $x < x_2$ will
occur. This integral is called the cumulative normal probability distribution, Fig. 2.9.
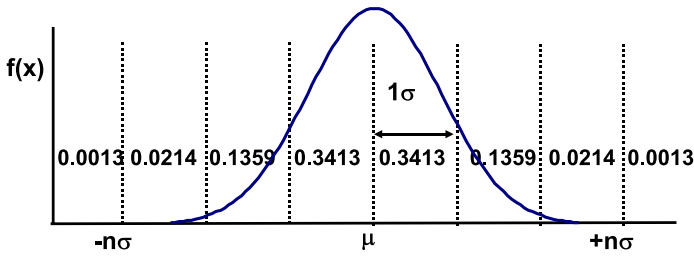
Fig. 2.8 The normal probability distribution with the probability values per one-σ interval

Table 2.10 Probability that an experimental value exceeds the $n\sigma$ limit on one or on two sides in a normal distribution [6–8]

n	Probability of single sided reject $P((x - \mu) > n\sigma)$	Probability of dual sided reject $P(\|x - \mu\| > n\sigma)$
1	0.159	0.317
2	0.0228	0.0455
3	1.3×10^{-3}	2.7×10^{-3}
4	3.1×10^{-5}	6.3×10^{-5}
5	2.9×10^{-7}	5.8×10^{-7}
6	1.0×10^{-9}	2.0×10^{-9}
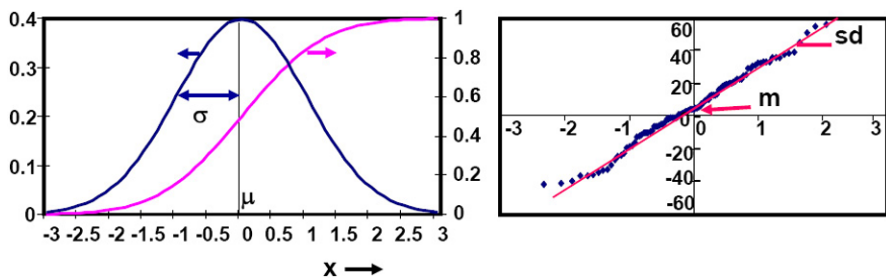
Fig. 2.9 The normal probability distribution and the cumulative probability distribution (*left*), and the normal-scaled cumulative probability distribution (*right*)

The characteristics of probability density functions are summarized in their n-th order moments or expected values:

$$E(x^n) = \sum_{i=0}^{\infty} x_i^n \, p(x_i)$$

$$E(x^n) = \int_{x=-\infty}^{\infty} x^n \, p(x) \, dx$$

(2.15)
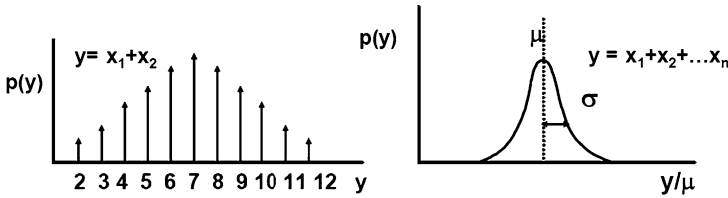
Fig. 2.10 The distribution function the sum of two dice and the sum of a large number of dice, illustrating the Central Limit Theorem

where the summation is used for a discrete probability density function and the integral for a continuous function. The mean value or the expectation value is reached with $n = 1$ and the variance is obtained via:

$$\mu = E(x)$$
$$\mathrm{Var}(x) = E(x^2) - (E(x))^2 \tag{2.16}$$

In electronics μ often equals the DC value and the variance estimates the AC-power (while considering the covariance). For a simple one-die experiment the values are $\mu = 3.5$ and the variance $= 2.92$. The normal probability distribution function gives $E(x) = \mu$ and the variance $= \sigma^2$.

The Central Limit Theorem says that every sum of independent random variables with the same probability distribution function converges to a normal distribution, Fig. 2.10. If a random variable y is formed by summing k instances of a random variable x: $y = x_1 + x_2 \cdots x_k$ the probability distribution for higher values of k converges to a normal distribution. The Central Limit Theorem requires that the random variables x_k are mutually independent: the outcome of one experiment cannot depend on the outcome of another experiment. However there is no requirement on the shape of the probability distribution of the random variables x_k. In nature and electronics many phenomena fulfill this requirement.

The probability distribution function and its describing parameters are theoretical models, e.g. the normal distribution is assumed to be valid for describing the length distribution of army cadets. In a series of experiments the validity of this assumption must be tested and its parameters must be estimated. The estimators for the expectation value and the square root of the variance are the mean m and the standard deviation "s.d." With an infinite number of experiments m will approach μ and the standard deviation will come close to $\sqrt{\text{variance}}$. It is a common habit to use the above formula's for the n-th order moment to estimate the variance and mean μ of a distribution:

$$\text{Estimation of } \mu(x) \quad m = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$\text{Estimation of variance } \mathrm{Var}(x) \quad (\text{s.d.})^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - m)^2 \tag{2.17}$$

The disadvantage of this method in practical applications is that unintended outliers that have no relation with the actual probability distribution, will influence the estimation in a disproportional manner. A more robust approach is to use rank-linear methods [9].

The cumulative normal probability distribution in Fig. 2.9 requires to create an ideal set of N points, where N corresponds to the number of data points in the real-life experiment. The vertical axis is subdivided in N equidistant intervals and for each interval the associated value on the horizontal axis is taken. This set of data is the inverse of the cumulative probability curve and corresponds to an ideal normally distributed data set with N samples. All N values obtained from the experiment are ranked in ascending order and paired to the ideal data points. A plot of these ideal N points on the horizontal axis and the experimental N points on the vertical axis, Fig. 2.9, should show a straight line. The intercept with the vertical axis is now an estimator for the mean value μ of the distribution. This zero value on the horizontal axis is the middle observation in the rank and its paired value on the vertical axis corresponds to the median value. The slope of the line estimates the $\sqrt{\text{variance}}$. A strong deviation of this plot with a straight line indicates that the assumption of normally distributed experimental values is not true.

In many situations the observed stochastic variables originate from one or more sources. Therefore between multiple stochastic parameters various relations can exist. Independence means that by no means the probability of stochastic variable x_1 is influenced by stochastic variable x_2 and vice-versa for any value of these random processes.

A measure for mutual influence is the covariance:

$$\text{Cov}(x_1, x_2) = E((x_1 - m_{x1})(x_2 - m_{x2})) = E(x_1 x_2) - m_{x1} m_{x2}$$

and the normalized value called the correlation coefficient:

$$\text{Correlation coefficient} = \frac{\text{Cov}(x_1, x_2)}{m_{x1} m_{x2}}$$

These terms indicate the amount of linear relationship between two random processes. Often correlation is misused to identify a "cause-effect" relation, the correlation coefficient is just a mathematical relation. If random processes are identical $x_1 = x_2 = x$ the covariance equals the variance of x. If x_1 and x_2 are independent the covariance equals "0". However $\text{Cov}(x_1, x_2) = 0$ is a necessary but not a sufficient condition for independence. It is possible to construct fully dependent stochastic parameters with probability density functions that result in $\text{Cov}(x_1, x_2) = 0$. The mere observation that the covariance between two stochastic variables equals zero is called "uncorrelated".

In circuit design the function $y = f(x)$ is often a transfer function of a current or voltage into another current or voltage. If y relates to x via a smooth and differentiable function, Fig. 2.11, the mean and variance of y can be approximated by the partial derivative using the Taylor series expansion [10]:

$$\mu_y = E(f(x)) \approx f(\mu_x)$$

$$\text{Var}(y) = E(f^2(x)) - (E(f(x)))^2 \approx \left(\frac{df(x)}{dx}\right)^2 \text{Var}(x)$$

Fig. 2.11 The transformation
of one stochastic variable in
another via a function
$y = f(x)$ uses the Taylor
expansion
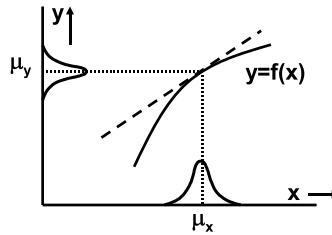
Equivalently the relation $g(x_1, x_2)$ which is a function of two random variables x_1 and x_2, can be calculated:

$$\text{Var}(g) \approx \left(\frac{dg(x_1)}{dx_1}\right)^2 \text{Var}(x_1) + \left(\frac{dg(x_2)}{dx_2}\right)^2 \text{Var}(x_2)$$

$$+ \left(\frac{dg(x_1)}{dx_1}\frac{dg(x_2)}{dx_2}\right)\text{Cov}(x_1, x_2)$$

For uncorrelated or independent variables this result reduces to:

$$\sigma_g^2 \approx \left(\frac{\partial g(x_1)}{\partial x_1}\right)^2 \sigma_{x1}^2 + \left(\frac{\partial g(x_2)}{\partial x_2}\right)^2 \sigma_{x2}^2 \tag{2.18}$$

The above equation can be easily expanded to three or more input variables. If g is a simple sum of terms, the variance of g equals the well-known sum of the variances of its composing terms, e.g. for the sum or differences of normal distributed variables:

$$g(x_1, x_2, x_3, \ldots) = a_1 x_1 \pm a_2 x_2 \pm a_3 x_3 \pm \cdots$$

$$E(g) = a_1 E(x_1) \pm a_2 E(x_2) \pm a_3 E(x_3) \pm \cdots \tag{2.19}$$

$$\sigma_g^2 = a_1^2 \sigma_{x1}^2 + a_2^2 \sigma_{x2}^2 + a_3^2 \sigma_{x3}^2 + \cdots$$

The squaring operation on the partial derivatives of the variance causes that the variance is independent of the "plus" or "minus" sign in front of the constituent terms.

2.2 Resistivity

Resistance is the property to obstruct the flow of current. If a voltage V is applied to a piece of material, Fig. 2.12, then Ohm's law expresses the current flowing through this piece of material as:

$$V = I \times R$$
$$I = G \times V \tag{2.20}$$

Fig. 2.12 Current flowing
through a piece of material
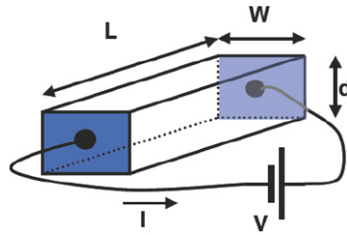
Black	Brown	Red	Orange	Yellow	Green	Blue	Violet	Grey	White
0	1	2	3	4	5	6	7	8	9
10^0	10^1	10^2	10^3	10^4	10^5	10^6	10^7		

Tolerance: gold=5%, silver=10%

Fig. 2.13 Color coding for discrete resistors. The first ring is closest to the extreme of the resistor.
The first two or three rings form the value ("47" and "560"). The last but one ring is the multiplier
(100× and 10×). The last ring indicates the accuracy and is at some distance of the other rings or
somewhat broader. The left-hand resistor is 4700 Ω, the right-hand is 5600 Ω

The value R describes the resistance formed by a specific piece of material. The
inverse notion of resistance is the conductance G. Resistors are used as discrete
elements in printed-circuit boards, see Fig. 2.13, and in integrated form. The resistor
is than constructed in a sheet of material.

Each material is characterized by the intrinsic property "resistivity" or as a sym-
bol: ρ in Ω m. Figure 2.12 shows a rectangular piece of material. Based on the
general material property of resistivity, the equivalent resistor value is determined
as:

$$R = \frac{L\rho}{W \times d} \; [\Omega] \tag{2.21}$$

With dimensions W for the width, L for length and d for thickness.

The resistivity of various materials in the field of semiconductors is given in
Table 2.11. In some cases the conductivity of material σ is specified, $\sigma = 1/\rho$. In
many semiconductor applications the thickness is determined by the process and is
therefore a fixed number in a process. Consequently a simplification is applied:

$$R = \frac{L R_\square}{W}, \quad \text{with } R_\square = \frac{\rho}{d}$$

where R$_\square$ is called the "square-resistance". This resistance is often quoted in pro-
cess specifications. The actual resistance of a track is now calculated by counting
the number of squares it contains and multiplying this by $R_\square$.

Table 2.11 Resistivity of (semi-)conductors at room temperature, the range indicates values from different sources, e.g. CRC handbook [11, p. e-78] and (2.64). See also Table 2.15

Material	Resistivity ρ Ω m	Temp. coeff. K^{-1}
Aluminum	2.82×10^{-8}	0.0039–0.0043
Copper	1.72×10^{-8}	0.0039
Gold	2.44×10^{-8}	0.0034–0.0037
Iron	9.7×10^{-8}	0.005–0.0056
Silicon doped with 10^{12} cm^{-3} As	4.4×10^{3}	0.007
Silicon doped with 10^{15} cm^{-3} As	4.6	0.007
Silicon doped with 10^{18} cm^{-3} As	2.2×10^{-2}	0.007
Poly crystalline silicon	$1\,\Omega$–1 k$\Omega/\square$	0.001 K^{-1}
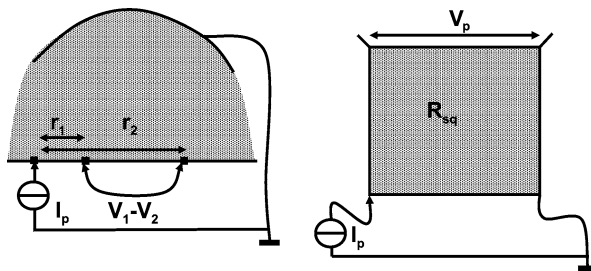
Fig. 2.14 van der Pauw theory and van der Pauw structure for measuring the sheet resistance

For general cases van der Pauw [12] has developed methods to determine the resistance value from general principles. Suppose that in an infinite sheet of resistive material at some point a current I_p is injected flowing into ground at infinite distance, Fig. 2.14 (left). This current will spread out uniformly through the material creating a circular potential distribution: on a distance r the potential is equal to $V(r)$. Adding another dr will decrease the potential by dV, which can be calculated by realizing that the total current I_p has to pass through a shell of the material with length dr and perimeter $2\pi r$

$$dV = \frac{-I_p \times R_\square \times dr}{2\pi r}$$

A voltage drop between two points on a distance from the injection point of r_1 and r_2, will now be equal to:

$$V_1 - V_2 = \int_{V2}^{V1} dV = \int_{r2}^{r1} \frac{-I_p R_\square}{2\pi r} dr = \frac{-I_p R_\square}{2\pi} \ln\left(\frac{r_2}{r_1}\right) \qquad (2.22)$$

Based on this theory in IC design the van der Pauw structure is used for measuring sheet resistances. After some mathematical manipulations, the voltage-to-current relation in Fig. 2.14 (right) is found as:

$$V_p = I_p R_\square \frac{\ln(2)}{\pi}$$

In a similar way the resistance of a semi-sphere contact in a substrate with a specific resistivity ρ is found:

$$V_1 - V_2 = \int_{V2}^{V1} dV = \int_{r2}^{r1} \frac{-I_p \rho}{4\pi r^2} \, dr = \frac{-I_p \rho}{4\pi} \left(\frac{1}{r_2} - \frac{1}{r_1} \right)$$

In a 10 Ω cm substrate a contact area with a radius of 1 μm will represent a resistance of approximately 8 kΩ to the substrate.

2.2.1 Temperature

Energy is the total amount of work that has to be done to carry out a task. Power is the amount of energy per unit time. Power and energy in a resistor are described as:

$$\text{Energy} = \int_{t=-\infty}^{\infty} P(t) \, dt, \quad P(t) = V(t) I(t) = I^2(t) R = \frac{V^2(t)}{R} \quad (2.23)$$

If a resistor consumes power, the accumulated energy must be removed. The resistor will turn the electrical power P into an equivalent amount of thermal power. The thermal power will raise the local temperature T_R and spread out to a region with lower temperature T_{ambient}. The temperature difference is found from:

$$T_R - T_{\text{ambient}} = P \times (K_{m1} + K_{m2} + \cdots) \quad (2.24)$$

Where K_{mi} represent the thermal resistances of the ith structure between the power consuming element and the ambient in degrees Kelvin per Watt. Consider as an example a chip consuming 0.8 W mounted in a package with a thermal resistance of $K_{mi} = 40$ K/W to the environment. The chip surface temperature will rise 32 K over the ambient temperature. The radiation of energy from the plastic package is limited. The thermal conductance of a package to the environment is dominated by the thermal properties of the leads of the package and by heat sinks.

The thermal properties of layers in an IC are dominated by the oxide layers. The thermal conductivity for silicon oxide κ_{ox} is two orders of magnitude lower than for bulk silicon A structure on a silicon oxide layer will experience a thermal resistance of

$$K_m = \frac{t_{ox}}{\kappa_{ox} W L} \quad (2.25)$$

where t_{ox} is the layer thickness, W and L the lateral dimensions. A heat source of 1×1 μm on top of a 1 μm silicon dioxide layer sees a thermal resistance of

Table 2.12 Thermal conductivity, the range indicates values from multiple sources, e.g. [13]

Material	Thermal conductivity W/K m (Watt per meter per Kelvin)
Aluminum	240
Copper	400
Silicon	130–150
Silicon dioxide	1.1–1.4

7×10^5 K/W. This thermal isolation is a major concern for high-power polysilicon resistors that are encapsulated in silicon dioxide and in Silicon-On-Insulator processes.

Next to a thermal resistance, every element on a chip will show a thermal capacity, or the ability to store thermal energy. Thermal capacitance and thermal resistance form a thermal time constant, similar to an electrical time constant. In Integrated Circuit manufacturing, this thermal time constant may be as low as several microseconds. This value implies that the temperature variation can follow signals in the audio frequency range. As the temperature rise and fall modulates the resistivity itself, this effect may result in distortion.

2.2.2 Voltage and Temperature Coefficient

In some resistors the applied voltage itself modulates the resistivity. In that situation a voltage dependence is defined:

$$R(V) = R(V_0)(1 + \text{VC}(V - V_0))$$
$$\text{VC} = \frac{1}{R}\frac{dR}{dV}$$

(2.26)

One of the consequences of a rising temperature is the change in resistivity, which in most materials is caused because of changes in the carrier mobility. Resistors will show a temperature dependency which is given by:

$$R(T) = R(T_0)(1 + \text{TC}(T - T_0))$$
$$\text{TC} = \frac{1}{R}\frac{dR}{dT}$$

(2.27)

T_0 is the reference temperature, and T the actual temperature. The temperature coefficient for some materials is given in Table 2.11.

2.2.3 Measuring Resistance

Measuring a resistance seems trivial as it requires the division of a voltage by a current. The fundamental problem is that measuring a current creates voltage drop and
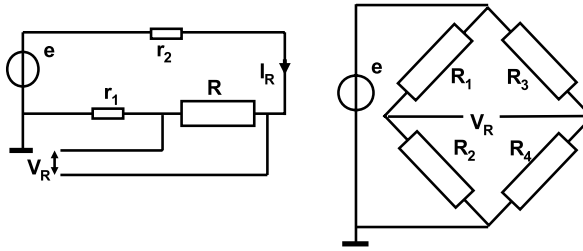
Fig. 2.15 Resistance measurement via a four-point method (*left*) and a Wheatstone bridge (*right*)

measuring voltage creates a current. Next to that all kind of parasitic resistances are present in the measurement loop. The measurement technique shown in Fig. 2.15 (left) supplies the current to the resistor via a path separate from the voltage measurement. The four-point technique or Kelvin measurement uses that fact that in practical situations a relatively small current is needed for a voltage measurement, while the additional resistances in the current loop may cause large measurement errors. This technique is applied in digital-to-analog converters based on resistor strings to avoid the influence of connection junctions.

The Wheatstone bridge is probably one of the first "circuits". It consists of four impedances and the output voltage equals:

$$V_R = e\left(\frac{R_2}{R_1 + R_2} - \frac{R_4}{R_3 + R_4}\right) \tag{2.28}$$

If one of the four elements is unknown, its value can be found by varying the other elements until $V_R = 0$ is reached. The unknown element is found because the ratios in both branches must be the same: $R_1 R_4 = R_2 R_4$. The elegance of this circuit certainly lies in the fact that the value ($\neq 0$) or form (AC/DC) of e is irrelevant at first glance. The circuit removes the DC offset and additional accuracy is obtained if V_R is amplified. This method is equally applicable to complex impedances and is very popular in sensor arrangements.

2.2.4 Electromigration

If a relatively large number of electrons flows through a small cross-section of a material, various side effects may occur. The electron flow in a conductor may become so strong that it displaces complete atoms. This effect is known as "electromigration". The displacement of an atom leaves less atoms to conduct the current. An exponential process starts up and removes more atoms resulting in a void. Electromigration is often modeled by the Black's equation also referred to as Black-Blech Equation [14]:

$$t_{50} = K_{em} J^{-n} e^{(E_a/kT)} \tag{2.29}$$

where:

$t_{50} =$ the median lifetime of the population of metal lines subjected to electromigration;

$K_{em} =$ a constant based on metal line properties such as bends and step coverage;

$J =$ the current density;

$n =$ the current exponent, many experts believe that $n = 2$;

$T =$ absolute temperature in K;

$k =$ the Boltzmann constant: $1.38 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$;

$E_a = 0.6\text{–}0.7 \text{ eV}$ for Aluminum and 0.9 for Copper.

With an activation energy of 0.7 eV a temperature shift from 100 °C to 113 °C will double the exponential term. A safe value for the current density through a cross section at 125 °C in Aluminum is 10^6 A/cm^2. With a wire thickness between 0.25 μm at the lowest wiring level and 1 μm at the top level, the DC current should be lower than 0.75 mA respectively 3 mA for every μm effective wire width. Higher margins are normally allowed for alternating and pulsed currents. Copper interconnects allow similar current densities with evidence of electromigration problems occurring close to via constructions [15].

2.2.5 Noise

In electronic devices the flow of current is associated with noise [16]. Only capacitors and ideal inductances are free of generating noise. The most common form of noise is thermal noise due to the Brownian motion of the charge carriers in a conductor. Thermal noise in a frequency band df is described by the thermal noise density S_{vv}:

$$\text{thermal noise density in 1 Hz:} \quad S_{vv}(f) = 4kTR$$

$$\text{or:} \quad S_v = \sqrt{4kTR} \; [\text{V}\sqrt{\text{Hz}}]$$

$$\text{rms noise voltage} \quad v_{\text{noise,rms}} = \sqrt{\int_{f=f_{\text{low}}}^{f=f_{\text{high}}} 4kTR \, df}$$

$$= \sqrt{4kTR(f_{\text{high}} - f_{\text{low}})} = \sqrt{4kTRBW} \; [\text{V}]$$

where k is Boltzmann's constant, T the absolute temperature and BW the bandwidth of interest. Thermal noise is assumed to have a "white noise" or flat frequency distribution: its magnitude is constant for all frequencies. The amplitude distribution is Gaussian.

A 1 kΩ resistor will give an effective noise voltage density of 4 nV$\sqrt{\text{Hz}}$ at room temperature and 1 MΩ results in 126 nV$\sqrt{\text{Hz}}$. In an audio bandwidth (20 Hz–20 kHz) the total rms noise voltage over 1 MΩ would sum up to 18 μV. Thermal noise is modeled by a voltage source in series or a current source in parallel to the resistor, Fig. 2.16. These two models are equivalent following the Norton theorem:

$$v_{\text{noise,rms}} = \sqrt{4kTRBW} \; [\text{V}]$$
$$i_{\text{noise,rms}} = \sqrt{4kTBW/R} \; [\text{A}] \tag{2.30}$$
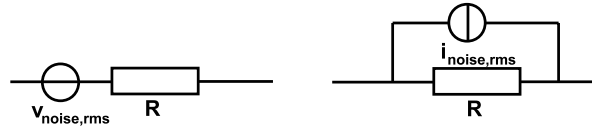
"$1/f$" noise is also called "Flicker noise" or "pink noise" and appears in many forms in nature ranging from semiconductor physics to the flooding of the river Nile. The origin or origins of flicker noise are disputed. The most common explanation assumes that carriers are trapped and released in charge traps, like interface states or oxide defects ("Mc Worther model"). An alternative by Hooge proposes mobility fluctuations as a contributor. Each carrier trap for $1/f$ noise can be modeled as random telegraph noise: a pulse train randomly switching between "+1" and "−1". The autocorrelation function of such a sequence has the form $R(\tau) = 1 - \alpha|\tau|$ and a spectral density function $S(f) \propto 1/f^2$. The location of the fluctuation or trap center with respect to the charge flow will influence the auto-correlation time constant. Combining the effect of a multitude of these traps with a uniform distribution of auto-correlation time constants requires the integration over the frequency domain, which will reduce the exponent of the frequency term to $S \propto 1/f$. Both flicker-noise mechanisms predict a relation with the area of the device. As $1/f$ noise is generated from carrier fluctuations, small size devices will show strong fluctuation in noise density due to the relatively large impact of random fluctuations.

The power spectral density and the amplitude density of the noise voltage source are given as:

$$S_{vv} = \frac{K_{1/f}}{\text{area} \times f}$$

$$S_v(f) = \sqrt{\frac{K_{1/f}}{\text{area} \times f}} \ [\text{V}\sqrt{\text{Hz}}]$$

(2.31)

where $K_{1/f}$ is the characteristic coefficient. S_{vv} is the power noise density and has as dimension Volt^2/Hz. Often S_{vv} is given at a certain frequency (e.g. 1 kHz), in a certain bandwidth (e.g. 1 Hz) and for a 1 μm^2 area, see e.g. Table 11.1. In order to estimate the impact of the noise on a system the above power density noise term must be integrated over the relevant frequency span.

An intriguing question is on the apparent explosion of "$1/f$" to infinity at zero frequency. Some authors consider the utilization of the electronic equipment as a lower frequency limit. As an example for the lower limit, the daily rhythm can be chosen with a frequency $f = 10^{-5}$ Hz. A fundamental property of $1/f$ noise is that the energy is equal in every decade of spectrum. This means that $1/f$ noise energy of a signal up to 1 GHz consists of 14 decades of which the 8–9 high-frequency decades or some 70% of the energy is of direct impact to a user. Stretching the argument to a lower boundary 30 years still means that 50% of the energy is of direct relevance to the user.

A form of noise common in tubes and in semiconductor devices with currents that are so low that the transition of individual electrons play a role, is shot noise. Its noise density is given by

$$i_{noise} = \sqrt{2q\,I_{bias}} \ [A\sqrt{Hz}] \tag{2.32}$$

Shot noise may appear in e.g. leakage currents in MOS gates.

"White noise" has a flat frequency spectrum, the noise density is constant for every frequency. Thermal noise is "white". The term "white noise" does not indicate what the amplitude distribution of the noise is. Thermal noise and shot noise are examples of white noise with a Gaussian amplitude distribution. Also $1/f$ noise shows a Gaussian amplitude distribution. All these noise effects are composed of the summation of many individual events. The Central Limit Theorem dictates that the sum of a large number of events becomes Gaussian distributed. Quantization errors in analog-to-digital conversion (Chap. 5) are basically distortion products that are modeled by a uniform distribution in the amplitude domain and a white noise spectrum. Random telegraph noise that is the basis for $1/f$ noise has typically two peaks in its amplitude distribution.

Noise is often characterized in systems in its relation to signals, as:

$$SNR = 10^{10}\log\left(\frac{\text{Signal power}}{\text{noise power}}\right) = 20^{10}\log\left(\frac{V_{signal,rms}}{V_{noise,rms}}\right) \tag{2.33}$$

In this equation the signal energy comes often from a single sine wave at one frequency. The noise has to be specified in a bandwidth. Therefore the SNR will depend on the signal strength, the spectral noise density and the relevant bandwidth.

2.3 Maxwell Equations

Electrostatic and electromagnetic events are described by Maxwell equations. These reflect the state of the art in electro-magnetic field theory in 1864. The Maxwell equations are the fundament for describing electronic elements and semiconductor devices, see e.g. [17].

Electromagnetic fields are represented by vectors.[7] $\vec{D}(x, y, z)$ is the electric displacement vector in Coulomb/m^2 at any point in space, which is related to the electric field vector $\vec{E}(x, y, z)$ as: $\vec{D} = \varepsilon_r\varepsilon_0\vec{E}$ in linear materials with components in the x, y, z directions $\vec{E} = [E_x, E_y, E_z]$. The surface A is represented by its normal vector $\vec{A}$.

The symbols $\nabla\cdot$ and $\nabla\times$ are standard vector operations. $\nabla\cdot$ is the gradient or divergence operator:

$$\nabla \cdot \vec{E}(x, y, z) = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} \tag{2.34}$$

---

[7]In this section the formal vector notation is applied. In the one-dimensional case the vector relation between the variables is obvious and the vector notation is omitted.

Its result is a single-valued function or a scalar proportional to the change of the vector. The rotation or curl operator $\nabla \times$ is defined on a vector $\vec{E}$ as:

$$\nabla \times \vec{E}(x, y, z) = \left[\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}, \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}, \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \right] \qquad (2.35)$$

In contrast to the divergence operator, the rotation operation results in a vector.

The first Maxwell equation is also known as Gauss law:

$$\nabla \cdot \vec{D} = \varepsilon_r \varepsilon_0 \nabla \cdot \vec{E} = \rho, \qquad \oint \varepsilon_r \varepsilon_0 \vec{E} \cdot d\vec{A} = \underset{\text{volume}}{\iiint} \rho(x, y, z) \, dx \, dy \, dz$$

$$(2.36)$$

A is an arbitrarily chosen surface fully surrounding the enclosed volume with charge density ρ. The dot $\cdot$ operator is the inner product. The inner product between E and A results in the net outgoing electrical field through that surface. Gauss law states that the total field passing perpendicularly through a surface is proportional to the enclosed charge.

Gauss law is a time-independent relation between the charge and its field.[8]

The electrical permittivity in vacuum ε_0 is multiplied with the relative permittivity ε_r. The permittivity in vacuum is $\varepsilon_0 = 8.8542 \times 10^{-14}$ F/cm $= 8.8542 \times 10^{-12}$ F/m, and the relative permittivity (also known as dielectric constant) ε_r varies between 1 and 20. Another notation uses the term electrical susceptibility χ_e, where:

$$\chi_e = \epsilon_r - 1$$

Example Suppose a point charge Q_p surrounded by a sphere with a radius r. The field $\vec{E}(x, y, z)$ is a vector formed by the line between the point charge and the location where the filed value E_p is measured. This vector has the same orientation as the normal vector to the local sphere area, reducing the vector operation into a simple multiplication.[9] The field strength E_p at distance r is found with (2.36) as:

$$E_p = \frac{Q_p}{4\pi r^2 \varepsilon_r \varepsilon_0}$$

This field will exercise a force F_{12} on a second point charge Q_s at a distance r_{12} of:

$$F_{12} = \frac{Q_p Q_s}{4\pi r_{12}^2 \varepsilon_r \varepsilon_0} \qquad (2.37)$$

This formula is known as "Coulomb's law". Gauss law and Coulomb's law are equivalent, one can be derived from the other.

---

[8]This formulation allows infinitely fast signaling, because the field will immediately disappear at any point in space if the charge is switched-off. Einstein's relativity theory does not permit this mode. The answer to this theoretical problem lies outside the scope of this book and is circumvented by applying Gauss law only in steady-state situations.

[9]A lot of vector manipulation consists of choosing a smart path or surface where only identical orientations or perpendicular orientations occur.

Closely related to Gauss law is the Poisson equation, which introduces the potential:

$$\nabla^2 \cdot V = -\nabla \cdot \vec{E} = \frac{-\rho}{\varepsilon_r \varepsilon_0} \tag{2.38}$$

The potential is the integral of the field. The potential difference in an electrical field associated with a point charge, is:

$$V(r_2) - V(r_1) = \int_{r=r1}^{r2} -E(r)\,dr = \frac{Q}{4\pi\varepsilon_r\varepsilon_0}\left(\frac{1}{r_2} - \frac{1}{r_1}\right) \tag{2.39}$$

The second Maxwell equation resembles mathematically the Gauss law but is defined for magnetics:

$$\nabla \cdot \vec{B} = 0, \qquad \oint \vec{B} \cdot d\vec{A} = 0 \tag{2.40}$$

$\vec{B}$ is the vector of the magnetic flux density in Tesla or $V\,s/m^2 = N/Am$ (Newton per Ampere-meter). The integral can be understood as the net flux passing through any closed surface (like the surface of a sphere or a cube) is zero. This law basically excludes the existence of magnetic monopoles, as known in the charge domain. In linear materials the magnetic flux density $\vec{B}$ is related to the magnetic field strength $\vec{H}$ as: $\vec{B} = \mu_r\mu_0\vec{H}$. The magnetic permeability in vacuum $\mu_0 = 4\pi \times 10^{-7}$ H/m or N/A^{-2} is multiplied by the relative permeability μ_r, which is specific for each material. An other notation uses the term magnetic susceptibility χ_m, where:

$$\chi_m = \mu_r - 1$$

Most materials in semiconductor have a relative permeability close to unity (Si: μ_r=0.9999963, Al: 1.000022, Cu: 0.9999938, O: 1.000019). Nickel, Iron and Ferrites can reach values of $\mu_r = 600\ldots10\,000$.

Maxwell's third law is Faraday's law of induction:

$$\nabla \times \vec{E} = -\frac{d\vec{B}}{dt}, \qquad \oint \vec{E} \cdot d\vec{s} = -\frac{d\Phi}{dt} \tag{2.41}$$

The rotation in the electrical field $\vec{E}$ equals the change in the flux. This flux equals the integral of the flux density $\vec{B}$ over the area $\vec{A}$ that is surrounded by the path $\vec{s}$. In integral notation the summation over an electrical gradient (resulting in a potential difference) equals the time change of the surrounded magnetic field.

Finally the fourth equation is known as Ampere's law:

$$\nabla \times \vec{B} = \mu_r\mu_0\vec{J} + \varepsilon_r\varepsilon_0\mu_r\mu_0\frac{d\vec{E}}{dt}$$
$$\oint \vec{B} \cdot d\vec{s} = \mu_r\mu_0\vec{I} + \varepsilon_r\varepsilon_0\mu_r\mu_0\frac{d}{dt}\oint \vec{E} \cdot d\vec{A} \tag{2.42}$$

Just as Coulomb's law for electric fields is related to Gauss law, the Biot-Savart law is an equivalent description for Ampere's law:

$$d\vec{B} = \frac{\mu_0}{4\pi r^2}\vec{I} \times \vec{r}_u\,dy \tag{2.43}$$

$\vec{B}(x, y, z)$ is a vector of the magnetic field due to a current I in a small portion of wire dy. $\vec{r}_u$ is the unit-length vector pointing from the wire fraction dy to the point where the field $\vec{B}$ is measured. $\vec{I} \times \vec{r}_u$ is the outer product between the vector $\vec{r}_u$ and the direction of the current flow. The $\times$ operator denotes the outer product and its result is a vector perpendicular to the plane formed by the two vectors $\vec{I}, \vec{r}_u$. Intuitively $I \, dy$ can be interpreted as a magnetic charge and dB is the resulting magnetic field on a sphere with radius r.

The four Maxwell's laws reduce in vacuum to the Heavyside formulation:

$$
\begin{aligned}
\nabla \cdot \vec{E} &= 0 \\
\nabla \cdot \vec{B} &= 0 \\
\nabla \times \vec{E} &= -\frac{d\vec{B}}{dt} \\
\nabla \times \vec{B} &= \varepsilon_0 \mu_0 \frac{d\vec{E}}{dt}
\end{aligned}
\tag{2.44}
$$

where $\vec{E}$ and $\vec{B}$ are functions of time t and space x, y, z. Substitution leads to:

$$
\varepsilon_0 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} - \nabla^2 \vec{E} = 0
$$

$$
\varepsilon_0 \mu_0 \frac{\partial^2 \vec{B}}{\partial t^2} - \nabla^2 \vec{B} = 0
$$

The combination of the magnetic $\vec{B}(t, x, y, z)$ field and the electric $\vec{E}(t, x, y, z)$ field define the electro-magnetic wave propagation. The solutions for $\vec{B}$ and $\vec{E}$ differential equations are based on sinusoidal functions.

$$
E(x, t) = E_0 \cos(\omega t - 2\pi x / \lambda)
$$

in one dimension, where $\omega = 2\pi f$ and $f\lambda = c$.

Electro-magnetic waves travel with the velocity of light c, which equals 2.99792×10^8 m/s for vacuum. The velocity of electro-magnetic waves in a material c_m is related to its permittivity and permeability via:

$$
c_m = \frac{1}{\sqrt{\varepsilon_r \varepsilon_0 \mu_r \mu_0}} = \frac{c}{\sqrt{\varepsilon_r \mu_r}}
$$

Unlike in circuit design the Maxwell's equations couple the magnetic and electric field also in another way. The magnetic field and the electrical field define a characteristic impedance:

$$
Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} = 377 \, [\Omega]
$$

This impedance is important as it defines the characteristic impedance of free space or the relation between voltage and current that an antenna (after correction for the transformation function) requires to get an optimum power transfer to the ether.

Fig. 2.17 Definitions in a coil
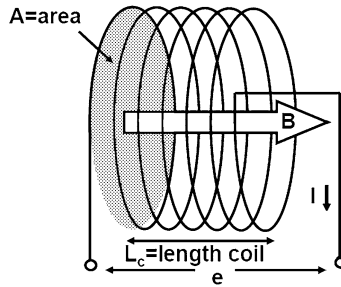
2.3.1 Inductors

The magnetic field in a coil in Fig. 2.17, is generated by the current flowing through the conductor. Using the fourth Maxwell equation, the magnetic flux density B is derived. As a path for ds the center line through the coil is chosen. The path is continued outside the coil in a direction perpendicular to the flux and than closed via a route with negligible flux. Only the path through the coil contributes to the integral. This path encircles the number of windings N_w times the current in the coil.

$$\oint \vec{B} \cdot d\vec{s} = B L_c = \mu_0 N_w I$$

The integral is unequal to zero only inside the coil, where the vectors align, so the vector notation is omitted and B simply equals the value of the field inside the coil.

$$B = \frac{\mu_0 N_w I}{L_c}$$

The total enclosed flux Φ is the product of the (constant) field B times the area A. The "inductance" is expressed in Henry:

$$L = \frac{N_w \Phi}{I} = \frac{N_w B A}{I} = \frac{\mu_0 N_w^2 A}{L_c} \tag{2.45}$$

with the magnetic permeability in vacuum $\mu_0 = 4\pi \, 10^{-7}$ H/m.

Faraday's law defines the electromagnetic induction or electromagnetic force[10] "emf" e in a conductor that encloses a magnetic flux (Φ) as:

$$e(t) = -\frac{d\Phi(t)}{dt} \tag{2.46}$$

---

[10]In physics an emf is a force that produces a current in a load. The term "voltage" is used for potential differences. Amongst electronics engineers both symbols "e" and "v" can denote a voltage source. This book follows the general practice, e.g. in MOS electronic circuits V_{DD} indicates the power supply.

The minus sign indicates that, if this electromagnetic force causes a current, its resulting magnetic field will oppose the generating field (Lenz's law). In the coil of Fig. 2.17 the number of windings N_w multiplies this potential to yield:

$$|e(t)| = N_w \frac{d\Phi(t)}{dt} = A N_w \frac{dB(t)}{dt} = \frac{A\mu_0 N_w^2}{L_c} \frac{dI(t)}{dt} = L \frac{dI(t)}{dt}$$

A sinusoidal current $i(t) = i_0 \sin(2\pi f t) = i_0 \sin(\omega t)$ results in a voltage over the coil as: $v(t) = L i_0 \cos(2\pi f t)$. In the frequency domain this relation is denoted as:

$$\frac{v(\omega)}{i(\omega)} = j\omega L$$

where the "j" operator indicates the 90° phase shift.

2.3.2 Energy in a Coil

The energy stored in a coil is:

$$E_L = \int_{t=0}^{\infty} e(t) I(t) \, dt = \int_{t=0}^{\infty} I(t) \frac{dLI(t)}{dt} \, dt = \int_{I=0}^{I_L} LI \, dI = 0.5 L I_L^2 \quad (2.47)$$

This equation assumes that the inductance remains constant. In case the inductance is a function of time, the overall energy equation becomes more complex.

2.3.3 Straight Wire Inductance

The inductance of a single wire is an uneasy theoretical problem as the inductance is only defined in a current loop[11] [18]. In integrated circuits this problem needs to be addressed for e.g. bond wire inductance. A common approximation considers a portion L_w of a wire.

Figure 2.18 shows a line of length L_w carrying a current I in the direction of the y-axis. The magnetic field B due to a current in wire length dy at position y on a distance x from the middle of the wire $x = 0$ is given by the Biot-Savart law:

$$dB(x, y) = \frac{\mu_0}{4\pi r^2} I \times r_u \, dy = \frac{\mu_0}{4\pi (x^2 + (y - y_B)^2)} I \sin(\theta) \, dy$$

$$= \frac{\mu_0 I}{4\pi} \frac{x \, dy}{(x^2 + (y - y_B)^2)^{3/2}}$$

The magnetic field due to a current I through a wire stretching from $y = -L_w/2$ to $y = L_w/2$ is found by integrating the section dy over $y = -L_w/2 \ldots L_w/2$ [7, Sect. 3], [11, Sect. A]:

---

[11]E.B. Rosa published an extensive paper in 1908, using the centimeter-gram-second (CGS) system for the symbols. This derivation follows his way of working in SI units: meter-kilogram-second.

Fig. 2.18 Definitions of
magnetic field along a wire

For an infinite long wire the field reduces to:
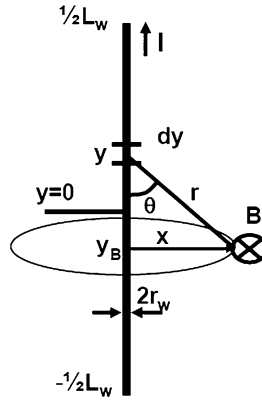
$$B(x) = \frac{\mu_0 I}{2\pi x}$$

$$B(x) = \frac{\mu_0 I x}{4\pi} \int_{y=-L_w/2}^{y=L_w/2} \frac{dy}{(x^2 + (y - y_B)^2)^{3/2}} = \frac{\mu_0 I x}{4\pi} \left. \frac{y - y_B}{x^2 \sqrt{x^2 + (y - y_B)^2}} \right|_{y=-L_w/2}^{y=L_w/2}$$

$$= \frac{\mu_0 I}{4x\pi} \left(\frac{y_B + L_w/2}{\sqrt{x^2 + (y_B + L_w/2)^2}} - \frac{y_B - L_w/2}{\sqrt{x^2 + (y_B - L_w/2)^2}} \right)$$

For an infinite long wire the field reduces to:

$$B(x) = \frac{\mu_0 I}{2\pi x}$$

which is more easily found by using Ampere's law and equating the constant B field
on a circle with radius x that encircles a current I: $2\pi x B = \mu_0 I$.

   The total magnetic flux Φ is found by integrating the magnetic field over the
surface the magnetic field penetrates. This surface is the rectangle formed by $y_B = -L_w/2 \ldots L_w/2$ and x from the boundary of the wire $x = r_w$ to infinity. The flux
integral is formed by multiplying this area with the magnetic field and integrating x
to infinity:

$$\Phi = \frac{\mu_0 I}{4\pi} \int_{x=r_w}^{x=\infty} \frac{1}{x} \int_{y_B=-L_w/2}^{y_B=L_w/2} \left(\frac{y_B + L_w/2}{\sqrt{x^2 + (y_B + L_w/2)^2}} \right.$$

$$\left. - \frac{y_B - L_w/2}{\sqrt{x^2 + (y_B - L_w/2)^2}} \right) dy_B \, dx$$

$$= \frac{\mu_0 I}{4\pi} \int_{x=r_w}^{x=\infty} \frac{2(\sqrt{x^2 + L_w^2} - x)}{x} \, dx$$

$$= \frac{\mu_0 I}{2\pi} \left[\sqrt{x^2 + L_w^2} - L_w \ln\left(\frac{L_w + \sqrt{L_w^2 + x^2}}{x} \right) - x \right]_{x=r_w}^{x=\infty}$$

$$= \frac{\mu_0 I}{2\pi} \left[L_w \ln\left(\frac{L_w + \sqrt{L_w^2 + r_w^2}}{r_w} \right) + r_w - \sqrt{L_w^2 + r_w^2} \right]$$

With $L_w \gg r_w$ and the inductance of the wire equal to the ratio between the flux and the current, the straight-wire inductance is found:

$$L = \frac{\mu_0 L_w}{2\pi} \left[\ln\left(\frac{2L_w}{r_w}\right) - 1 \right] \text{[H]}$$

with L_w and r_w as the length and radius of the wire in meter. Assuming[12] that these two will have a ratio of a few hundred, a rough approximation for the inductance of a wire is 1 nH/mm.

2.3.4 Skin Effect and Eddy Current

At higher frequencies conductors will no longer behave as uniform conduction paths: the laws of electromagnetism force the high frequency currents to run at the perimeter of the conductor. This effect is known as the "skin" effect. The current density decays exponentially from the surface with a coefficient δ called the skin depth:

$$\delta = \frac{1}{\sqrt{\pi \sigma f \mu}} \tag{2.48}$$

where $\sigma = 1/\rho$ is the conductivity of the conductor, f the frequency and μ the permeability. For 50 Hz the skin depth is in the order of 1 cm, at 10 MHz the skin depth is 20 μm, and for a 1 GHz signal this effect results in a skin depth of roughly 2 μm for copper and aluminum. A typical 33 μm diameter bond wire presents a four times higher resistance at this frequency.

Another phenomena associated with magnetic fields is the "eddy-current". A changing or moving magnetic field induces currents in conductive layers it penetrates. The transformer is based on this principle as well as the permanent-magnetic motor and dynamo. For normal coils the eddy currents show up as a resistive component in the total coil impedance and are considered a parasitic lossy element.

2.3.5 Transformer

In a transformer two or more inductors share their magnetic fields. Via this shared field energy can flow from one pair of terminals to the other. The relation between the currents and voltages in an ideal transformer is:

$$\frac{V_p}{V_s} = \frac{I_s}{I_p} = \frac{\text{number of primary windings}}{\text{number of secondary windings}} = \frac{1}{N} \tag{2.49}$$

In a transformer the input side is often called the primary side and the output the secondary side. N equals the number of secondary windings divided by the number

---

[12]Implicitly the return path is estimated at a distance of a few hundred times the wire radius.
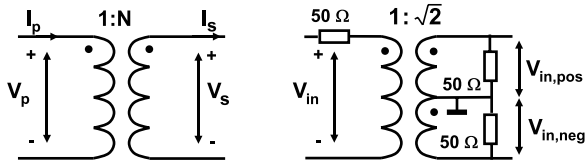
Fig. 2.19 Notation conventions in a transformer and a transformer used to create differential signals. The *dots* indicate the side of the coils with mutually in-phase signals

of primary windings. As voltages are multiplied by the transformer ratio N and currents are divided by N the impedance transformation is with N^2, $Z_s = N^2 Z_p$.

In the field of analog-to-digital conversion the transformer is mostly used for deriving differential signals from a single ended source. Figure 2.19 (right) shows a single-to-differential transformer. The ratio between primary and secondary side is chosen in a manner that preserves the 50 Ω impedance levels on both sides. The middle tap is shown at ground level however as a transformer does not couple any DC, another bias level is equally suited. RF transformers[13] are typically used for single-to-differential conversion in the frequency range between 1 MHz and 1 GHz.

The coupling of the primary to secondary side is not perfect. The coupling coefficient $0 \le k \le 1$ determines the contribution of the mutual inductance:

$$M = k\sqrt{L_p L_s} \tag{2.50}$$

M is the net useful inductance for the transformation of the coils L_p and L_s. In good transformers k is close to 1.

The coupling coefficient is hopefully much less than 1 in the second appearance of transformers in converters. Every wire is in fact a single winding of a transformer. Wires that run in parallel at a close distance form parasitic transformers. Bond wires that run from a package to bondpads are a notorious example of parasitic transformers. A typical coupling factor of two neighboring wires is $k = 0.3$–0.5. The effect of this factor is:

$$V_{\text{wire}} = L_p \frac{I_p}{dt} - M \frac{I_s}{dt} \approx L_p \left(\frac{I_p}{dt} - k \frac{I_s}{dt} \right)$$

A current spike will cause a voltage drop over a wire and half of that voltage over the neighboring wire. It is therefore important to check carefully the neighboring signals of sensitive inputs.

---

[13]Many manufacturers offer RF transformers: e.g. Macom, Minicircuits, Coilcraft and Pulse.

Fig. 2.20 Parallel plate capacitor
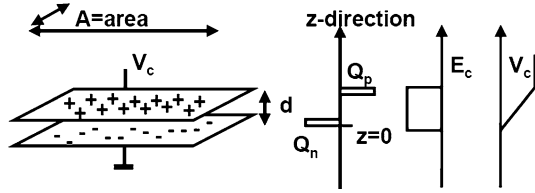
2.3.6 Capacitors

Capacitance is the amount of charge that can be stored in a device and is expressed in Farad (F).

$$C = \frac{dQ}{dV} \ [F]$$

which reduces in the linear situation to: $C = Q/V$.

In Fig. 2.20 two parallel plates have been charged with a charge Q. This charge on the top plate induces an opposite charge on the bottom plate. Using Gauss equation in (2.36) the integral over the volume formed by the lower plate area A and the z coordinate is taken. The electrical field E is perpendicular to the area A, which reduces the inner-product to a simple multiplication, so:

$$E = \iiint\limits_{\text{volume}} \frac{\rho}{A\varepsilon_r\varepsilon_0}\,dx\,dy\,dz = \begin{cases} 0, & z < 0 \\ \dfrac{Q}{\varepsilon_r\varepsilon_0 A}, & 0 < z < d \\ 0, & z > d \end{cases}$$

The electrical field outside the plates is zero, when passing the lower plate in the vertical z-direction, the charge on the bottom plate is included. This charge is seen as an infinitely thin layer and builds up a constant electrical field between the plates. The field reduces to zero when passing the upper plate, which contains an equal but opposite charge quantity. Now the potential can be derived by integrating the field using Poisson's equation in the z-direction:

$$V = \int -E\,dz = \begin{cases} 0, & z < 0 \\ \dfrac{Qz}{A\varepsilon_r\varepsilon_0}, & 0 < z < d \\ \dfrac{Qd}{A\varepsilon_r\varepsilon_0}, & z > d \end{cases}$$

resulting in a relation between the plate potential V and the charge Q. Now the capacitance of this plate capacitor is:

$$C = \frac{Q}{V} = \frac{A\varepsilon_r\varepsilon_0}{d} \tag{2.51}$$

The capacitance is proportional to the area and inversely proportional to the distance between the plates. The current is the derivative with time of the charge:

$$I(t) = \frac{dQ(t)}{dt} = \frac{dCV(t)}{dt} = C\frac{dV(t)}{dt} + V\frac{dC(t)}{dt} \tag{2.52}$$
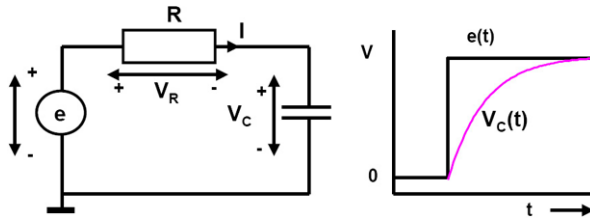
Fig. 2.21 Charging a capacitor

For time-invariant capacitors only the first term remains.

With a sinusoidal voltage $v(t) = v_0 \sin(2\pi f t) = v_0 \sin(\omega t)$ the current through is given as $i(t) = Cv_0 \cos(2\pi f t)$. In the frequency domain this relation is denoted as:

$$\frac{v(\omega)}{i(\omega)} = \frac{1}{j\omega C} = \frac{-j}{\omega C}$$

The 90° phase shift of the voltage-to-current relation in the coil is opposite to the phase shift of the voltage-to-current relation of the capacitor.

2.3.7 Energy in Capacitors

If a capacitor is charged from a voltage source with a step-shaped pulse of amplitude V as in Fig. 2.21, the voltage over the capacitor can be calculated as follows:

$$I(t) = C\frac{dV_C(t)}{dt} \quad \text{and} \quad I(t) = \frac{V - V_C(t)}{R} \quad \text{result in}$$
$$V_C(t) = V(1 - e^{-t/RC}) \tag{2.53}$$
$$I(t) = \frac{V}{R}e^{-t/RC}$$

The energy that is delivered by the voltage source $e(t)$ and that is consumed in the resistor and stored in the capacitor is:

$$E_e = \int_{t=0}^{\infty} e(t)I(t)\,dt = CV^2$$
$$E_R = \int_{t=0}^{\infty} V_R(t)I(t)\,dt = 0.5CV^2 \tag{2.54}$$
$$E_C = \int_{t=0}^{\infty} V_C(t)I(t)\,dt = 0.5CV^2$$

Half of the energy that has been delivered, is still present and stored in the capacitor while the other half is consumed in the series resistance. The size of the resistance determines the timing but not the outcome of the energy transfer.
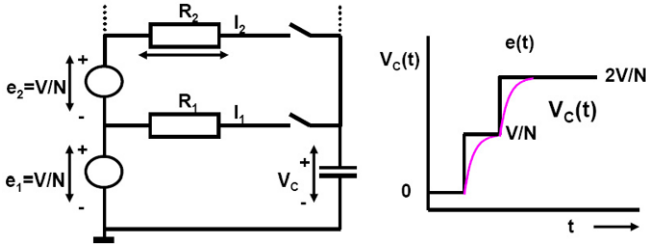
Fig. 2.22 Charging a capacitor from multiple sources

2.3.8 Partial Charging

The previously found energy equations are not easy to circumvent. One method is
the use of multiple supplies as shown in Fig. 2.22. Suppose that the original potential
source e is split in N equal sources, which will load sequentially the capacitor. In
the first step the above equations hold

$$E_{e1} = CV^2/N^2$$
$$E_{R1} = 0.5CV^2/N^2$$
$$E_C = 0.5CV^2/N^2$$

In the next cycle the second source will deliver the same energy and the second
resistor will consume the same energy as in the first cycle. However, the first source
has to deliver energy as the charging current also flows through this source. So the
total energy delivered, dissipated and stored is:

$$E_{e1} = CV^2/N^2 + CV^2/N^2$$
$$E_{e2} = CV^2/N^2$$
$$E_{R1} = 0.5CV^2/N^2$$
$$E_{R2} = 0.5CV^2/N^2$$
$$E_C = 0.5C(2V/N)^2$$

After the second cycle the delivered energy is $3CV^2/N^2$ of which CV^2/N^2 is burnt
in the resistors and $2CV^2/N^2$ stored in the capacitor. If this procedure is extended
for N cycles, the energy balance is:

$$E_{e1...eN} = \frac{(N+1)}{2N} CV^2$$
$$E_{R1...RN} = \frac{1}{2N} CV^2$$
$$E_C = \frac{1}{2} CV^2$$

For large N this method allows to reach a nearly lossless charging of the capacitor.
The energy analysis excludes the measures that have to be taken to create the volt-
ages and control the timing. In audio processing a similar method for transferring

Table 2.13 An indication of the active digital power consumption

Core	Power mW per MHz	Notes/source
Cortex M3	0.3	0.18 μm CMOS, 1.8 V, ARM
1 mm^2 standard logic	0.2	C90 test, NXP

efficiently energy to a load is called "class-G". Also some charge based analog-to-digital converters use this basic idea.

2.3.9 Digital Power Consumption

The energy needed for charging a capacitive load in digital circuitry is equal to CV^2. Half of this energy is consumed during the rising edge on the capacitive load, while the other half is burnt during the discharge. If there are N_d nodes in a digital circuit that are charged and every node has an average capacitance of C_d, while the switching frequency is f_d, then the energy per second or the active power is:

$$P_d = f_d N_d C_d V^2 \tag{2.55}$$

In digital circuitry not all nodes switch at every clock cycle. The "activity" factor α denotes the fraction of transitions that occur per clock cycle. If every next clock cycle the data changes polarity the activity is 1. Depending on the type of circuit the practical activity factor is between 0.1 and 0.5. An extreme case is a ripple-carry adder, where due to the rippling of data multiple transitions can occur on one clock edge. Locally the activity factor can exceed 1. f_d now equals the product of the activity and the clock frequency: αf_c.

Using the linear relation between the digital power consumption and the frequency some general rule-of-thumb numbers can be derived from literature, Table 2.13. Next to the active power the leakage power becomes a major contribution in digital design. This power stems from the remaining conduction of MOS transistors. Some more remarks can be found in Fig. 2.38 and in Table 11.1.

2.3.10 Coaxial Cable

In high-speed integrated circuits (flash analog-to-digital converters) on chip structures are used for transporting signals that bear resemblance to coaxial cables. Therefore as an example of the Maxwell laws the coaxial cable as shown in Fig. 2.23 is examined. The charge on the inner conductor with diameter d and length l is equal to Q. The electrical field on a distance r from the center line can now be found with Gauss law. A cylindrical surface with radius r and length L_c is chosen for determining the electrical field E. From symmetry the electrical field is normal to the

Fig. 2.23 Definitions for a coaxial cable
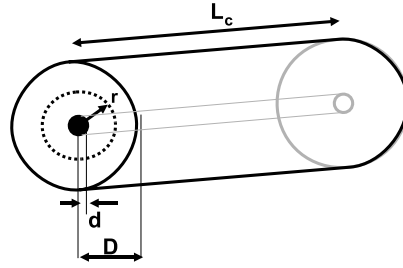
cylindrical surface and constant at a distance r from the center line. On both head-ends the field is in the plane of the enclosing surface and the inner product is zero. With (2.36):

$$2\pi L_c r E = \frac{Q}{\varepsilon_r \varepsilon_0}$$

Integrating E from the inner conductor to the outer conductor gives the voltage:

$$V = \int_{r=d}^{r=D} E \, dr = \int_{r=d}^{r=D} \frac{Q}{2\pi L_c r \varepsilon_r \varepsilon_0} \, dr = \frac{Q \ln \frac{D}{d}}{2\pi L_c \varepsilon_r \varepsilon_0}$$

The capacitance of a unit length of coaxial cable is now:

$$C_{\text{unit}} = \frac{Q}{V} = \frac{2\pi \varepsilon_r \varepsilon_0}{\ln \frac{D}{d}} \ [\text{F/m}] \tag{2.56}$$

Referring to the coaxial cable in Fig. 2.23 the inductance of a coaxial cable per unit length carrying a current I can be derived:

$$\oint B \times ds = \mu_r \mu_0 I$$

Choosing a circular path around the conductor with a radius r and assuming from circular symmetry that B is constant along that perimeter:

$$B(r) = \frac{\mu_r \mu_0 I}{2\pi r}$$

The total magnetic flux in between both conductors is now:

$$\Phi = \int_{r=d}^{r=D} B(r) \, dr = \frac{\mu_r \mu_0 I \ln \frac{D}{d}}{2\pi}$$

When $\Phi = L \times I$ defines the inductance parameter L, this results in an inductance per unit length of:

$$L_{\text{unit}} = \frac{\Phi}{I} = \frac{\mu_r \mu_0 \ln \frac{D}{d}}{2\pi} \ [\text{H/m}] \tag{2.57}$$

Combing the inductance per unit length with the formula for the capacitance per unit length, the specific impedance ($\epsilon_r = \mu_r = 1$) is:

$$Z_0 = \sqrt{\frac{L}{C}} = \frac{\ln \frac{D}{d}}{2\pi} \sqrt{\frac{\mu_r \mu_0}{\epsilon_r \epsilon_0}} = 59.988 \times \ln \frac{D}{d} = 138 \times {}^{10}\log \frac{D}{d} \ [\Omega] \tag{2.58}$$

In electronics the signals propagate either through wires or conductors on a substrate. The associated electromagnetic wave travels also through the dielectric surrounding the conductor. From the specific capacitance and inductance per unit length the propagation speed is:

$$c_{\text{coax}} = \frac{1}{\sqrt{L_{\text{unit}} C_{\text{unit}}}} = \frac{c}{\sqrt{\varepsilon_r \mu_r}} \tag{2.59}$$

The properties of the dielectric determine the velocity of the wave. In a coaxial cable the conductors are separated by polyethylene ($\epsilon_r = 2.25$, $\mu_r = 1$), which will allow a velocity for the signals of 2/3 of the vacuum velocity. In silicon the conductor is surrounded with silicon dioxide ($\epsilon_r = 3.9$, $\mu_r = 1$) which halves the speed of electromagnetic waves.

2.4 Semiconductors

In a semiconductor both holes and electrons are available to conduct current. Their relation to the material constants is given by:

$$n \times p = n_i^2 = N_C N_V e^{-E_G/kT} \propto T^3 e^{-E_G/kT} \tag{2.60}$$

N_V and N_C are the effective density of states in the valence and conduction bands. The valence band is the highest energy level defined by quantum mechanics in which electrons are present at 0 K. The conduction band requires even more energy and is separated from the valence band by the bandgap energy. This energy difference between the valence and conduction bands, is a material property. For conductors the bandgap energy is so small that electron can be excited by thermal energy to go from the valence to the conduction band. In insulators the band gap prevents electrons to go to the conduction band. In silicon E_G equals 1.12 electronVolt (eV) at 0 Kelvin and in germanium the band gap is 0.67 eV. In the design of a bandgap reference circuit the bandgap of silicon is determined by extrapolating back the pn-junction voltage to 0 K. The "back-extrapolated" band-gap voltage is found at 1.205 eV.

The above equation describing the pn product also holds in case of doped materials. In n-doped material where electrons form the majority carriers the electron concentration equals: $n_n = N_D$. There is a remaining fraction holes, called the minority concentration: p_n. In p-doped material the majority concentration is: $p_p = N_A$. The minority carrier concentration n_p and p_n are:

$$p_n = \frac{n_i^2}{N_D} \qquad n_p = \frac{n_i^2}{N_A} \tag{2.61}$$

For thermal equilibrium the electron and hole concentration can be related to the intrinsic energy level as:

$$n = n_i e^{(E_F - E_i)/kT} \qquad p = n_i e^{(E_i - E_F)/kT} \tag{2.62}$$

where n and p are the electron and hole concentrations at the Fermi energy level E_F under equilibrium. If the electron concentration is increased by dopants, the Fermi level will deviate from the intrinsic level. In turn this will cause the hole concentration to be reduced by the same factor.

A loose definition of the Fermi level states that at this energy level half of the potential states are occupied by electrons. E_i is the intrinsic energy level halfway between the conduction and valence band with n_i as the corresponding electron concentration. k is the Boltzmann's constant and equals: 1.38×10^{-23} J/K. T is the absolute temperature with T Kelvin equal to $(T - 273.15)$ Celsius.[14]

The above mathematical description is derived from Maxwell Boltzmann statistics and is most fundamental to semiconductor physics. Nearly all occurrences of exponential relations in semiconductor descriptions are derived from these statistics. It is important to realize how this concept differs from Netwon's physics. In a basket full of apples the ones below the rim (low energy) stay in the basket, while the ones with more energy will roll out. At 0 K this picture applies too for a potential well filled with electrons. At higher temperatures the thermal energy will statistically spread over the electrons and a distribution according the above description is the result. Referred back to the apples: there will be empty spaces in the basket and some apples will jump in and out of the basket.

Some specific habits apply to the semiconductor world. Semiconductor engineers calculate in centimeters, thin layers are expressed in Ångström (Å) or nanometer (10 Å = 1 nm). In older U.S. literature the term "mil" is used. 1 mil equals 1/1000 inch or 25.4 μm.

2.4.1 Semiconductor Resistivity

The conductivity of silicon being a semiconductor, depends on the amount of free electrons and holes. The intrinsic electron and hole concentration at room temperature is $n_i = p_i = 1.4 \times 10^{10}$ cm^{-3}. More free electrons and holes are supplied by the dopant of the semiconductor. In silicon the elements Phosphorus (P), Arsenic (As) and Antimony (Sb) act as electron suppliers or donors. The resulting material is "n-type". Binding of electrons or equivalently supplying holes is performed by acceptors like Boron (B) or Aluminum (Al). If a concentration of N_D donor dopants are present in the material the excess electron concentration, called the majority carriers, is equal to $n = N_D$ (assuming $N_D \gg n_i$). Equivalently an acceptor dopant generates holes: $p = N_A$. The conductivity is proportional to the charge: $q \times n$ or $q \times p$ for n-type or p-type material, where n is the free electron concentration and p is the free hole concentration in cm^{-3}. The resistivity is the inverse of the conductivity:

$$\rho_{R,n} = \frac{1}{qn\mu_n} \qquad \rho_{R,p} = \frac{1}{qp\mu_p} \tag{2.63}$$

---

[14]Better forget about Fahrenheit, Réaumur, and Rankine conversion.

Table 2.14 Some electrostatic properties of semiconductors. The permittivity in vacuum $\varepsilon_0 = 8.85 \times 10^{-12}$ F/m [13]

Material	Relative permittivity ε_r	Bandgap (eV)	Breakdown field (V/µm)	n/p mobility cm^2/Vs
Silicon Si	11.7	1.1	30	1500/450
Silicon dioxide SiO$_2$	3.9	9	600	
Silicon nitride Si$_3$N$_4$	7.5	5	1000	
Germanium Ge	16.0	0.67	8	3900/1900
Gallium arsenide GaAs	13.1	1.42	35	8500/400
Gallium nitride GaN	9.5	3.36	330	380/
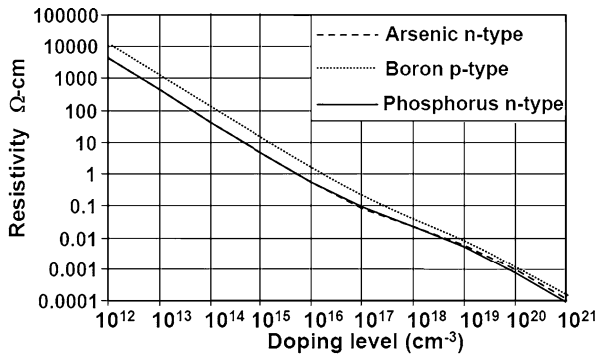
Fig. 2.24 Resistivity of doped silicon: *upper line* is boron (p-type), *lower two lines* represent phosphorus and arsenic (n-type)

q is the charge of an electron $q = 1.6 \times 10^{-19}$ Coulomb. The proportionality constants $\mu_n \approx 1500$ cm^2/V s and $\mu_p \approx 450$ cm^2/V s are the mobility for electrons and holes respectively, see Table 2.14. Mobilities tend to change due to temperature, pressure and doping levels. Moreover the mobility near the surface of devices can reduce to a fraction of the bulk value.

The relation between resistivity and doping level is plotted in Fig. 2.24 [13, p. 32].

2.4.2 Voltage and Temperature Coefficient

In doped silicon the temperature coefficient, see (2.27), is due to the temperature dependence of the mobility that is due to impurity scattering (obstruction by fixed charges) and lattice scattering (vibration of the crystal lattice also called: phonon interaction). Theoretically these two effects limit the temperature dependence between

Table 2.15 An indication of resistor characteristics in a semiconductor process (0.18 μm to 90 nm generations) from ITRS [19] and various publications

Material	Square resistance $\Omega/\square$	Voltage coeff. V^{-1}	Temp. coeff. K^{-1}	Matching A_R %μm
n/p diffusion	$75\ldots125$	1×10^{-3}	$1\ldots2 \times 10^{-3}$	0.5
n well diffusion	1000	80×10^{-3}	4×10^{-3}	
n-polysilicon	$50\ldots150$		$-1+\cdots+1 \times 10^{-3}$	2
p-polysilicon	$50\ldots150$		0.8×10^{-3}	2
Polysilicon (silicide)	$3\ldots5$		3×10^{-3}	
Aluminum	$0.03\ldots0.1$		3×10^{-3}	

$T^{1.5}$ and $T^{-1.5}$. Experimentally $\mu \propto T^{\alpha}$, with $\alpha = -2.0\ldots-2.5$. As a consequence the temperature coefficient of silicon is in the range:

$$\text{TC} = \frac{1}{\rho_R}\frac{d\rho_R}{dT} = \frac{-\alpha}{T} \approx 0.007\,^{\circ}\text{C}^{-1} \qquad (2.64)$$

A minor dependence with doping is present. In poly crystalline silicon the effects of grain boundaries play a role, strongly influencing the temperature coefficients in poly silicon [20].

It is obvious that resistors with voltage dependency affect the distortion behavior of an electronic circuit. These effects are modeled with (2.26). Table 2.15 lists voltage dependencies of integrated resistors.

2.4.3 Matching of Resistors

If multiple resistors of the same size are used, the mutual differences will be due to variations in the geometrical definitions and in the composition of the material. Also environmental effects, such as the proximity of other structures, contacts, heat sources and stress, will affect the value of each individual resistor. Special lay-out techniques as in Fig. 7.9 reduce these effects. Similar to MOS devices, see Sect. 11.4, the mismatch due to the granularity of the material is defined as:

$$\frac{\sigma_{\Delta R}}{R} = \frac{A_R}{\sqrt{\text{area}}} \qquad (2.65)$$

where A_R is expressed in %μm and area in μm^2.
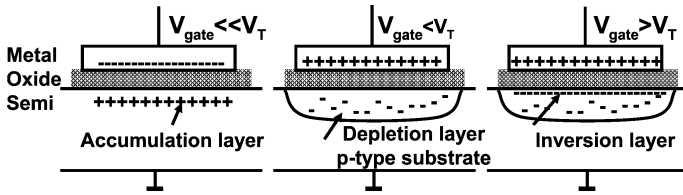
Fig. 2.25 The MOS capacitor: in accumulation, depletion and inversion
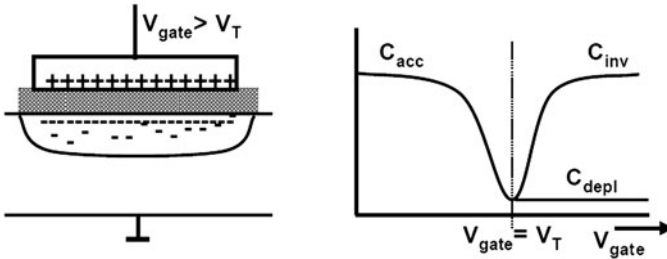
Fig. 2.26 MOS capacitance and *C-V* curve

2.4.4 MOS Capacitance

Figure 2.25 shows the construction of a Metal-Oxide-Semiconductor capacitor. In this example a p-type semiconductor substrate contains an excess of holes, which will be pulled towards the interface (accumulation) if the gate voltage is made negative with respect to the substrate. If the gate voltage is close to zero (the flat band voltage), the holes will be repelled from the interface and a layer without free charge carriers will be created: the depletion layer.

In this depletion layer only the negative ions from the p-type doping remain and form a space charge. After some time, remaining electrons will find their way into the depletion layer. The electrical field will pull them towards the interface, where the electrons form the inversion layer. This layer can be contacted by an adjacent n-type doped region.

If the capacitance of this structure is measured as a function of the applied gate voltage, the curve shows a dip in the region where depletion occurs, Fig. 2.26. If an inversion layer is formed and contact to ground is made, the inversion capacitance of this structure is basically identical to the capacitance in accumulation. Without such a contact, or if the measurement is carried out at high frequency, the capacitance remains at the depleted value.
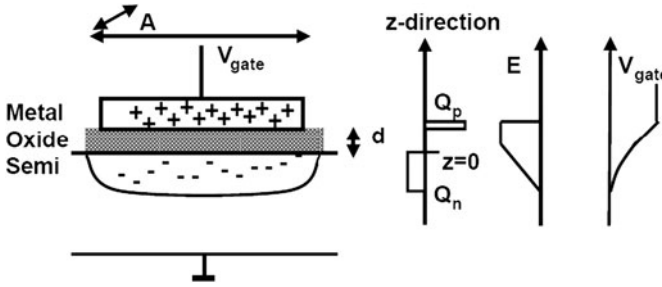
Fig. 2.27 The MOS capacitance consisting of a metal plate, an oxide insulating layer and a depleted region in the semiconductor

The structure in Fig. 2.27 contains a space charge instead of a sheet of charge. Assuming that the bulk is at the ground potential, the gate voltage is determined by using the Gauss equation (2.36) in the a similar way as in the parallel plate example:

$$
E(z) = \begin{cases} 0, & z < -z_d \\ \dfrac{qN(z+z_d)}{\varepsilon_r \varepsilon_0}, & -z_d < z < 0 \\ \dfrac{qNz_d}{\varepsilon_r \varepsilon_0}, & 0 < z < z_g \\ 0, & z > z_g \end{cases}
$$

A one-dimensional description is used, resulting in capacitance per unit area. N is the doping density per unit volume. In the depletion region (from $z = z_d$ to $z = 0$) the charge density ρ is equal to the dopant concentration N times the electron charge. While passing through this charge region, the electrical field increases. At the oxide interface the build-up of the electrical field stops and the electrical field remains constant as no net charge quantity is present. Only when reaching the gate plate, the opposing charge is met and the electrical field goes to zero.

Having determined the electrical field, a second integration step results in the gate voltage with respect to the substrate potential:

$$
V = \int -E \, dz = \begin{cases} 0, & z < -z_d \\ \dfrac{qN(z+z_d)^2}{2\varepsilon_r \varepsilon_0}, & -z_d < z < 0 \\ \dfrac{qNz_d^2}{2\varepsilon_r \varepsilon_0} + \dfrac{qNz}{\varepsilon_r \varepsilon_0}, & 0 < z < z_g \\ \dfrac{qNz_d^2}{2\varepsilon_r \varepsilon_0} + \dfrac{qNz_g}{\varepsilon_r \varepsilon_0}, & z > z_g \end{cases}
$$

The potential on the gate subdivides in a part over the depletion layer and a part over the gate-oxide. The voltage drop over the depletion layer V_d is proportional to the square of the depletion layer thickness.

$$z_d = \sqrt{\frac{2\varepsilon_r \varepsilon_0 V_d}{qN}}$$
$$Q_d = qNz_d = \sqrt{2qN\varepsilon_r \varepsilon_0 V_d} \tag{2.66}$$

Where Q_d is the depletion charge per unit area. Applying the parallel plate capacitance formula, the capacitance per unit area is:

$$C_d = \frac{\varepsilon_r \varepsilon_0}{z_d} = \sqrt{\frac{qN\varepsilon_r \varepsilon_0}{2V_d}} \tag{2.67}$$

This result is of course identical to the result obtained from differentiating the depletion charge for the depletion voltage. The voltage over the depletion region is not equal to the terminal voltage as various additional components are still involved, like the work-function differences and fixed oxide charges. The work-function ϕ_{MS} is the energy (in electron Volt) to extract an electron from a metal or semiconductor into the vacuum. As each material has a different work-function, a junction of two metals will result in a potential difference. At the flat-band voltage all energy levels in the semiconductor are flat, which means that there is no depletion region. Seen from the terminals this means that the flat-band voltage accumulates all additional components. So the depletion capacitance as measured from its terminal voltage V is:

$$C_d = \frac{\varepsilon_r \varepsilon_0}{z_d} = \sqrt{\frac{qN\varepsilon_r \varepsilon_0}{2(V - V_{FB})}}$$

The Mott-Schottky method in Fig. 2.28 rearranges the terms in the basic capacitance voltage relation to:

$$C_d^{-2} = \frac{2(V - V_{FB})}{qN\varepsilon_r \varepsilon_0} \tag{2.68}$$

From the intercept with the x-axis the flat band voltage can be derived, while the slope of the line is proportional to the dopant level in the substrate.
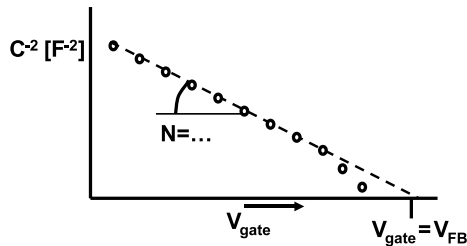
Fig. 2.28 The Mott-Schottky curve of a C-V measurement

2.4.5 Capacitance Between Layers

In a semiconductor process various capacitors are available through the construction
of the layers in the process, Fig. 2.14.

- Diffusion capacitances are based on a depleted semiconductor region, see
  Sect. 2.4.4 and (2.67). Semiconducting capacitances suffer from leakage and
  nonlinearities. They are asymmetrical in the sense that the properties change
  fundamentally if the voltage is reversed. Table 2.16 compares diffusion capaci-
  tors in two processes. Due to scaling the impact of edge capacitors in a design is
  more significant in an advanced process.
- The gate-to-channel capacitor, see Sect. 2.4.4, is formed by the MOS transistor
  gate plate, the dielectric and a conductive layer in the semiconductor material.
  This layer can be an inversion layer or an accumulation layer.

  In a 90-nm process the specific gate capacitance is around 12 fF/μm^2, see
  Table 2.17. This device requires a significant turn-on voltage in excess of the
  threshold voltage. Even at a significant turn-on voltage a slight non-linearity be-
  tween charge and voltage will remain. Using voltages of opposite polarity is not
  possible. Therefore this capacitor is mostly applied for decoupling supplies and
  bias lines. In older technologies the potential resistance of the poly-silicon gate
  must be taken into account.

Table 2.16 Diffusion capacitances in a 0.25 μm process and a 65 nm CMOS process. The numbers
are indicative for a node and can vary per foundry [19]

Area capacitance in fF/μm^2	0.25 μm	65 nm
n+ diffusion to substrate bottom	0.4	1.3
p+ diffusion to N-well bottom	0.6	1.1
N-well to substrate bottom	0.2	0.14
Edge capacitance in fF/μm	0.25 μm	65 nm
n+ diffusion to substrate thick-oxide edge	0.4	0.08
p+ diffusion to N-well thick-oxide edge	0.6	0.07
N-well to substrate thick-oxide edge	0.4	0.7

Table 2.17 The gate
capacitance in various
processes. The numbers are
indicative for a node and can
vary per foundry (e.g. ITRS
[19] and various publications)

CMOS process node	Oxide thickness nm	Capacitance fF/μm^2
0.8 μm	15	2.3
0.5 μm	12	2.9
0.35 μm	8	4.3
0.25 μm	5	6.9
0.18 μm	4	8.3
90 nm (LP)	2.4	11.7
65 nm	2.2	12.6

Table 2.18 An indication of capacitor characteristics in a semiconductor process ranging from 0.18 μm to 90 nm generations (sources e.g. ITRS [19] and various publications)

Material stack	Capacitance fF/μm^2	Voltage coeff. V^{-1}	Temp. coeff. C^{-1}	Matching %/$\sqrt{\text{fF}}$
Diffusion	0.5		3×10^{-4}	
MOS gate 0.18 μm	8.3	3–5×10^{-2}		
Fringe capacitors	1.5			0.3
MIM capacitors	4–15	10^{-5}		0.3
Plate capacitors (7–9 layers)	0.8			0.5
Poly-poly 0.35 μm	0.8	5–10×10^{-4}	-8×10^{-5}	

- Interconnects form plate capacitors and fringe[15] capacitors, see Fig. 7.23. The capacitance is defined by (2.51). In devices with thin insulators the effective thickness is determined by the distance between the conducting plates and the effective thickness of the charge sheet (normally around 1 nanometer). In low-doped polysilicon plates partial depletion can occur, which dramatically increases the effective insulator thickness. Also tiling patterns can affect the capacitor value, see Sect. 11.3.2. Stacking various layers of interconnect, where odd and even numbered layers form the plates of a capacitor are generally a good solution. The plate capacitance between two successive layers is in the order of 0.1 fF/μm^2. Stacking 9 layers brings a capacitance of 0.8 fF/μm^2. This capacitor requires no bias, has a good linearity and low parasitics. Fringe capacitors use a finger structure and combine lateral and vertical capacitances. Densities of around 1.5 fF/μm^2 are obtainable.
- In some process variants a so called Metal-Insulator-Metal capacitance option is present. With an additional mask the isolator layer between two levels of metal is thinned or even replaced by a layer with a higher dielectric constant allowing a high specific capacitance. This construction can allow the implementation of a high-density low-parasitic device with a specific capacitance of 4 fF/μm^2 up to 15 fF/μm^2. It remains however questionable whether real gain is achieved over the before-mentioned plate and fringe capacitances.
- In older processes a double poly-silicon capacitor option is offered. The depletion of the poly-silicon can lead to voltage non-linearities.

A simple example: a 10×10 μm^2 conductor is deposited on a 3700 Å (=370 nm) silicon oxide on a conductive substrate. The permittivity in vacuum $\varepsilon_0 = 8.85 \times 10^{-14}$ F/cm, so the resulting capacitor is:

$$C = \frac{A\varepsilon_r\varepsilon_0}{d} = \frac{10^{-3} \times 10^{-3} \times 3.9 \times 8.85 \times 10^{-14}}{370 \times 10^{-7}} = 9.3 \times 10^{-15} \text{ F} = 9.3 \text{ fF}$$

---

[15]There are many fringe capacitor lay-outs. For lawyers patent 7.170.178 [21] is a good starting point, engineers might prefer the "examples" section of [22].

2.4.6 Voltage and Temperature Coefficient

Voltage and temperature coefficients are defined in a similar way as for resistors:

$$C(V) = C(V_0)(1 + \text{VC}(V - V_0)) \quad \text{with} \quad \text{VC} = \frac{1}{C}\frac{dC}{dV} \tag{2.69}$$

$$C(T) = C(T_0)(1 + \text{TC}(T - T_0)) \quad \text{with} \quad \text{TC} = \frac{1}{C}\frac{dC}{dT} \tag{2.70}$$

The temperature coefficient of plate capacitors is mostly negligible. Capacitances of depleted regions may show significant temperature sensitivity. For large voltage variations the basic equation (2.52) must be used.

2.4.7 Matching of Capacitors

Multiple capacitor structures of the same size are not necessarily equal. Mutual differences will be due to variations in the edge definitions and in the dielectric thickness [23–25]. The variation in dielectric thickness is in many processes dominant, causing an area dependency of the mismatch. In Sect. 11.4, the mismatch is described as:

$$\frac{\sigma_{\Delta C}}{C} = \frac{A_{C,\text{area}}}{\sqrt{\text{area}}} = \frac{A_{C,\text{area}}\sqrt{\varepsilon_0\varepsilon_r/d_{ox}}}{\sqrt{C}} = \frac{A_C}{\sqrt{C}} \tag{2.71}$$

With area $= Cd_{ox}/\varepsilon_0\varepsilon_r$. The area is expressed in μm^2 and $A_{C,\text{area}}$ in %μm. With $A_C = 0.5\%\sqrt{\text{fF}}$, see Table 2.18, a 400 fF capacitor has a mismatch of 0.025% or $\sigma_C = 0.1$ fF.

2.4.8 The pn-junction

A pn-junction is formed by two semiconductor regions of opposite dope. The acceptor concentration N_A generates a majority concentration p_p and a minority concentration n_p. Similarly the donor concentration N_D generates a majority concentration n_n and a minority concentration p_n. In the junction region the excess electrons in the n-doped semiconductor will see a p-doped semiconductor with a much lower electron concentration. Majority carriers will diffuse into the p-doped region. Similarly the excess holes from the p-doped material will diffuse into the n-doped region. So on both sides of the junction the mobile charge will disappear and create a depleted zone of fixed charge. The charge exchange creates a potential difference over the junction that will ultimately stop the flow of holes and electrons, see Fig. 2.29. This potential is called the "built-in" voltage:

$$V_{bi} = \frac{kT}{q}\ln\left(\frac{N_A N_D}{n_i^2}\right) = 2\phi_F \tag{2.72}$$

Fig. 2.29 Cross-section of a
pn-junction in equilibrium
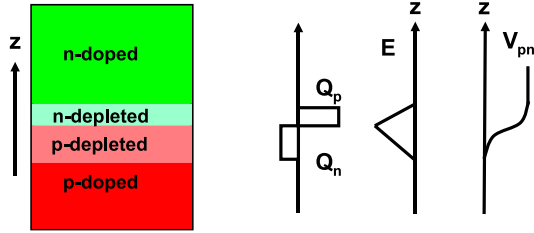
which can be understood by splitting the formula in two portions representing the voltage ϕ_F needed to go from N_A to n_i and a portion ϕ_F to go from n_i to N_D. The field over the junction keeps the majority carriers on both sides separated. Only a small fraction will have sufficient energy from the diffusion process to successfully pass the depletion region. The minority carriers experience the polarity of the field assisting them to cross the junction. The field drives these carriers: the drift component. In equilibrium the drift and diffusion components will cancel each other for both electrons and holes. So the total current densities in the general current equations are zero.

$$J_n = q\mu_n n E + q D_n \frac{dn}{dz} = 0$$
$$J_p = q\mu_p p E - q D_p \frac{dp}{dz} = 0 \tag{2.73}$$

D_n is the diffusion constant for electrons that is related to the mobility via Einstein's relationship:

$$D_n = \frac{kT}{q}\mu_n, \qquad D_p = \frac{kT}{q}\mu_p \tag{2.74}$$

An external voltage V_{pn} over the junction modifies the internal energy levels. Consequently the electron and hole concentration will adapt to these energy levels.

$$p_n = \frac{n_i^2}{N_D} e^{-qV_{pn}/kT}$$
$$n_p = \frac{n_i^2}{N_A} e^{qV_{pn}kT} \tag{2.75}$$

The levels of the minority carriers at the edges of the depletion region are given by these equations, see Fig. 2.30.

The current in a pn-junction is determined by the minority carrier charge transport to the depleted region. This transport of minority carriers depends on diffusion in the neutral parts towards the depleted zone as there is zero field in a neutral zone. In the neutral zone recombination of minority carriers with the majority carriers takes place. The minority concentration builds up over a recombination distance

$$L_n = \sqrt{D_n \tau_n} \tag{2.76}$$

τ_n is the generation-recombination time constant. In modern processes the diffusion length is in the order of one millimeter. In that case most recombination takes place
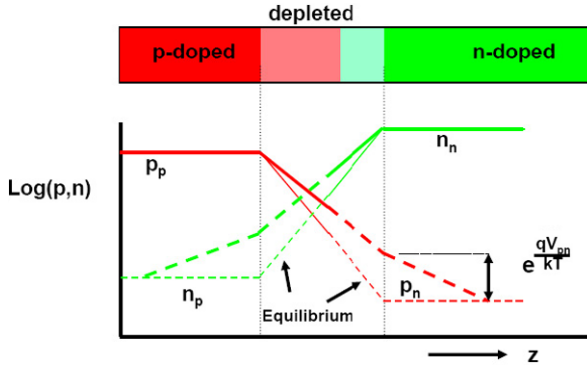
Fig. 2.30 The electron an hole concentrations in a pn-junction in equilibrium and during forward bias

at the contact areas. The current in a pn-junction is now found by applying the general diffusion equation for the carrier density $n(z, t)$ that is a function of the position z and the time:

$$\frac{\partial n(z, t)}{\partial t} = D_n \frac{\partial n(z, t)}{\partial z} \tag{2.77}$$

with the current I being equal to the time derivative of the electron concentration multiplied by the area A of the junction. The diode current is represented by:

$$I = q A \frac{dn}{dt} = q A D_n \frac{dn}{dz}$$

The current over the pn-junction consists out of two components: hole current and electron current. So:

$$I_{pn} = q A \left(D_n \frac{dn}{dz} + D_p \frac{dp}{dz} \right) = q A \left(D_n \frac{n_p}{L_n} + D_p \frac{p_n}{L_p} \right)$$

Substitution of the equation for levels of minority carriers leads to the first order model for a diode:

$$I_{pn} = q A \left(D_n \frac{n_i^2}{N_A L_n} + D_p \frac{n_i^2}{N_D L_p} \right) (e^{q V_{pn}/kT} - 1) \tag{2.78}$$

Often a pn-junction consists of one heavily doped area and one lightly doped area. In that case the term before the exponential is dominated by lightly-doped side and the other term is ignored. In the lightly doped side there is less dope, a lower majority carrier level and consequently a higher minority level. The resulting term before the exponential is summarized with the symbol I_s. If N_A is the level of the lowest doping:

$$I_{pn} = \frac{q A D_n n_i^2}{N_A L_n} (e^{q V_{pn}/kT} - 1) = I_s (e^{q V_{pn}/kT} - 1) \tag{2.79}$$

Increasing N_A means that for a constant current density, the required forward voltage V_{pn} has to increase.

Equation (2.79) is the general equation for the diode or rectifier: the exponential term leads to a large forward-current and a small reverse current. The exponential change of the current at room temperature equals a factor 10 for every 60 mV of applied voltage. This behavior will return in the bipolar and MOS transistors as the "subthreshold slope".

At a certain combination of current I_{pn} and voltage V_{pn} the pn-junction will behave for small excitations as a resistance:

$$\frac{1}{r_{pn}} = \frac{dI_{pn}}{dV_{pn}} = \frac{qI_{pn}}{kT}$$

A practical rule of thumb is to divide the thermal voltage of 26 mV by the amount of current. The noise in a pn-junction is easily derived from the small-signal resistance:

$$v_{\text{noise,rms}} = \sqrt{4kTBWr_{pn}}$$

Substituting (2.60) for n_i^2, the forward voltage at current I_{pn} is given:

$$V_{pn} \approx \frac{E_G}{q} - \frac{kT}{q} \ln \frac{I_{pn}N_aL_n}{qAD_nN_CN_V} = \frac{E_G}{q} - \frac{kT}{q} \ln \frac{I_{pn}n_i^2}{I_sN_CN_V} \qquad (2.80)$$

The first term is the bandgap voltage and the second term represents a negative temperature coefficient part in the diode forward voltage. This relation is relevant for bandgap reference circuits.

The capacitance of a pn-junction with one heavily doped side is approximated by the MOS-capacitance expression:

$$C_{pn} = \frac{\varepsilon_r\varepsilon_0}{z_d} = \sqrt{\frac{qN\varepsilon_r\varepsilon_0}{2(V_{bi} \pm |V_{pn}|)}} \qquad (2.81)$$

The positive sign applies for reverse-bias diode voltages, the negative sign for forward biasing.

Many parasitic phenomena are associated with diodes [13]:

- Avalanche breakdown appears in reverse bias mode. In an avalanche pn-junction the free electrons that cross the depletion region are strongly accelerated. Due to their large kinetic energy every collision will generate an electron–hole pair (impact ionization), giving rise to an exponential increase in current. The current caused by this impact ionization must be limited in order to avoid damage.
- If the reverse voltage over a pn-junction is used to allow electrons to jump directly from the valence band to the conduction band the Zener effect appears. In lightly doped pn-junctions the bandgap is too large. However in junctions with extreme doping levels, the energy bands are less separated. Zener voltages of 3–6 V can be created.
- If both the n- and the p-side of a junction are heavily doped the depletion region is extremely thin. At zero-bias the energy levels in the valence band on one side of the junction are aligned with the energy levels in the conduction band of the other side. A tunnel mechanism creates an increasing current with increasing bias voltages. At some forward voltage the energy levels align more and more

Fig. 2.31 The bipolar transistor: schematic cross section and the symbol for an npn transistor

Depleted regions
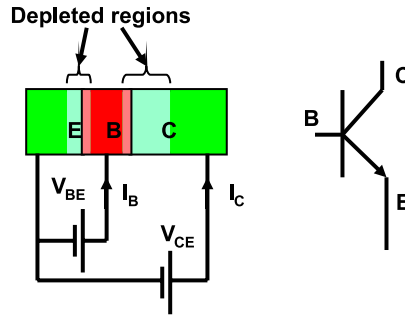
Fig. 2.32 The bipolar transistor: a typical set of characteristics
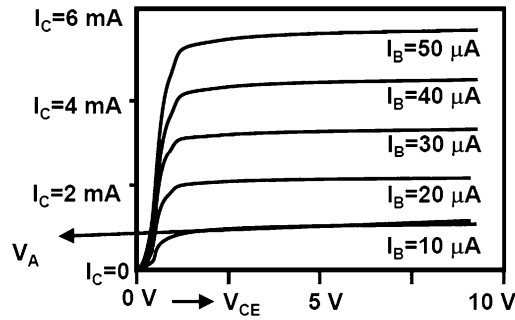
with the forbidden energy levels in the band gap: the current decreases. This is a region with a negative small-signal impedance. Finally the normal thermal diffusion conduction takes over. This type of diodes are called Esaki or tunnel diodes and are used as high-frequency oscillators [13, Chap. 9].

2.4.9 The Bipolar Transistor

In a bipolar transistor a voltage between the base and the emitter will cause electrons from the n- emitter to cross the forward biased base-emitter diode. As the base region is rather thin most of the charge carriers will not leave the base region towards the base voltage source, but diffusion causes the electrons to move into the n-type collector region, see Fig. 2.31. Only a small current is needed to keep a forward bias of the base-emitter junction. This current is due to the recombination of the hole current in the emitter. This small hole current I_B results in a much larger electron collector current: I_C. The ratio between the two is called the current gain h_{fe}:

$$I_C = h_{fe} \times I_B \tag{2.82}$$

The collector current relates to the V_{BE} voltage as:

$$I_C = I_0(e^{\frac{q V_{BE}}{kT}} - 1) \tag{2.83}$$
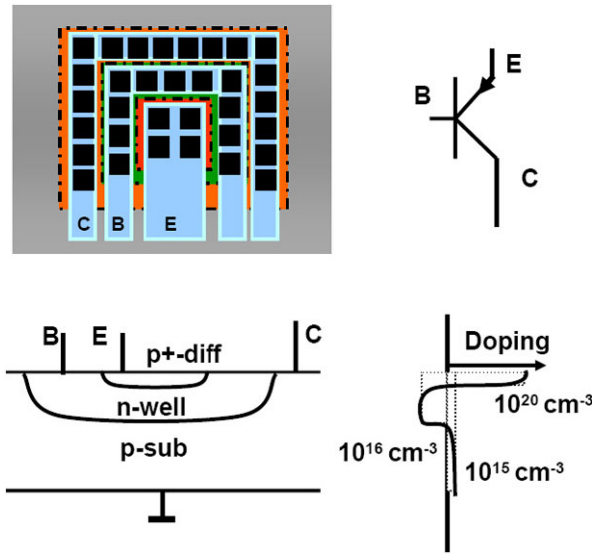
Fig. 2.33 The bipolar pnp-transistor can be formed in a CMOS process from the n-well and the p-diffusion: lay-out, symbol, cross-section and doping profile

Consequently the transconductance is found:

$$g_m = \frac{dI_C}{dV_{BE}} = \frac{qI_C}{kT} \tag{2.84}$$

The transconductance is independent of the transistor size, of course the maximum currents simply scale with the emitter area. As can be observed from the $I_C - V_{CE}$ characteristics in Fig. 2.32, the output current shows a slight dependence on the applied collector-emitter voltage. The effect, called the Early effect, is due to the narrowing of the neutral base due to the expansion of the base-collector depletion region. If the slope of the $I_C - V_{CE}$ curve is extrapolated to the $I_C = 0$ point, the Early voltage V_A is found. This Early voltage gives a first order estimate for the effective output impedance of the bipolar transistor:

$$r_o = \frac{dv_{ce}}{di_c} \approx \left| \frac{V_A}{I_C} \right| \tag{2.85}$$

In a CMOS process a parasitic bipolar pnp transistor can be formed by the p-source or drain diffusion, the n-well and the p-type substrate, Fig. 2.33. The built-in voltage of the pnp base-emitter junction is rather high due to the high doping levels of the source and drain implants in advanced CMOS processes. See for detailed analysis [13, Chap. 3], [26]. A second collector can be created through a diffusion next to the base. This collector creates an additional pnp transistor parallel to the vertical pnp transistor. The emitter current will divide over both the neighboring collector and the collector in the substrate. The split ratio depends on the current level and is normally a problem if an exact current density ratio is required as in a bandgap circuit.

Table 2.19 Data for vertical
pnp transistors

h_{fe}	5–10
V_A	150 V^{-1}
V_{be} at $J_E = 1$ µA/µm^2	0.8 V
σ_{Vbe} emitter = 4 µm^2	0.1 mV
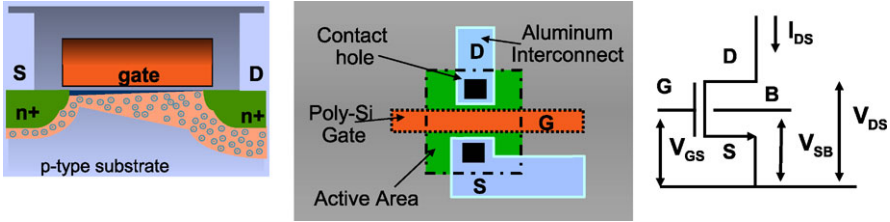
Fig. 2.34 Cross-section, lay-out and symbol of a MOS transistor

The mismatch between a pair of bipolar transistors can be expressed in a relative
collector current mismatch or an absolute base-emitter mismatch. The mismatch is
in ancient processes due to lithography deviations of the emitter patterning. Over
the last 20 years however, the mismatch is determined by fluctuations in the base
doping.

$$\sigma_{\Delta Vbe} = \frac{A_{Vbe}}{\sqrt{\text{area of emitter}}} \qquad (2.86)$$

For a high-speed bipolar process with thin base regions there is a relatively large im-
pact of base dope fluctuations and $A_{Vbe} = 0.3$–0.5 mV µm. The base formed by an
n-well of a parasitic vertical pnp in a CMOS process exhibits much better matching
properties: $A_{Vbe} = 0.1$–0.2 mV µm. These numbers are much better than for thresh-
olds of MOS devices. If however, the relative current mismatch of bipolar and MOS
devices differs less. This device with parameters as in Table 2.19 is used in reference
circuits, see Sect. 6.2.1.

2.5 The MOS Transistor

The Metal-Oxide-Silicon (MOS) transistor in Fig. 2.34 is a charge-controlled de-
vice.[16] The charge under the interface can form an electrical connection between
the source and drain terminal of the MOS transistor. The source is chosen as the
voltage reference terminal. The dimensions of the channel area between the source
and drain are specified by the width of the source to channel edge W and the length
L is the distance between source and drain. There are many specifications of these

---

[16]In this paragraph an NMOS transistor is described, for a PMOS transistor behavior the polarities
of charges and voltages need to be inverted.

dimensions: drawn on the mask, as expected on silicon and as measured electrically. Differences between these specifications can amount to a considerable fraction of a micron. The insulator (gate dielectric) between the channel in the silicon substrate and the gate electrode is mostly silicon-oxide with a thickness d_{ox}. In many older processes the oxide-thickness is of the order of 0.02 of the smallest line-width, e.g. a 0.25 μm process will have a 5–6 nm gate oxide thickness. The specific oxide capacitance is related to the oxide-thickness via the dielectric capacitor formula:

$$C_{ox} = \frac{\varepsilon_r \varepsilon_0}{d_{ox}} = \frac{3.9 \times 8.86 \times 10^{-14}}{d_{ox}} \text{ F/cm}^2$$

For an 0.18 μm process $d_{ox} = 4$ nm $= 4 \times 10^{-7}$ cm and $C_{ox} = 8$ fF/μm^2. In more advanced processes, two phenomena blur this simple picture. The assumption is that the gate material is a perfect metal plate, however, even highly doped polysilicon material will always be slightly depleted, thereby increasing the effective gate thickness. Also the use of additional layers in the gate stack cause different values of the dielectric constant. Therefore the term "effective oxide thickness" or EOT is introduced.

The charge under the gate is built up if the gate voltage is such that carriers from the source terminal are pulled under the gate. The electric field of the gate attracts the electrons towards the insulator: this sheet of electrons forms the "channel" between the two terminals of the MOS transistor. This marks the start of inversion, because the concentration of the charge in the channel region exceeds the substrate doping. If the gate voltage of an NMOS transistor V_{GS} exceeds a threshold voltage V_T (with a few kT/q) the "strong-inversion regime" applies. The simplest condition for the definition of inversion is therefore:

$$V_{GS} > V_T \tag{2.87}$$

Nearly all gate voltage variation is reflected by the inversion charge variation. Just a little change in the surface potential of the inversion charge is needed to increase the charge concentration as demanded by the charge equation (2.62). Without sufficient gate voltage, there is no inversion charge and the structure behaves as an MOS capacitance in depletion. The transistor is in the off state, Fig. 2.35.

The threshold voltage is composed of some fixed interface potential contributions and of the voltage that is necessary to build up the depletion charge (2.66):

$$\begin{aligned}
V_T &= V_{FB} + 2\phi_F + \frac{\sqrt{2\varepsilon_r \varepsilon_0 q N_a (2\phi_F + V_{SB})}}{C_{ox}} \\
&= V_{FB} + 2\phi_F + \gamma \sqrt{2\phi_F + V_{SB}} \\
&= V_T(V_{SB} = 0) + \gamma(\sqrt{2\phi_F + V_{SB}} - \sqrt{2\phi_F})
\end{aligned} \tag{2.88}$$

where ϕ_F represents the band-bending voltage or the potential that is needed to change the concentration from the doped level N_a to an intrinsic level n_i:

$$\phi_F = \frac{kT}{q} \ln\left(\frac{N_a}{n_i}\right) \tag{2.89}$$
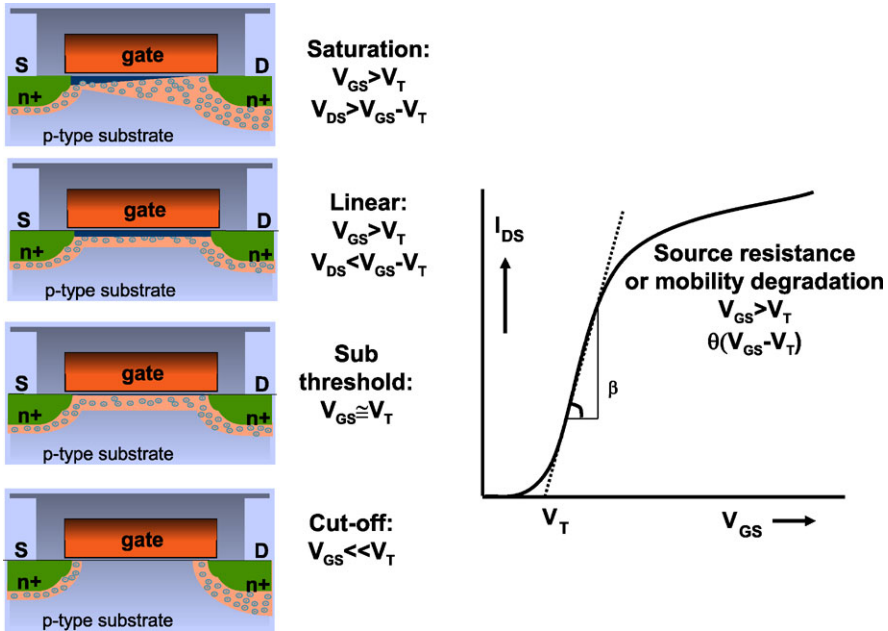
Fig. 2.35 *Left*: various regimes in a NMOS transistor. *Right*: in a I_{DS} versus V_{GS} plot in the linear regime the threshold voltage and current factor can be easily identified

The flat-band voltage V_{FB} is composed of material related potential differences. Next to fixed charges in e.g. the oxide layer, the most important contribution comes from the work-function difference ϕ_{MS}.

The back-bias factor γ is an indication of the effect that a change in back-bias voltage V_{SB} has on the threshold voltage.

$$\gamma = \frac{\sqrt{2\varepsilon_r \varepsilon_0 q N_a}}{C_{ox}} \tag{2.90}$$

Now the equivalent current can be calculated from the average conductance under the entire gate area (the linear regime):

$$I_{DS} = \frac{W C_{ox} \mu}{L}\left(V_{GS} - V_T - \frac{1}{2}V_{DS}\right)V_{DS} \tag{2.91}$$

In saturation mode the drain voltage rises to a level where at the drain side no longer inversion exists[17]: $V_{GD} < V_T$. The perpendicular electrical field from the gate to the channel pulls the electrons to the interface only in that part of the channel where the potential is low enough. At some point along the channel the potential reaches a level where the field reverses. The electrons travel along the substrate-insulator

----

[17]For PMOS transistor the voltages invert and the carriers are holes.
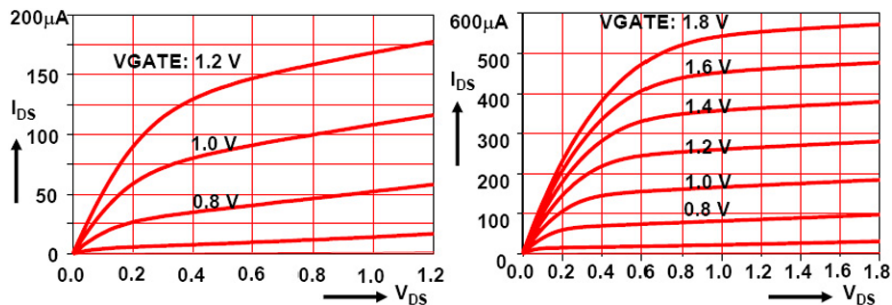
Fig. 2.36 Two drain current I_{DS} versus drain-source voltage V_{DS} plots. The *left set of curves* is a 0.3/0.065 NMOS transistor in a 65-nm process. The *right set of curves* are for a 1/0.18 NMOS transistor in a 0.18 μm process. The gate voltage runs from 1.2 resp. 1.8 V with decrements of 0.2 V

interface up to this point, but from there onwards move through the substrate into the drain terminal. Referring the voltages to the source, the saturation condition is:

$$(V_{GS} - V_{DS}) < V_T, \qquad V_{DS} > (V_{GS} - V_T) \qquad (2.92)$$

This is the critical drain-source voltage (or "saturation voltage") for a transistor to operate efficiently as a current source. Below this level the transistor operates in the linear regime and a considerable parallel resistor impairs the performance. The current is in saturation mode given by the classical "square law MOS equation":

$$I_{DS} = \frac{W C_{ox} \mu}{2L} (V_{GS} - V_T)^2 \qquad (2.93)$$

The quantity $W C_{ox} \mu / L$ is replaced by the current factor β.[18] For a square ($W = L$) transistor β_{square} or $\beta_\square$ is used. In real MOS devices the quadratic law can still be applied, however many secondary effects impair the ideal behavior. Figure 2.36 shows two sets of characteristics of 65 nm and 0.18 μm processes.

The mobility is part of the current factor and a significant difference will occur between NMOS and PMOS transistor currents. With a ratio of approximately three between the ideal values of $\mu_n \approx 1500 \text{ cm}^2/\text{V s}$ and $\mu_p \approx 450 \text{ cm}^2/\text{V s}$ for the mobility for electrons and holes, the current factors will differ accordingly. Due to many impairments of the current flow under the gate, the real values for the mobilities are normally much lower, see Fig. 2.37. Other choices in the process, such as buried channels, also affect the NMOS/PMOS current factor ratio. Table 2.20 gives an indication of the threshold voltage, current factors and mismatch of various processes.

---

[18] Some authors prefer K.
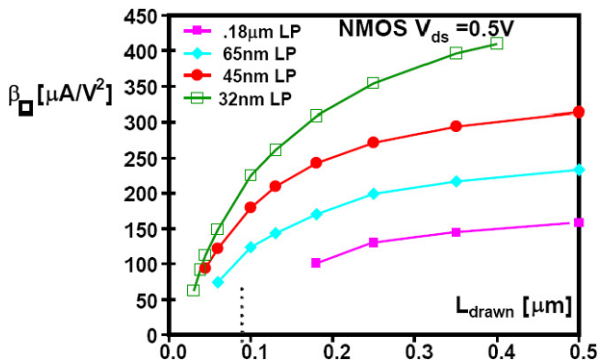
Fig. 2.37 The current factor in a MOS transistor for different processes is much lower than the nominal current factor. A strong dependence on the length can be observed, $V_{GS} - V_T \approx 0.2$–0.3 V [27]

Table 2.20 The values of NMOS and PMOS parameters are indicative for what is used in industry. Especially the 130, 90 and 65 nm values depend on many technological effects and bias settings. See for a discussion of the current factor in these processes Fig. 2.37 (*source*: ITRS [19] and various publications)

Process	V_{DD} (V)	d_{ox} (nm)	$V_{T,n}$ (V)	$V_{T,p}$ (V)	$\beta_{\square,n}$ (μA/V^2)	$\beta_{\square,p}$ (μA/V^2)	$A_{VT,n}$ (mV μm)	$A_{VT,p}$ (mV μm)	$A_{\beta,np}$ (% μm)
0.8 μm	3.3	15	0.6	−0.65	125	55	10.7	18.0	4
0.6 μm	3.3	13	0.65	−0.8	150	50	11.0	8.5	
0.5 μm	3.3	12	0.6	−0.6	130	36	9	10	1.8
0.35 μm	3.3	7.7	0.63	−0.6	190	46	8	7.4	2
0.25 μm	2.5	6	0.57	−0.53	235	53	6	6	1.5
0.18 μm	1.8	4	0.48	−0.5	300	80	6	5	1.6
0.13 μm	1.2	2.5	0.34	−0.36	590	135	5	5	1.6
90 nm (LP)	1.2	2.3	0.37	−0.39	550	160	4.5	3.5	
65 nm (LP)	1.2	2.2	0.32	−0.36	450	200	5	3.5	1.2

2.5.1 Weak Inversion

When the gate voltage is depleting the substrate and is not exceeding the threshold voltage by at least a few times the thermal voltage, the weak inversion regime applies. The regime above the threshold voltage between weak and strong inversion is called "moderate inversion". In this transition regime simulation tools are needed for proper analysis. In both the strong and weak inversion regimes the gate-source voltage can be subdivided between the gate-to-channel potential and the channel-to-source potential. In strong inversion an increase in gate potential mainly results in an increase in inversion charge. Therefore most of the gate voltage variation will be over the oxide capacitance and will control the inversion charge as if it was a
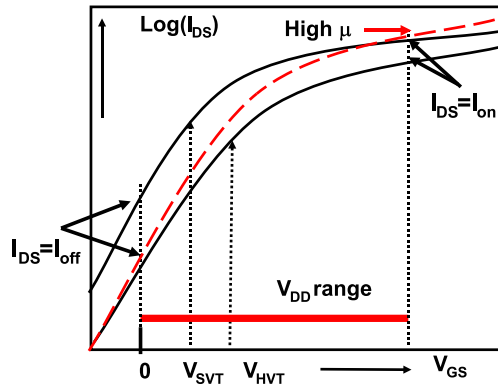
Fig. 2.38 The current in a MOS transistor is determined by weak inversion for gate-source voltages lower than the threshold voltage, while for higher gate-source voltages the strong-inversion regime applies. In digital CMOS processes multiple threshold voltages are available. Here V_{SVT} and V_{HVT} indicate the standard and high threshold voltage that are used for respectively high-speed and low-leakage performance. The *dotted line* indicates potential improvements in the subthreshold slope and the mobility

capacitor charge. However the Boltzmann relation for the concentration difference between the source and the inversion channel must be met too. So if the inversion layer concentration varies e.g. a factor 10, a source-to-channel voltage variation of 60 mV will be needed. This has been ignored in the above derivation.

In weak inversion the gate voltage controls the channel-to-source potential via the capacitive ratio between gate capacitance and bulk capacitance. The channel potential determines the concentration in the inversion layer with respect to the source via the Boltzmann's relation. The inversion layer charge is not yet large enough to reduce significantly the effective gate voltage variations as in the strong inversion regime. As the inversion charge is determined by the Boltzmann equation, the current is given by:

$$I_{DS} = \frac{W I_o}{L} e^{q(V_{GS} - V_T)/mkT} \tag{2.94}$$

Where $m = (C_{ox} + C_{sub})/C_{ox} \approx 1.0 \ldots 1.3$ defines the capacitive division between the gate voltage and the channel potential. In Fig. 2.38 the current in a MOS transistor on a logarithmic scale is shown as a function of the gate-source voltage. The shape of the $\log(I_{DS})$ curve is determined by the choices in the process. The intersection of the slope in the weak-inversion regime with $V_{GS} = 0$ determines the leakage current in the off-state. The log curve has a slope between 60 and 100 mV/decade, depending mainly on the choice of the gate isolation structure. The current at $V_{GS} = V_{DD}$ gives the saturation current and is essential for maximum speed in digital circuits. In modern technology the process control is such that various flavors of threshold voltages can be supplied. In Fig. 2.38 a standard and a high threshold voltage are shown. The first is used for active high-speed circuits such as processors. The high threshold voltage is no more than 100 mV higher and results

in a decrease of leakage currents in the order of $30\times$ and is used for memories. In digital design the only option on transistor level to reduce leakage and increase the drive current is in improving the mobility factor and reducing the effective gate dielectric. The subthreshold slope may improve slightly, and the saturation current can be increased between 30% and 80%.

2.5.2 Matching

Circuit design and especially analog-to-digital design heavily relies on the assumption that equally designed components will behave the same. This assumption is limited by systematic and random variations, see Chap. 11. The largest contribution in random offset in the threshold definition of MOS transistors is the fluctuation of the number of fixed charged atoms in the depletion region. These charged atoms (dopants, dislocations, oxide charges, interface states, etc.) are implanted, diffused or generated during the manufacturing process, but not in an atom-by-atom controlled manner. The average value is controlled by implantation dope levels or average substrate dopes. The actual number of carriers in a particular depletion region differ from this average. The difference between the threshold voltages of two transistors is statistically characterized by a normal distribution with a mean value zero and a variance:

$$\sigma_{\Delta VT} = \frac{A_{VT}}{\sqrt{WL}} \tag{2.95}$$

All technological constants are merged into one parameter A_{VT} [28].

Table 2.20 compares the A_{VT} coefficients as used in industry.

The proportionality factor for the current factor $\beta = \mu C_{ox} W/L$ is defined as:

$$\frac{\sigma_{\Delta\beta}}{\beta} = \frac{A_\beta}{\sqrt{WL}} \tag{2.96}$$

The relative matching of the current factor is also proportional to the inverse square root of the area.

Considering only the threshold and current factor variations, the variance of the difference in drain currents ΔI between two equally sized MOS devices can be calculated. With the help of (2.18):

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{dI}{dV_T}\right)^2 \sigma_{\Delta VT}^2 + \left(\frac{dI}{d\beta}\right)^2 \sigma_{\Delta\beta}^2 \tag{2.97}$$

For strong inversion this equation can be written as:

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{2\sigma_{\Delta VT}}{V_{GS} - V_T}\right)^2 + \left(\frac{\sigma_{\Delta\beta}}{\beta}\right)^2 \tag{2.98}$$

In weak inversion the current is modeled as an exponential function. The current factor mismatch is due to the associated low current levels of less importance:

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{q\sigma_{\Delta VT}}{mkT}\right)^2 \tag{2.99}$$

Table 2.21 Transistor parameters for 0.18 μm CMOS as specified by ITRS [19] and various publications

	NMOS	PMOS
Threshold voltage V_T	0.39 V	−0.45 V
Current factor (3/3) $\beta_\square$	300 μA/V^2	80 μA/V^2
Output conductance λ	0.05 at L_{minimum}	0.05 at L_{minimum}
Output conductance λ	0.01 at $L = 1$ μm	0.05 at $L = 1$ μm
Back-bias factor γ	0.60 V$^{0.5}$	0.66 V$^{0.5}$

A more extensive analysis of random matching is given in Sect. 11.4 and systematic offsets are discussed in Sect. 11.3.

2.5.3 Drain Voltage Influence

The square-law formula is useful for hand calculations. For circuit simulation a lot of secondary effects influence the behavior and a more elaborate model is required. The square law equation (2.93) implies that the MOS transistor in saturation is an ideal current source, as there is no dependence of the current with the drain voltage. However, due to the increase in drain voltage less charge in the depletion region needs to be depleted by the gate voltage. The drain voltage reduces the threshold voltage and increases the current. This Drain-Induced-Barrier-Lowering (DIBL) is the major cause for an increased drain current in a sub-micron transistor. In older processes an increased drain voltage pushes back the point where inversion still exist, and reduces the effective gate length. Some authors use the term "static feedback" to describe these effects. In the first models the output impedance is described by reducing the effective gate length:

$$I_{DS} = \frac{W C_{ox} \mu}{2L(1 - \alpha V_{DS})}(V_{GS} - V_T)^2 \approx \frac{W C_{ox} \mu}{2L}(V_{GS} - V_T)^2(1 + \alpha V_{DS}) \quad (2.100)$$

Here α has the dimension 1/Volt. In more recent models all drain related effects that affect the output impedance (static feedback, drain induced barrier lowering) are added, using the dimensionless parameter λ, that modulates the effective gate drive voltage resulting from drain voltage variations:

$$I_{DS} = \frac{W C_{ox} \mu}{2L}(V_{GS} - V_T + \lambda V_{DS})^2 \quad (2.101)$$

The effect of the drain voltage modulating the transistor current, results in a limited output impedance of the transistor.

In order to reduce the impact of the drain field on the channel, a lightly-doped drain (LDD) is often applied. This shallow implant between the channel and the normal drain diffusion allows to improve the output impedance of the MOS transistor without introduction too much series resistance.

Fig. 2.39 Small signal linearization in a bias point
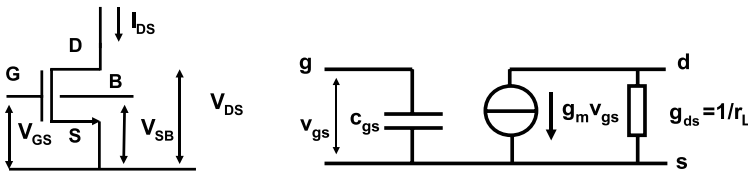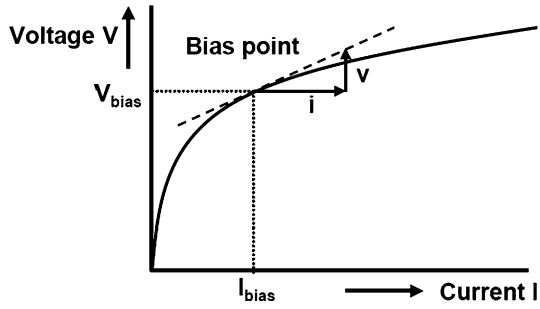
Fig. 2.40 The MOS transistor as a small-signal equivalent circuit

2.5.4 Large Signal and Small Signal

Until now voltages, charges and currents have been measured with respect to their zero-values, these values are called large signals or bias values. As can be seen in the $V-I$ curve shown in Fig. 2.39 the relations between voltages and currents are not linear over the entire signal range and bias. In analog design small signals are used for obtaining better linearity in the processed signals. Just using a small portion of the transfer curve allows higher gains and better linearity. Analog circuits are designed and analyzed using small-signal equivalent circuits. Based on the DC bias of each transistor a small-signal equivalent circuit is used to calculate the behavior of small voltage and current excursions in the bias point, Fig. 2.40. In small-signal mode all voltages and currents are written in lower-case notation, while the bias conditions are indicated with capitals. The transconductance of a MOS transistor g_m is the change in current due to a change in gate voltage:

$$g_m = \frac{i_d}{v_{gs}} = \frac{dI_{DS}}{dV_{GS}} = \frac{WC_{ox}\mu}{L}(V_{GS} - V_T + \lambda V_{DS}) = \sqrt{2\beta I_{DS}}$$
$$= \sqrt{\frac{2WC_{ox}\mu I_{DS}}{L}} \tag{2.102}$$

In some analysis it is necessary to include the modulation due to the substrate voltage as a controlled current source parallel to $g_m v_{gs}$:

$$g_{mb} = \frac{dI_{DS}}{dV_{SB}} = \frac{-WC_{ox}\mu(V_{GS} - V_T)}{L}\frac{dV_T}{dV_{SB}}$$
$$= \frac{-WC_{ox}\mu}{L}\frac{\gamma(V_{GS} - V_T)}{2\sqrt{2\phi_F + V_{SB}}} \tag{2.103}$$

Fig. 2.41 The voltage gain
of a transistor with a fixed
ratio $W/L = 2$ as a function
of gate length and technology
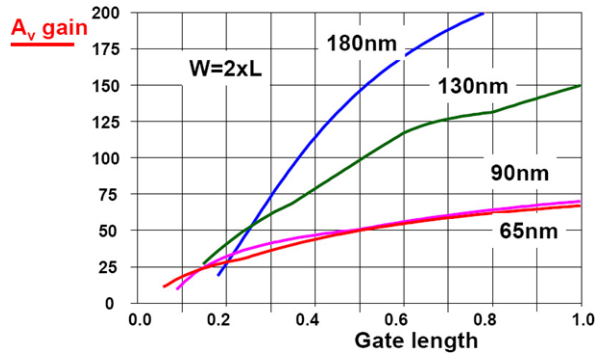
Typically the bulk transconductance amounts 10% to 20% of the gate transconductance.

The output conductance is physically caused by Drain-Induced barrier Lowering and Static Feedback. In the electrical domain the output conductance is the change in current due to the change in drain voltage and is found in a similar way:

$$g_{ds} = \frac{1}{r_{ds}} = \frac{dI_{DS}}{dV_{DS}} = \frac{\lambda W C_{ox} \mu}{L}(V_{GS} - V_T + \lambda V_{DS}) = \lambda g_m \qquad (2.104)$$

With this definition the maximum voltage amplification A_V of a transistor is given.

The current variation due to an input signal v_{in} is $i_{ds} = g_m v_{in}$. This current experiences on the drain side a load impedance in the form of the output conductance g_{ds}, which limits the maximum voltage amplification to:

$$A_V = \frac{v_{out}}{v_{in}} = \frac{i_{ds}r_{ds}}{v_{in}} = \frac{1}{\lambda} \qquad (2.105)$$

Practically A_V is limited to the range 20–50, see Fig. 2.41.

2.5.5 High-frequency Behavior

Figure 2.42 shows the capacitors surrounding a MOS transistor. The intrinsic speed limitation in a MOS transistor is caused by the traveling time of the charge carriers from source to drain. This transit time is in most applications so short that the resulting speed effects are ignored. This approximation is called: "quasi-static behavior". Non quasi static behavior (NQS) in transistors with long gates can often be circumvented by splitting the transistor in a number of shorter devices [30]. The transit time limits the highest frequencies obtainable with MOS devices to some 1000 GHz [31].

The high-frequency behavior of MOS transistors in analog design is mostly controlled by the current drive capability of the device (the transconductance g_m) and the surrounding capacitors. Figure 2.42 shows the most relevant capacitors in a MOS device and Table 2.22 shows some values. The detailed behavior of these capacitors
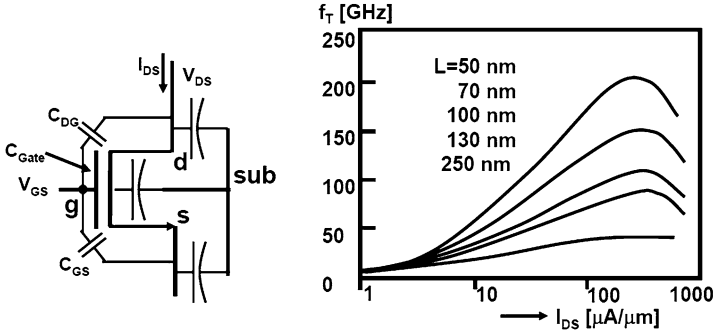
Fig. 2.42 The capacitances around a MOS device in 90 nm technology determine the high-frequency behavior. A characteristic set of performance curves shows a strong dependence for the maximum frequency on gate-length and current (after [29])

Table 2.22 Some parasitic transistor capacitors in 90 nm technology (source: ITRS [19] and various publications)

Gate capacitance	$C_{\text{Gate}} = WLC_{ox}$	$C_{ox} = 7$ fF/μm^2
Gate-drain capacitance	$C_{GD} = WC_{\text{olap}}$	$C_{\text{olap}} = 0.3$ fF/μm
Drain-substrate capacitance	$C_{\text{Dsub}} = WC_{\text{diff}}$	$C_{\text{diff}} = 0.6$ fF/μm

is subject to a lot of study [26]. The most common measure of high-frequency behavior is the cut-off frequency:

$$f_T = \frac{g_m}{C_{\text{gate}} + C_{gs} + C_{gd} + C_{\text{parasitcs}}} \tag{2.106}$$

In popular terms this measure is defined by that frequency where the current needed to drive a transistor equals the current that the transistor itself can generate. Using the transconductance formula $g_m = W\beta(V_{GS} - V_T)/L$ and only looking at the largest capacitor:

$$f_T \approx \frac{\mu_{n,p}(V_{GS} - V_T)}{L^2} \tag{2.107}$$

it is clear that especially a short effective gate-length increases the cut-off frequency. In Fig. 2.42 (right) a typical set of curves shows the behavior of the cut-off frequency. The V_{GS} values where the maximum frequencies occur are mostly close to the power supply. A problem for analog-to-digital designers is that their options to use those biasing points are limited.

Next to the cut-off frequency a number of other high-frequency indicators are used. The f_{MAX} frequency [29] defines the highest frequency at which an oscillation can be maintained. f_{MAX} is linked to the power gain and normally 20–40% lower than the current gain defined by f_T.

The lay-out of a transistor can easily impact the maximum frequency. Wide and short gates result in large resistances of the gate. This gate resistance is reduced by splitting the transistor in a number of parallel connected devices with a fraction of

the gate width. In processes with relatively high-ohmic gate material even medium frequency applications can suffer from gate resistance.

2.5.6 Gate Leakage

In electronic circuits the ideal MOS switching characteristics are jeopardized by leakage phenomena. The remaining drain to source conductivity is described by the weak-inversion regime. Also some gate current may leak through the isolation dielectric layer. Especially in transistors with gate oxide thickness less than 2.5 nm this effect can lead to considerable currents. At 2.5 nm the ITRS roadmap expects a gate leakage current density of $10^{-3}\,\mathrm{A/cm^2}$, rising at 2 nm to $10^{-1}\,\mathrm{A/cm^2}$ and at 1.5 nm to $10\,\mathrm{A/cm^2}$.

The basic mechanisms for the leakage current are Fowler-Nordheim tunneling and direct tunneling. The first component comes from charge carriers that tunnel through the thin isolator due to a high electric field over the isolator. The corresponding current density is a function of the electric field over the isolator:

$$J_{FN} = C_1 E_{ox}^2 e^{\frac{-C_2}{E_{ox}}} \tag{2.108}$$

Direct tunneling occurs at lower voltage differences, but requires thin dielectrics (<2 nm) to generate a relevant current. The mathematical description is similar to the Fowler-Nordheim formula. Large tunnel currents will in the end lead to oxide breakdown and failure of the device.

2.5.7 Temperature Coefficient

The temperature dependence of a MOS transistor is related to the threshold voltage and the current factor. The temperature dependence of the threshold voltage is found by examining (2.88). The contributions to the threshold temperature dependence come from the work-function ϕ_{MS} and from the thermal potential ϕ_F [32]. Both are related to the Boltzmann equation which describes the carrier concentration as a function of potential difference and temperature. For n-channel:

$$\frac{dV_T}{dT} = \frac{d\phi_{MS}}{dT} + \frac{2d\phi_F}{dT} + \frac{\gamma}{\sqrt{(2\phi_F + V_{SB})}} \frac{d\phi_F}{dT} \tag{2.109}$$

The temperature dependence of the workfunction depends on the doping polarity of gate and substrate. Equal polarity give 1 mV/K, opposite gives -3 mV/K. For a p-channel transistor the sign of both ϕ_F terms reverses. The temperature dependence of the threshold voltage is thereby determined by the temperature dependence of the work-function and of the band-bending voltage ϕ_F. This last quantity can be analyzed with the help of (2.60):

$$\phi_F = \frac{kT}{q} \ln\left(\frac{N_a}{n_i}\right) = \frac{kT}{q} \ln\left(\frac{N_a}{\sqrt{N_V N_C}}\right) - \frac{E_G}{2q}$$

As the bandgap energy and the band-bending voltage are well-known quantities the logarithmic term is replaced and the temperature dependence is rewritten as[19]:

$$\frac{d\phi_F}{dT} = \frac{k}{q} \ln\left(\frac{N_a}{\sqrt{N_V N_C}}\right) = \frac{1}{T}\left(\phi_F - \frac{E_G}{2q}\right)$$

Evaluating the last formula with $E_G/q = 1.205$ V and $\phi_F \approx 0.35$ V gives a temperature coefficient of the band-bending voltage of -0.9 mV/°C, which brings the contribution of the band-bending in the threshold voltage to a value of $-2\ldots-2.4$ mV/°C.

The temperature dependence of the work function depends on the choice of the gate and substrate materials [32]. An n-type gate on a p-substrate creates a similar work-function behavior as the inversion layer of an n-channel MOS in a p-type substrate. So these two effects on source and gate largely compensate. If a p-type gate is used this compensation effect does not take place and a much larger temperature coefficient is reached. In advanced processes the work function of the total gate construction must be considered. A good rule-of-thumb starting value for the threshold temperature dependence is

$$\frac{dV_T}{dT} = \pm 2 \text{ mV/°C} \tag{2.110}$$

where the negative value is for the n-channel and the positive value for the p-channel. Both threshold voltages reduce in absolute sense with increasing temperature.

Next to the threshold also the current factor is affected by an increase in temperature. The temperature dependent factor is the current factor is the mobility $\mu \propto T^\alpha$, with $\alpha = -2.0\ldots-2.5$. This effect results in a current factor temperature dependence:

$$\beta = \left(\frac{T}{T_0}\right)^\alpha \beta(T = T_0) \tag{2.111}$$

At higher temperatures the current factor decreases.

The overall current depends on temperature as:

$$\frac{dI_D}{dT} = \frac{dI_D}{d\beta}\frac{d\beta}{dT} + \frac{dI_D}{dV_T}\frac{dV_T}{dT} = \frac{\alpha}{T}I_D - \frac{2I_D}{V_{GS} - V_T}\frac{dV_T}{dT} \tag{2.112}$$

where the temperature dependence of the current factor will reduce the current while the lowering of the threshold will increase the current for a fixed gate-source voltage. It is not difficult to see that with $\alpha = -2.0$ and a threshold temperature coefficient of -2 mV/°C, a temperature cancellation occurs at $V_{GS} - V_T = 0.6$ V. In a digital CMOS circuit with a threshold voltage of around $V_T = 0.4$ V a supply voltage of 1 V will lead to a temperature insensitive circuit.

---

[19]Ignoring a small temperature dependence in N_V, N_C.

Fig. 2.43 Various noise regimes in a MOS transistor
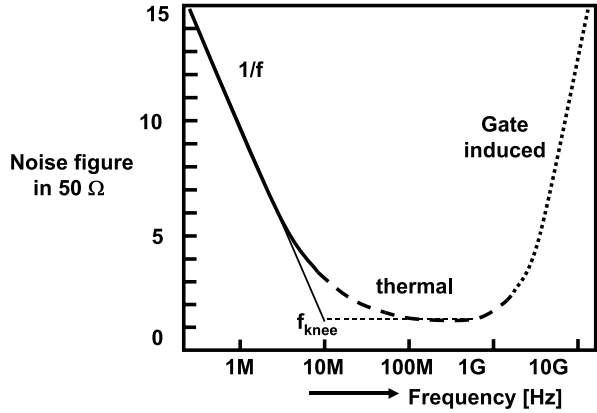
2.5.8 Noise

Figure 2.43 shows the dominant noise contributions in a MOS transistor, see [16, 33]. At low frequencies the noise is dominated by the $1/f$ noise. The spectral density of the voltage source in series with an MOS gate is:

$$S_{vv,\text{MOS}} = \frac{K_{1/f}}{WLf}$$

$$S_v(f) = \sqrt{\frac{K_{1/f}}{WLf}} \ [\text{V}\sqrt{\text{Hz}}] \tag{2.113}$$

where $K_{1/f} \approx 10^{-10}$ [V^2 μm^2] and the MOS device dimensions in μm for an 0.18 μm process. In advanced processes with smaller dimensions this coefficient rises to $K_{1/f} \approx 10^{-9}$ [V^2 μm^2]. S_{vv} is a power noise density and has as dimension V^2/Hz. Often the $1/f$ noise is given as the value of S_{vv} at 1 kHz in a 1 Hz bandwidth. The total relevant noise is obtained by integrating the noise over the bandwidth of interest.

Depending on the biasing and dimensions of the transistor, the thermal noise takes over at the noise knee frequency. The thermal noise is caused by the Brownian motion of the carriers in the channel. The general description for thermal noise applies to the MOS channel as well:

linear regime $\quad S_{vv} = 4kTR_{\text{channel}} \qquad S_{ii} = 4kT/R_{\text{channel}}$
saturated regime $\quad S_{vv} = 4\alpha kT/g_m \qquad S_{ii} = 4\alpha kTg_m$

In the saturated regime the transconductance is not evenly distributed over the channel. Classical long-channel theory requires to take into account a correction factor of $\alpha = 2/3$. Measurements on short channel devices suggest that the various additional contributions may require to set $\alpha = 1$ [33].

While the knee frequency for 0.5 μm CMOS processes was in the 100 kHz range, this cross-over point can be found at frequencies of over 200 MHz for a minimum channel length 65-nm transistor. Increasing the gate length to 1 μm moves the knee
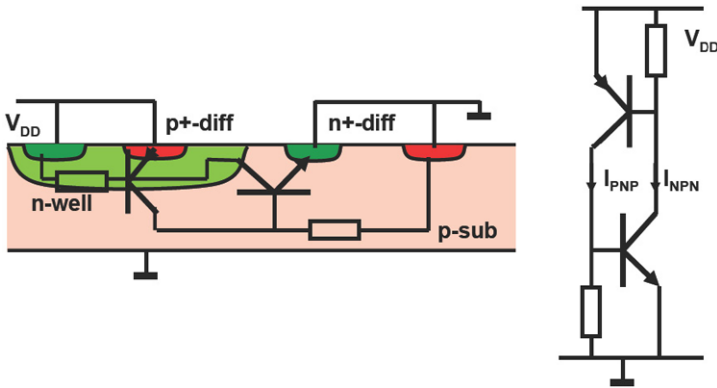
Fig. 2.44 Cross-section of a basic latch-up situation in CMOS and a circuit representation

frequency back to the 1 MHz range. At high frequencies this noise can couple capacitively through the gate terminal to the outside world. If the gate is connected to a high impedance, the gate-induced noise can be observed over this impedance. This noise stems from the same source as the thermal noise; therefore these two are correlated.

In close proximity to the intrinsic transistor, the resistance of the well and the gate structure can contribute significantly to the total observed noise in the transistor. Splitting up transistors in many parallel sections with many well connections reduces this problem.

2.5.9 Latch-up

Figure 2.44 shows a cross-section through a CMOS die. In the n-well the p-diffusion is the source or drain of a PMOS transistor. Here only an isolated diffusion is drawn. This p-type region forms together with the surrounding n-well and the substrate a pnp bipolar device. Similarly an n-diffusion is shown that is part of an NMOS transistor. Together with the n-well and the p-substrate an npn transistor is formed. The pnp and npn devices create a latching circuit. If a current I_{PNP} is present this current can act as a base current for the npn bipolar transistor. This device will amplify the base current and induce a larger base current for the pnp device. This process continues until external parasitic elements limit the total current or because the temperature rises to a level that creates damage.

This scenario is a nightmare for each CMOS circuit. A number of measures can be taken to prevent latch-up. A low-ohmic substrate will short-circuit the base of the npn transistor. Many contacts between the p-diffusion and the n-well and between the substrate and the n-diffusion also prevent that a voltage over the resistors becomes to big. Furthermore each design should avoid to inject any current into the substrate that could act as the start of latch-up. Bandgap circuits from Sect. 6.2.1,
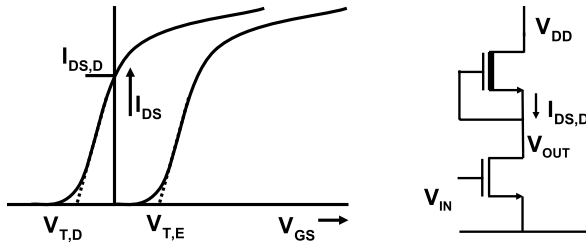
Fig. 2.45 The threshold voltage of an NMOS transistor can be positive or negative. A negative threshold voltage leads to a current when source and gate are connected together. The *right-hand plot* shows a characteristic inverter arrangement

use these pnp bipolar transistors and therefore create deliberately a substrate current. It is important that this current is diverted as close to the device as possible. In advanced CMOS the forward voltages of the junctions have become higher while the power supply voltage is lower. It is sufficient to use these processes within the allowed operating range to eliminate the risk of latch-up.

2.5.10 Enhancement and Depletion

In digital circuits a MOS transistor is mostly used as a switch. For that purpose a positive threshold voltage for an NMOS transistor and a negative threshold voltage for a PMOS transistor allow to switch on and switch off the flow of current with voltages that are within the supply range. Transistors with thresholds that aim at zero current at $V_{GS} = 0$ are called "normally-off". The carriers have to be brought into the channel to get conduction. Also the term "enhancement MOS" is used for this type of devices.

Threshold-adjust ion implantations in both NMOS and PMOS transistors can shift the threshold to a value, where the transistor conducts at $V_{GS} = 0$, see Fig. 2.45. These depletion-mode transistors are "normally-on". The implant that creates the normally-on behavior causes a buried channel between source and drain. The physics of the buried-channel transistor is slightly different from surface-channel enhancement devices, as the current dominantly flows through the implanted buried layer and not along the insulator interface [13, p. 456]. A simple threshold voltage shift will do for most hand calculations. With the source connected to the gate the transistor behaves as a current source. In the early years of semiconductor manufacturing large circuits were built based on enhancement NMOS transistors as pull-down devices and depletion NMOS transistors as load devices.

Today still some applications require a depletion device in order to have a conductive path available at the moment when no supply voltages are present. An example is a pass-on switch in a video recorder. The antenna signal has to pass the video recorder on its way to the television set even when the recorder is switched off. The switch will divert the antenna signal into the recorder when the recorder is active.

In some processes it is allowed to use transistors without threshold modulation implantations. Some designers refer to these devices as "natural transistors". With the common levels of wafer doping, these devices show a threshold which is close to zero. Some low-voltage track-and-hold circuits use these devices.

2.5.11 Models

The complex nature of MOS transistor physics can only be handled in sufficient precision by means of models. Starting from the simple square-law model the art of modeling MOS physics has developed over the years. It is necessary to carefully distinguish between the inherent mathematical description of the model and the necessary parameters that are supplied by the process foundry: the parameter characterization. The mathematical description is often freely available, however accurate parameters describing the typical process outcome and its accepted deviations (often called: worst-case and best case, or the slow and fast corners of the process) is more cumbersome. The quality of the model description together with accurate parameters determines the outcome of the simulation. This is a list of the main MOS models in use:

- The BSIM model originates from modeling work at the University of California in Berkeley. The transistor is described starting from the source terminal, which creates a physically non-existent asymmetry with the drain. This model originally aimed at digital circuit design and was gradually extended with all kind of analog-relevant phenomena. The result is a 300 item parameter list which is only fully understood by a handful of specialists. A starting point for reading is offered in [34]. The BSIM model is mainly in use in U.S. companies.
- The EKV model is named after the fathers: Christian Enz, Francois Krummenacher and Eric Vittoz. This model takes the substrate as a reference and is well-suited for analog circuit design problems where transistors are used in various regimes. Source and drain are equivalent terminals and an excellent model behavior is achieved for resistor-like applications. Also the subthreshold behavior is well-described. The model is popular in academia as it allows a good insight on the underlying physics, see [35].
- The PSP-model is the successor of the Philips Research developments. This model aims at a description that is based on the device physics "compact modeling". The model takes the surface-potential in the channel as a starting point and allows with 40 parameters an accurate description in all regimes. This model is preferred by the modeling council. The PSP model predicts the standard analog parameters well: distortion, noise, mismatch, output conductance, [36].
- Next to the above models there are many alternatives. Sakurai describes a model where the quadratic term is replaced by a parameter α, [37]. This is a relatively simple approach to a sufficient model for digital applications. Redman-White and his research group described Silicon-on-Insulator transistors and the design consequences [38].

Despite all modeling effort a designer should always be careful in believing what a model-based simulation predicts. Most simulated effects in circuits can be understood with physics, other effects require knowledge of simulator artifacts.

2.6 Network Theory

Electronic functions are built from networks of passive and active components. A complex impedance contains resistive elements and reactive elements: $Z = R + jX$. The reactance X can be positive corresponding to an inductive behavior, or negative for capacitive behavior. There are various ways to analyze, classify or synthesize networks. Kirchhoff's laws form the basis for all these methods.

2.6.1 Kirchhoff's Laws

Kirchhoff has formulated a set of rules for analyzing networks of components. For every node in the network the summation of all currents must yield zero, there is no charge storage in a summation point:

$$\Sigma I = 0 \tag{2.114}$$

where the direction of the current (in or out of the summation node) determines the sign. In the example of Fig. 2.46: $+I_1 - I_2 - I_3 = 0$.

Equivalently for every loop in a circuit the sum of all voltages must be zero:

$$\Sigma V = 0 \tag{2.115}$$

where again the sign of a voltage contribution depends on whether the positive or the negative terminal is met first in the loop. So in Fig. 2.46: $-e + V_1 + V_{23} = 0$.

Adding to these two equations the simple Ohm's equations per element: $V_i = I_i \times R_i$ results in a set of five equations with five unknowns: $V_1, V_{23}, I_1, I_2, I_3$. This set can be solved to yield the currents and voltages as a function of the impedances and voltage source.
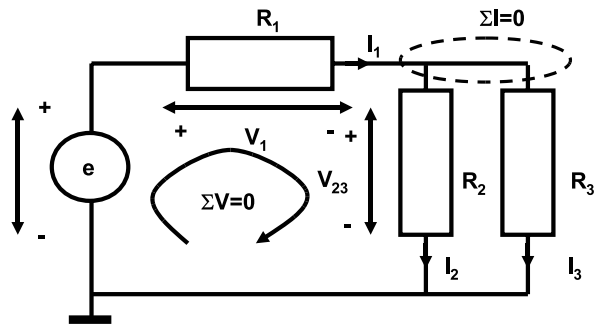
Fig. 2.46 Kirchhoff's law for voltage and current
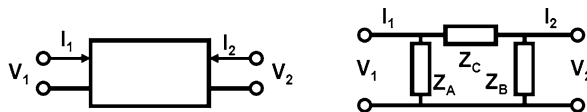
Fig. 2.47 A two-port circuit

2.6.2 Two-port Networks

In electronic systems signals are manipulated by circuits. These circuits can often be reduced to so-called two-port networks where an input signal is applied to one port and an output signal becomes available at a second port. In the left-hand side of Figure 2.47 a general two-port model with its voltages and currents is shown. In order to describe the signal transfer from one terminal pair to another several parameter sets are used. A classical description is the set of z-parameters[20]:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \begin{pmatrix} i_1 \\ i_2 \end{pmatrix}$$

where each of the z-parameters is defined as:

$$z_{11} = \left.\frac{v_1}{i_1}\right|_{(i_2=0)} \qquad z_{12} = \left.\frac{v_1}{i_2}\right|_{(i_1=0)}$$
$$z_{21} = \left.\frac{v_2}{i_1}\right|_{(i_2=0)} \qquad z_{22} = \left.\frac{v_2}{i_2}\right|_{(i_1=0)} \tag{2.116}$$

As an example the z-parameters of Fig. 2.47 (right) are found as:

$$z_{11} = \frac{Z_A(Z_B + Z_C)}{Z_A + Z_B + Z_C} \qquad z_{12} = \frac{Z_A Z_B}{Z_A + Z_B + Z_C}$$
$$z_{21} = \frac{Z_A Z_B}{Z_A + Z_B + Z_C} \qquad z_{22} = \frac{Z_B(Z_A + Z_C)}{Z_A + Z_B + Z_C}$$

Note that $z_{12} = z_{21}$ which is always the case in linear networks and referred to as: "reciprocity". Now the inverse calculations give the following relations:

$$Z_A = \frac{z_{11}z_{22} - z_{12}^2}{z_{22} - z_{12}}$$

$$Z_B = \frac{z_{11}z_{22} - z_{12}^2}{z_{11} - z_{12}}$$

$$Z_C = \frac{z_{11}z_{22} - z_{12}^2}{z_{12}}$$

---

[20]z-parameters have no relation with the z-transform.

2.6.3 Energy and Power

Energy and power are the key metrics in any form of design. Energy is expressed in Joule (1 Joule = 1 Watt-second = 1 Volt-Ampere-second = 1 Coulomb-Volt = 1 kilogram-meter). The first law of thermodynamics states that energy is conserved at all times. The implication is that energies from different origins are related to each other via summation. In the electrical domain energy (the ability to perform work) is defined as:

$$E = \int_{t=-\infty}^{\infty} V(t) \times I(t)\, dt \tag{2.117}$$

Just as in any domain, energy conservation also holds in the electrical domain. Capacitors and coils can store respectively electrical and magnetic energy, while only resistors can dissipate energy or in a more formal way: convert electrical energy in heat.

Power is the energy flow per unit time:

$$P = \lim_{T \to \infty} \frac{1}{T} \int_{t=-T/2}^{T/2} V(t) \times I(t)\, dt \tag{2.118}$$

In DC conditions, the voltage and current are constant and the above formula reduces to: $P = V \times I$.

In order to facilitate design without the need for calculating power integrals, many signals are expressed by their Root-Mean-Square value:

$$\text{Root-Mean-Square value:} \quad V_{\text{rms}} = \sqrt{\frac{1}{T} \int_{t=0}^{T} V^2(t)\, dt} \tag{2.119}$$

This RMS value represents the equivalent DC voltage with the same power level.

If a sinusoidal voltage $V(t) = \hat{V} \sin(\omega t) = \hat{V} \sin(2\pi f t)$ with $T = 1/f$ is applied over a resistor R the resulting power is:

$$P = \frac{1}{T} \int_{t=-T/2}^{T/2} \frac{(\hat{V} \sin(\omega t))^2}{R}\, dt = \frac{1}{T} \int_{t=-T/2}^{T/2} \frac{\hat{V}^2 (1 - \cos(2\omega t))}{2R}\, dt = \frac{\hat{V}^2}{2R}$$

For a sinusoidal signal the power equivalent DC-voltage is $\hat{V}/\sqrt{2}$ of the peak voltage.

Assume two voltages sources e_1 and e_2 in series with load R_{load}. Now the power consumed in resistor R_{load} over a certain period of time, is:

$$P_R = \frac{e_1^2}{R_{\text{load}}} + \frac{e_2^2}{R_{\text{load}}} + \frac{2\text{Cov}(e_1, e_2)}{R_{\text{load}}}$$

In case where e_1 and e_2 are the RMS values of two signals and their covariance (see Sect. 2.1.6) is zero, the total power consists of the sum of the powers of both voltage sources. In this case the RMS voltage over the resistor is defined as:

$$V_{\text{load,rms}} = \sqrt{e_1^2 + e_2^2} \tag{2.120}$$

This form of summing is called: "root-mean-square" sum and is applicable to un-correlated signals.

If both voltages are fully correlated the covariance equals $e_1 \times e_2$, in which case the power equation reduces to Kirchhoff's law: $V_{\text{load}} = e_1 + e_2$. If the polarity of one voltage source is reversed, a full negative correlation appears: $V_{\text{load}} = e_1 - e_2$. In general energy can always be summed in a linear way considering all applicable correlations. Only over time periods, where there is full correlation between voltages and currents over elements, the Kirchhoff voltage laws apply. In all other cases the correlated contributions of each voltage and current source must be taken into account before their energies are summed.

In Fig. 2.48 a voltage source is shown with an internal impedance R_{in}. The internal impedance R_{in} can be derived from the open terminal voltage and the short circuit current: $R_{\text{in}} = V_{\text{open terminal}}/I_{\text{short circuit}}$. This method results in the "Thevenin equivalent circuit" and can be determined for any pair of terminals connected to a circuit composed of linear voltage and current sources and linear impedances. In that case $V_{\text{open terminal}}$ and R_{in} form a full equivalent for the internal circuit. An alternative circuit consists of a current source with the value $I_{\text{short circuit}}$ loaded with a parallel resistor R_{in} which is known as "Norton's equivalent circuit".

Networks with multiple voltage and current sources can be analyzed with the help of the "superposition theorem". This theorem applies to linear networks. It states that the voltage or current in an impedance due to multiple sources can be calculated by summing the effect of each source individually, while replacing all other voltage sources with a short circuit and all other current sources by an open connection.

If the voltage source of Fig. 2.48 is connected to an impedance R_{load}, power will be transferred to this load. The choice of the value of R_{load} determines the transfer of power from the source to the load. Three "extreme" regimes can be distinguished: see Table 2.23.

An infinite load and a zero-load result in no power transfer into the load. The optimum is reached if the real part of the load impedance is equal to the real part
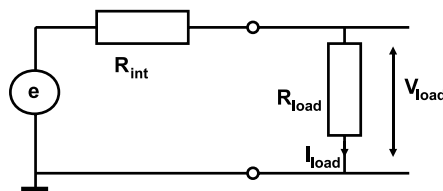
Fig. 2.48 The voltage source has an internal resistance and is loaded with a impedance R_{load}

Table 2.23 Various modes of transfer

Maximum output voltage	$R_{\text{load}} = \infty$, $V_{\text{load}} = V_{\text{open terminal}}$
Maximum output current	$R_{\text{load}} = 0$, $I_{\text{load}} = I_{\text{short circuit}}$
Maximum power	$R_{\text{load}} = R_{\text{in}}$, $V_{\text{load}} = 0.5 V_{\text{open terminal}}$
	$I_{\text{load}} = 0.5 I_{\text{short circuit}}$, $P_{\text{load}} = 0.25 V_{\text{open terminal}} \times I_{\text{short circuit}}$

Fig. 2.49 The general
scheme of a feedback system
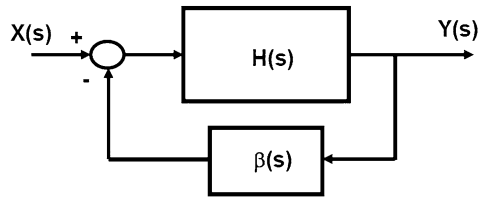
of the internal impedance and the imaginary parts are of opposite sign, also called
complex conjugated. This is the maximum power transfer theorem or the "Jacobi
law".

2.6.4 Feedback

A generally used technique in many systems and of course in analog IC design
is feedback. In Fig. 2.49 the transfer function $H(\omega)$ is connected in a feedback
loop via the feedback path β. Analysis of the resulting transfer results in a Laplace
description:

$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1 + \beta(s)H(s)} \tag{2.121}$$

In the frequency range where $\beta H(\omega) \gg 1$ the transfer reduces to $1/\beta$. So the feed-
back function determines the transfer, independent of the properties of the initial
transfer function $H(\omega)$.

The feedback can be added to the input signal or subtracted from the input sig-
nal, assuming that there is no inversion present in $\beta H(s)$. Negative feedback or
degenerative feedback as drawn in Fig. 2.49 counteracts the signal and is useful to
control the overall amplification of a system and to reduce artifacts and distortion
in H. Positive feedback or regenerative feedback is in phase with the input signal.
The feedback signal takes over the role of input signal and is only limited by the
physical boundaries of H, such as the power supply voltages.

The main criterion for negative feedback systems is the stability of the loop. The
signal passing through $H(s)$ and $\beta(s)$ is delayed and amplified. For low-frequency
signals the delay can be ignored and the signal is subtracted at the input node. The
high gain of the loop will try to equalize the return signal to the input signal thereby
defining the output at $Y = X/\beta$. At higher frequencies a point is reached where the
delay equals half of the signal's period time. Now the inversion at the summation
point will create a reinforcing signal. If the total gain at that frequency equals "one"
this signal will feed itself and an oscillator is created. A gain larger than "one" will
lead to exponential growth, and a gain smaller than "one' to an exponential decay.

Figure 2.50 shows one of the analysis methods for feedback systems: the "Bode-
plot". For an open loop analysis the gain and the phase are plotted versus the fre-
quency. Proper analysis requires to take the loading of the loop at the position where
the loop is opened into account. In this example the open loop transfer function has
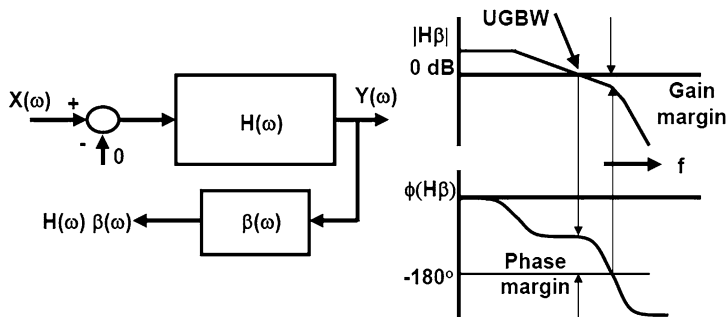
Fig. 2.50 Open loop analysis of stability using the Bode plots

a low frequency pole and two closely spaced poles at higher frequencies. At DC
the transfer function is characterized by the DC-gain, which after the first pole turns
into an attenuation of a factor of 2 for every doubling of the frequency. In control
and communication theory a logarithmic measure is used for the power gain and
amplitude gain:

$$A(\text{in dB}) = 10^{10}\log\left(\frac{P_Y}{P_X}\right) = 10^{10}\log\left(\frac{Y^2}{X^2}\right) = 20^{10}\log\left(\frac{Y}{X}\right) \quad (2.122)$$

Assuming equal impedances for X and Y. This relation results in a decay of 6 dB
per octave or 20 dB per decade. The bandwidth given by the 0 dB crossing ($1\times$
amplification) of the gain curve is called the Unity Gain Bandwidth (UGBW). If a
system is first order before crossing 0 dB, the UGBW equals the gain-bandwidth
product.

The form of the denominator in the transfer function is essential. The main fea-
tures of a transfer function can often be characterized by the following second order
function[21]:

$$\frac{1}{1 + \beta H(s)} = \frac{1}{s^2 + 2\alpha s + \omega_0^2} \quad (2.123)$$

If this denominator is close to "0" the transfer function mathematically goes to an
unbounded state. Physically the system will force its output to a limit (e.g. sup-
ply voltage) or it will oscillate. A zero denominator corresponds to the condition[22]
$|\beta H(s)| = 1$ and $\arg(\beta H(s)) = -180°$.

In the frequency domain $s = \alpha + j\omega$ is replaced by $j\omega$, which allows inspec-
tion of the frequency behavior. The resulting frequency amplitude and phase plots
("Bode-plot") allow to characterize the feedback system by two numbers: the phase
margin is the remaining amount of phase with respect to $-180°$ at the point where

[21] When facing higher order functions, it is advisable to try to reduce first to the two most important
terms (mostly the two low-frequency poles) and subsequently consider the effect of the higher order
terms.

[22] Here the minus sign at the addition point is crucial.

the amplitude gain is "1". The gain margin is the amount of attenuation below the "0 dB" level at the frequency where the phase is $-180°$. The abstraction into phase and gain margin is the result of a more complex analysis in the Laplace domain and allows a graphical interpretation of the instability. If the amplification of the loop is larger than "1" at the frequency with $180°$ phase, the loop is theoretically stable. Any loss of amplification will turn the loop instable. This situation is called "conditionally stable" and not advised.

A more formal inspection of stability requires to find the time domain response of the transfer. This denominator equation must be factorized using Table 2.2, which leads in the Laplace domain to:

$$s_{1,2} = -\alpha \pm \sqrt{\alpha^2 - \omega_0^2}$$

This factorization can result in three solutions depending on the polarity of the term under the square-root.

$$
\begin{aligned}
\alpha^2 - \omega_0^2 > 0 &\rightarrow \text{overdamped} \qquad & f(t) &= c_1 e^{\gamma_1 t} + c_2 e^{\gamma_2 t} \\
& & \gamma_{1,2} &= -\alpha \pm \sqrt{\alpha^2 - \omega_0^2} \\
\alpha^2 - \omega_0^2 = 0 &\rightarrow \text{critically damped} & f(t) &= c_1 e^{-\alpha t} + c_2 t e^{-\alpha t} \\
\alpha^2 - \omega_0^2 < 0 &\rightarrow \text{underdamped} & f(t) &= c_1 e^{-\alpha t} \cos(\omega_0 t + \phi)
\end{aligned}
\tag{2.124}
$$

$f(t)$ shows the basic terms in the time response. Most practical loops are over-damped, simply because they need to be stable. The phase margin is somewhere between $90°$ and $120°$. A critically damped feedback loop has the fastest time response that still monotonically moves from one level to the other. An underdamped feedback loop shows oscillatory behavior with an overshoot or undershoot.

The frequency response of the denominator is found by the substitution of $s \rightarrow j\omega$. Its absolute value is:

$$\left| \frac{1}{-\omega^2 + 2j\alpha\omega + \omega_0^2} \right|^2 = \frac{1}{\omega^4 + \omega^2(4\alpha^2 - 2\omega_0^2) + \omega_0^4}$$

If $2\alpha^2 > \omega_0^2$ the frequency response of the closed loop transfer due to the denominator is monotonically decreasing. For the transfer of the complete loop the denominator must be multiplied with the numerator.

When $2\alpha^2 < \omega_0^2$ the frequency response will show peaking at $\omega^2 = \omega_0^2 - 2\alpha^2$. The damping factor α shifts the effective resonance to a lower frequency. The amplitude peaking due to the denominator at the resonance frequency with respect to the value at DC is:

$$
\sqrt{\frac{(\omega^4 + \omega^2(4\alpha^2 - 2\omega_0^2) + \omega_0^4)|_{\omega=0}}{(\omega^4 + \omega^2(4\alpha^2 - 2\omega_0^2) + \omega_0^4)|_{\omega^2=\omega_0^2-2\alpha^2}}} = \frac{\omega_0^2}{2\alpha\sqrt{\omega_0^2 - \alpha^2}}
$$

$$
\approx \frac{\omega_0}{2\alpha} = Q \tag{2.125}
$$

For small damping the last term becomes equal to the quality factor Q as defined for filters, see Sect. 2.6.8. With $Q = \omega_0/2\alpha$ the transfer function $R(\omega)$ in the frequency domain for a second order resonance becomes:

$$R(\omega) = \frac{1}{\omega^2 + 2\alpha\omega + \omega_0^2} = \frac{1}{\omega^2 + \omega_0\omega/Q + \omega_0^2} \tag{2.126}$$

The derivation for these negative feedback cases equally applies to positive feedback situations, as occurs in cross-coupled latches. The transfer function in the Laplace domain

$$\frac{Y(s)}{X(s)} = \frac{H(s)}{1 - \beta H(s)}$$

is solved in a similar way. If the function $H(s)$ is composed of two integrator stages $1/s\tau$ with unity feedback, the analysis results in:

$$f(t) = \frac{1}{2}c_1 e^{+t/\tau} - \frac{1}{2}c_2 e^{-t/\tau} \tag{2.127}$$

where the first term will cause an exponential growth of the starting value c_1.

2.6.5 Opamps and OTAs

In electronic circuits amplification is often reduced to the abstract concept of an operational amplifier (opamp). An opamp has a positive and negative input terminal and one output or a pair of positive and negative output terminals. Ideally an opamp has an infinite gain, no input current and zero output impedance. Figure 2.51 shows a general operational amplifier configuration. Under the above conditions, the analysis gives:

$$V_{\text{out}} = \frac{R_4}{R_1}\frac{R_1 + R_2}{R_3 + R_4}V_{\text{in2}} - \frac{R_2}{R_1}V_{\text{in1}} \tag{2.128}$$

Specific topologies are found by setting resistances to zero or infinite. In Fig. 2.52 the three classical topologies are given: inverting, non-inverting and unity gain.

The inverting opamp uses two impedances to define the transfer. Two resistors create a negative amplification. Complex impedances will lead to frequency dependent transfers. The choice of the signal-ground level at the positive terminal of the
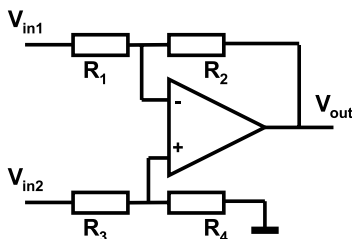
Fig. 2.51 A general operational amplifier configuration
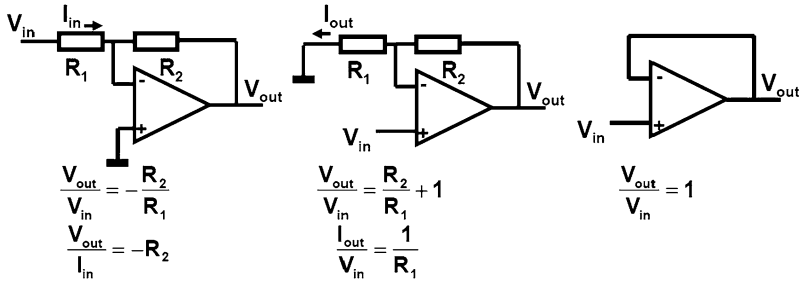
Fig. 2.52 Three classical opamp configurations: inverting, non-inverting and unity gain
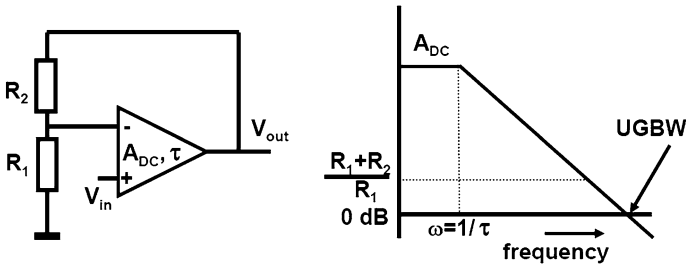
Fig. 2.53 An operational amplifier with limited gain and bandwidth. Open loop transfer (*bold line*) and closed loop transfer (*dotted line*)

opamp determines the voltage level around which the input and output will move. In older opamp designs with considerable input currents a resistor of the same value as R_1 is placed between the positive input and ground to compensate offset.

The non-inverting opamp configuration has the advantage of zero load on the input voltage, however the input stage of the opamp is moving through the entire input range, requiring a good Common-Mode Rejection-Ratio (CMRR) for the frequency range of interest. The feedback factor is determined by the ratio of the impedances. A low feedback factor reduces the stability requirements.

The most simple but also the most demanding topology is the unity-gain amplifier. CMRR is an issue and on top of that this configuration has a unity feedback factor which requires stability at the highest frequencies.

Figure 2.53 (left) shows an opamp with two non-idealities: a limited gain A_{DC} and a first order pole. The transfer is: $H(\omega) = A_{DC}/(1 + j\omega\tau)$. Straight-forward analysis or applying the feedback formula with H and a feedback factor $\beta = R_1/(R_1 + R_2)$ gives:

$$
\begin{aligned}
\frac{V_{\text{out}}}{V_{\text{in}}} &= \frac{R_1 + R_2}{R_1} \frac{1}{\frac{R_1+R_2}{A_{DC}R_1} + j\omega\tau \frac{R_1+R_2}{A_{DC}R_1}} \\
&\approx \frac{R_1 + R_2}{R_1} \frac{1}{1 + j\omega\tau \frac{R_1+R_2}{A_{DC}R_1}}
\end{aligned}
\tag{2.129}
$$

Inspection of the transfer function shows that the overall transfer equals the inverse of the feedback factor. The brute force amplification of the amplifier is tailored to the users needs by a simple impedance ratio. Another relevant observation is that the overall speed is not determined by the opamps physical pole described by τ but by this time constant divided by the total loop gain. The performance of an opamp in a circuit is therefore best characterized by monitoring the unity-gain feedback frequency $\omega_{\mathrm{UGBW}} = A_{DC}/\tau$ and the inverse of the feedback factor β. These two parameters set the performance curve. At first glance it may seem strange that a circuit can react much faster than its physical pole allows. If a step voltage is applied on the input, the circuit will start charging the physical pole. After a short time period t_r a voltage change on the pole of $V_{\mathrm{step}} \times t_r/\tau$ occurs. With the large amplification factor A_{DC} this small voltage change on the pole is already sufficient to create an output voltage equal to the input swing.

In CMOS circuit design a Miller-opamp, see Sect. 2.7.11, shows a first-order opamp behavior. Its output impedance is rather low. Many other topologies resemble an operational transconductance amplifier (OTA). This type of amplifier has a relatively large output impedance and can therefore better be characterized by the current difference at the output as a result of an input difference.

$$i_{\mathrm{out}} = g_m v_{\mathrm{in}} \tag{2.130}$$

Loading an OTA with a DC current inevitably means an offset voltage on the input terminals. In applications with a full capacitive load or applications where the transconductance is a functional part of the circuit, as in $g_m C$ filters, OTAs replace opamps.

2.6.6 Differential Design

In CMOS technology signals are mostly represented by voltages. In Fig. 2.54 three common methods are shown to process these signals. The simplest form uses a signal defined by the voltage difference between a signal wire and ground. This "single-ended" format is sensitive to changes in the potential of the ground line, e.g. due to voltage drops over resistive wires. In order to circumvent this problem "differential design" defines the signals between two separate wires. The second scheme in Fig. 2.54 shows two amplifiers each processing a single-ended signal, however in
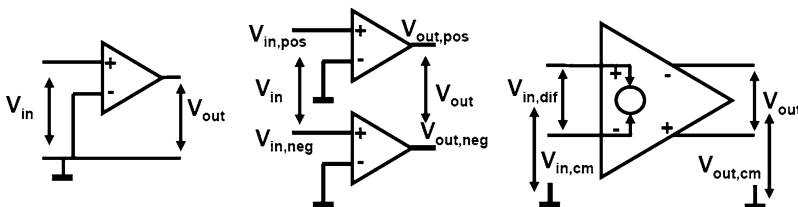

Fig. 2.54 Single-ended amplifier configuration, pseudo-differential and full-differential design

their combination processing the difference between both signal lines. This topology is called "pseudo-differential". A change in ground potential will affect both signals in the same way and will not affect their difference. If both input signals are in perfect anti-phase another advantage appears. If an input signal $0.5V_a \sin(\omega t)$ is being distorted by the amplifier, distortion products at odd and even multiples of the input frequency will show.

$$v_{\text{out,pos}} = 0.5V_a \sin(\omega t) + bV_a^2 \sin(2\omega t) + cV_a^3 \sin(3\omega t) \qquad (2.131)$$

As the lower amplifier receives the inverse input signal $-0.5V_a \sin(\omega t)$, its output will look like:

$$v_{\text{out,neg}} = -0.5V_a \sin(\omega t) + bV_a^2 \sin(2\omega t) - cV_a^3 \sin(3\omega t) \qquad (2.132)$$

Consequently the differential amplification of the signal $V_a \sin(\omega t)$ is the difference between these two:

$$v_{\text{out}} = V_a \sin(\omega t) + 2cV_a^3 \sin(3\omega t) \qquad (2.133)$$

The even harmonics disappear in this configuration. Full cancellation is achieved if both signals are in perfect opposite phase. If the positive side of the signal is e.g. $0.55V_a \sin(\omega t)$ and the negative side: $0.45V_a \sin(\omega t)$ it is clear that a fraction of the even order distortion will show up in the output signal.

   The pseudo-differential design has the advantage of a simple design structure and defined DC-levels at its outputs. Yet any difference between both paths will directly affect the overall signal-processing quality. A full-differential design is shown in the third scheme of Fig. 2.54. Both, the positive signal component and the negative signal component are processed by the same hardware. A typical construction uses a differential transistor pair ("long-tailed pair") to subtract both components from each other, and to avoid to have to process the common part of both signals. This construction allows a high differential gain A_{diff}, with a low common gain A_{com}. The overall amplification can now be written as:

$$v_{\text{out}} = A_{\text{com}}(v_{\text{in,pos}} + v_{\text{in,neg}}) + A_{\text{diff}}(v_{\text{in,pos}} - v_{\text{in,neg}}) \qquad (2.134)$$

The first part reflects the effect of common input voltage $V_{\text{in,cm}}$ in the output common voltage $V_{\text{out,cm}}$, while the second part is the differential gain. The ratio between both is called "the common mode rejection ratio" or CMRR:

$$\text{CMRR} = 20^{10}\log\left(\frac{A_{\text{com}}}{A_{\text{diff}}}\right) \qquad (2.135)$$

Differential design suppresses the impact of supply voltages changes on the output. The parameter defining the impact is the "Power Supply Rejection Ratio" or PSRR:

$$\text{PSRR} = 20^{10}\log\left(\frac{\Delta V_{\text{out}}}{\Delta V_{\text{supply}}}\right) \qquad (2.136)$$

This ratio defines the transfer from power supply variations to the output of the amplifier. Both the PSRR and the CMRR are frequency dependent.
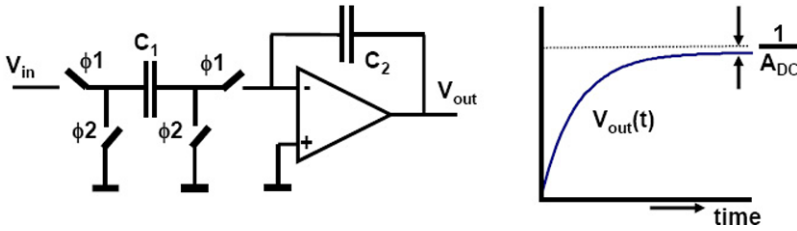
Fig. 2.55 An elementary switched-capacitor circuit

2.6.7 Switched-capacitor Circuits

CMOS technology allows to create excellent switches and capacitors. The switched-capacitor technique utilizes this advantage to implement time-discrete filters and data converter processing. Various aspects are discussed in [39–45]. Figure 2.55 shows a basic integrator configuration. After every clock cycle a charge sample $C_1 V_{in}$ is transferred to capacitor C_2. The switch signals ϕ_1 and ϕ_2 represent the normal and the inverted phase of the sample frequency with some precaution to avoid a time overlap. In the drawn positions, the output will be inverted with respect to the input. Interchanging ϕ_1 and ϕ_2 on one right-hand side of C_1 will create a non-inverted output. This switch and capacitor arrangement is not sensitive to parasitic capacitance on a node. The parasitic capacitors on the left-hand side of the capacitor are charged and discharged by the input and ground. The node connected to the input of the opamp moves around shortly after switching but returns to its (virtual) ground level. Therefore there is no net charge exchange with the charge stored on the input capacitor.

At the moment the input capacitor C_1 is connected to the virtual ground of the opamp the voltages around the opamp will change instantaneously. The charge on C_1 is redistributed over the connected capacitors. If the opamp has a zero-ohm output (or is loaded with a large capacitor), the charge will be divided over $C_1 + C_2$, creating a voltage change on the opamp input of $V_{in} C_1/(C_1 + C_2)$. After this initial phase the opamp will act to reduce the input offset to zero and transfer the remaining charge in C_2. On the other hand, if the opamp is more of a transconductance type with a high-ohmic output, the charge on C_1 will see only the parasitic capacitors and develop a large input swing on the input. This swing can result in opening substrate diodes, opamp slewing or other undesired opamp behavior. A method to reduce these effects is to connect a capacitor between the virtual ground and the real ground.

In z-domain notation the n-th sample at the output equals:

$$V_{out}(z)z^n = V_{out}(z)z^{n-1} + \frac{C_1}{C_2} V_{in}(z)z^n \rightarrow \frac{V_{out}(z)}{V_{in}(z)} = \frac{C_1/C_2}{1 - z^{-1}} \quad (2.137)$$

In case the gain of the amplifier is limited to A_{DC} the output voltage will show a small deviation of the ideal case [39]:

$$\frac{V_{\text{out}}(z)}{V_{\text{in}}(z)} = \frac{C_1/C_2}{1 - z^{-1}} \frac{1}{1 + \frac{C_1 + C_2}{A_{DC} C_2}} \approx \frac{C_1/C_2}{1 - z^{-1}} \left(1 - \frac{C_1 + C_2}{A_{DC} C_2}\right) \qquad (2.138)$$

The fraction V_{out}/A_{DC} is found as an offset at the input of the amplifier. Two major consequences for the design of analog-to-digital conversion must be considered:

- The output does not reach the full signal swing. When no corrective measures are taken, the first order resolution of the converter will be limited to $1/A_{DC}$.
- A small fraction of the output voltage remains present at the input. The next charge packet is not formed as the difference between V_{in} and (virtual) ground, but related to $V_{\text{in}} - V_{\text{out}}/A_{DC}$. Moreover charge must be supplied to any parasitic capacitance on the input node of the opamp.

A high DC-gain will reduce both problems.

The transfer characteristics of the switched-capacitor circuit are determined by the switching frequency and the capacitor ratio. The dependence on the ratio and not the absolute capacitor value allows to design technology independent analog processing structures such as filters and analog-to-digital converters. Mismatch in capacitor ratio affects the overall performance and, just as the gain error, must be taken into account in designing a circuit.

$$\left| \frac{C_{1,\text{real}}}{C_{2,\text{real}}} - \frac{C_{1,\text{ideal}}}{C_{2,\text{ideal}}} \right| < \frac{1}{A_{DC}} \qquad (2.139)$$

Another aspect of the choice of capacitor values is the accumulation of noise. Every switching action on a capacitor introduces an independent packet of kT/C noise, see Sect. 3.1.3. One charge transfer cycle with two switch configurations results in an input referred noise of $v_{\text{noise}}^2 = 2kT/C$. Next to this, the noise of the opamp in the form of opamp referred input noise is sampled into the system and processed by the switched-capacitor circuit. For a short-hand analysis, the contributions of the switched-capacitors can be added up in the energy domain. For an extensive analysis see e.g. [45].

The bandwidth of the switched-capacitor integrator depends on the unity gain bandwidth of the opamp loaded with the relevant capacitors (C_2 and succeeding stages). During the transfer of a charge packet from C_1 into C_2 the feedback configuration of the opamp is determined by the ratio $\beta = C_2/(C_1 + C_2)$. Normally the integration capacitor is the largest, so $\beta \approx 1$. The settling time constant is thereby determined by the unity gain bandwidth: $\omega_{\text{UGBW}} = A_{DC}/\tau$ and feedback factor β.

$$\tau_{sc} = \frac{\tau \beta}{A_{DC}} = \frac{\beta}{\omega_{\text{UGBW}}} \qquad (2.140)$$

In order to obtain a settling error smaller than $1/A_{DC}$, A_{DC} should be in the range of ≈ 1000–$10\,000$ for a $N = 10$–14 bit analog-to-digital converter. A time period T_{sc} is needed of some 7 to 10 time constants.

$$e^{-T_{sc}/\tau_{sc}} < \frac{1}{A_{DC}} < 2^{-N} \qquad (2.141)$$

Depending on this choice and the duty cycle of the charge transfer clock, the unity-gain bandwidth should exceed the charge transfer clock rate by 30–100%.

A second speed aspect is the slew-rate of the circuit. Most likely the output combines the largest voltage swing with the largest load capacitance. These two elements define a charge that has to be supplied within the time constant τ_{sc}. The slew rate current is:

$$I_{\text{out,slew}} = \frac{V_{\text{out,max}}(C_2 + C_{\text{load}})}{\tau_{sc}} \tag{2.142}$$

If the biasing of the output stage does not allow to deliver this current in both directions, distortion will occur. This distortion can be accepted if sufficient settling time allows to reach the linear mode of operation. If not, the slew-rate will turn into real distortion. For a noise discussion see Fig. 4.20.

2.6.8 Filters

A filter selects certain portions of a signal and separates them from the rest. An electronic filter can select in the amplitude domain (level dependent), in the time domain (like a multiplexer) or in the frequency domain. In the frequency domain two major classes of filters exist: time-continuous and discrete-time. Discrete-time filters are discussed after the sampling theory is introduced and can be found in Sect. 3.2.

The class of linear time-invariant (LTI) filters is by definition invariant for signal amplitudes and for time events. This is a dominant class of filters in microelectronics. The transfer function $h(\tau)$ describes the behavior in the time domain. This function specifies the output of the filter for a unity step input. The parameter τ reflects the delay between the input excitation and the resulting output. In causal filters and systems the output changes in time after the input has changed and the delay time is positive. The transfer function for negative τ equals zero in causal systems:

$$h(\tau) = 0, \quad \forall \tau < 0 \tag{2.143}$$

The response to an arbitrary input signal $x(t)$ is found by considering that a past input signal $x(t - \tau)$ arrives at time t at the output because it is delayed by the filter by a delay τ. The corresponding multiplication coefficient for the signal with a delay τ is $h(\tau)$. The total output signal $y(t)$ equals the summation of the delayed input signals multiplied by their coefficient $h(\tau)$ for all possible (positive) values of τ:

$$y(t) = x(t) * h(t) = \int_{\tau=0}^{\tau=\infty} h(\tau)x(t - \tau)\, d\tau$$

$$= \int_{\tau=0}^{\tau=\infty} h(t - \tau)x(\tau)\, d\tau \tag{2.144}$$

This integral is a convolution function and is most easily evaluated in the Laplace domain:

$$Y(s) = H(s)X(s) \tag{2.145}$$

where $Y(s)$, $H(s)$ and $X(s)$ are the Laplace transforms of the originating time functions $y(t)$, $h(t)$ and $x(t)$. Therefore most linear time-invariant filters are analyzed and synthesized in the Laplace domain.

Coils and capacitors implement the integration function in the electrical domain. Frequency-domain filters consist of combinations of passive (R, L and C) and active elements (g_m). These elements are considered linear and time invariant. In a properly designed electronic filter each coil and capacitor will contribute to the filtering of the signals. With N coils or capacitors the transfer curve will show a decay proportional to ω^N. N is called the order of the filter. A properly designed third-order filter suppresses the unwanted frequency components with 18 dB per octave frequency shift. A filter with higher suppression needs a higher order.

The inherent property of signal integration and time delay is used in filters to enhance certain frequencies and suppress others. The simple parallel connection of a coil and capacitor shows that behavior. The energy in this circuit moves from the capacitor to the coil and back again. The time needed is the resonance period and is inversely proportional to the resonance frequency, see (2.126):

$$\omega_0 = 2\pi f_0 = \frac{1}{\sqrt{LC}} \tag{2.146}$$

At the resonance frequency the impedances of the capacitance and the inductor are equal in magnitude $\omega_0 L = 1/\omega_0 C$. If an external source drives this circuit with a frequency equal to the resonance frequency, the input signal will be in phase with the signal in the circuit. Its energy will increase and the amplitude will grow. This process continues till a physical limit is reached, such as the break-down of an isolator. Normally the LC circuit will also contain some resistor, in the example of Fig. 2.56 a resistance in series with the coil. The resistor will dissipate some energy, thereby limiting the total energy available. The ratio

$$\alpha = \frac{R}{2L} \tag{2.147}$$

is the damping factor of the circuit and determines the decay of energy. If the circuit is driven with a signal an equilibrium will appear at the point where the source supplies exactly the amount of energy that is dissipated in the lossy element, in this case the series resistor.

The ratio between the resistor and impedance of the coil or capacitor at the resonance frequency is called the quality factor:

$$Q = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 RC} = \frac{1}{R}\sqrt{\frac{L}{C}} = \frac{\omega_0}{2\alpha} \tag{2.148}$$

Compare (2.125) to (2.148). The quality factor also reflects the ratio between the energy swinging between coil and capacitor and the dissipated energy in the resistor. As the resistor adds thermal noise energy, a high Q is a requirement for low-noise
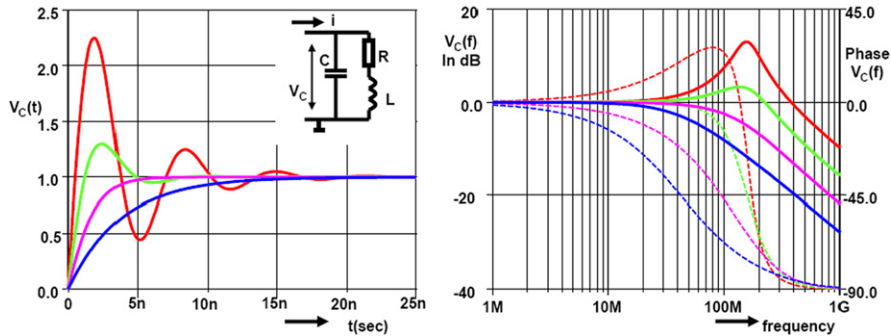
Fig. 2.56 A parallel connection of a capacitor with a coil and resistor with $L = 1\,\mu H$, $C = 1$ pF, and $R = 500, 1000, 2000$ and $4000\,\Omega$. The quality factor varies from 2, 1, 0.5 to 0.25. In the time domain a high quality factor creates ringing in the frequency domain the transfer curve peaks, $Q = 0.5$ corresponds to critical damping

operation (e.g. of an oscillator). Signals that have a different frequency from the resonance frequency will not fit to the signal in the circuit at the resonance frequency and partially extinguish. The larger the frequency difference the stronger the signal will be suppressed.

The equivalent impedance of the *RLC* circuit in Fig. 2.56 is calculated in the Laplace domain:

$$Z(s) = \frac{R + sL}{LCs^2 + sRC + 1} = \frac{1}{LC}\frac{R + sL}{s^2 + 2\alpha s + \omega_0^2} \tag{2.149}$$

The condition for critical damping is: $\alpha = \omega_0$ which reduces the denominator to $(s + \omega_0)^2$ and the time response to an exponential decaying function $e^{-\alpha t}$. If $\alpha > \omega_0$, or the condition of "overdamping", the time domain response of the filter is composed of two exponential decaying terms. If $\alpha < \omega_0$, or the condition of "underdamping", the time domain response is composed of an exponentially decaying sine wave see (2.124).

Combinations of coils and capacitors implement a second order resonance function as in (2.126). In integrated circuits realizations exist that do not require the use of coils. Figure 2.57 shows a popular example of a resonator constructed with a transconductance stage and a capacitor.

Many filter types exist, depending on their function (low-pass, high-pass or band-pass) or a specific property. Three types of filters cover most applications:

• Butterworth filters are characterized by having a maximum flat response near the point of optimization, which is in a low-pass design 0 Hz. This definition is mathematically equivalent with:

$$|H(j\omega)|^2 = \frac{1}{1 + \omega^{2N}} \tag{2.150}$$

where the target is to have a normalized low-pass band up to $\omega = 1$ rad/sec. The definition of maximum flatness means that all high order derivatives of $|H(j\omega)|^2$
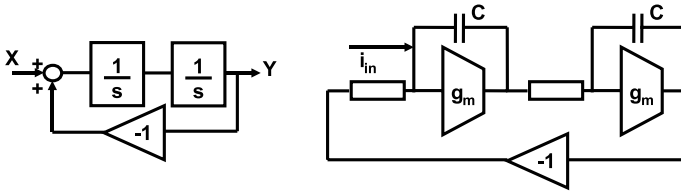
Fig. 2.57 The ideal resonator consists of two integration functions in a unity feedback configuration. *Right*: a practical $g_m - C$ resonator

are zero. In case an additional term proportional to ω^{2K} would be present in the denominator the $2K$th derivative would be non-zero.

The complex poles of a second order filter $s_{1,2} = -a \pm jb$ can now be calculated for a 1 rad/sec cut-off frequency:

$$H(s) = \frac{1}{(s+a)^2 + b^2}$$

$$|H(j\omega)|^2 = \frac{1}{(a^2 + b^2)^2 + (2a^2 - 2b^2)\omega^2 + \omega^4}$$

$$s_{1,2} = -\sqrt{0.5} \pm j\sqrt{0.5}$$

$$H(s) = \frac{1}{s^2 + s\sqrt{2} + 1}$$

Moving the cut-off frequency from 1 rad/sec to e.g. 5 MHz means multiplying the poles with 2π and with this frequency. Comparison of the transfer function of a simple RLC low-pass filter gives:

$$H(s) = \frac{1}{s^2 + s \times 2\pi \times 7.07 \times 10^6 + 4\pi^2 \times 25 \times 10^{12}}$$

$$H(s) = \frac{1}{s^2 + s\frac{R}{L} + \frac{1}{LC}}$$

$$R = 50 \quad \rightarrow \quad L = 1.13 \,\mu\text{H}, \quad C = 909 \,\text{pF}$$

As there are three component values to choose and only two equations to fulfill, the remaining degree of freedom is used to choose the impedance level at 50 Ω.

- Bessel filters optimize the flatness of the group delay. The group delay is defined as:

$$\text{group delay} = -\frac{\partial \arg(H(\omega))}{\partial \omega} = -\frac{\partial \phi(\omega)}{\partial \omega} \tag{2.151}$$

In loose terms a constant group delay preserves the time shape of a signal. As the group delay is the derivative of the phase of the transfer function, a "linear phase" of the transfer is a sufficient condition to preserve the shape of the signal:

$$\frac{\phi(\omega)}{\omega} = \text{constant} \tag{2.152}$$

The requirement of linear phase may apply only in a part of the frequency range, depending on the application.

Table 2.24 Transfer functions for low-pass filters with a cut-off frequency at 1 rad/s

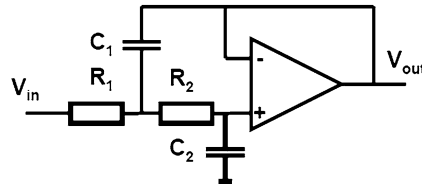| Order	Butterworth	Bessel	Chebyshev
1	$\dfrac{1}{s+1}$	$\dfrac{1}{s+1}$	$\dfrac{1}{s+1}$
2	$\dfrac{1}{s^2+1.41s+1}$	$\dfrac{1}{s^2+3s+3}$	$\dfrac{1}{1.41s^2+0.911s+1}$
3	$\dfrac{1}{s^3+2s^2+2s+1}$	$\dfrac{1}{s^3+6s^2+15s+15}$	$\dfrac{1}{3.98s^3+2.38s^2+3.7s+1}$
4	$\dfrac{1}{s^4+2.61s^3+3.41s^2+2.61s+1}$	$\dfrac{1}{s^4+10s^3+36s^2+95s+105}$	$\dfrac{1}{5.65s^4+3.29s^3+6.6s^2+2.3s+1}$

Fig. 2.58 The Sallen-Key active filter topology

- Chebyshev filters aim to make a steep roll-off of the pass-band into the stop band. Type-1 filters allow ripple in both the pass and stop-bands, while type-2 (or the inverse Chebyshev filter) provides a flat pass band. The mathematical derivations can be found in specialized literature.
- Elliptic (also called Cauer filters) have equal ripple in the pass and stop-bands. These filters have an optimum fast transition between pass and stop band.

From the perspective of the transfer function, filter design is a matter of coefficient choice. Table 2.24 summarizes the transfer functions for Butterworth, Bessel and Chebyshev filters of the first till fourth order.

Some of these denominators can be factored, e.g. the third order Butterworth can be rewritten as $s^3 + 2s^2 + 2s + 1 = (s+1)(s^2+s+1)$. The last formulation has only coefficients "1" and is popular in exercises.

Next to these transfer functions the choice for the topology must be made. A popular scheme is the Sallen-Key realization, see Fig. 2.58 [46]. The transfer function is:

$$H(s) = \frac{1}{s^2 R_1 R_2 C_1 C_2 + s C_2 (R_1 + R_2) + 1}$$
$$= \frac{1}{R_1 R_2 C_1 C_2 (s^2 + s(R_1 + R_2)/R_1 R_2 C_1 + 1/R_1 R_2 C_1 C_2)}$$

The damping and resonance frequency are:

$$\omega_0^2 = \frac{1}{R_1 R_2 C_1 C_2}, \qquad 2\alpha = \frac{R_1 + R_2}{R_1 R_2 C_1} \tag{2.153}$$

If the resistor values change, both damping and resonance frequencies will alter. These component values can be mapped on the second order transfer function in
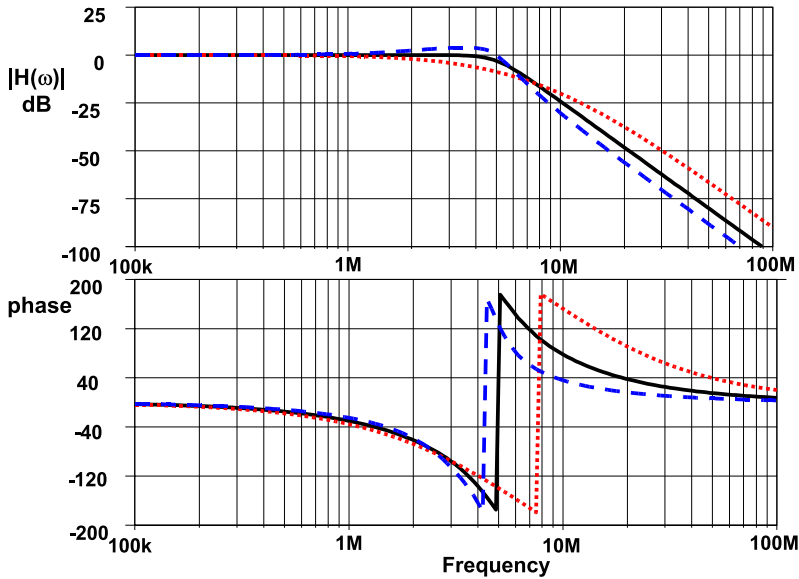
Fig. 2.59 The Sallen-Key transfer function for a fourth order Butterworth (*solid*), Bessel (*dotted*) and Chebyshev (*dashed*) filter. The filter is composed of two cascaded sections of Fig. 2.58

Fig. 2.60 A filter realized with transconductances and capacitors. The first section is a first order filter, the second section a second order filter
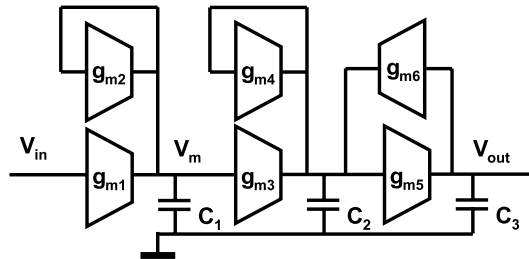
Table 2.24. A result for a fourth order filter is given in Fig. 2.59. Changes in these parameters directly affect the filter performance. The circuit topology for the high-pass Sallen-Key filter is obtained by interchanging the capacitors and the resistors. Of course new component values have to be calculated.

A popular filter technique in the field of integrated circuits is the $g_m - C$ filter [47]. A transconductance g_m can be easily realized as a differential pair and high-quality capacitors are available. Various filter types use a $g_m - C$ filter technique. Figure 2.60 shows a third order filter consisting of a first order and a second order section, with a transfer function:

$$H(s) = \frac{V_{\text{out}}(s)}{V_{\text{in}}(s)} = \frac{g_{m1}/C_1}{(s + g_{m2}/C_1)} \frac{g_{m3}g_{m4}/C_2C_3}{(s^2 + sg_{m5}/C_3 + g_{m4}g_{m6}/C_2C_3)} \quad (2.154)$$

The term g_{m3} determines the damping and is mostly implemented as the load of a differential pair. The terms g_{m4}, g_{m6} relate to the resonance frequency and are to a

Fig. 2.61 The filter coefficients of this feed-forward gm-C filter are implemented with weighted transconductances
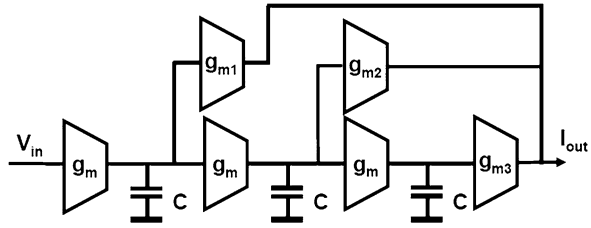
Fig. 2.62 Single transistor amplifiers: grounded source, grounded-gate and grounded drain configurations
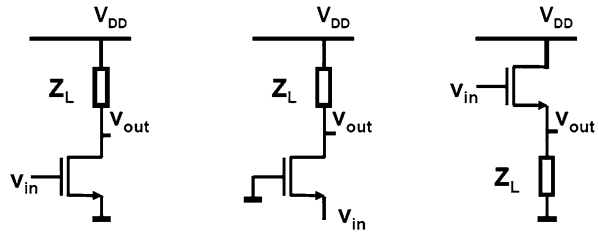
certain extent independent of the damping implementation. With equal g_m and C the transfer is:

$$H(s) = \frac{V_{\text{out}}(s)}{V_{\text{in}}(s)} = \frac{1}{(1 + sC/g_m)(1 + sC/g_m + (sC/g_m)^2)} \qquad (2.155)$$

This is the standard transfer function for a third order Butterworth filter.

Another $g_m - C$ architecture is known as the feed-forward filter as shown in Fig. 2.61. Now the filter transfer function is:

$$H(s) = \frac{I_{\text{out}}(s)}{V_{\text{in}}(s)} = g_{m1}\left(\frac{g_m}{sC}\right) + g_{m2}\left(\frac{g_m}{sC}\right)^2 + g_{m3}\left(\frac{g_m}{sC}\right)^3$$

$$= \frac{(g_{m3} + g_{m2}sC g_{m2}/g_m + g_{m1}s^2C^2/g_m^2)}{(sC/g_m)^3} \qquad (2.156)$$

2.7 Electronic Circuits

2.7.1 Classification of Amplifiers

Electronic circuit design is extensively described in e.g. [48, 49]. Electronic components are called "active" if they allow to increase the signal power. A transistor as an active element forms a simple amplifier stage, Fig. 2.62.[23] Single transistor amplification stages are classified along the terminal that is connected to signal

---

[23]In many circuits a small horizontal bold line indicates the reference ground node. In principle a ground node just indicates the reference terminal that is by definition 0 V, it should not supply any current and appear only once. This requires all circuit diagrams to show all signal and power sources. For better readability these sources are often omitted resulting an improper use of the ground symbol in this book.
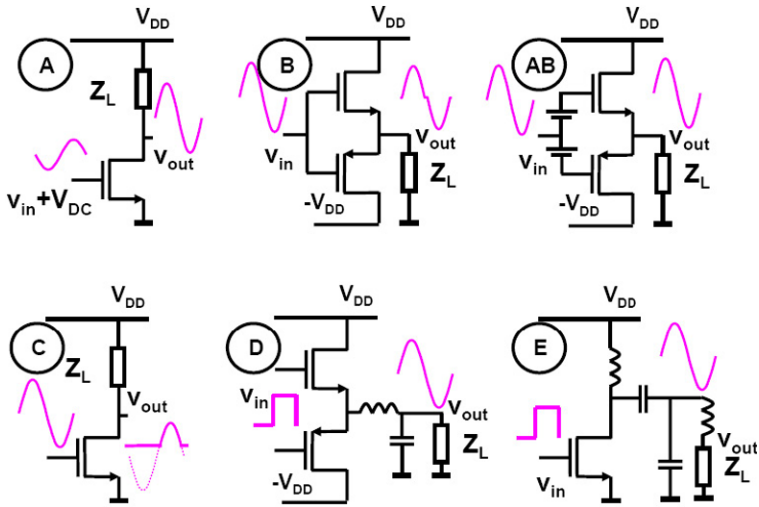
Fig. 2.63 The modes of operation in transistor amplifiers are described by letters: class-A, B, AB, C, D, E, etc.

ground. The left configuration is therefore also called "grounded-source" or "common source". Alternative topologies are called are "grounded-gate" and "grounded-drain". The last circuit is better known as source follower. A grounded-source configuration allows the modulation of current and due to the high output impedance on the drain, creates voltage gain. The disadvantage is the potential feedback from the drain-gate capacitor. The grounded-gate configuration circumvents this problem. RF-transistors in the first transistorized tuners had to be operated in this mode. The input impedance of this configuration equals roughly the transconductance. In tuners this transconductance is used as termination resistor. Finally the grounded-drain configuration modulates the current, the inherent feedback of the source voltage limits the voltage gain to less than one.

The mode of operation provides another more general classification. The circuits in Fig. 2.63 are drawn with one or two transistors. In reality complex circuits implement the general ideas that are common to these simple amplifier stages.

A "class-A" amplifier consumes a DC-current on which a smaller amplitude signal is modulated. Even with a full amplitude sine wave the ratio between the power of the sinusoidal signal and the DC power is theoretically no more than:

$$\frac{\int_{t=0}^{t=T} \hat{V} \sin(\omega t) I \sin(\omega t)\, dt}{\int_{t=0}^{t=T} 2\hat{V} I_{\text{load}}\, dt} = 25\% \tag{2.157}$$

with T as the signal period.

In power amplifiers a whole scale of configurations exist: Class-B used pull-push devices and can obtain an efficiency of $\pi/4$ or 78.5%. At the zero crossing both MOS transistors will be off and a "cross-over" distortion will appear. This

Fig. 2.64 A transistor as amplifier
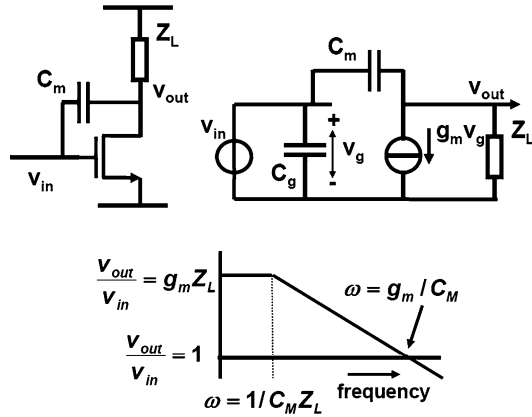
even-order distortion is not appreciated by the human ear.[24] Class-AB biases both push-pull devices in an operating regime where a smooth transition remains possible between positive and negative excursions. This quiescent current reduces the theoretically obtainable power efficiency.

Class-C conducts only part of the signal period and is used to generate harmonics of the (RF) signal. This stage is normally followed by some form of filter. Class-D uses switches and produces a block shape signal, containing the fundamental signal. The signal is obtained after low-pass filtering. The efficiency (85–95%) of this amplifier is limited by the switching: both the $CV^2 f$ switching power as the on-resistance of the switches. The switching frequency is chosen as low as possible, limited by the overall distortion and interference issues.

Class-E operates in combination with a resonant circuit on RF frequencies.

More modes of operation exist if various power supply schemes are used, e.g. class-G uses additional power rails to feed extreme excursions.

2.7.2 One-transistor Amplifier

A first approach to the single-stage amplifier is to analyze a scheme using a simple transistor model as given in Sect. 2.5 and Fig. 2.64. The input voltage to the gate is composed of a DC-voltage with superimposed a signal voltage: $V_{GS} + v_{gs}$. On the drain side also a DC-component and a signal component will appear: $V_{DS} + v_{ds}$. Expanding the square-law equation (2.101) gives the total output current:

$$\frac{WC_{ox}\mu}{2L}(V_{GS} + v_{gs} - V_T + \lambda(V_{DS} + v_{ds}))^2$$
$$= \frac{WC_{ox}\mu}{2L}(V_{GS} - V_T + \lambda V_{DS})^2 + \frac{WC_{ox}\mu}{L}(V_{GS} - V_T + \lambda V_{DS})v_{gs}$$

---

[24]In contrast to odd-order distortion as generated by transformers in tube amplifiers.

$$+ \frac{WC_{ox}\mu}{L}(V_{GS} - V_T + \lambda V_{DS})\lambda v_{ds} + \frac{WC_{ox}\mu}{2L}\left(v_{gs}^2 + \lambda^2 v_{ds}^2 + 2\lambda v_{gs}v_{ds}\right)$$
$$= I_{DS} + g_m v_{gs} + g_{ds}v_{ds} + i_{\text{second order}}$$

The power and bias voltages and currents are denoted with capitals: V_{GS}, while their small signal counterparts are denoted as v_{gs}. In the resulting equation I_{DS} is the DC-current of the circuit. The terms $g_m v_{gs}$ and $g_{ds}v_{ds}$ represent the first-order signal transfer components, which are normally an order of magnitude lower in amplitude than the DC-components. v_{gs} and v_{ds} are treated here as independent variables. In the single-transistor amplifier loaded with an impedance Z_L the voltage drop over this load due to the signal current will generate the v_{ds} component, which then has the inverse polarity of v_{gs} and can be seen as an additional load impedance for the signal current source $g_m v_{gs}$. The quadratic nature of the MOS current formula results also in a number of second order terms that create the distortion in the circuit. In small-signal analysis the DC-terms are considered part of the bias design. The second-order terms are ignored until the distortion analysis becomes relevant. The above equations reduce to the first-order small-signal signal transfer equation:

$$i_{ds} = g_m v_{gs} + g_{ds}v_{ds} \tag{2.158}$$

which is used for hand-calculation. In the above diagram all load components of the stage are combined in an impedance Z_L. The dominant degradation in frequency response is due to a capacitor C_M between drain and gate, feeding back the output signal to the input. In many amplifier topologies this capacitor is deliberately applied to modify the frequency response and is called: "Miller-capacitor".

The schematic diagram can be evaluated by using the small-signal equivalent diagram of Fig. 2.64 (right). The transfer is found as[25]:

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{Z_L(g_m - j\omega C_M)}{1 + j\omega Z_L C_M} \tag{2.159}$$

Inspection of the transfer shows that for low-frequencies the amplification equals the transconductance-output impedance product $g_m Z_L$. The output impedance consists of the load but can also incorporate the output conductance of the transistor. At medium frequencies the second term in the denominator becomes larger than unity and the amplifier enters the first-order decay regime. This regime will ultimately lead to the frequency where the amplification is unity. This frequency is the Unity gain frequency or Unity Gain Bandwidth (UGBW):

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \left| \frac{g_m}{j\omega_{\text{UGBW}}C_M} \right| = 1, \qquad f_{\text{UGBW}} = \frac{g_m}{2\pi C_M} \tag{2.160}$$

Obtaining a large UGBW means that short channel transistors must be used and that capacitive loading is minimized. After that only increasing the current will lead to an improvement in UGBW, although the improvement goes ideally only with a

---

[25]The easiest start to the mathematical analysis is to list Kirchhoff's current equations and start eliminating variables.

Fig. 2.65 Large and small
signal transfer for an inverter
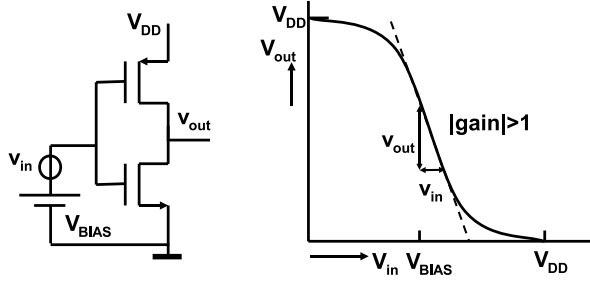
square-root function. Often the speed analysis of a circuit is best approached by
looking at the UGBW point and extrapolating back.

At high frequencies the second term in the numerator of equation (2.159) be-
comes relevant. The current flowing from v_{in} through the capacitor C_M exceeds the
amplified current from the transistor. Ultimately this will cause the phase of the out-
put signal to turn back to 0°, causing stability problems. A popular method to avoid
the current through the capacitor to exceed the amplified current is to put a resistor
in series with the Miller-capacitor of a value $R \geq 1/g_m$.

Another consequence of C_M is seen when the effect on the input impedance of
this circuit is considered. Next to the gate capacitance, the Miller capacitance creates
an input impedance of:

$$\frac{v_{in}}{i_{in}} = \frac{1 + j\omega Z_L C_M}{j\omega C_M (Z_L g_m + 1)} \tag{2.161}$$

For low frequencies the numerator can be set to "1" and the effective input capac-
itance equals the value of the Miller capacitor times $(1 + \text{DC-amplification})$. This
allows to create large load capacitors without the need for spending many pico-
Farads. This effect is only applicable in small-signal situations. If the amplifier is
used as a digital buffer where input and output swing between the power supply
V_{DD} and ground, the Miller effect is reduced to a factor of 2, as the charge on the
capacitor changes from $+C_M V_{DD}$ to $-C_M V_{DD}$.

2.7.3 The Inverter

The large-signal amplification of the inverter in Fig. 2.65 is obviously -1 because
the input swing V_{in} from "ground" to V_{DD} will result in an output swing from V_{DD}
to "ground". In a suitable mid-range bias point, however, a small input signal v_{in} is
amplified via the transconductance of the NMOS transistor and the PMOS transistor
in a small signal current. This current sees the drain conductances of both MOS
transistors as a load, so the amplification is:

$$A_V = \frac{v_{out}}{v_{in}} = -\frac{g_{mN} + g_{mP}}{g_{dsN} + g_{dsP}} \approx -6 \ldots -10 \tag{2.162}$$

In large signal mode the voltage over the Miller capacitor of an inverter (the ca-
pacitor between the input and output terminal) is changing its polarity. Effectively
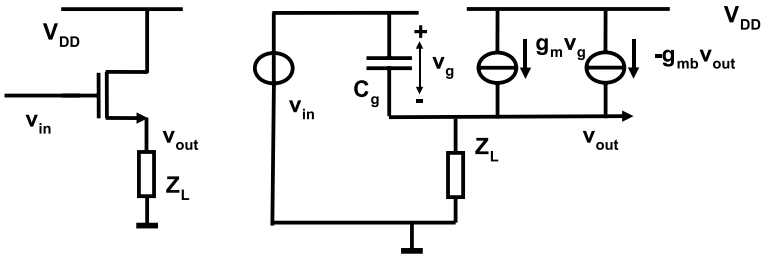
Fig. 2.66 The source follower circuit and its small-signal equivalent diagram

this capacitance is seen on the input as twice its physical size. During the transition through the input region the small-signal contribution of the Miller capacitor is equal to is physical size when the output is on ground or supply level, but can increase to $1 + A_V$ times its physical size during the steepest part of the transition.

2.7.4 Source Follower

In a source-follower configuration the transistor has its drain connected to the power supply. Now the source is connected to the load impedance and the circuit will produce a near copy of the input signal on the output terminal. Power gain is achieved because the output current is larger than the input current, Fig. 2.66. In a source follower the source carries signal. Therefore the equivalent transistor diagram contains also the source-bulk transconductance. In a similar way as in the previous example, the transfer is found:

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{Z_L(g_m + j\omega C_g)}{1 + j\omega Z_L C_g + g_m Z_L + g_{mb} Z_L} \tag{2.163}$$

At low frequencies the transfer is close to $1 - 1/g_m Z_L$ which in a practical circuit is 0.9. The input impedance is calculated as:

$$\frac{v_{\text{in}}}{i_{\text{in}}} = \frac{1 + j\omega Z_L C_g + g_m Z_L + g_{mb} Z_L}{j\omega C_g} \tag{2.164}$$

For low frequencies the input capacitance is reduced to $C_g/(1 + g_m Z_L)$. If the load of the source follower contains a capacitance, the input impedance will show a strong increase at the frequency where the load capacitance reduces the $g_m Z_L$ term. The phase of the overall transfer will rapidly turn with 180°: 90° due to rapidly increasing input capacitance and another 90° in the load capacitor. These two close poles makes a source follower a difficult circuit in a feedback loop. The output impedance is:

$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{Z_L}{1 + j\omega Z_L C_g + g_m Z_L} \tag{2.165}$$

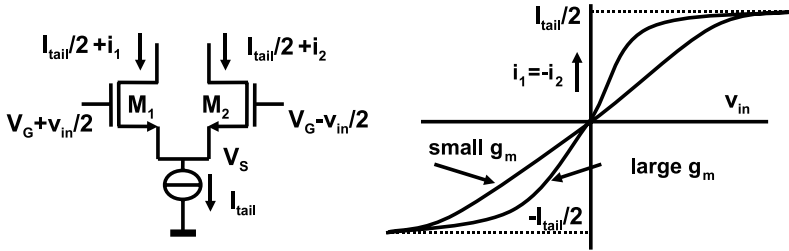which goes to $1/g_m$ in practical design.

Fig. 2.67 The differential pair and the DC transfer curve for two values of the transconductance

Fig. 2.68 The small-signal equivalent scheme of the differential pair. The ideal tail-current source behaves for small signals as a circuit open
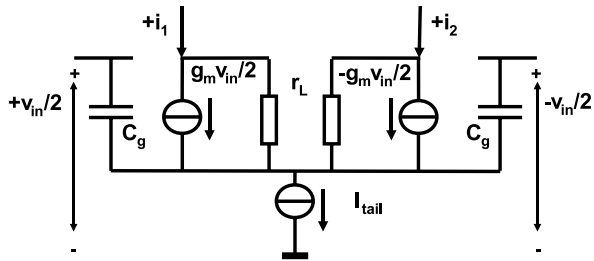
2.7.5 The Differential Pair

The differential pair is a fundamental building block in many electronic circuits. It provides the means for subtraction of signals and for separation of the signals from the DC-voltages needed for operating the devices in the correct biasing points. Figure 2.67 shows the basic topology. The drains of the transistor pair are connected to other elements like resistors or current mirrors. These elements have been left out, to concentrate on the basic properties of this topology.

The symmetry and the ideal tail-current source make that with equal gate voltages the currents will split up in equal parts and that any difference in gate voltages will route a part of the current from one branch into the other. A small input signal v_{in} causes the relatively small current deviations i_1 and i_2. In this mode the differential pair can be replaced by its small-signal equivalent Fig. 2.68. In linear approximation the small-signal current is related to the small-signal input via the transconductance of the input transistors.

$$i = i_1 - i_2 = \frac{dI_{DS1}}{dV_{G1}} v_{in}/2 - \frac{dI_{DS2}}{dV_{G2}} (-v_{in}/2) = g_m v_{in} \qquad (2.166)$$

Assuming equal transistor sizes and bias conditions. In advanced design this approximation is not sufficient and a non-linear analysis is used. The transfer is linear up to a certain input voltage excursion, then saturation will occur. Figure 2.67 (right) shows the current deviations i_1 and (in perfect symmetry) its inverse i_2. When the voltage difference on the gate nodes is close to zero, the transfer is close to linear. At higher voltage excursions the linearity disappears and saturation versus half of the tail current occurs. In bipolar circuits there is a fixed relation between current and

base voltage (2.83). Therefore the relation between the differential collector current and the input voltage is also independent of sizes:

$$\frac{i}{I_{\text{tail}}} = \frac{I_{C1} - I_{C2}}{I_{C1} + I_{C2}} = \frac{e^{qV\text{in}/2kT} - e^{-qV\text{in}/2kT}}{e^{qV\text{in}/2kT} + e^{-qV\text{in}/2kT}}$$
$$= \tanh(q/kT \times V_{\text{in}}/2) \tag{2.167}$$

This equation shows that the transfer of bipolar differential pairs becomes rather non-linear if the differential voltage approaches the thermal voltage of $kT/q = 26$ mV at room temperature. In MOS design there is an additional degree of freedom: the width over length W/L ratio of the gate. Figure 2.67 (right) shows two transfer curves at the same tail current. For one curve a large W/L causes a high transconductance but also results in a small linear v_{in} window. If the W/L ratio is chosen smaller, the transconductance reduces, but the linear window increases.

Further analysis requires to define the relation between the current in the transistor and the control voltage in general terms as a function f: $I_{DS} = f(V_G - V_S)$. This function is in an extended model a complex relation between the controlling voltages, the technology parameters and the currents. For hand calculations on MOS devices the function can take the form of (2.101) or for bipolar devices of (2.83).

For the differential pair controlled by a fully symmetrical signal this results in the set of equations:

$$I_{DS1} = I_{\text{tail}}/2 + i_1 = f(V_{G1} - V_S) = f(V_G + v_{\text{in}}/2 - V_S)$$
$$I_{DS2} = I_{\text{tail}}/2 + i_2 = f(V_{G2} - V_S) = f(V_G - v_{\text{in}}/2 - V_S) \tag{2.168}$$

V_G and I_{tail} are fixed values. V_S and $i_{1,2}$ will vary with v_{in}. Eliminating V_S requires to invert function f to a function h with $V_G - V_S = h(I_D)$, yielding:

$$v_{\text{in}}/2 = h(I_{\text{tail}}/2 + i_1) - h(I_{\text{tail}}/2 - i_2) \tag{2.169}$$

i_1 is (much) smaller than the tail current, so a Taylor expansion can be used for the function h:

$$h(I_{\text{tail}}/2 + i_1) = h(I_{\text{tail}}/2) + \frac{i_1}{1!}h'(I_{\text{tail}}/2) + \frac{i_1^2}{2!}h''(I_{\text{tail}}/2) + \frac{i_1^3}{3!}h'''(I_{\text{tail}}/2) + \cdots$$

In an ideal differential structure with $i = i_1 - i_2$ the even terms are canceling each other while the odd components are doubling:

$$v_{\text{in}} = \frac{i}{1!}h'(I_{\text{tail}}/2) + \frac{i^3}{3!}h'''(I_{\text{tail}}/2) + \cdots$$

This series of terms can be reversed [7, Eq. 3.6.25]:

$$i = \frac{1}{h'(I_{\text{tail}}/2)}v_{\text{in}} - \frac{h'''(I_{\text{tail}}/2)}{3!(h'(I_{\text{tail}}/2))^4}v_{\text{in}}^3 + \cdots$$

When only the third-order term is taken into account and some elementary calculus is applied [7, Eqs. 3.3.9–3.3.11]:

$$h'(I_D) = 1/f'(V_G - V_S) \quad \text{and} \quad h'''(I_D) = \frac{-[f'''f' - 3(f'')^2]}{(f')^5}$$

the differential output current is written as a function of the small-signal input voltage swing:

$$i = f'(V_G - V_S)v_{\text{in}} + \frac{f'''(V_G - V_S)f'(V_G - V_S) - 3(f''(V_G - V_S))^2}{3!f'(V_G - V_S)}v_{\text{in}}^3$$

If a simple bipolar model is used:

$$I_C = f(V_{be}) = I_o e^{qV_{be}/kT}$$

$$\text{with the } n\text{-th derivative:} \quad f^{(n)} = \left(\frac{q}{kT}\right)^n I_C \qquad (2.170)$$

the resulting expression is:

$$\frac{i}{I_{\text{tail}}} = \frac{v_{\text{in}}}{2kT/q} - \left(\frac{v_{\text{in}}}{kT/q}\right)^3 \qquad (2.171)$$

With a first order square-law model for the MOS transistors, the MOS current I_{DS} equals:

$$\begin{aligned}
I_{DS} &= f(V_G - V_S) = \frac{\beta}{2}(V_G - V_S - V_T)^2 \\
f'(V_G - V_S) &= \beta(V_G - V_S - V_T) = g_m \\
f''(V_G - V_S) &= \beta \\
f'''(V_G - V_S) &= 0
\end{aligned} \qquad (2.172)$$

resulting in:

$$\begin{aligned}
i &= g_m v_{\text{in}} - \frac{\beta^2}{2g_m}v_{\text{in}}^3 \quad \text{or} \\
\frac{i}{I_{\text{tail}}} &= \frac{v_{\text{in}}}{(V_G - V_S - V_T)} - \left(\frac{v_{\text{in}}}{V_G - V_S - V_T}\right)^3
\end{aligned} \qquad (2.173)$$

The first term agrees with the linearized model. All terms have been written as a ratio between excursion and bias. This ratio is often called the small-signal modulation factor. A low modulation factor results in low distortion, or otherwise stated: it takes a lot of current to avoid distortion.

In Fig. 2.67 V_S is the common node of the transistors and the current source. Due to the quadratic nature of the MOS characteristic, an increase in gate-source voltage on the left side creates more additional current than an equal decrease on the right side will let go. As the total current remains constant the potential V_S will move up. The same reasoning applies to an increase on the other side. Applying two equal but anti-phase sine waves on both gates results therefore in a second harmonic order signal on V_S. Any form of capacitive or resistive load on this node will cause a current carrying the second harmonic component. Returning to (2.168) the voltage V_S can be found as:

$$V_S = V_G - (h(I_{\text{tail}}/2 + i_1) + h(I_{\text{tail}}/2 - i_2)) \qquad (2.174)$$
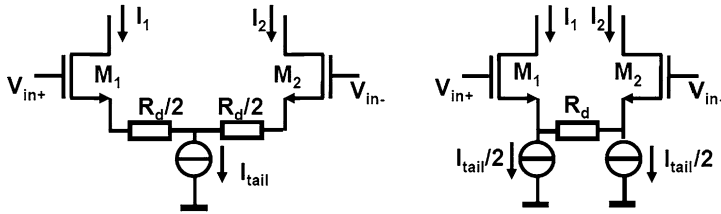
Fig. 2.69 The differential pair is linearized by degeneration

The addition of the *h* functions now leads to elimination of the odd terms and appearance of the even-order terms. In a similar analysis the magnitude of the second order component $v_{s,2}$ in V_S is found:

$$\frac{v_{s,2}}{(V_G - V_S - V_T)} = \frac{v_{in}^2}{2(V_G - V_S - V_T)^2} \qquad (2.175)$$

This analysis shows the relation between large signals and small signals. In most designs only the linearized terms are considered. In case the harmonic distortion component become relevant the analysis in this section can serve as a starting point.

2.7.6 Degeneration

The distortion of the differential pair can be reduced by applying a degeneration resistor. The signal between the gates of the MOS transistors creates a drive voltage in the transistors, but falls partly over the series resistor between the two sources, Fig. 2.69. The transfer is:

$$v_{in} = (i_1 - i_2)(R_d/2 + 1/g_m) \qquad (2.176)$$

The degeneration resistor helps to linearize the voltage to current transfer and allows to process larger input voltages. In the topology in Fig. 2.69 (right) the degeneration resistor does not carry any DC current. This arrangement avoids problems with DC biasing due to the relatively large voltage drop over the degeneration resistor. A degenerated input stage can be used in a voltage-gain amplifier by feeding the output currents in load resistors. The ratio between load and degeneration resistor determines the amplification, see [50].

2.7.7 Current Mirror

The current mirror copies current from one branch into another, see Fig. 2.70. This circuit is used for biasing and signal transfer purposes. In the first application the
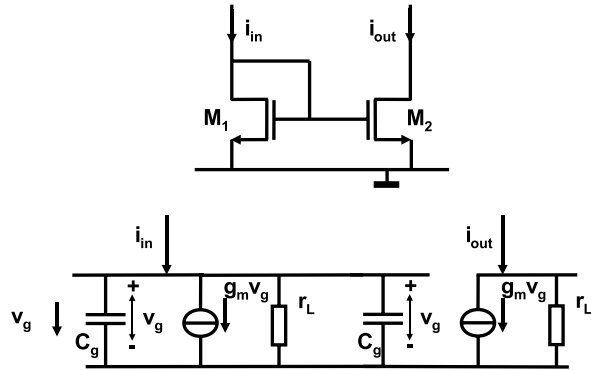
input current is DC and applied to the gate-drain of the input transistor M_1 (also called diode-connected). One or more transistors M_2 copy the input current and provide the bias currents to the circuits. Different W/L ratios between input and copier transistors allow choosing various levels of bias current. Correct operation of the copier transistors requires their drain-source voltages to remain higher than the saturation voltage ($V_{GS} - V_T$). Therefore W/L ratios are often chosen at a level where the saturation voltage is between 150 and 300 mV, leading to W/L ratios considerably larger than 1. With the length being the smallest value, the larger width is used as the variation parameter, thereby minimizing the effects of parameter spread. In case exact copying of currents is needed, copying is performed by adding only unit transistors, which allows only ratios of integers.

While in DC mode equal terminal voltages will lead to equal DC currents, the small-signal behavior is jeopardized by various parasitic elements, like gate and load capacitors and finite output conductance. This conductance is due to the drain-voltage modulation on the current as expressed by the factor λ, see (2.101).

Based on the small-signal schematic diagram, the input-output transfer is (for equally sized transistors):

$$\frac{i_{\text{out}}}{i_{\text{in}}} = \frac{g_m}{g_m + 1/r_L + 2j\omega C_g} - \frac{v_{\text{out}}}{i_{\text{in}}r_L} \tag{2.177}$$

In a first approximation the transfer is close to unity as the transconductance is the largest term. The transfer shows a first-order frequency dependency with a roll-off frequency at $f = g_m/4\pi C_g$. Decreasing the length of the transistors is beneficial for higher frequency operation: the capacitance decreases linearly, while the transconductance increases with a square-root at equal current, see the equations in Sect. 2.5. Unfortunately the output impedance decreases, as λ is length dependent. Increasing the width provides more transconductance, however the increase in gate capacitance is stronger so the roll-off frequency will lower. A higher DC-current will increase the roll-off frequency, however at the cost of a higher saturation voltage, thereby leaving less signal amplitude.

Figure 2.71 shows two alternatives to the regular current mirror. The left schematic is popular in bipolar configurations. The base current creates an offset
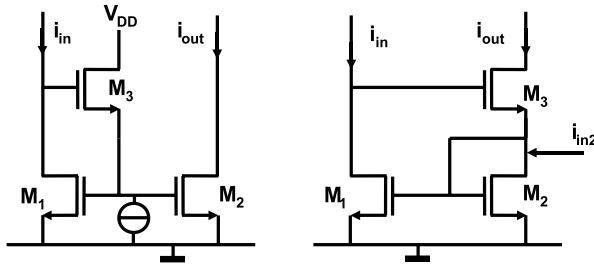
Fig. 2.71 The buffered current mirror and the Wilson mirror

in a standard current mirror, now the base current is supplied by the third transistor. In MOS this configuration may sometimes help to reduce the pole caused by the two gate capacitances.

The Wilson mirror (right) provides a high-ohmic output impedance. Next to the cascode operation of device M3 also the feedback in the circuit helps to improve the output impedance. To visualize this effect assume that a current is forced into the output terminal. This current will see the output impedance of device M3, but will also be copied by device M2 into device M1. As a result this device will pull down the gate of device M3, which strongly opposes the injected current.

The source node of device M3 forms a low-ohmic input terminal. A small negative voltage change Δv on this node leads to a current change in device M1 and a positive voltage change on its drain of $g_{m1} R_{\text{drain1}} \Delta v$, where R_{drain1} represents the total impedance on that drain node. This voltage change results in an additional current $\Delta i_{\text{source3}} = g_{m3} \times g_{m1} R_{\text{drain1}} \Delta v$ in device M3, if the drain of this device is connected to a current supplying terminal. The ratio of this current and the originating voltage change gives a first order indication of the impedance on this node:

$$\frac{\Delta v}{\Delta i_{\text{source3}}} = \frac{1}{g_{m3} g_{m1} R_{\text{drain1}}} \tag{2.178}$$

There is similarity between the Wilson mirror and the regulated cascode in Fig. 2.74 (right). If the gate of device M2 of the Wilson-mirror is disconnected, the devices M2 and M3 form the right-hand branch of the regulated cascode where device M1 is a simple amplifier with its own threshold voltage featuring as V_{cas} in Fig. 2.74.

Both current mirrors in Fig. 2.71 have as a disadvantage the stack of two threshold voltages in their input terminals, which makes them impractical for low supply voltage applications.

2.7.8 Darlington Pair

The base current of a bipolar transistor is in the order of $1/h_{fe} = 1\%$ of the collector-emitter current. In many applications this base current creates some loading of the input source. Sidney Darlington [51] proposed in 1953 a combination

Fig. 2.72 The Darlington
circuit
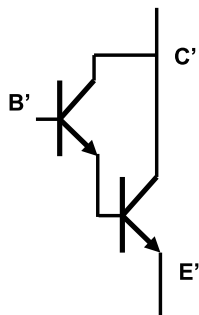
Fig. 2.73 The cascode circuit
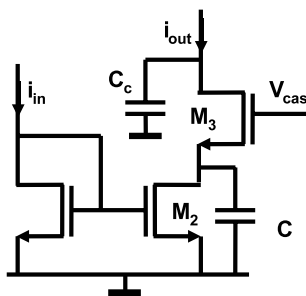
of two transistors, being one of the first integrated circuits see Fig. 2.72. Now the
left transistor generates the base current for the right transistor and the compound
transistor shows a collector current to base current ratio of h_{fe}^2. This compound
transistor shows a double $V_{b'e'}$ value, but is popular in bipolar design to mitigate the
undesired consequences of base currents.

2.7.9 Cascode and Regulated Cascode

The finite output impedance of a MOS transistor often limits the achievable voltage
gain in high-precision design. Cascoding[26] transistors as shown in Fig. 2.73, is a
potential remedy. An output impedance means that voltage variations on the drain
terminal of the current copier transistor cause a current fluctuation in the right-hand
branch. A cascode transistor and the gain-boost technique both aim at reducing these
voltage variations on the drain of the current copier transistor.

Assuming a constant bias current in the cascode transistor, a voltage variation
on the drain v_D results in a reduced voltage variation $\lambda_3 v_D$ on the drain of the
current copier transistor M_2. The current variation in this transistor is then $\lambda_3 \times$

---

[26]There is some ambiguity between the terms "cascade" and "cascode". Here a cascade will refer
to a series connection of e.g. amplifiers, while a cascode is a set of transistors stacked on top of
each other.

Fig. 2.74 The regulated cascode (gain boost [52]) and a simple realization

lower. Adding a cascode transistor increases the output impedance r_L of the current copier M_2 to roughly r_L/λ_3 for the combination of both transistors. A more rigorous analysis gives the overall output impedance:
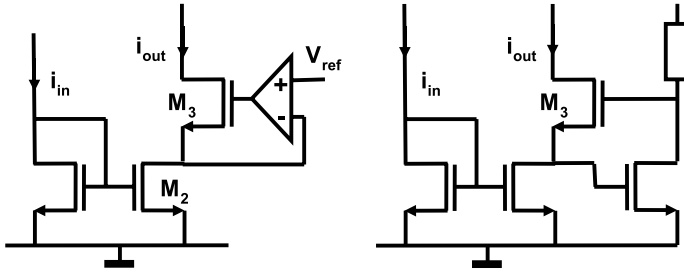
$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{g_{m,3}r_{L,3}r_L + r_{L,3} + sr_{L,3}r_L C + r_L}{sr_L C + 1} \tag{2.179}$$

where $g_{m,3}r_{L,3}r_L = r_L/\lambda_3$ is the dominant factor. The output impedance for a single cascode transistor can be further increased by stacking more cascode transistors. This approach requires for each transistor some additional voltage head room. In parallel to this impedance, the capacitive load at the drain of the cascode transistor C_3 dominates the impedance at high frequencies.

Another approach for increasing the output impedance is the regulated-cascode or gain-boost technique [52], see Fig. 2.74. Now an amplifier is used to reduce the voltage swing of the node between the current copier transistor M_2 and the cascode transistor M_3. In first order this technique increases the output impedance to $(A + 1)r_L/\lambda_3$, where A is the amplification factor of the feedback amplifier. Figure 2.74 (right) shows a simple implementation of the amplifier where the threshold voltage of the amplifier MOS serves as the reference voltage. Figure 2.75 shows the output impedance that can be obtained from a standard cascode output stage (dotted line $0\times$). A first order roll-off behavior starts at 10 MHz and unity-gain is reached at 10 GHz. Applying the additional feedback loop with a roll-off frequency at 10 kHz and gains of 10, 100, 1000 and 10 000 times improves the overall gain. The lower amplification curves allow to identify the complex output impedance of this circuit. The transfer description is:

$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{g_{m,3}r_{L,3}r_L(A + 1 + s\tau_{fb})}{1 + s\tau_{fb} + sg_{m,3}r_{L,3}r_L C_3(A + 1 + s\tau_{fb})} \tag{2.180}$$

At low frequencies the DC amplification is $g_{m,3}r_{L,3}r_L(A + 1)$. If the frequency is increased, first the pole corresponding to the roll-off of the feedback amplifier will dominate:

$$\frac{v_{\text{out}}}{i_{\text{out}}} \approx \frac{g_{m,3}r_{L,3}r_L(A + 1)}{1 + s\tau_{fb}} \tag{2.181}$$
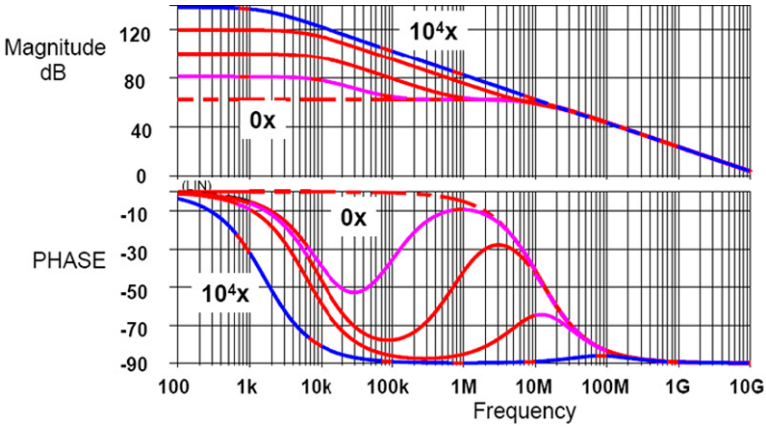
Fig. 2.75 Magnitude and phase of the output impedance of a 90-nanometer designed cascode stage (dotted line) and of a regulated cascode with amplification factors of: 10, 10^2, 10^3, 10^4 and an internal pole at 10 kHz. The gain in the feedback path translates in more gain at the output node of the circuit

At the moment this roll-off has reached the its unity-gain frequency ($\omega \tau_{fb}/A = 1$), the zero flattens the curve to the original cascode impedance.

$$\frac{v_{\text{out}}}{i_{\text{out}}} \approx \frac{g_{m,3}r_{L,3}r_L(A+1)(1 + s\tau_{fb}/(A+1))}{s\tau_{fb}} \approx g_{m,3}r_{L,3}r_L \qquad (2.182)$$

At the cascode roll-off frequency the final first order behavior due to the output capacitor C_3 starts, the s^2 term in the denominator and the s term in the numerator dominate:

$$\frac{v_{\text{out}}}{i_{\text{out}}} = \frac{g_{m,3}r_{L,3}r_L}{s g_{m,3}r_{L,3}r_L C_3} = \frac{1}{sC_3} \qquad (2.183)$$

Now the amplifier only provides a DC-bias voltage and the circuit operates with a DC-biased cascode transistor.

This design shows a pole-zero doublet in (2.180). This doublet is formed by the unity-gain frequency of the feedback amplifier (forming the zero) and the pole of the cascode. In the Laplace domain equation (2.180) can be rewritten in the form of a formal pole-zero doublet:

$$H(s) = \frac{1 + s\tau_z}{(1 + s\tau)(1 + s\tau_p)} \qquad (2.184)$$

where τ is the time constant of the dominant low-frequency pole of the feedback amplifier, and τ_z and τ_p form the doublet and are derived from the unity-gain frequency of the feedback amplifier and the original pole of the cascode. Their mutual relation is: $\tau \gg \tau_z > \tau_p$. Note that the unity-gain frequency of the feedback amplifier itself can be pushed beyond the original cascode pole, however in the overall transfer function this will not change the order of the poles and zero. The decompo-
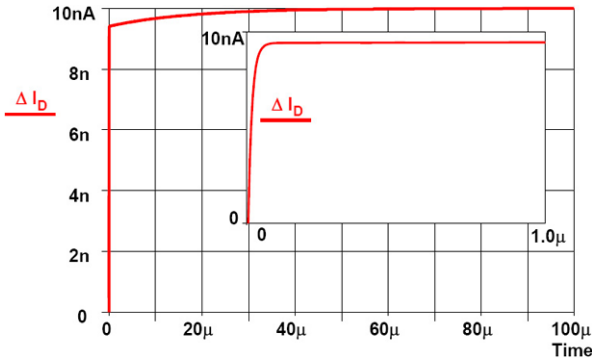
Fig. 2.76 Example of the response of a doublet in a regulated cascode. The plot shows the steep initial reaction to a change (see also the *insert*), followed by a much slower final settling. This regulated cascode circuit corresponds to Fig. 2.75 with an amplification factor of 100 and a time constant in the feedback loop of 16 µs

sition of the above expression in case of a step wise current excitation, leads to two exponential decay terms [53]:

$$i_{\text{out}}(t) = i_{\text{step}}\left(1 - \frac{\tau - \tau_z}{\tau - \tau_p}e^{-t/\tau} - \frac{\tau_z - \tau_p}{\tau - \tau_p}e^{-t/\tau_p}\right) \tag{2.185}$$

In Fig. 2.76 the slow and fast settling components can be observed. The amplitude of the slow settling component is proportional to the relative difference between the pole and zero with respect their mutual distance to the dominant pole.

2.7.10 Single-stage Amplifier

A class of high-gain amplifiers is based on the gain-enhancement techniques as cascoding and gain-boosting. Figure 2.77 shows a telescopic fully differential amplifier. The current driving part is formed by the two input transistors and the gain is made by creating a high output impedance by cascoding. It is obvious that this is a current efficient technique, be it that sufficient power supply voltage is needed. The capacitive loading of the amplifier creates the dominant pole and the unity gain bandwidth is at the frequency where the input voltage v_{in} is equal to the output voltage over the capacitor, resulting in:

$$f_{\text{UGBW}} = \frac{g_m}{2\pi C} \tag{2.186}$$

The capacitive loading on the intermediate nodes between the transistors creates higher frequency poles. The two resistors on the output generate the common voltage of the output. Feeding back this voltage to the bias line of the top side PMOS transistors results in a stable operating point. The resistors in this graph are indeed used by some designers, alternatives are source followers or switched capacitor

Fig. 2.77 The telescopic fully differential amplifier. The two resistors provide the common mode feedback path
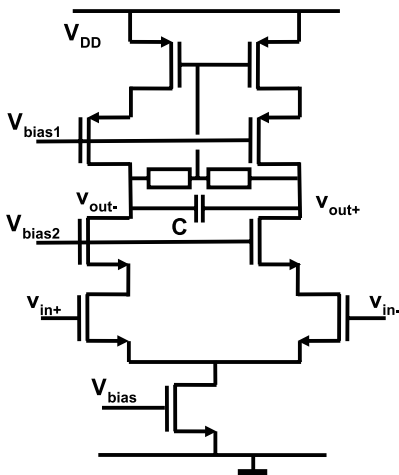
Fig. 2.78 The folded cascode amplifier
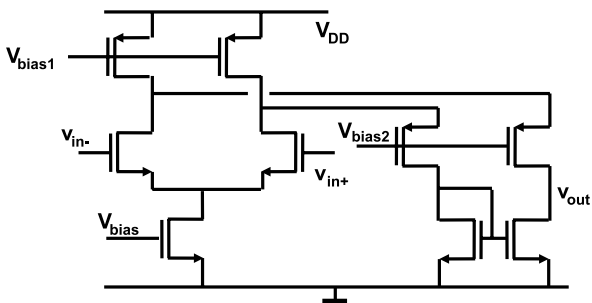
equivalents of the resistors. The common-mode level at the input can differ considerably from the output common mode level. High-threshold input MOS transistors help to avoid the cascode stages to go out of saturation. This type of opamp is not well-suited for high signal-swing application.

An alternative topology allowing lower power supplies is the folded cascode amplifier. Figure 2.78 shows the circuit. The top-most PMOS transistors merely serve as current sources, supplying both the left and right hand branch of NMOS devices. Any current modulation due to a differential voltage on the input pair is mirrored in the load branch. Next to the possibility to operate this amplifier at lower supply voltages also an excellent PSRR can be obtained.

2.7.11 Miller Amplifier

Figure 2.79 shows a simple two stage "Miller-amplifier" [54–57]. The first stage consists of a differential pair (M_4 and M_5) terminated with a current mirror (M_6 and M_7). The second stage is a class-A amplifier with a constant current source as a load (M_3 and M_8).

Fig. 2.79 The Miller opamp

Fig. 2.80 The Miller opamp small-signal equivalent circuit

The NMOS bias transistors M_2 and M_3 will not carry any signal, however the noise contribution can be significant: the noise of the output transistor is directly visible, whereas the noise of the differential pair current source becomes visible during input imbalance. Relatively long transistors will improve the current source behavior, although the resulting drain capacitors must be taken into account for optimum speed. Moreover the 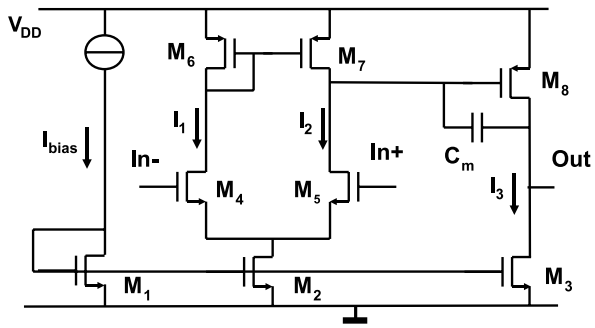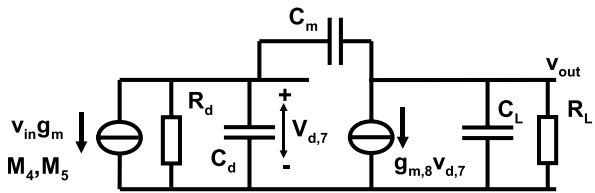saturation voltage should be monitored as this saturation voltage will determine the minimum common mode voltage at the input.

The input pair transistors (M_4 and M_5) need a high transconductance: wide but short transistors are used. Most opamp topologies use a single transistor pair for forming the input differential signal. In standard CMOS processes the available transistors have a considerable threshold voltage that together with the necessary turn-on voltage and the minimum voltage drop over the current source, leads to a significant loss of input range in e.g. unity-gain feedback topologies. A solution is a rail-to-rail input stage. This circuit employs two pairs of complementary transistors that allow to use the full power supply range for an input signal. An elegant solution is given in [58].

The current mirror and the output driver carry signals and a near[27] minimum gate length is used. This will reduce the parasitic pole in the current mirror.

The Miller amplifier shows two dominant poles. The capacitive load of the gate of the driver transistor and the output impedance of the current mirror normally create the lowest-frequency pole at the gate of M_8. The second pole is associated with the output stage and its capacitive loading at the output node (drain of M_8). The third pole in importance is the common gate node of the PMOS current mirror. In Fig. 2.80 a small-signal equivalent circuit is shown. The input pair M_4, M_5

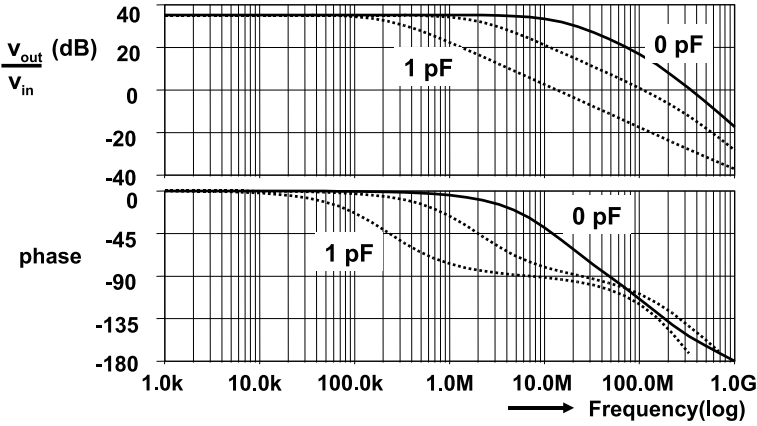[27]Compare the current factor versus gate length in Fig. 2.37.

Fig. 2.81 The amplification and phase behavior of the Miller amplifier for Miller capacitors of 0 (*solid line*), 0.1 and 1 pF

is replaced by a voltage controlled current source g_m feeding the second stage. Applying Kirchhoff's current law to the two nodes of the small-signal diagram of Fig. 2.80 with the Miller capacitance, gives:

$$g_m v_{in} + \frac{v_d}{Z_d} - \frac{v_{out} - v_d}{Z_m} = 0$$

$$g_{m,8} v_d + \frac{v_{out} - v_d}{Z_m} + \frac{v_{out}}{Z_L} = 0$$

All symbols with suffix d refer to the drain of M_7 that connects to the gate terminal of the driver transistor M_8. $g_{m,8}$ is the transconductance of M_8. Z_L is the equivalent load impedance at the output and $Z_m = 1/j\omega C_m$ represents the Miller feedback path. Solving the above equations results in:

$$\frac{v_{out}}{v_{in}} = \frac{g_m Z_d Z_L (g_{m,8} Z_m - 1)}{Z_d Z_L g_{m,8} + Z_L + Z_d + Z_m} \tag{2.187}$$

For frequencies where Z_m is infinitely large the equation reduces to the simple dual pole transfer, and for DC to the amplification factor:

$$\frac{v_{out}}{v_{in}} = g_m Z_d Z_L g_{m,8} \rightarrow A_{DC} = g_m R_d g_{m,8} R_L \tag{2.188}$$

The frequency behavior defined by the two poles, are shown as the upper lines in the Bode plots of Fig. 2.81. The DC-impedances R_d and R_L are the result of the output impedances of the connected transistors. This limits the overall gain to $A_{DC} = 1/\lambda_{M7}\lambda_{M8}$ when the output impedances of the remaining transistors are ignored.

A Miller compensation capacitor is applied to control the stability. The first stage sees a higher capacitive load on the gate terminal of the driver transistor. Any voltage change on this terminal is amplified by the gain of the driver $g_{m,8} Z_L$, and therefore the Miller capacitor will load the preceding stage with $C = C_m(g_{m,8} Z_L + 1)$, which

is a multiple of the Miller capacitor value. Consequently the application of a Miller capacitor will shift the first pole of the amplifier to a lower frequency. The low-frequency transfer is dominated by the Miller capacitor $Z_m = 1/j\omega C_m$:

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{g_m Z_d Z_L (g_{m,8} Z_m - 1)}{Z_d Z_L g_{m,8} + Z_L + Z_d + Z_m}$$

$$\approx \frac{g_m Z_d Z_L (g_{m,8} Z_m - 1)}{Z_d Z_L g_{m,8}} = g_m \left(Z_m - \frac{1}{g_{m,8}} \right) \approx \frac{g_m}{j\omega C_m}$$

The last expression can be understood by considering that the input pair generates a signal current that is fed in a single-transistor amplifier (M_8) with C_m as a feedback. Assuming that this amplifier is dominated by its feedback factor, the overall transfer in this frequency range is the product of the current and the capacitor. In a properly designed Miller amplifier this pole determines the 0 dB amplification point. Therefore the unity-gain frequency of the Miller amplifier is:

$$\text{UGBW} = \frac{g_m}{2\pi C_m} \tag{2.189}$$

Next to lowering the first pole also the output pole is affected: at higher frequencies the Miller capacitor connects the gate to the drain of the driver transistor and puts its transconductance in parallel to the output load. Any change in voltage on the output node is now also countered by the transconductance of the driver transistor. This will effectively lower the output time constant and shift the output pole to higher frequencies.

The Miller principle creates a dominant pole inside the amplifier that moves to lower frequencies for increasing Miller capacitance values. A larger Miller capacitance also lowers the output impedance, thereby shifting the pole on the output terminal to a higher frequency. The combination of these two effects is called: "pole-splitting".

At high frequencies the Miller capacitor creates an undesired side effect. Its impedance is lower than the transconductance and current coming from the input stage flowing via the Miller capacitor, exceeds the current of the driver stage. This current via the Miller capacitor has an opposite polarity compared to the driver output current. The sign of the output signal will flip, causing instability. This effect is visible in the transfer description in the $(g_{m,8} Z_m - 1)$ term, causing a so-called "right half-plane zero" in the complex s-plane. A well-known trick to avoid the current through the Miller capacitor to exceed the driver current is to add in series a resistor $Z_m = R_m + 1/j\omega C_m$. With $R_m > 1/g_{m,8}$ the right half-plane zero moves to the left half plane. A second option is to connect a buffer to the output node and feed from there the Miller capacitance.

The current is determined by the dynamic requirements such as the output slew-rate.

2.7.12 Choosing the W/L Ratios in a Miller Opamp

The initial design of a Miller opamp starts by inspecting the required performance parameters. The two gain stages can deliver a DC-gain in the order of $1/\lambda^2$, where λ is the transistor drain-voltage feedback factor, see Sect. 2.5.4. Without further additions of cascode stages this limits the gain in practical situations to 40–60 dB.

The second performance parameter concerns the load and bandwidth. A DC-output current in a resistive load should be avoided as the output stage acts partly as a current source. A DC output current will result in a considerable input offset. With a capacitive load C_L and a required signal V_{pp} a first estimate for the current I_3 in the output branch is found from the slew rate equation:

$$I_3 T_{\text{slew}} = V_{pp}(C_L + C_M + C_{\text{parasitic}}) \tag{2.190}$$

where T_{slew} is the time allowed to build up a signal of the amplitude V_{pp} over the total capacitive load. $T_{\text{slew}} \approx 1/(2BW)$ is linearly related to the bandwidth for large amplitude signals. A first estimate for the currents in the differential pair I_1 and I_2 is a fraction between 5% and 10% of the current in M_3. These currents must also be sufficient to create the required unity-gain bandwidth, as they contribute via a square-root function to the UGBW.

$$\text{UGBW} = \frac{g_m}{2\pi C_M} = \frac{1}{2\pi C_M}\sqrt{\frac{2I_{1,2}W\beta_\square}{L}} \tag{2.191}$$

The DC levels for the signal swing at the inputs and output of the amplifier must be chosen. The input range depends on the opamp topology, see Fig. 2.52. In case of a unity-gain configuration the output range must take into account that the input range cannot span the entire power supply range. The input swing is on the low side limited by the drive voltage of the differential pair transistor, its threshold voltage and the saturation voltage of the current source. Higher voltages reduce the effects of mismatch and noise but limit the signal swing. At the high side of the signal range the output driver transistor M_8 must remain in the saturation regime. Next to that switches around the opamp may require to adapt the swing requirements.

With some target values for the currents above, the choice of the transistor sizes can start. The choice of the saturation voltage $V_{GS} - V_T$ for each transistor is between increasing the signal swing (low value) and less sensitivity to noise, mismatch (high value). A practical value ranges between 150 and 300 mV in a 0.18 μm process. Proper sizing helps to avoid DC-offset at the input:

- Obviously M_4 and M_5 are identical in width and length and the same holds for M_6 and M_7.
- The current mirror M_1, M_2 and M_3 is built with transistors of identical gate length. As the length is the smallest dimension, this parameter is kept constant to minimize length dependent parameter variations. The length is chosen somewhat larger than minimum size thereby improving the output impedance. For a 0.18 μm process 0.3 μm is a proper start.
- Current ratios between M_1, M_2 and M_3 are obtained by varying the width or even better by using a standard width and multiply the number of unit transistors.

- The lengths of the PMOS transistors M_6, M_7 and M_8 are identical and close to the minimum value as they form the signal path.
- The width ratio between the bias transistors M_2 and M_3 determines the width ratios between the PMOS transistors M_6, M_7 and M_8. In equilibrium the current of M_2 is split in halves, which results in a gate voltage for M_6 and M_7 that consumes these bias currents. The drain voltage of M_6 is equal to the gate voltage and because M_7 carries the same current and sees the same source and gate voltages as M_6 its drain voltage must be equal to the drain voltage of M_6. In equilibrium the drain voltages of M_6 and M_7 together act as a diode load for the bias current of M_2. The widths of M_3 and M_8 must therefore be multiples of respectively the width of M_2 and the sum of width of M_6 and M_7. In input equilibrium transistor M_8 will now show the same source, gate and drain voltages as M_6. The actual widths are found from combination of the current and the allowable drive voltages.
- The length of the transistors in the differential pair $M_{4,5}$ is chosen close to minimum in order to get a maximum transconductance. The width is determined by the required bandwidth and additional constraints such as input swing, mismatch, etc.

2.7.13 Dominant-pole Amplifier

In some applications a dominant pole on the output is present at a low frequency e.g. in circuits where the output drives a large capacitive load. In that case the amplifier design has to be based on the dominant pole.

Figure 2.82 shows the Bode plots of an amplifier that is topologically identical to the Miller amplifier without the Miller capacitor. The output load is in this sim-
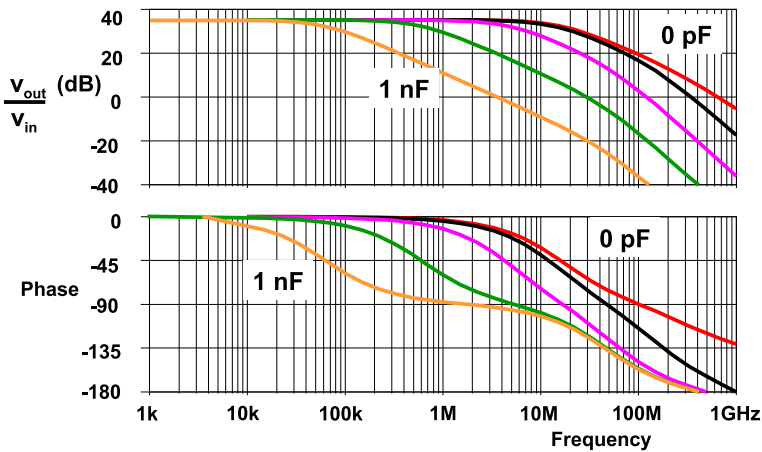

Fig. 2.82 The amplification and phase behavior of the amplifier for load capacitors of 0, 1 pF, 10 pF, 100 pF and 1 nF. A larger load capacitor creates a low-frequency character

ulation increasing from 0 to 1 nF. Without output load the amplifiers phase margin (phase surplus till 180°) is still 60° and a stable feedback at unity gain can be obtained. The effect of 1 pF loading is more visible in the phase than in the amplitude and reduces the phase margin to below 30°. Also at 10 pF there is insufficient phase margin due to the proximity of the amplifier pole and the output pole. If the load capacitance is further increased to 100 pF or 1 nF the output capacitance dominates the low-frequency characteristic and stability requirements. A stable feedback path exists for unity gain. As the roles of the two poles in a Miller opamp are interchanged, it is crucial to reduce in a dominant pole amplifier all capacitances that can add to a parasitic Miller capacitor.

2.7.14 Feedback in Electronic Circuits

In Sect. 2.6.4 the general theory of feedback systems was summarized. In electronic circuits signal are represented by currents or voltages. Both voltages and currents are used as inputs or outputs of a circuit.

Table 2.25 lists the four transfer modes. The feedback classification as used in Fig. 2.83 is derived from electrical motor design. Signals can be extracted or injected in two fashions: in series or as a shunt (parallel) connection. Gray and Meyer [48] use this terminology extensively. The terms "series" and "shunt" may cause some confusion because at the input currents are added by a shunt connection and voltages are added via a series connection. At the output a shunt extracts a voltage and a series connection a current. Yet this classification is useful for designing optimum signal transfer. E.g. feeding the circuit topology in Fig. 2.83 (right) from a voltage source, is not effective as all feedback current drains into the input source. Realizing whether inputs and outputs are current or voltage driven, leads to a stable amplification determined by the feedback path. Moreover a correct feedback improves the implicit characteristics of the amplifier: a low-ohmic current input node

Table 2.25 Transfers in an electronic circuit

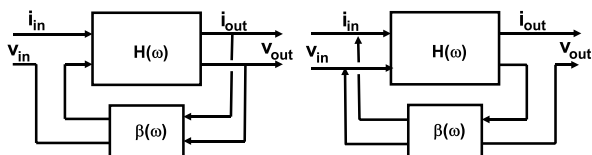| Name	Input	Output	Transfer name	Symbol	Feedback
Voltage amplifier	v_{in}	v_{out}	Voltage gain	A_V	Series-shunt
Current amplifier	i_{in}	i_{out}	Current gain	A_I	Shunt-series
Transconductance amp	v_{in}	i_{out}	Transconductance	g_m	Series-series
Transimpedance amp (TIA)	i_{in}	v_{out}	Transimpedance	Z_m	Shunt-shunt

Fig. 2.83 A series-shunt and a shunt series configuration

Fig. 2.84 An example of a series-series circuit and a shunt-shunt circuit
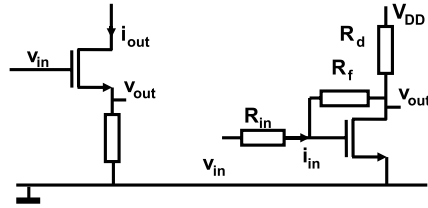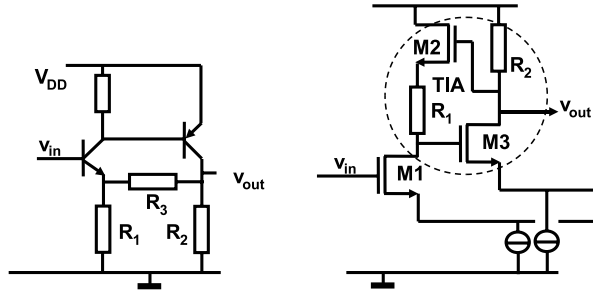

Fig. 2.85 Two circuits using the Cherry-Hooper technique

will be further reduced by a shunt feedback, a low-ohmic output voltage node will show a lower impedance through a shunt feedback.

Figure 2.84 (left) shows a series-series feedback circuit. The input variable is v_{in} and the output variable is i_{out}, so the circuit is a transconductance amplifier. The output current generates a voltage swing over the resistor. This voltage is connected in series with the input voltage. Effectively the transistor is driven by $(v_{in} - i_{out}R_s)$ showing the impact of the feedback.

The circuit in Fig. 2.84 (right) is a shunt-shunt amplifier. The input voltage is turned into a current via R_{in} and the output current of the transistor is converted in a voltage via R_d. Feedback in this circuit uses the output voltage via a shunt connection and converts it via R_f into a current. This current is added to the input current via an input shunt.

Cherry and Hooper [59] already in an early stage realized that connecting an output that behaves as a current source to an input that expects a current is the efficient way to design high performance circuits. The classical series-shunt circuit in Fig. 2.85 (left) shows an npn-transistor as a voltage-to-current transconductance input stage. Its output current flows into the low-ohmic pnp base. This stage operates as a current-to-voltage amplifier. The feedback is formed by resistors R_3 and R_1. The shunted output voltage is fed back and connected in series with the input voltage. The amplification of this stage is in first order given by the inverse feedback factor $(R_1 + R_3)/R_1$. In this circuit, the input impedance is increased due to the feedback which also lowers the output impedance. The intermediate node formed by the collector of the npn and the base of the pnp is low-ohmic and shows a minimum voltage swing. Capacitive parasitic loading does not impact the performance.

In CMOS similar ideas are used in e.g. broadband optical receiver front-ends [60]. A popular circuit, see Fig. 2.85 (right) is built with an input transconductance stage and an output formed by a feedback trans-impedance amplifier (TIA).
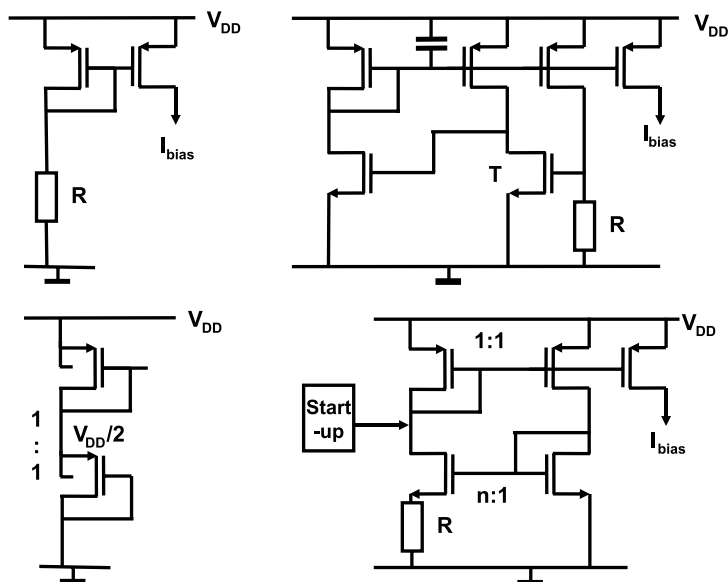
Fig. 2.86 Four bias circuits, *top*: simple resistors based current and a feedback circuit that will set the current to V_T/R. *Bottom*: a voltage divider and a beta-resistor circuit

Again the similarities with the Wilson mirror Fig. 2.71 and the gain boost technique Fig. 2.74 (right) are instructive.

2.7.15 Bias Circuits

Figure 2.86 shows four schemes to provide bias voltages and currents for circuits. In the upper left circuit schematic a resistor determines the current:

$$I = \frac{V_{DD} - V_T}{R} \tag{2.192}$$

The direct influence of the power supply voltage translates in a poor power supply rejection of the output current. Decoupling between the transistor gate voltage and the power supply will improve the PSRR for higher frequencies.

The stack of PMOS transistors in Fig. 2.86 (lower left) creates a middle voltage equal to half of the power supply. The n-well connection of PMOS transistors is available in a p-substrate based technology and allows to eliminate in this circuit the effect of threshold changes due to differences in substrate voltage.

On the right hand side of Fig. 2.86 two more complicated bias circuits are depicted. In the upper circuit the current is defined by transistor T and resistor R. If the voltage over this resistor is low, the transistor T is off and this will cause the succeeding transistor to generate a large current. This current is mirrored into the resistor causing the voltage drop over the resistor to increase until the threshold

voltage of T is reached and T starts to conduct. This negative feedback results in a simple first order approximation of the current in the resistor:

$$I = \frac{V_T}{R} \tag{2.193}$$

The feedback in this circuit can result in an oscillatory behavior. A simple way of stabilization and improving the PSRR is to connect a large capacitor between the gate connection of the PMOS transistors and the power supply. For high frequencies the capacitor will keep the gate-source voltage of the bias transistors constant. An issue with this circuit is the situation where no current flows. If the drain of transistor T remains under the NMOS threshold voltage the entire circuit will draw no current. Therefore a start-up facility must be provided (e.g. a high-ohmic resistor to pull-up the drain). Threshold variation will inevitably influence the current level, and substrate noise will penetrate into the bias current through threshold modulation.

It is also possible to link the bias current to the current factor of a MOS device, Fig. 2.86 (lower-right). The two NMOS transistors and a resistor determine the current. The differential arrangement of the transistors eliminates the influence of the threshold, except for some back-bias modulation. Now a simple analysis gives:

$$I = \sqrt{\frac{2I}{R^2 \beta} \left(1 - \frac{1}{\sqrt{n}}\right)} \tag{2.194}$$

Where $n > 1$ is a transistor size ratio. In order to avoid the $I = 0$ condition, a start-up circuit is necessary and care has to be taken for stability.

An example: a 0.18 μm technology shows an NMOS transistor current factor $\beta_\square = 350$ μA/V^2. With NMOS transistor W/L ratios of 2/0.4 and 8/0.4 ($n = 4$) and $R = 10$ kΩ the above formula yields 3 μA. Unfortunately the current factor is a less well determined factor in advanced CMOS processes, which makes this circuit less suited for use with minimum gate lengths.

2.7.16 Oscillators

Oscillators generate timing signals for all kinds of purposes in an electronic system. The basis for any oscillator is a feedback circuit in which the fed-back signal is in-phase with the input signal. The total delay must equal one oscillation period or the phase of the oscillation frequency in the chain must add up to 360° or 2π rad. Moreover the total gain must be equal to 1. The conditions are known as the "Barkhausen Criterions".[28] Lower gain will create extinction of the oscillation, higher gain will create saturation. The design of an oscillator has to meet these two requirements.

---

[28]Heinrich Barkhausen (1881–1956) was appointed in 1929 as the world's first professor in Electrical Engineering in Dresden. He discovered the noise generated by changing magnetic walls under a varying magnetic field.
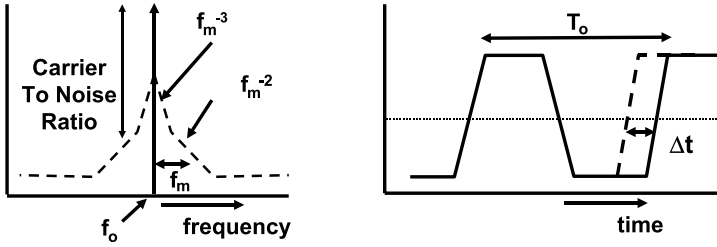
Fig. 2.87 *Left*: the frequency spectrum of an oscillator, *right*: time jitter

The delay in the feedback loop can be designed with various combinations of active and passive elements: R-C, L-C, I-C, etc.

Oscillators based on resonant circuits are popular. The reason is that in a resonant circuit like an LC circuit or a mechanical resonator as a quartz crystal the resonant energy swings between two domains (electrical and magnetic, or electrical and mechanical) without much disturbance. In resonators with a high quality factor Q, by definition, only a small fraction of the energy is lost and has to be replaced.

$$Q = 2\pi \frac{\text{Energy stored}}{\text{Energy dissipated/cycle}} \tag{2.195}$$

Consider as an example a sinusoidal current $\hat{i} \sin 2\pi f_0 t$ flowing in an oscillating RLC tank, Fig. 2.56. The maximum energy stored in the inductor is $\hat{i}^2 L/2$, the energy dissipated per cycle is $\hat{i}^2 R/2 f_0$. Substitution of these quantities results in (2.148).

The supplied energy is prone to noise disturbance, which causes the oscillator to generate unwanted components. Noise picked up on the output terminal is simply added to the oscillation signal. If noise is generated in the resonating components of the oscillator, this noise will create additional oscillatory signals with random phases and frequencies. These signals will modulate the wanted oscillation thereby creating phase-noise see Fig. 2.87 (left). The phase-noise shows in the frequency domain as side lobes. The quality of the oscillation frequency depends on the total noise power in the side lobes and the distribution. This power is characterized by its noise spectral density relative to the carrier power: $S_{ff}(\omega)$. This quantity is measured in terms of "carrier power-to-noise power" CNR ratio or the inverse form: single sideband noise-to-carrier ratio $\mathcal{L}(\omega) = 10 \log S_{ff}(\omega)$ at frequency ω measured in 1 Hz bandwidth expressed in dBc/Hz. The term "jitter" is used in the time domain. Jitter is concentrating on the zero-crossing moment of the timing signal.

Mathematically the ideal output frequency of an oscillator is modeled as a time-invariant quantity $\omega_0 = 2\pi f_0$. The phase of the output signal at time t is $\theta(t)$ and the disturbing factors are represented as an excess phase shift $k(t)$ causing frequency modulation. The total phase is now written:

$$\theta(t) = \int_{\tau=0}^{\tau=t} \omega_0 + k(\tau) \, d\tau \tag{2.196}$$

as the argument of a sine wave description:

$$v_0(t) = V_A \cos(\theta(t)) = V_A \cos\left(\omega_0 t + \int_{\tau=0}^{\tau=t} k(\tau)\,d\tau\right) \qquad (2.197)$$

As an example the phase distortion is chosen as a discrete tone: $k(t) = k_0 \sin(\omega_m t)$, where $m = k_0/\omega_m$ is called the modulation index. $\omega_m = 2\pi f_m$ is the modulating frequency of the disturbance, which appears at $\omega_0 \pm \omega_m$ in the spectrum of the oscillation. Applying goniometric decomposition:

$$v_0(t) = V_A \cos(\omega_0 t)\cos(m\sin(\omega_m t)) - V_A \sin(\omega_0 t)\sin(m\sin(\omega_m t)) \quad (2.198)$$

The nested sine and cosine terms can be rewritten with Bessel functions:

$$v_0(t) = V_A [J_0(m)\cos(\omega_0 t) - J_1(m)(\cos(\omega_0 t - \omega_m t) - \cos(\omega_0 t + \omega_m t))$$
$$+ J_2(m)(\cos(2\omega_0 t - 2\omega_m t) - \cos(2\omega_0 t + 2\omega_m t)) - \cdots]$$

where the Bessel function is defined as:

$$J_n(m) = \frac{1}{\pi} \int_{\theta=0}^{\theta=\pi} \cos(m\sin(\theta) - n\theta)\,d\theta \qquad (2.199)$$

For small values of m only the first two terms are relevant: $J_0(m) \approx 1$, $J_1(m) \approx m/2$, which results in an output signal:

$$v_0(t) = V_A\left(\cos(\omega_0 t) - \frac{m}{2}\cos(\omega_0 t - \omega_m t) + \frac{m}{2}\cos(\omega_0 t + \omega_m t)\right) \quad (2.200)$$

In first order the phase disturbance shows up as amplitude modulation of the oscillation frequency. In this example the carrier-to-noise ratio is the ratio between the power of the desired frequency and the power of the two modulation carriers:

$$\frac{\text{Power of side frequencies}}{\text{Power of carrier}} = (m/2)^2 + (m/2)^2 = \frac{k_0^2}{2\omega_m^2} \qquad (2.201)$$

The modulation index is m proportional to the inverse of the modulation frequency ω_m. This simple situation with a discrete sine wave as disturbance results in a disturbance amplitude that shows a second order roll-off with increasing distance from the carrier frequency.

   Leeson [61] showed that disturbing the phase of an oscillation with a flat spectrum from thermal kT noise, also creates a second order slope in the output power spectrum. A real phase disturbance results in a symmetrical frequency disturbance around the oscillation frequency. The power spectral density relative to the power in the oscillation frequency ω_0 is found [61]:

$$S_{ff}(\omega) = \frac{2FkT}{P_0} \frac{\omega_0^2}{(2Q)^2(\omega - \omega_0)^2} \qquad (2.202)$$

P_0 is the oscillator power level, F the noise figure of the oscillator and kT the thermal energy. Q is the quality factor which boosts the oscillation signal over the noise. This theoretical description shows that a high quality factor allows a low phase-noise.

Figure 2.87 shows the typical behavior of phase noise in an oscillator. Depending on the design the $1/f$ noise density can dominate the thermal noise density close to the carrier. The exponent of the term $(\omega - \omega_0)$ term in the numerator will increase to 3. Outside the $1/f$ region the thermal noise in the oscillator components creates the above described second-order roll-off. Far away is the flat noise generated by elements following the oscillator, such as drivers.

For sampling systems the jitter is an important parameter. An instantaneous phase deviation $\Delta\theta$ offsets the zero-crossings of a sinusoidal signal of frequency ω_0 with a time deviation (jitter) Δt. Relative to the frequency period T_0:

$$\frac{\Delta t}{T_0} = \frac{\Delta\theta}{2\pi} \tag{2.203}$$

This definition specifies the cycle-to-cycle jitter. In some systems a long-term jitter component can be relevant, e.g. for a display on a screen the jitter between two samples above each other is relevant. This jitter is specified over a line-time.

A first-order indication of the relation between jitter and phase-noise can be obtained by considering that the power of the phase noise is also responsible for generating the jitter. In the time domain the instantaneous jitter Δt is replaced by its time-averaged root-mean-square value: $\sigma_{t,\mathrm{rms}}$. The spectral noise density $S_{ff}(f)$ must be integrated over both side lobes to give the total equivalent phase noise:

$$\left(\frac{\sigma_{t,\mathrm{rms}}}{T_0}\right)^2 = \frac{2\int_{f_{\mathrm{low}}}^{f_{\mathrm{high}}} S_{ff}(f)\,df}{(2\pi)^2} \tag{2.204}$$

The integration cannot start at $\omega = \omega_0$ Hz due to the singularity in the spectral density. For a second-order roll-off in the phase-noise slope, the contribution of the frequency bands beyond an offset frequency Δf is:

$$\sigma_{t,\mathrm{rms}} \approx \sqrt{\frac{(\Delta f)S_{ff}(f_0 - \Delta f)}{2\pi^2 f_0^2}} \tag{2.205}$$

With $f_0 = 1$ GHz, $\Delta f = 1$ MHz and $S_{ff}(999 \text{ MHz}) = 10^{-12}$ equivalent to $\mathcal{L}(999 \text{ MHz}) = -120$ dBc/Hz, a jitter of 0.2 ps is found. Taking the bands beyond $\Delta f = 100$ kHz into account raises the jitter to 0.7 ps.

The choice for f_{low} depends on whether the cycle-to-cycle jitter is calculated or longer-term jitter variations. The energy in the low-frequency second-order lobes of the phase spectrum is responsible for the increase of long-term jitter over cycle-to-cycle jitter. In the extreme case of only white phase noise the contribution of the low-frequency band would be negligible and the long-term jitter would be comparable to the cycle-to-cycle jitter. Translating various forms of phase-noise densities in time jitter clearly requires an assumption of spectral density function for the phase noise [61–64].

The quality of the oscillator design is determined by comparing the noise level in the second-order roll-off region. This level is a measure for the noise in the core of the oscillation process.

Fig. 2.88 A Colpitts oscillator and a cross-coupled LC oscillator
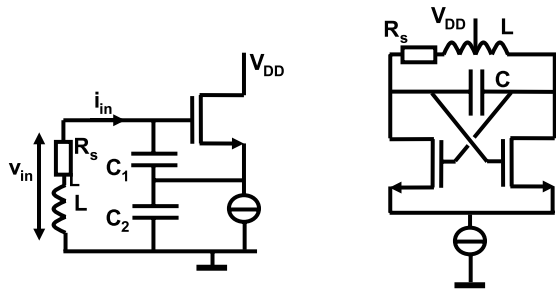
$$\text{Oscillator number} = 10\log(S_{ff}(\omega_m)) + 20\log\left(\frac{\omega_m}{\omega_0}\right) = \text{CNR}(\omega_m) - 10\log\left(\frac{\omega_0^2}{\omega_m^2}\right)$$

$$\text{Figure of Merit} = \text{Oscillator number} + 10\log\left(\frac{\text{Power}}{1\,\text{mW}}\right) \tag{2.206}$$

A typical LC oscillator shows an oscillator number in the range of -180 to -190 dB.

The oscillator number is often extended with a power term, to represent the power efficiency. This Figure of Merit compares the oscillator number normalized to 1 milliWatt. Good figures of Merit are between -160 dB for relaxation oscillators and -180 dB for high-Q LC oscillators [65, pp. 177–179].

Figure 2.88 (left) shows the basic circuit of a "Colpitts oscillator". The LC circuit resonance frequency is given by the coil and the series connection of the capacitors.[29]

$$\omega_0 = \frac{1}{\sqrt{L\frac{C_1 C_2}{C_1 + C_2}}} \tag{2.207}$$

The active element provides current into the capacitor connection. The resonance effect generates a voltage amplification at the gate that feeds current into the tank. A stable oscillation is achieved if the energy lost in the series resistor in the coil is compensated by the energy provided by the transistor. Equating these energies results in the handy observation that (seen from the branch with the coil and resistor) the capacitor-transistor circuit must behave as a negative resistor in amplitude equal to R_s. Analysis of that part of the circuit starts by looking at the currents on the node formed by the two capacitors and the transistor with the voltage v_s:

$$(v_{\text{in}} - v_s)j\omega C_1 + g_m(v_{\text{in}} - v_s) = v_s j\omega C_2 \tag{2.208}$$

This equation leads to an expression for v_s and the equivalent input impedance (seen from the branch with the coil and resistor) is found as:

$$\frac{v_{\text{in}}}{i_{\text{in}}} = \frac{1}{j\omega C_1} + \frac{1}{j\omega C_2} + \frac{g_m}{j\omega C_1 j\omega C_2} \tag{2.209}$$

---

[29]Ignoring the small effect that R_s has on the oscillation frequency.

The last term shows a negative impedance for positive values of g_m and the capacitors. If the total transconductance term at the oscillation frequency equals $-1/R_s$ the oscillation can be sustained. Huang derived the phase-noise relation for a Colpitts oscillator in [66].

Starting up the oscillator requires a much higher transconductance. A high quality factor also means that injecting energy into the resonant circuit requires many oscillation cycles: the order of magnitude is comparable to Q. After the oscillation has been built up, some form of limitation has to reduce the overall gain till the point is reached where the phase and amplification criteria are exactly met. This is a design challenge, as many forms of non-linear phenomena introduce specification loss. E.g. using the supply as a limiting factor introduces power supply coupling and operating the active device in an extreme regime can add unwanted noise contributions.

Simulation of oscillators by means of regular circuit simulators is tedious. Again the quality factor of the resonant element(s) is related to the (large) number of simulation cycles. Where thermal noise or switching noise creates the starting conditions in real life, a deliberate incident must be created to start the oscillation in a simulation.

The Colpitts oscillator exists in many variants, e.g. in the way the active element is connected to the resonant L-C circuit. Just as the Colpitts oscillator, topological variants carry the names of their inventors:

- An additional capacitor in series with the coil creates a "Clapp-oscillator". This capacitor can tune the frequency without changing the feedback ratio of C_1 and C_2.
- Changing the capacitors into coils and the coil into a capacitor results in the "Hartley oscillator". The capacitor in a Hartley-oscillator for the same frequency equals $C_h = C_1 C_2/(C_1 + C_2)$, which consumes considerably less chip area than $(C_1 + C_2)$. This fact can be an advantage because of less area, but also increases the sensitivity for parasitics.
- Figure 2.88 (right) depicts a cross-coupled LC oscillator. The resonant coil and capacitor circuit is hardly limited by voltage constraints. Also the loading by the transistors is minimum. Therefore this type of oscillator can reach a good phase-noise performance. The amplitude of the swing is controlled by the resistance of the coil and the current source in the tail of the circuit. Two LC-oscillators can be coupled to give a 90° phase shift.
- Replacing the coil in a Colpitts oscillator by a crystal leads to a "Pierce-oscillator".

The electrical components for implementing inductive and capacitive behavior limit the maximum achievable quality factor to the 100–1000 range. Higher quality factors require to use mechanical resonators like a quartz crystal as a resonator. Figure 2.89 shows the electrical equivalent circuit of a quartz crystal. A series circuit formed by C_1, L_1 and R_1 is shunted with a capacitor C_0. This last capacitor represents the electrical capacitance of the leads and package. C_0 dominates the frequency characteristic of the impedance, see Fig. 2.89 (right). C_1 and L_1 represent
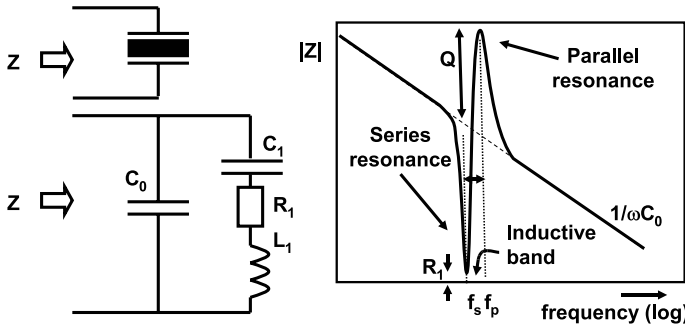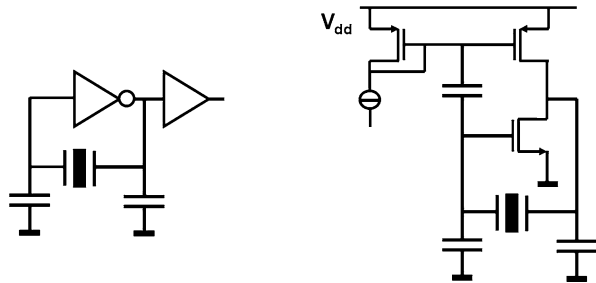
Fig. 2.89 The electrical equivalent circuit for crystal contains a series branch shunted with C_0. *Right*: A schematic view of the impedance of a crystal (not to proportion!)

Table 2.26 Parameters for the equivalent circuit of a quartz crystal

$C_0 = 4$ pF	$C_1 = 16$ fF
$H_1 = 5$ mH	$R_1 = 50\ \Omega$
$f_s = 18$ MHz	$Q = 10\,000$

Fig. 2.90 A Pierce crystal oscillator and a push-pull arrangement [68]

the motional capacitance and inductance. At the series resonance frequency f_s these two elements cancel and the total equivalent impedance of the crystal drops to R_1. At the parallel resonance frequency (also called anti-resonance frequency) the capacitor C_0 is in series with C_1, forming is slightly smaller capacitor than C_1. These two resonance frequencies and the quality factor are specified as:

$$f_s = \frac{1}{2\pi}\sqrt{\frac{1}{L_1 C_1}} \qquad f_p = \frac{1}{2\pi}\sqrt{\frac{C_0 + C_1}{L_1 C_1 C_0}} \qquad Q = \frac{1}{2\pi f_s R_1 C_1} \quad (2.210)$$

A high impedance occurs at a frequency $f_p = f_s(1 + C_1/2C_0)$. In between f_s and f_p the series circuit dominates the overall impedance and shows an inductive behavior. At frequencies above f_p the overall behavior follows again the shunt capacitor. Some simulation values are given in Table 2.26.

Quartz crystals and resonators are available from 32.768 Hz for watches up to 30 MHz. Crystals can show parasitic oscillation modes, called overtones. These overtones act like additional series branches in the equivalent circuit.
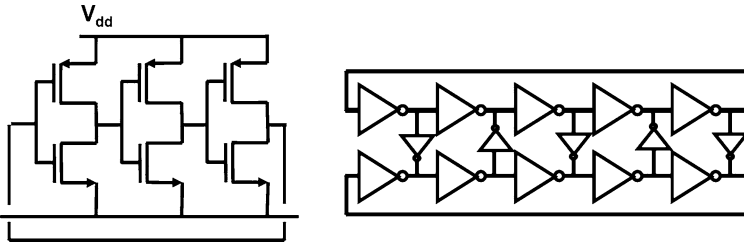
Fig. 2.91 A inverter ring as an oscillator and pseudo-differential ring oscillator

The stable mechanical construction allows to create high-quality oscillators. Comparison of Fig. 2.90 [67] with the Colpitts oscillator shows that the crystal must be operated in the inductive regime. This oscillator is mostly built from a simple digital inverter, with the crystal between in and output and the capacitors between input and ground and output and ground. A popular design extends the normal Pierce oscillator via a simple capacitor to a push-pull device, Fig. 2.90.

Coils in integrated circuits have disadvantages:

- In order to achieve a reasonable quality factor (e.g. $\omega L/R > 10$ at 1 GHz) wide interconnect tracks are needed and quite some area is used. Special lay-out tricks improve the quality factor somewhat.
- As the coil lies on the substrate, most of the magnetic field penetrates into the substrate causing eddy currents. These substrate currents act as resistive losses to the signal.
- All coils on one substrate have the same orientation, which can cause mutual coupling.

Oscillators without coils will have to resupply the full energy needed for every oscillation cycle. This will introduce noise contributions and reduce the quality of the signal. Nevertheless inductorless oscillators are popular for their low area consumption. Figure 2.91 (left) shows the standard digital ring oscillator. In its most primitive form it consists of a ring with an odd number of inverters. The delay of a switching inverter gate is the basis for the frequency definition.[30] This inverter delay is determined by the (dis)charging time of the intermediate node by the transistors:

$$t_{d,\mathrm{inv}} = \alpha_{\mathrm{inv}} \frac{C_{\mathrm{node}} V_{DD}}{I_{DS,\mathrm{sat}}} \tag{2.211}$$

where $\alpha_{\mathrm{inv}} \approx 0.5$ depending on e.g. the NMOS/PMOS ratio. If $I_{d,\mathrm{sat}} \propto (V_{DD} - V_T)^2$ as the quadratic equation promises, the delay of an inverter is inversely proportional to the supply voltage minus the threshold voltage, while the frequency is linearly proportional. This property is extensively used in voltage controlled oscillators (VCOs) for phase-locked loops. In order to improve the performance of the

---

[30]People advertising the virtues of technologies reverse this definition: the gate delay is the inverse fraction of the measured oscillation frequency.

oscillator a pseudo-differential approach is applied as is seen in Fig. 2.91 (right). The two inverter loops run in anti-phase. The anti-phase behavior is guaranteed by the small coupling inverters between the taps of the lines. The current consumption is relatively constant.

A mulitivibrator of Fig. 2.92 (left) was a popular timing circuit, when these circuits were built from discrete devices. The resistors are dimensioned to operate the transistors in a digital regime. If M_1 is fully conductive and M_2 is not, the only activity in the circuit is the loading of the right hand side of C_1 by R_2. If the gate voltage of M_2 passes the threshold the drain voltage will start dropping and will drive the gate voltage of M_1 down. The drain voltage of M_1 will increase towards the power supply, thereby pushing via C_1 the gate voltage of M_2 even further up. When the drain of M_2 reaches a low level and the drain of M_1 is close to the power supply, the opposite situation has been reached. Some precautions must be taken to clip the gate voltages to ground level. The frequency of this multivibrator is determined by the time constants on the gate nodes:

$$f = \alpha_{\mathrm{mul}} \frac{1}{R_1 C_2 + R_2 C_1} \tag{2.212}$$

with $\alpha \approx 0.7$. Recently multivibrators that are more adopted to integrated circuit design as in Fig. 2.92 (right) have been developed. Now two current sources charge and discharge a capacitor. The amplitude of the oscillation is determined by the hysteresis of the comparator. The noise sources in this circuit have been analyzed and the topology has been improved [69] to yield comparable results with practical LC oscillators.

The main problem with integrated oscillators based on charging of capacitors, is the reproducibility and the stability of the current due to process variations and the temperature dependence of the mobility. Referencing the current to high quality resistors requires extra processing steps. Compensating the temperature coefficient allows accuracies in the order of 1% [70].

It is not always desired to generate a free-running timing signal. Often some form of time reference is present in a system and a fraction or multiple of that frequency

Fig. 2.92 A traditional astable multivibrator and a current-source multivibrator
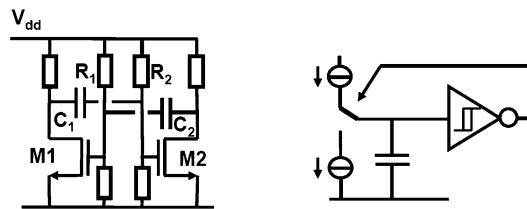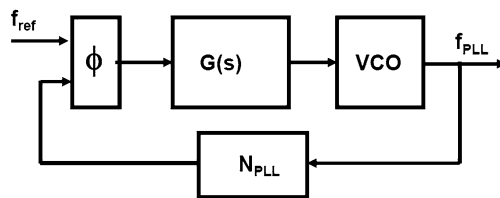

Fig. 2.93 A phase-locked loop circuit

is needed. A circuit that locks the oscillation frequency to a reference frequency is a "Phase-Locked Loop" (PLL). Figure 2.93 shows the basic set up of such a loop. Central is the voltage-controlled oscillator. In a fully integrated solution the oscillator is often implemented as a ring oscillator. This oscillator is controlled by a signal that is derived from comparing the reference signal to a fraction N_{PLL} of the oscillator signal. If the loop is properly designed the output frequency will be equal to $f_{out} = N_{PLL} f_{ref}$.

The transfer of the phase demodulator can be modeled as: $v_d = K_d(\theta_{ref} - \theta_{out}/N_{PLL})$. The transfer of the filter is $H(s)$. The voltage controlled oscillator will generate an output frequency proportional to the input voltage $f_{out} = K_v v_{osc}$. The phase is the time integral of the frequency $f_{out}(s) = s\theta_{out}(s)$, therefore an additional $1/s$ factor must be added in the transfer for the oscillator. This results in an overall transfer function of:

$$\frac{\theta_{out}(s)}{\theta_{ref}(s)} = \frac{K_v K_d H(s)/s}{1 + K_v K_d H(s)/N_{PLL}s} \qquad (2.213)$$

There are three main noise sources in a Phase-locked loop: the noise of the reference signal, the noise contributed by the phase detector and the noise in the oscillator itself. The reference noise and the detector noise are filtered by the loop filter and can be reduced by choosing a narrow band loop filter. However in the pass band of the loop-filter these components dominate. The noise of the VCO itself is suppressed by the loop. At higher frequencies this noise becomes dominant. In order to suppress this noise source, the loop-filter should be wide band.

Chapter 3
Sampling

3.1 Sampling in Time and Frequency

Sampling is a common technical process. A movie consists out of a sequences of photographs (the samples), a newspaper photograph has been grated and a television broadcast consists out of a sequence of half pictures, etc. The sampling process determines the value of a signal on a predetermined frame of time moments. The sampling frequency f_s defines this frame and determines the sampling moments as:

$$t = \frac{n}{f_s} = nT_s, \quad n = -\infty, \ldots, -3, -2, -1, 0, 1, 2, 3, \ldots, \infty \quad (3.1)$$

Between the sampling moments there is a time frame T_s, where strictly mathematically speaking no value is defined. In practice this time period is used to perform operations on the sample sequence. The various operations (summation, multiplication, delay) that are possible form the class of time-discrete signal processing, e.g. [71, 72]. In this book the value of T_s is considered constant, resulting in a uniform sampling pattern. Generalized non-uniform sampling requires extensive mathematical tools.

An analog-to-digital converter samples the signal and delivers a sequence of values on which these subsequent operations can be performed. The mathematical description of this process uses the "Dirac" function. This function $\delta(t)$ is a strange mathematical construct as it is only defined within the context of an integral.[1] The integral is defined in such a way that the result of the integral equals the value of the integral function at the position of the integration variable that is given by the Dirac function's argument.

$$\int_{t=-\infty}^{\infty} f(t)\delta(t - t_0)\, dt = f(t_0) \quad (3.2)$$

---

[1] Strange in the sense that many normal mathematical operations cannot be performed, e.g. $\delta^2(t)$ does not exist.

The dimension of the Dirac-function is the inverse of the dimension of the integration variable. A more popular, but not exact, description states that the integral over a Dirac-function approximates the value "1":

$$\delta(t) = \begin{cases} 0, & -\infty < t < 0 \\ \frac{1}{\epsilon}, & 0 < t < \epsilon \\ 0, & \epsilon < t < \infty \end{cases} \qquad \Rightarrow \qquad \int_{t=-\infty}^{\infty} \delta(t)\, dt = 1 \tag{3.3}$$

with $\epsilon \to 0$.

A sequence of Dirac pulses mutually separated by a time period T_s defines the time frame needed for sampling:

$$\sum_{n=-\infty}^{n=\infty} \delta(t - nT_s)$$

This repetitive sequence of pulses with a mutual time spacing of T_s can be equated to a Fourier series. This Fourier series will have frequency components with a base frequency $f_s = 1/T_s$ and all integer multiples of f_s. The multiplication factor for each frequency component kf_s is C_k. Equating both series:

$$\sum_{n=-\infty}^{n=\infty} \delta(t - nT_s) = \sum_{k=-\infty}^{\infty} C_k e^{jk2\pi f_s t} \tag{3.4}$$

The coefficients C_k of the resulting Fourier series are found by integrating over a period, see the example (2.5). The additional factor 2 in (2.5) is due to the single sided range.

$$C_k = \frac{1}{T_s} \int_{t=-T_s/2}^{T_s/2} \sum_{n=-\infty}^{n=\infty} \delta(t - nT_s) e^{-jk2\pi f_s t}\, dt \tag{3.5}$$

Within the integration interval there is only one active Dirac pulse at $t = 0$, so the complicated formula reduces to:

$$C_k = \frac{1}{T_s} \int_{t=-T_s/2}^{T_s/2} \delta(t) e^{-jk2\pi f_s t}\, dt = \frac{1}{T_s} e^{-jk2\pi f_s \times 0} = \frac{1}{T_s} \tag{3.6}$$

Now the substitution of C_k results in the mathematical description of the Fourier transform from the sequence of Dirac pulses in the time domain.

$$\sum_{n=-\infty}^{n=\infty} \delta(t - nT_s) = \frac{1}{T_s} \sum_{k=-\infty}^{\infty} e^{jk2\pi f_s t}$$

$$= \frac{1}{T_s} \int_{f=-\infty}^{\infty} \sum_{k=-\infty}^{k=\infty} \delta(f - kf_s) e^{jk2\pi f_s t}\, df \tag{3.7}$$

The last term is the standard inverse Fourier transform with a summation of frequencies as argument. This sum of Dirac functions in the Fourier series is therefore the

Fig. 3.1 Sampling an analog signal (**a**) in the time domain results in a series of signal samples (**b**). In the frequency domain the analog signal (**c**) is folded around the sampling frequency and its multiples (**d**)
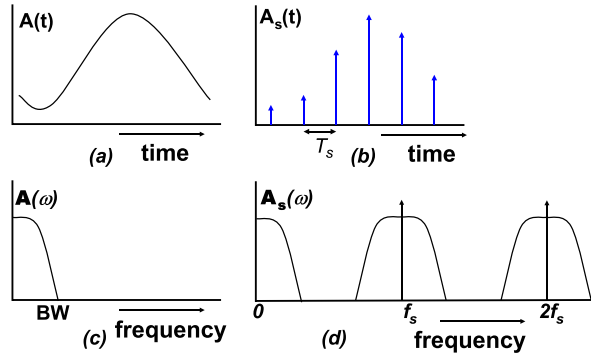
frequency domain counterpart of the time Dirac series.

$$\sum_{n=-\infty}^{n=\infty} \delta(t - nT_s) \quad \Leftrightarrow \quad \frac{1}{T_s} \sum_{k=-\infty}^{k=\infty} \delta(f - kf_s) \tag{3.8}$$

The infinite sequence of short time pulses corresponds to an infinite sequence of frequency components at multiples of the sampling rate.

In Fig. 3.1 an example is presented of the sampling of a signal $A(t)$ corresponding to[2] $\mathbf{A}(\omega) = \mathbf{A}(2\pi f)$ with a bandwidth from $f = 0$ Hz to $f = BW$.

$$\mathbf{A}(\omega) = \int_{t=-\infty}^{\infty} A(t)e^{-j2\pi ft}\, dt \tag{3.9}$$

Mathematically the sampling is performed by multiplying the time-continuous function $A(t)$ of Fig. 3.1(a) with the sequence of Dirac-pulses, resulting in a time discrete signal Fig. 3.1(b). The product of the time-continuous function and the Dirac sequence is defined for those time moments equal to the multiples of the sampling period T_s

$$A_s(t) = \sum_{n=-\infty}^{n=\infty} A(t) \times \delta(t - nT_s) \quad \Rightarrow \quad \sum_{n=-\infty}^{n=\infty} A(nT_s) \tag{3.10}$$

Using the previous description of the sampling signal in the frequency domain, the time sequence of samples $A_s(t)$ of the continuous function $A(t)$ is described in the frequency domain as $\mathbf{A_s}(\omega)$:

$$\mathbf{A_s}(\omega) = \int_{t=-\infty}^{\infty} \left[\sum_{n=-\infty}^{n=\infty} A(t)\delta(t - nT_s) \right] e^{-j2\pi ft}\, dt$$

$$= \int_{t=-\infty}^{\infty} A(t) \frac{1}{T_s} \sum_{k=-\infty}^{\infty} e^{jk2\pi f_s t} e^{-j2\pi ft}\, dt$$

---

[2]To keep in this section time-continuous and sampled sequences and their spectra apart, time domain signals use normal print, while their spectral equivalents use bold face. The suffix s refers to sample sequences.
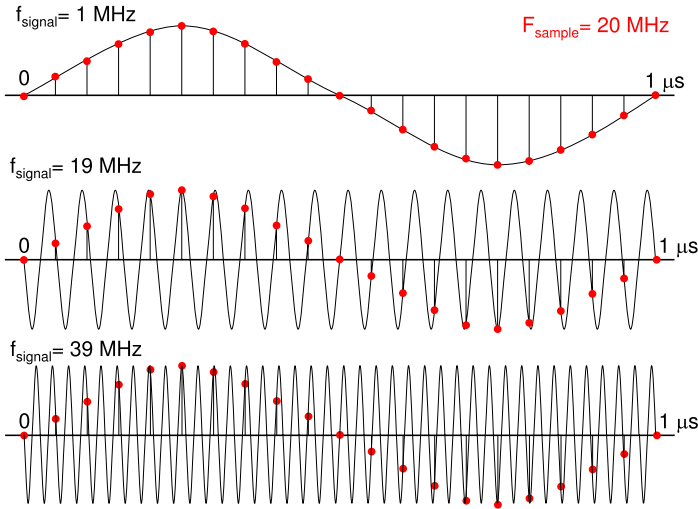
Fig. 3.2 Aliasing: Three analog signals: 1 MHz, 19 MHz and 39 MHz sine waves result after sampling with 20 Ms/s in the same sampled data sequence (*dots*)

$$= \sum_{k=-\infty}^{\infty} \frac{1}{T_s} \int_{t=-\infty}^{\infty} A(t)e^{-j2\pi(f-kf_s)t}\,dt \qquad (3.11)$$

Comparing the last integral with the previous transform for $\mathbf{A}(\omega)$, results in the observation that the integral equals the Fourier transform with a frequency shift of $k \times \omega_s$. The total spectrum $\mathbf{A_s}$ can be written as:

$$\mathbf{A_s}(\omega) = \sum_{k=-\infty}^{\infty} \frac{1}{T_s}\mathbf{A}(\omega - k\omega_s) = \sum_{k=-\infty}^{\infty} \frac{1}{T_s}\mathbf{A}(2\pi(f - kf_s)) \qquad (3.12)$$

The time continuous signal $A(t)$ is connected to only one spectrum band in the frequency domain $\mathbf{A}(\omega)$. By sampling this signal with a sequence of Dirac-pulses with a repetition rate ($f_s = 1/T_s$) a number of replica's of the original spectral band $\mathbf{A}(\omega)$ are created on either side of each multiple of the sampling rate f_s. Figures 3.1(c) and 3.1(d) depict the time-continuous signal and the sampled signal in the frequency domain. In the frequency domain of the sampled data signal, next to the original signal, also the upper bands are shown. The closest upper band directly adjacent to the original band is called: "the alias band".

Another consequence of sampling is that the useful signal band is only $0.5f_s$ wide. Every signal in the analog domain is mapped into this bandwidth. Frequency components in the analog domain that have an equal frequency distance to an arbitrary multiple of the sampling frequency will end up on the same frequency location in the sampled data band. Figure 3.2 shows three different sine wave signals that all result in the same sampled data signal (dots).

The idea that from one spectrum an infinite set of spectra is created seems to contradict the law on the conservation of energy. If all spectra where mutually uncorre-

Fig. 3.3 The analog signal contains higher frequency components and does not satisfy the Nyquist criterion. Both the sample series in the time domain (**b**) as well as the sampled spectrum that overlaps with the components around the sample frequency (**d**), do not allow reconstructing uniquely the original analog signal
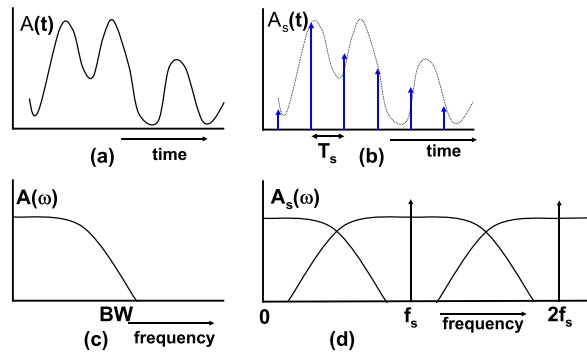
lated, there would indeed be a contradiction. However, as all spectra are correlated and hence mutually dependent, the summation of the energies remains equal to the energy of the original signal.

While describing the signals in the previous paragraphs, implicitly the band of interest was assumed to be a base-band signal, starting at 0 Hz with a bandwidth BW. This is the situation that exists in most data-acquisition systems. The alias bands will appear around the sample rate and its harmonics. This choice for this location of the band of interest is by no means obligatory. A band of interest located on a higher frequency, or even beyond the sample rate, is equally sampled. The signal band can be regarded as being sampled by the closest multiple of the sample rate. This band is again copied to all harmonics of the sample rate, including "0 Hz". This process is called "under-sampling" or "sub-sampling". If there are components of the signal lying above and below a harmonic of the sample rate, both of these will be sampled into the same frequency region. The consequence is then an overlap of signals. Deliberate forms of sub-sampling are used in communication applications, where sub-sampling is used as demodulation, see Fig. 3.6. Undeliberate forms of sub-sampling occur if undesired signals are present in the signal band. Examples are:

- Harmonic distortion products of the base-band signal.
- Thermal noise in the entire input band, see Sect. 3.1.3.
- Interference signals from other parts of the equipment or antenna.

3.1.1 Folding Back of Spectra

Figure 3.3 shows a signal that spans a much larger bandwidth than the signal in Fig. 3.1. The samples of this signal are valid values of the signal at that sample moment, however, it is not possible to reconstruct uniquely the signal based on these values. Figure 3.3 shows that a bandwidth larger than half of the sample rate will cause the alias-band to mix up with the base band. In a correctly designed analog-to-digital conversion system the bandwidths of the incoming signal is therefore limited by means of an "alias-filter", so that no mixing can take place.

This limitation during the sampling of signals is known as the "Nyquist" criterion. Indicated already in a paper by H. Nyquist [73, Appendix 2-a], it was C.E. Shannon who extended his mathematical theory of communication [74] in 1949 with a paper dealing with communication in the presence of noise. In that paper [75] Nyquist's criterion is formulated as:

> "If a function contains no frequencies higher than BW cycles per second, it is completely determined by giving its ordinates at a series of points spaced 1/2 BW seconds apart".

This criterion imposes a simple mathematical relation on the bandwidth BW and the sample rate f_s:

$$f_s > 2BW \tag{3.13}$$

This criterion is derived assuming ideal filters and an infinite time period to reconstruct the signal. These constraints are in practical circumstances never achieved. An example coming close is the compact-disc music recording format where the sample rate[3] represents a desired signal bandwidth of 20 kHz. This combination leaves only a small transition band between 20 and 24.1 kHz to suppress the alias band by some 90 dB. The expensive filter required to achieve this suppression needs some 11 to 13 poles. Moreover such a filter will create a non-linear phase behavior at the high base band frequencies. Phase distortions are time distortions (Δphase equals signal frequency $\times \Delta$time) and have a strong audible effect. Fortunately the use of "oversampling" allows to separate base band and alias band sufficiently, see Sect. 9.1.

An interesting discussion on present insights in the mathematical aspects of the Nyquist theorem was published by Unser [76].

The Nyquist criterion specifies that the bandwidth is limited by the sample rate. The only constraint on where this limited bandwidth is positioned in the time-continuous spectrum is that this bandwidth does not include a multiple of half of the sample rate. That would lead to overlap. However there is no need to specify the bandwidth starting from 0 Hz. If it is known that the original signal in Fig. 3.2 is in the bandwidth between 10 and 20 MHz, the samples can be reconstructed to the originating 19 MHz time-continuous sine wave. The property that also a frequency range above the sample rate is properly sampled and generates copies around all multiples of the sample rate, is in some communication systems used to down-modulate or down-sample signals, see Sect. 3.1.2.

An implicit assumption for the Nyquist criterion is that the bandwidth of interest is filled with relevant information. This is not necessarily true in all systems. Video signals are by their nature sampled signals: a sequence of images consisting of sequences of lines. The spectral energies are concentrated around multiples of the video line frequency. The intermediate frequency bands are empty and sampling mechanisms in video use this property.

---

[3]The only storage in the early days of CDs were video recorders. The 44.1 ks/s sample rate was chosen such that the audio signal exactly fits to a PAL video recorder format (25 fields of 588 lines with 3 samples per line) of 44.1 ks/s.
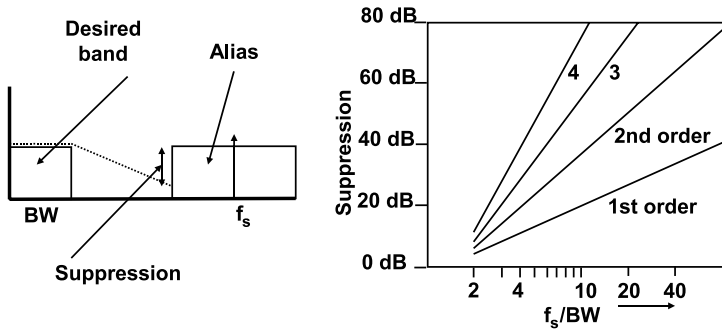
Fig. 3.4 The attainable suppression of the anti-alias filter depends on the number of poles in the filter and the ratio of the bandwidth to the sample rate

An even more advanced approach is observed in the theory of compressive sensing or compressive sampling [77]. In communication systems often only a limited number of carriers are simultaneously active. A sparse signal in a relatively wide bandwidth can be reconstructed after sampling by a non-uniform sampling sequence. Such a sequence can be generated by a high-frequency random generator. The information from the few active carriers is spread out over the band and theoretically it is possible to design algorithms that recover this information. A first intuitive approach is to assume a high uniform sampling pattern, from which only a few selected samples are used. In a higher sense the Nyquist criterion is still valid: the total amount of relevant bandwidth is still (far) less than the effective sample rate. The theory is promising. Whether a real advantage can be obtained remains to be proven.

The Nyquist criterion forces the input signals to be band-limited. An analog-to-digital converter is therefore preceded by a band-limiting filter: the anti-aliasing filter. This filter prevents the components outside the desired frequency range to be sampled and to mix up with the wanted signals. In practical system design it is recommended to choose a higher sample rate than prescribed by the Nyquist criterion. The fraction of frequency spacing between the extremes of the base and its alias with respect to the sample rate determines the number of poles needed in the anti-alias filter. A filter will suppress signals at a rate of 6 dB per octave per filter pole, Fig. 3.4.

Sharp band-limiting filters require many accurately tuned poles. Additional amplification is needed, and therefore these filter tend to become expensive and hard-to-handle in a production environment. On the other hand there are some good reasons not to choose for an arbitrary high sample rate: the required capacity for storing the digital data will increase linear with the sample rate, as well as the power needed for any subsequent data processing.

Anti-alias filters are active or passive time-continuous filters: time-discrete filters, such as switched-capacitor filters, sample the signal themselves and require consequently some alias filters. An additional function of the anti-alias filter can be the suppression of unpredictable interference in the system. Of course there should be an equal interest in suppressing any interference on supply lines. Some systems
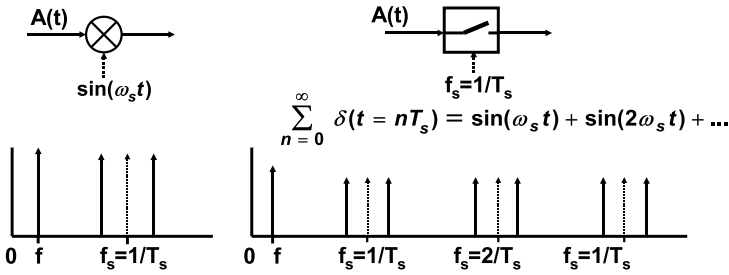
Fig. 3.5 Modulation and sampling of signals. Ideally the modulation and sampling frequencies disappear from the resulting spectrum. Here they are indicated for reference
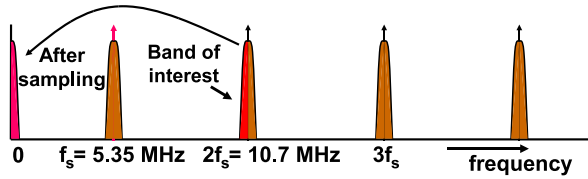

Fig. 3.6 Demodulation and sampling of an IF-FM signal at 10.7 MHz by a 5.35 Ms/s sample pulse

are band-limited by construction. In a radio, the IF filters of a heterodyne radio architecture may serve as anti-alias filters, and in a sensor system, the sensor may be band limited.

3.1.2 Sampling and Modulation

Sampling of signals does bear some resemblance to modulation of signals. In both cases the operation results in the creation of frequency shifted bands of the original signal. A modulator multiplies the base band signal with a sine wave, resulting in a pair of upper bands around the carrier frequency, see Fig. 3.5.

In contrast to modulation, sampling results in upper bands around every multiple of the sample rate. The sequence of Dirac-pulses is equivalent to a summation of sine waves with frequencies at multiples of the sample rate.

$$D_s(\omega) = \frac{2\pi}{T_s} \sum_{k=-\infty}^{k=\infty} \delta\left(\omega - \frac{2\pi k}{T_s}\right) \tag{3.14}$$

Therefore sampling can be viewed as a summation of modulations. The intrinsic similarity between sampling and modulation can be used in certain systems: an example is found in down-mixing of radio frequency signals.

For example an FM-radio IF signal of 100 kHz at a carrier frequency of 10.7 MHz, can be down modulated and sampled at the same time by a 5.35 Ms/s signal, Fig. 3.6.
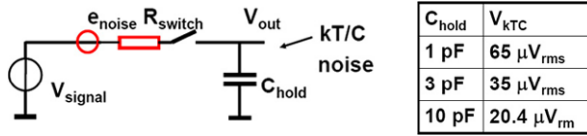
C_{hold}	V_{kTC}
1 pF	65 μV_{rms}
3 pF	35 μV_{rms}
10 pF	20.4 μV_{rm}

Fig. 3.7 Switched capacitor noise sampling: the series resistances act as a noise sources

3.1.3 Sampling of Noise

Figure 3.7 shows an equivalent schematic of a basic sampling circuit consisting of a switch and storage capacitor. Compared to the ideal situation two non-ideal elements have been added to the switch: the switch resistance R combining all resistive elements between source and capacitor. The resistor is impaired with thermal noise, consequently a noise source is added e_{noise} whose spectrum reaches far beyond the sampling rate of the switch.

$$e_{noise} = \sqrt{4kTRBW} \tag{3.15}$$

with Boltzmann's constant $k = 1.38 \times 10^{-23}$ m^2 kg s^{-2} K^{-1} and the absolute temperature T in Kelvin. This formulation expresses the noise in the positive frequency domain $f \in [0, \infty]$. As a consequence each multiple of the sampling frequency will modulate the adjacent noise back to the base band, where all the noise accumulates.

When the switch connects to the capacitor, a low pass filter is formed by the resistor and the capacitor. The average noise energy on the capacitor is therefore a filtered version of the noise energy supplied by the resistor and is filtered by the complex conjugated transfer function of the RC network. Using the integral in Table 2.6:

$$v_{C,noise}^2 = \int_{f=0}^{f=\infty} \frac{4kTR\,df}{1+(2\pi f)^2 R^2 C^2} = \frac{kT}{C} \quad \Rightarrow \quad v_{C,noise} = \sqrt{\frac{kT}{C}} \tag{3.16}$$

The simple and well-known expression for the sampled noise on a capacitor is called: kT/C-noise. The magnitude of the resistor (the origin of the noise) is not part of this first-order expression. On one hand an increase of the resistor value will increase the noise energy proportionally, however, that same increase in resistor value will reduce the relevant bandwidth also proportionally.

Equation (3.16) holds for the time-continuous case, where the switch is permanently conductive, but holds equally for the sampled situation. Although the signals looks completely different in the time domain, both the time continuous and the sampled noise signal have values taken from a normal distribution with mean value zero and a $\sigma^2 = kT/C$.

This kT/C noise can be interpreted as a flat spectrum in the band from "0" to $f_s/2$ as long as the RC cut-off frequency largely exceeds the sample rate. If the RC cut-off frequency is low the noise bandwidth must be treated in a similar fashion as a normal signal band, see Fig. 3.8.

Fig. 3.8 Band-limited noise
is sampled in a similar
manner as normal signals.
Top: this noise has a finite
analog bandwidth, *bottom*:
after sampling. The power
spectra add up and are
mirrored
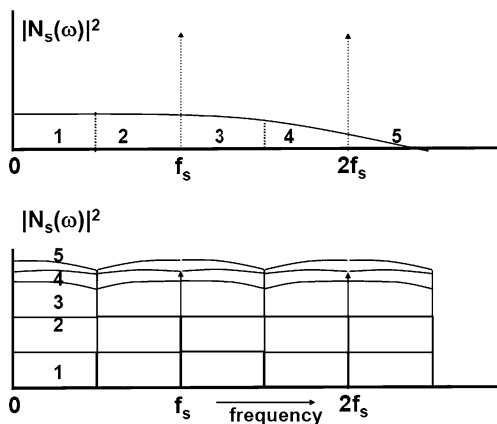
This kT/C noise term presents a lower boundary in choosing the value for a sampling capacitance. An analog-to-digital converter is signal-to-noise limited because of this choice. A circuit with a total sampling capacitance of 1 pF will be limited with a noise voltage floor of 65 μV_{rms} at room temperature. Unfortunately, a large capacitance value will require IC area, and will directly impact the power budget.

Example In a process with a supply voltage of 1.2 V a sinusoidal signal of 500 mV (peak-peak) corresponds to an RMS voltage of $500/2\sqrt{(2)} = 177$ mV. With a signal-to-noise ratio of 4000 (corresponding to a 12 bit ADC) the kT/C noise must be lower than 44 μV_{rms}, and a minimum capacitor of 2.15 pF is needed. A sinusoidal signal with a frequency 100 MHz requires a current of $+/-0.34$ mA. The charge on the capacitor has to be supplied, and bias current of 1 mA must be supplied to avoid slew-rate behavior. In first order the circuit will consume 1.2 mW.

The spectral power noise density (power per Hz) of kT/C noise in a sampled system is equal to kT/C over half of the sample rate:

$$S_{ff,SH} = \frac{2kT}{Cf_s} \tag{3.17}$$

The power noise density of the time-continuous network with the same resistor and capacitor having a cut-off frequency of $f_{RC} = 1/2\pi RC$ in its pass-band is:

$$S_{ff,rc} = 4kTR = \frac{2kT}{\pi Cf_{rc}} \tag{3.18}$$

The comparison of these two noise densities shows that in the sampling process the noise density increases by a factor $\pi f_{rc}/f_s$. This factor corresponds to the number of bands that stack up in Fig. 3.8. This considerable increase in noise density causes major problems when designing high-resolution converters.

The switching sequence can influence the total noise accumulated in the circuit. In switched capacitor circuits every switch cycle will add one portion of noise. As these noise portions are uncorrelated, they will sum in and root-mean-square way.
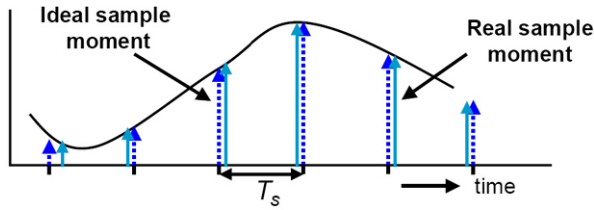
Fig. 3.9 The ideal sampling moments (*dashed*) shift in an arbitrary fashion in time if the sample clock is disturbed by jitter

Also in situations where a switch discharges the charge of a capacitor into a fixed voltage kT/C noise will appear ("reset-noise").

3.1.4 Jitter of the Sampling Pulse

In the previous analysis it was assumed that the sample moments are defined with infinite precision. In practice all signals that define time moments have limited bandwidths, which means that there is no infinitely sharp rising edge. Oscillators, buffers and amplifiers are all noisy devices, so consequently they add noise to these edges in Fig. 3.9. The noise changes the switching level of the digital buffer in Fig. 3.10 therefore the outgoing edge will have no fixed delay with respect to the incoming edge. This effect is called: jitter. Jitter causes sample moments to shift slightly from their position, and consequently sample the signal at another position. Next to noise-like components also signal-related components may influence the clock edge. Jitter from noisy sources will result in noise contributions to the signal, jitter from deterministic sources leads to tones (from fixed carriers) or to distortion (if the jitter source is correlated to the signal). Examples of systematic offsets in timing are: skews due to unequal propagation paths of clocks, interference from subdivided clocks, loading of clock lines and clock doubling by means of edge detection. Random "jitter" variations occur during the generation of clock signals in noise-sensitive oscillators, PLLs, long chains of clock buffers fed by noisy digital power supplies, etc. A practical value for jitter on a clock edge in a digital CMOS environment is 30–100 ps$_{\text{rms}}$.[4]

See Sect. 2.7.16 for some theoretical background and the relation to phase-noise.

Figure 3.9 shows the effect of shifting sample moments. If a sinusoidal signal with a radial frequency ω is sampled by a sample pulse with jitter, the amplitude error is estimated as:

$$A(nT_s + \Delta T(t)) = \hat{A}\sin(\omega \times (nT_s + \Delta T(t))) \tag{3.19}$$

---

[4]A peak-peak value is often used for jitter, but peak-peak values for stochastic processes have no significance if the process and the corresponding number of observations is not identified.
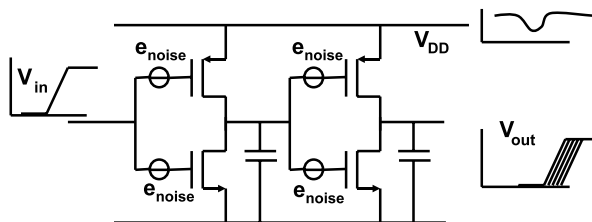
Fig. 3.10 A clock buffer for generating the digital sample signal can add to an ideal sample signal some noise of the buffer transistors. Also fluctuations on the power supply will affect the switching behavior of the buffer, causing uncertainty on the edges and jitter in the sampling

Table 3.1 Jitter specifications of some commercially available parts

Part	Description	Jitter
"2011"	Quartz 50–170 MHz	3 ps rms
"8002"	Programmable oscillator	25 ps rms
"1028"	MEMS+PLL combi 100 MHz	95 ps rms
"6909"	RC oscillator 20 MHz	0.2%
"555"	RC oscillator/timer	>50 ns rms

$$\Delta A(nT_s) = \frac{d\hat{A}\sin(\omega t)}{dt} \times \Delta T(nT_s) = \omega \hat{A}\cos(\omega nT_s)\Delta T(nT_s) \qquad (3.20)$$

The amplitude error is proportional to the slope of the signal and the magnitude of the time error. If the time error is replaced by the standard deviation σ_{jit} describing the timing jitter, the standard deviation of the amplitude σ_A is estimated as:

$$\sigma_A = \sqrt{\frac{1}{T}\int_{t=0}^{T}(\omega \hat{A}\cos(\omega t)\sigma_{jit})^2\, dt} = \frac{\omega \hat{A}\sigma_{jit}}{\sqrt{2}} \qquad (3.21)$$

Comparing this result to the root-mean-square value of the sine wave $\hat{A}/\sqrt{2}$ over the time period T results in the signal to noise ratio:

$$\text{SNR} = \left(\frac{1}{\omega\sigma_{jit}}\right)^2 = \left(\frac{1}{2\pi f \sigma_{jit}}\right)^2 \qquad (3.22)$$

or in deciBel (dB):

$$\text{SNR} = 20^{10}\log\left(\frac{1}{\omega\sigma_{jit}}\right) = 20^{10}\log\left(\frac{1}{2\pi f \sigma_{jit}}\right) \qquad (3.23)$$

For sampled signals the above relations hold for the ratio between the signal power and the noise in half of the sampling band. This simple relation estimates the effect of jitter, assuming no signal dependencies. Nevertheless it is a useful formula to make a first order estimate.

Figure 3.11 shows the signal to noise ratio as a function of the input frequency for three values of the standard deviation of the time jitter. As a reference Table 3.1 indicates some jitter numbers from commercial timing components.
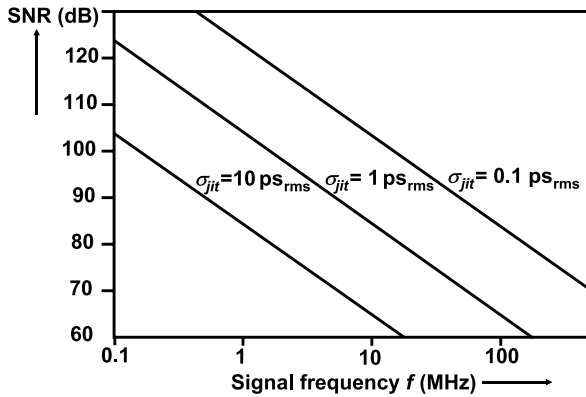
Fig. 3.11 The signal-to-noise ratio depends on the jitter of the sampling signal and the frequency of the analog signal
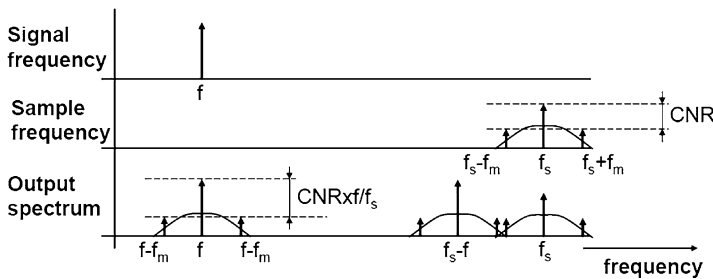

Fig. 3.12 Jitter around the sampling frequency will produce side spectra around the input tone

The linear dependence of jitter noise to the input frequency often allows a rapid identification of jitter in a time-discrete system. Another point of recognition is the flattening of the SNR versus input amplitude curve [78].

Jitter is here described as a random time phenomena. Mostly jitter shows a multitude of frequency components. In a Phase-Locked Loop circuit some typical components can be observed, see Fig. 2.87, such as

- White noise in the output (no dependency on the frequency).
- White noise that modulates the oscillator shows up in the power spectrum with a decreasing slope of $1/f^2$ from the oscillation frequency.
- PLLs multiply a reference frequency. Often spurious tones are visible on both sides of the generated frequency at a distance equal to the reference frequency.
- Undesired tones entering the PLL via substrate coupling and modulate the output.

From a spectral point of view, the jitter spectrum modulates the input tone. Therefore the jitter spectrum around the sampling pulse will return around the input frequency as in Fig. 3.12.   Translated to a lower frequency the time error due to jitter will produce a proportionally smaller amplitude error. Therefore the carrier-to-noise ratio (CNR) improves.

If jitter is caused by delay variations in digital cells as shown in Fig. 3.10, the jitter can contain signal components and strong spurious components e.g. linked to periodic processes in the digital domain. These contributions are demodulated similar as in Fig. 3.12 and are the source for spurious components and signal distortion. Also the digital circuits that generate and propagate the sample pulse must be treated as if these were analog blocks.

3.2 Time-discrete Filtering

Time-discrete filtering is used for up- and down sampling of spectra in digital-to-analog converters of Sect. 9.1 and forms a sub set of time-discrete filtering e.g. [71, 72]. In analog-to-digital converters, especially the sigma-delta modulators of Sect. 9.4 apply time-discrete filters, both in their architecture as well as in the necessary post-processing.

3.2.1 FIR Filters

Sampled signals can easily be delayed in the time-discrete domain. Switched-capacitor techniques in various implementation styles are the most common examples. In the analog time-discrete domain, charge packets are transferred from one capacitor into another capacitor by means of switching sequences. After amplitude quantization samples can also be delayed in the digital domain via digital delay cells, such as flipflops. Frequency filters are realized by combining the time-delayed sampled with specific weighting factors or multiplying coefficients.

The most simple approach to filtering in the time-discrete domain is the comb filter. The addition of a time-discrete signal to an m sample periods delayed signal gives a frequency transfer that can be evaluated in the Laplace domain:

$$H(s) = 1 \pm e^{-smT_s} = e^{-smT_s/2}(e^{+smT_s/2} \pm e^{-smT_s/2})$$

$$|H(\omega)| = 2|\cos(\omega m T_s/2)|, \quad \text{addition} \tag{3.24}$$

$$|H(\omega)| = 2|\sin(\omega m T_s/2)|, \quad \text{subtraction}$$

where the sign at the summation point determines whether the cosine response (with equal signs) or the sine response (with opposite signs) applies, see Fig. 3.13. Comb filters are mostly applied in systems where signals have to be separated. An example is the analog composite video signal, where the frequency carriers with the color information are interleaved between the carriers for the luminance signal.

The comb-filter adds signals to their delayed versions. A more general approach uses multiple delays, Fig. 3.14. A filter with this structure is known as a "Finite Impulse Response" filter (FIR-filter). The term "finite" means that any input disappears from the filter after passing through the N delay elements. In the summation the signals from the different delay elements can enhance or extinguish each other
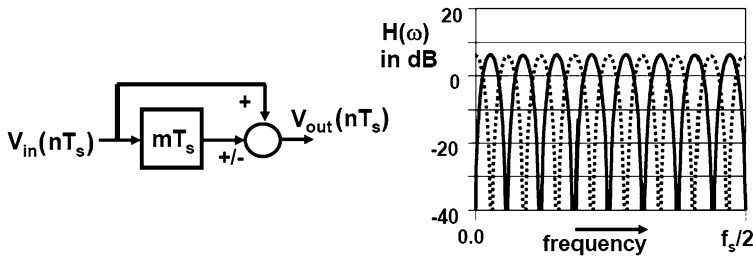
Fig. 3.13 The comb-filter and its frequency response. The *solid line* is the sine response, while the *dotted line* represents the cosine response
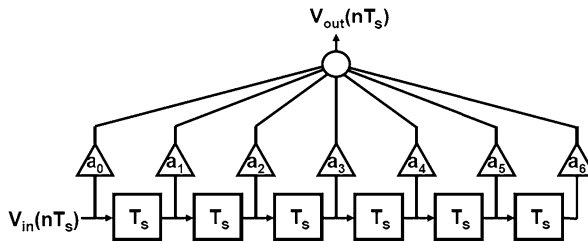

Fig. 3.14 The general structure of a finite impulse filter

depending on the periodicity of the signal with respect to the delay time and the multiplication factor. The filter designer can adapt the filter characteristic through these multiplication coefficients or weight factors. Similar to the time-continuous case in Sect. 2.6.8 the discrete-time transfer function defines the relation between input, filter transfer function and output in the time domain with a convolution:

$$y(nT_s) = \sum_{k=0}^{k=\infty} h(k)x(nT_s - kT_s) \tag{3.25}$$

Applied to the filter in Fig. 3.14, this gives:

$$V_{\text{out}}(nT_s) = \sum_{k=0}^{k=N-1} a_k V_{\text{in}}((n - k)T_s) \tag{3.26}$$

A formal mathematical treatment requires to use the z-transform, see Sect. 2.1.5 and [71, 72]. z^{-1} corresponds to a delay of one sample period. The transfer of a FIR filter results in the function:

$$H(z) = \sum_{k=0}^{k=N-1} a_k z^{-k} \tag{3.27}$$

Table 3.2 Coefficient values
for the low-pass FIR filter of
Fig. 3.15

Coefficient	Value
$a_0 = a_6$	−0.06
$a_1 = a_5$	0.076
$a_2 = a_4$	0.36
a_3	0.52

In order to transform this transfer function from the discrete time domain to the frequency domain, the term z^{-1} is substituted by $e^{-j\omega T_s}$ which results in:

$$H(\omega) = \sum_{k=0}^{k=N-1} a_k e^{-jk\omega T_s} \tag{3.28}$$

An important property of this filter is related to the choice of the weighting factors. Suppose the values of the coefficients are chosen symmetrical with respect to the middle coefficient. Each symmetrical pair will add delayed components with an average delay equal to the middle position. The delay of each pair and consequently the total filter equals $NT_s/2$. The same arguments holds if the coefficients are not of equal magnitude but have an opposite sign ("antisymmetrical"). This "linear phase" property results in an equal delay for all (amplified or attenuated) signal components and is relevant for the quality of e.g. audio processing.[5] Mathematically this can be shown by substitution of the Euler's relation[6]:

$$e^{-j\omega T_s} = \cos(\omega T_s) - j \sin(\omega T_s) \tag{3.29}$$

After moving the average delay $NT_s/2$ out of the summation, real and imaginary terms remain:

$$H(\omega) = e^{-j\omega NT_s/2} \sum_{k=0}^{k=N/2-1} (a_k + a_{N-k}) \cos(k\omega T_s/2)$$
$$- j(a_k - a_{N-k}) \sin(k\omega T_s/2) \tag{3.30}$$

Without violating the general idea, N has been assumed here to be even. If the coefficients a_k and a_{N-k} are equal as in the symmetrical filter the sine term disappears. The cosine term is removed by having opposite coefficients in an asymmetrical filter. Both filters have a constant delay. Depending on the symmetry and the odd or even number of coefficients the filters have structural properties, e.g. an asymmetrical filter with an even number of coefficients has a zero DC-transfer.

Figure 3.15 shows the time response and the frequency transfer function from Fig. 3.14 with the coefficients of Table 3.2. The time response is an impulse response: the input for the filter is a pulse with an amplitude of "1" during one sample

---

[5]The human ear is sensitive for delay variations in sound.

[6]The definition for Euler's relation is: $e^{j\pi} + 1 = 0$. According to Feynman this is the most beautiful mathematical formula as it is relates the most important mathematical constants to each other.
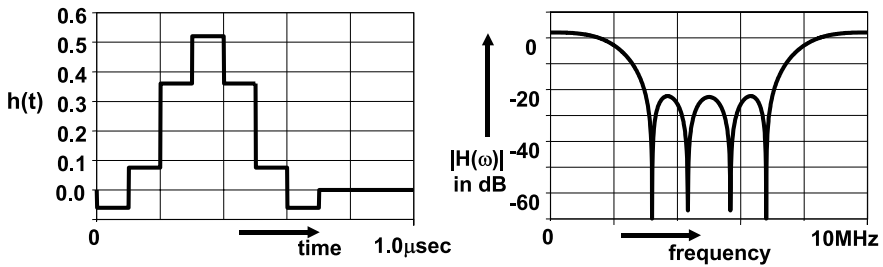
Fig. 3.15 The impulse response and the frequency transfer function of a seven coefficient filter from Fig. 3.14 at a 10 MHs/s sample rate
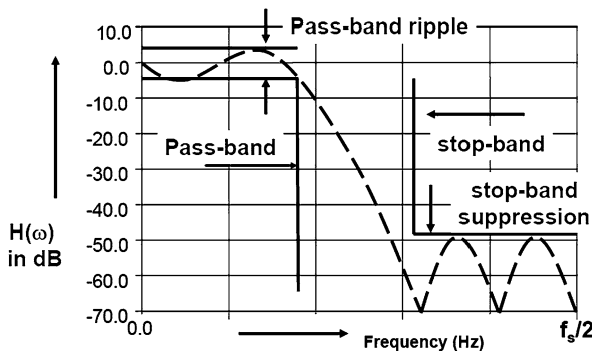

Fig. 3.16 Definition scheme of a filter response

period. The time response in a FIR filter therefore displays all coefficients subsequently. In this example of a time discrete filter the frequency transfer is symmetrical with respect to half of the sampling rate, which was chosen at 10 Ms/s. The spectrum repeats of course at multiples of the sampling rate.

The transfer characteristic can be determined experimentally for a small number of coefficients. A filter that averages over N samples is designed with coefficients of value $1/N$. More complex filters require an optimization routine. A well-known routine was proposed by McClellan, Parks and Rabiner (MPR or the "Remez exchange algorithm") [79]. This routine optimizes the transfer based on a number of specifications.

Figure 3.16 shows a number of terms to define various specification points. Next to that the number of delay elements N, the number of pass and stop bands and the form of the transition between the bands are required. Some variants of filter design programs allow to include the compensation of alias filters or hold-effects.

Figure 3.17 shows the impulse response for a somewhat more elaborate filter with 48 coefficients. The transition between the pass band and the stop band is more steep than in the 7-tap filter. An RLC filter transfer function with a quality factor of 0.5 is drawn in dotted lines as a comparison. The delay time is of course much
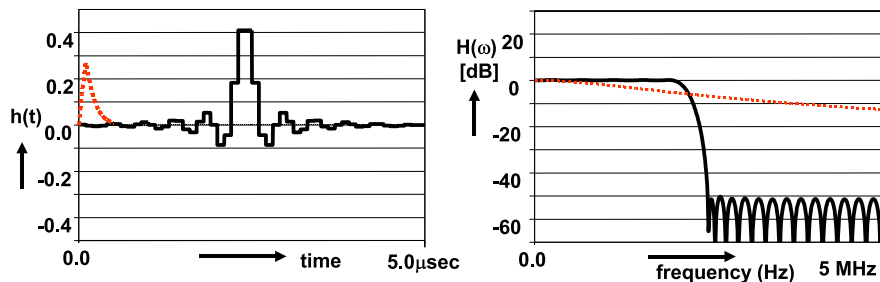
Fig. 3.17 A low-pass FIR filter with 48 coefficients, *left* is the impulse response of the filter and its analog realization (*dashed*). *Right* is the frequency response of both
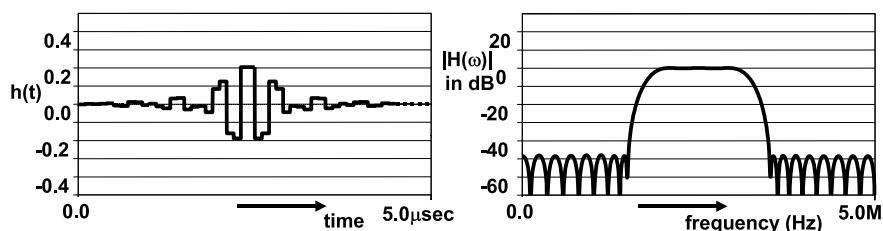

Fig. 3.18 A band-pass FIR filter with 48 coefficients, *left* is the impulse response of the filter and its analog realization. *Right* is the frequency response of both

shorter than the 24 cycles of the FIR filter. However the suppression of the digital filter is superior.[7]

Redesigning this filter with the same 10 Ms/s sample rate and 48 coefficients creates a band pass filter, Fig. 3.18.

The digital time response curve highly resembles the ringing of a high-Q analog filter of the same specification. The accuracy in which required filter characteristics can be defined with FIR filters is clearly illustrated here. However, in practical realizations the price for an accurately defined filter is the large hardware cost of the delay elements and the coefficients and the associated power consumption.

The FIR filter has been described in this section as a mathematical construction and no relation was made with the physical reality. Most FIR filters are implemented in the digital domain: from IC building blocks and FPGA to software modules. Some examples of fully analog FIR realizations are found in Charge-Coupled Devices [80].[8] In digital-to-analog conversion the semi-digital filter uses digital delays with analog coefficients, see Sect. 7.2.4.

---

[7] An equivalent analog filter would require 10–12 poles.

[8] In the period 1970–1980 the charge-coupled device was seen as a promising candidate for storage, image sensing and signal processing. Analog charge packets are in this multi-gate transistor shifted, split and joint along the surface of the semiconductor. Elegant, but not robust enough to survive the digital era.

Fig. 3.19 Four half-band filters, with 3, 5, 7, and 9 non-zero coefficients (*courtesy*: E.E. Janssen)
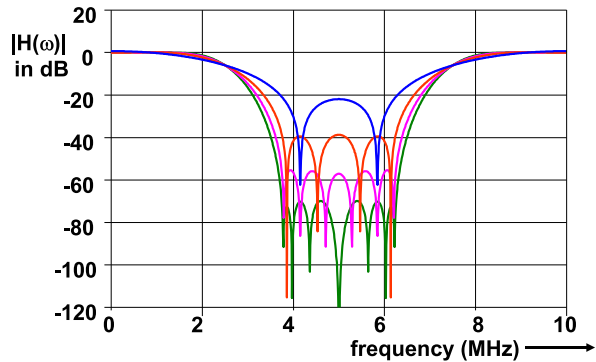
3.2.2 Half-band Filters

In order to reduce the complexity of digital FIR filters additional constraints are needed. Introducing the symmetry requirement:

$$H(\omega) + H(\omega_s/2 - \omega) = 1 \qquad (3.31)$$

leads to such a complexity reduction. At a frequency $\omega = \omega_s/4$ this constraint results in $H(\omega_s/4) = 0.5$, while the simplest fulfillment of the symmetry requirement around $\omega_s/4$ forces a pass band on one side and a stop band on the other side of this quarter sample rate. Consequently these filters are known as "half-band" filters. Substitution of the transfer function for symmetrical filters with an odd number of N coefficients $k = 0, 1, \ldots, m, \ldots, N - 1$ and with the index of the middle coefficient equal to $m = (N - 1)/2$, leads to:

$$a_m = 0.5$$
$$a_{m+i} = a_{m-i} = C_i, \quad i = 1, 3, 5, \ldots$$
$$a_{m+i} = a_{m-i} = 0, \quad i = 2, 4, 6, \ldots$$

Half of the filter coefficients are zero and need no hardware to implement. Optimizing the filter transfer for a minimum deviation of an ideal filter results in a $\sin(x)/x$ approximation, Fig. 2.1, for its coefficient values:

$$a_{m+i} = \frac{\sin(i\pi/2)}{i\pi}, \quad i = -m, \ldots, -2, -1, 0, 1, 2, \ldots, m \qquad (3.32)$$

Table 3.3 lists the coefficients for four half band filters designed for a pass band from 0 to $f_s/8$ and a stop band from $3f_s/8$ to $f_s/2$. Figure 3.19 compares these four half-band filter realizations. The filter with the least suppression has three non-zero coefficients increasing to nine for 72 dB suppression.

In order to obtain a small-area implementation the coefficients are rounded integers. With integer filter coefficients no full multiplier circuit is needed but dedicated shift and add circuits create the weighting of the signal samples.

Table 3.3 The non-zero coefficients for four half band filters in Fig. 3.19 (*courtesy*: E.E. Janssen)

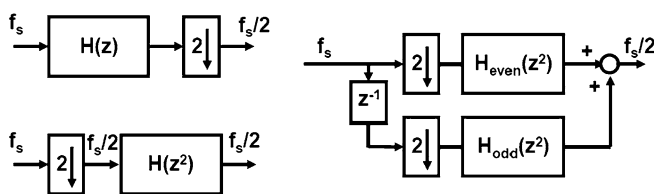| Coefficients	Suppression	Ripple
$a_m = 0.5$	20 dB	0.8 dB
$a_{m-1} = a_{m+1} = 0.2900$		
$a_m = 0.5$	38 dB	0.1 dB
$a_{m-1} = a_{m+1} = 0.2948$		
$a_{m-3} = a_{m+3} = -0.0506$		
$a_m = 0.5$	55 dB	0.014 dB
$a_{m-1} = a_{m+1} = 0.3016$		
$a_{m-3} = a_{m+3} = -0.0639$		
$a_{m-5} = a_{m+5} = 0.0130$		
$a_m = 0.5$	72 dB	0.002 dB
$a_{m-1} = a_{m+1} = 0.3054$		
$a_{m-3} = a_{m+3} = -0.0723$		
$a_{m-5} = a_{m+5} = 0.0206$		
$a_{m-7} = a_{m+7} = -0.0037$		

Fig. 3.20 Down-sampling can be realized by selecting every second sample (*upper-left*). Sample selection and decimation can be interchanged (*lower-left*). More advanced methods split the filter in an even-sample and odd-sample portion

3.2.3 Down Sample Filter

One of the applications of half-band filters is in sigma-delta modulation where the sample rate of the output bit stream must be reduced. A half-band filter cleans the frequency range between $f_s/4$ and $3f_s/4$ of unwanted signal components such as quantization noise. After that the sampling rate can be reduced from f_s to $f_s/2$. This process is called "down-sampling". A cascade of various down-sample sections achieves a large overall down-sample factor.

High-speed down-sampling can be achieved via half-band filtering at full-speed and selecting every second sample, see Fig. 3.20 (upper-left). This method is not efficient, half of the calculated output is unused. In filter theory the "Noble-identities" allow to reverse the order of filtering and down-sampling and create an energy efficient solution to the problem of implementing energy-efficient down-sampling,

Fig. 3.21 The structure of an infinite impulse response (IIR) filter containing a feedback path from output to the summation nodes
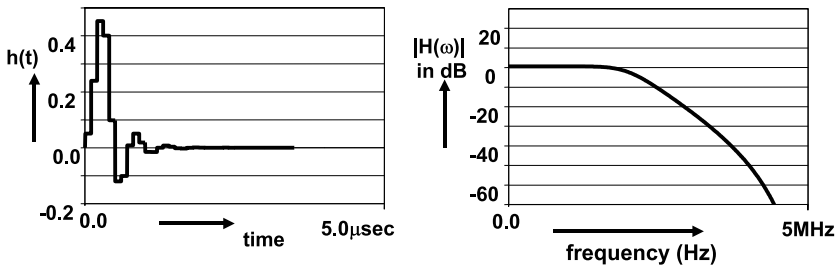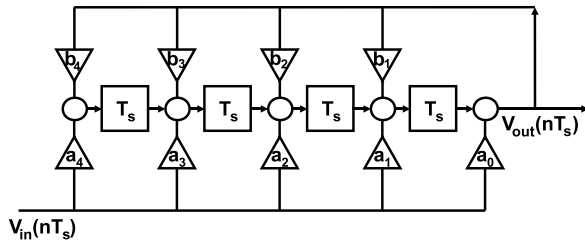

Fig. 3.22 The impulse response of this 4-tap IIR continues beyond four sample periods. The frequency response of this 4-tap filter is much steeper than the response of the 7-tap FIR filter

Fig. 3.20 (lower-left) [72]. An efficient alternative is to split the filter in two sections for the odd and even samples, Fig. 3.20 (right).

3.2.4 IIR Filters

A drastic solution to the hardware problem of FIR filters is the "infinite impulse response" IIR filter.

Figure 3.21 shows the general or canonical form of a digital IIR filter. Coefficients a_0 to a_4 perform the same function as in a FIR filter. In addition the coefficients b_1 to b_4 feed the output back into the delay chain. This feedback modifies the FIR transfer. A similarity to the RLC filter (see Sect. 2.6.8) is that in both filter types the signal is circulating in the filter. In the RLC filter the signal swings between the electrical energy in the capacitor and the magnetic energy in the coil. In an IIR filter the signal circulates via the feedback path. The signal frequency in relation to the delay of the loop and the coefficients will determine whether the signal is amplified or attenuated and for how long. The transfer function of an IIR filter is (for the mapping from z-domain to frequency domain the approximation $z = e^{j\omega T_s}$ is applied):

$$H(z) = \frac{\sum_{k=0}^{k=N-1} a_k z^{-k}}{1 - \sum_{k=1}^{k=N-1} b_k z^{-k}} \quad \Leftrightarrow \quad H(\omega) = \frac{\sum_{k=0}^{k=N-1} a_k e^{-jk\omega T_s}}{1 - \sum_{k=1}^{k=N-1} b_k e^{-jk\omega T_s}} \quad (3.33)$$

The denominator is the feedback path and is formulated as a polynomial in z. For absolute stability (a bounded input results in a bounded output signal) the zeros

Table 3.4 Comparison of discrete filter realization techniques

Implementation	Switched capacitor	Semi-digital filter Sect. 7.2.4	Digital hardware
Delay	Analog	Digital	Digital
Coefficients	Analog	Analog	Digital
Most used	As IIR/resonator	As FIR	Both FIR and IIR
Noise	Accumulates in signal range	Only from coefficients	Related to word width
Tolerance	Capacitor matching	Current source matching	Unlimited
Alias-filter	Required	Depends on system requirements	Required
Power	Moderate	Output related	High
Performance	Limited by noise	Limited by noise	Limited by word width

of the denominator polynomial must be smaller than 1. A signal that experiences a feedback factor $1 - \Delta$ will pass in the order of $1/\Delta$ times through the filter. In theory the signal will never fully extinguish, this is a practical approximation. Such a filter is called a resonator and resembles a high-Q RLC filter. An IIR construction where the denominator has a zero term equal to "1" will oscillate. If all coefficients b_k equal zero, again a FIR filter will result.

A sharp low-pass filter with just four delay elements as in Fig. 3.22 realizes between 2 and 4 MHz a suppression of 40 dB, which is comparable with a seventh order analog filter.

Time-discrete filters can be realized in various implementation technologies. Table 3.4 compares three realization forms of time-discrete filters. The switched capacitor filters are mostly used in medium specification circuits. The realization is practically limited to cascaded resonator type filters aiming at 40–50 dB signal-to-noise levels.

Chapter 4
Sample and Hold

4.1 Track-and-Hold and Sample-and-Hold Circuits

In Chap. 3 the theory of sampling is described. A designer of a complex system will try to concentrate the limitations of the sampling process and the optimization of the function into one circuit block of the system. Often this function is realized as a "Track-and-Hold" (T&H) circuit, which creates a stable signal for a part of the sample period of the ADC. The most elementary T&H circuit consists of a switch and a capacitor, Fig. 4.1. During the conducting phase of the switch, the signal on the capacitor follows the input signal, while in the isolating phase (the hold phase) the signal value remains fixed at its value at the moment of opening the switch. This moment is the theoretical sampling point in time. Two T&H circuits connected in cascade form a sample-and-hold circuit (S&H). The second T&H circuit is triggered by an inverted sampling signal. Figure 4.2 shows the input signal and the output of a T&H and a S&H circuit during track and hold operation. A S&H circuit will hold the signal over the full period of the sampling clock. This allows more processing to be done on the output signal.

Track-and-hold and sample-and-hold circuits are used for performing the sampling operation on the analog input signal at a sample moment in an analog-to-digital converter. The T&H circuit can keep the signal at that level for a time period thus allowing repeated use of the signal during the analog-to-digital conversion, see Fig. 4.3 (left). The output value of the track or sample-and-hold remains here associated with the original sampling moment. The fact that the output value is internally used at slightly delayed time moments will not create any effect in signal transfer characteristics, because in the end the resulting value will be assigned to the original sampling moment. As the output value of the T&H or S&H circuit is used only at specific time moments, some compromises may be acceptable in obtaining a high quality output signal. E.g. slewing during the initial phase of the settling will not affect the overall performance as long as a stable value is reached at the required time moment.

Another application of specifically the sample-and-hold circuit is restoration of the Dirac sequence into a time-continuous signal in digital-to-analog conversion,

Fig. 4.1 A switch and a storage element form a track-and-hold circuit
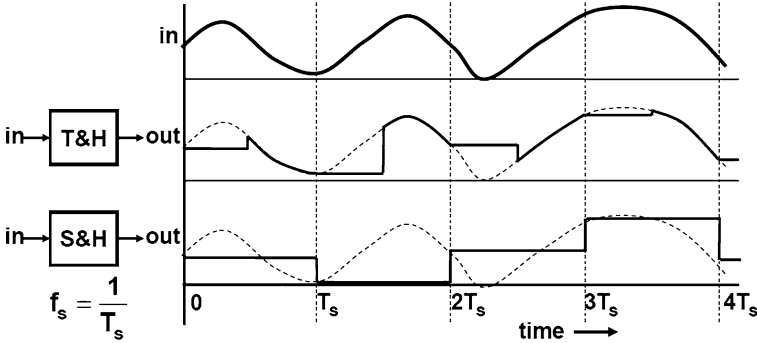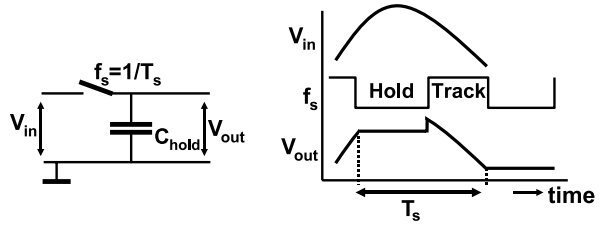

Fig. 4.2 The input signal (*above*) is tracked by a track-and-hold circuit during one phase of the clock period and held during the other phase. A sample-and-hold circuit holds the signal during the full period of the clock signal

Fig. 4.3 A sample and hold circuit is used as an input sampler for an analog-to-digital converter or as a de-glitch circuit in an digital-to-analog converter
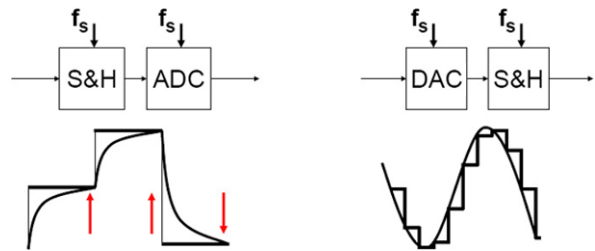
see Fig. 4.3 (right). The sequence of samples (after analog-to-digital conversion and any form of digital signal processing) that arrives at the input of a digital-to-analog converter is still a set of numerical values corresponding to the frame of sample moments. A spectral analysis would result in an ideal sampled data spectrum, where all multiples of the signal band are equivalent. In the time domain the value of the signal between the sample moments is (mathematically spoken) not defined. The most common implementation to deal with this problem is to simply use the value of the signal at the sample moment and to keep it for the entire sample period. Figure 4.4 (left) shows this "zero-order hold" mode. A more sophisticated mechanism interpolates between two values as in Fig. 4.4 (middle). An elegant form of interpolation uses higher-order or spline-fit algorithms, Fig. 4.4 (right).
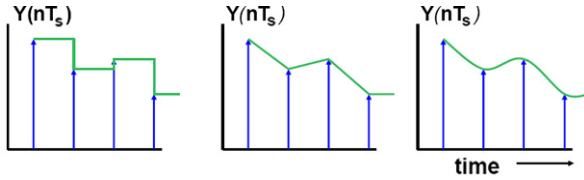
Fig. 4.4 A Dirac sequence can be reconstructed via a zero-order hold, a first-order interpolation or higher order reconstruction algorithms

In most digital-to-analog converters a zero-order hold function is sufficient because the succeeding analog filters perform the interpolation. Moreover a zero-order hold operation is often for free as the digital input signal is stored during the sample period in a set of data latches. The conversion mechanism (ladders or current sources) simply converts at any moment whatever value the data latches hold. In case of algorithmic digital-to-analog converters the output signal has to be constructed during the sampling period (see e.g. Chap. 7.5.4). Then a sample-and-hold circuit is needed to avoid that incomplete conversion results appear at the output. Whenever the output signal of a digital-to-analog converter contains glitches, an explicit sample-and-hold circuit will remove the glitches and improve the quality of the conversion. In this application the output signal of the S&H circuit can be used by the subsequent signal processing at any moment in time. Consequently much higher demands are put on the S&H circuit.

Holding the signal during a period T_h, however, changes the shape of the signals passing through a zero-order hold operation. Holding of the signal creates a transfer function. The impulse response of the hold transfer function is found by considering that the Dirac sequence is multiplied by a function consisting of a constant term "1" over the hold period T_h.

$$h(t) = \begin{cases} 1, & 0 < t < T_h \\ 0, & \text{elsewhere} \end{cases} \tag{4.1}$$

The frequency-domain transfer function $H(\omega)$ of a zero-order hold function or a sample-and-hold circuit is calculated via the Fourier transform. The result of this transform has the dimension time.[1] In order to obtain a dimensionless transfer function, a normalization to T_s is introduced:

$$H(\omega) = \int_{t=0}^{t=\infty} h(t) \times e^{-j\omega t}\, dt = \frac{1}{T_s} \int_{t=0}^{t=T_h} 1 \times e^{-j\omega t}\, dt$$

$$= \frac{\sin(\pi f T_h)}{\pi f T_s} e^{-j\omega T_h/2} \tag{4.2}$$

Figure 4.5 shows the time and frequency response of the S&H circuit. A delay T_h is introduced as the value of the signal that was first concentrated in the sample

---

[1] Formally the result of a Fourier transform reflects the intensity of a process or signal at a frequency. Therefore the result has the dimension "events per Hz" or "Volt per Hertz".
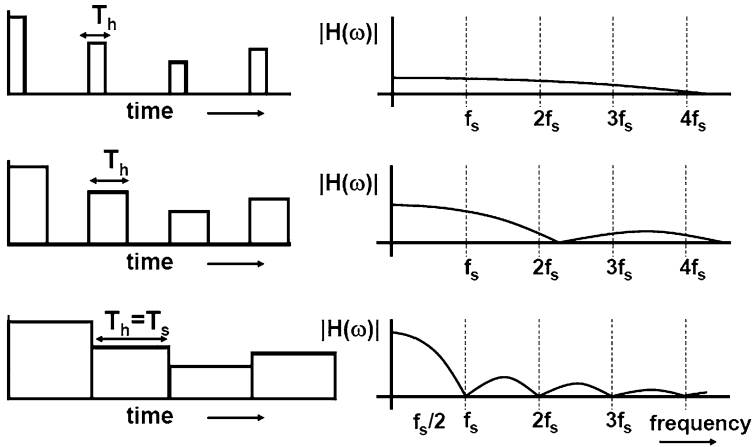
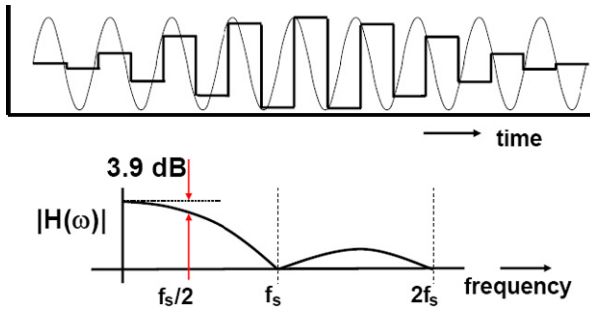**Fig. 4.5** The hold time determines the filter characteristics of a sample and hold

Fig. 4.6 Time and frequency response of a sample-and-hold signal close to half of the sampling rate

moment, is now distributed over the entire hold period. The average value moves from the sampling moment (the edge of the clock pulse) to the middle of the hold period. A zero response occurs at frequencies equal to multiples of the inverse of the hold time. Obviously signals at those frequencies exactly average out over the hold time period. For short hold periods approximating a Dirac function, this zero is at high frequencies and the transfer of the sample-and-hold circuit is flat over a large frequency range. If T_h becomes equal to the sample period T_s the transfer function shows a zero at the sample rate and its multiples. The mathematical formulation of the transfer curve is often summarized to "$\sin(x)/x$" behavior. Some authors use "$\mathrm{sinc}(x)$".

The amplitude response in the Fourier/frequency domain is a representation of the average energy over infinite time. In the time domain T&H output signals can occur with amplitudes close to the analog input amplitude. However, a signal close to half of the sampling rate can show in one time period a small amplitude while achieving a value close to the full range input signal at another time instance de-
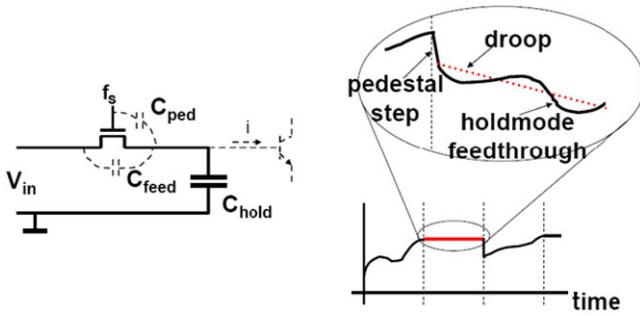
Fig. 4.7   Three artifacts in a sample and hold: droop, hold-mode feed through and pedestal step

pending on the phase relation of the signal and the sample rate. Still this signal has over infinite time an averaged attenuation of 3.9 dB. In that sense the attenuation in Fig. 4.6 is different from a frequency transfer function of e.g. an R-C network, where the attenuation at a certain frequency is independent of the phase.

4.2 Artifacts

The sample-and-hold circuit has to fulfill the requirements of maximum speed in the analog domain as well as in the sampled data domain. In order to achieve maximum signal-to-noise ratio the S&H circuit has to handle the maximum amplitude. These two requirements make that many designers see the S&H circuit as one of the most critical components in realizing an optimum conversion performance. The specification of the performance of a S&H circuit consists of the standard analog ingredients: distortion (THD), signal-to-noise ratio (SNR), power consumption (P), etc. Next to these standard requirements, a number of specific specifications exists for an S&H circuit, see Fig. 4.7:

- Cross-talk from the switching pulse may cause an instantaneous drop in the voltage of the hold capacitor, called the "pedestal step". Another contribution to the pedestal step comes from the inherent removal of the conductive layer of the MOS switch or the base charge in a bipolar transistor. This charge will flow back into the source and into the hold capacitor.
- During the hold phase of the sample-and-hold, charge can leak from the hold capacitor, the signal will show "droop". In a bipolar design this leakage is caused by the base current. This effect results in a minimum sample rate of e.g. a few hundred ksamples/s. In deep-submicron processes gates may become so leaky that again droop will become a relevant parameter.
- The hold-mode feed-through describes the transfer from source to the output of the circuit in hold-mode. Next to trivial reasons, such as unintended coupling via wiring or power supply, some residual coupling may occur due to the source-drain capacitor in a MOS switch. Although normally hold-mode feed-through

is minimum in integrated solutions, a T-switch may further reduce this effect, Fig. 4.12.

- The aperture time is the relatively small time period in which the sample pulse makes the transition to switch off. During this clock edge still a some tracking of the signal may continue. A constant aperture time is therefore mostly just a small additional fixed delay of the sampling with respect to the start of the switch-off process. More devastating are changes in aperture time. Aperture jitter is equivalent to the jitter discussed in Sect. 3.1.4. Deviations of the aperture time can also arise from changes related to the signal amplitude causing distortion.

4.3 Capacitor and Switch Implementations

4.3.1 Capacitor

The sample switch and the hold capacitor are the prime components of a track-and-hold or sample-and-hold circuit. The value of the hold capacitor is for signal-to-noise ratios in excess of 40 dB determined by kT/C noise. The requirements on the switch transistor depend on the way the capacitor is used: as a voltage buffer or as a charge store. A simple capacitor with the plate with most parasitic components (the bottom plate) connected to ground, can be read out as a voltage buffer and can tolerate some voltage dependence of the capacitance. A capacitor that is used as a voltage-to-charge converter, as in most switched capacitor implementations, must be linear.

In the implementation of the capacitor in CMOS technology a few options are present, see also Sect. 2.3.6:

- Diffusion capacitances are not suited for most specifications. Leakage, nonlinearities, low capacitance per unit area generate too many problems.
- The gate-to-channel capacitor, see Sect. 2.4.4, gives the highest capacitance per unit area. However, this device requires a significant turn-on voltage in excess of the threshold voltage. Using voltages below the threshold voltage is not possible as the inversion layer will disappear.
- Interconnects allow top design plate capacitors and fringe capacitors. Stacking various layers of interconnect, where odd and even numbered layers form the plates of a capacitor, is generally a good solution. This capacitor requires no bias, has a good linearity and low parasitics, see Fig. 7.23.
- In some process variants a so called Metal-Insulator-Metal capacitance option is present with excellent properties.
- In older processes a double poly-silicon capacitor option is offered. The depletion of the polysilicon can lead to voltage non-linearities.

Designers will typically choose for a fringe or plate capacitance in an advance process or a double-poly in an older process. Most of the capacitor constructions show some form of asymmetry because one layer is closest to the substrate. Depending

on the circuit topology the capacitor terminal least susceptible to e.g. substrate noise must be implemented in the layer that is close to the substrate. Next to that also horizontal coupling must be avoided. Clocked interconnect lines should be kept away of the capacitor or even shielded by placing grounded lines in between.

4.3.2 Switch Topologies

The track-and-hold switch is characterized by its on/off impedances. In a T&H circuit the on-resistance must be small and constant and the off-impedance must be infinite. In the on-state the conductivity of a single MOS transistor as a switch depends on the gate-to-channel voltage minus the threshold voltage, see Sect. 2.5:

$$R_{\mathrm{on,NMOS}} = \frac{1}{(W/L)_N \beta_{N\square}(V_{DD} - V_{\mathrm{in}} - V_{T,N})}$$
$$R_{\mathrm{on,PMOS}} = \frac{1}{(W/L)_P \beta_{P\square}(V_{\mathrm{in}} - |V_{T,P}|)} \tag{4.3}$$

The NMOS switch is non-conductive if the input voltage is less than a threshold voltage below the gate-voltage. The maximum gate voltage often equals the power supply voltage. The PMOS switch is non-conductive at an input voltage a threshold voltage above ground. At low supply voltages and large signal excursions this voltage-dependent resistance of the switch can lead to aperture time differences with distortion. Figure 4.8 shows a track-and-hold circuit simulation in a 1.2 V 90-nm CMOS process. The sinusoidal input signal is sampled with 100 Ms/s on a 10 pF capacitor. The simulation with a wide switching transistor (50/0.1 μm) shows nearly no artifacts. However, the tenfold narrower transistor shows clearly that the delay
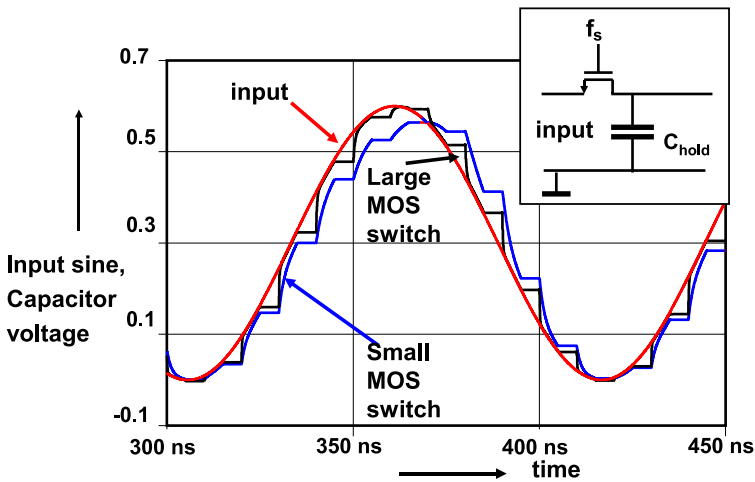

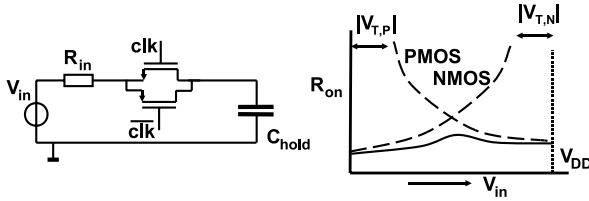**Fig. 4.8** The resistance of a small NMOS switch causes significant distortion in the hold signal

Fig. 4.9 The complementary PMOS-NMOS switch

between input and output increases at higher signal voltages. At those higher signal voltages there is less channel charge in the NMOS transistor and the RC-time increases. The net effect of this signal-dependent aperture delay time is distortion.

With a simple approximation the magnitude of the distortion is estimated. Assume that the total variation of the resistor over the signal range is ΔR:

$$R(V_{in}(t)) = R_0 + \frac{V_{in}(t)}{V_{in,peak-peak}} \Delta R$$

With an input signal $V_{in}(t) = 0.5 V_{in,peak-peak} \sin(\omega t)$ the current is mainly determined by the capacitor: $I(t) \approx \omega C 0.5 V_{in,peak-peak} \cos(\omega t)$. The voltage drop over the resistor is composed of linear and second order terms. This last term is equal to the second order term that will appear over the capacitor. The HD2 is then estimated as:

$$HD2 = \frac{\omega \Delta R C}{4} \tag{4.4}$$

Some fine tuning in a simulator is needed as the switch resistance also shows higher order artifacts.

The signal-dependent impedance is the main problem with a simple switch controlled by fixed voltages. Also in more complicated switched capacitor structures there is always one critical switch in this respect.

Figure 4.9 shows a popular implementation of a low-resistance switch. An NMOS transistor and a parallel connected PMOS transistor compensate each others poor conductivity regions. The overall resistivity of the switch over the input voltage range is considerably more flat. The use of two transistors implies that the controlling clock pulses come exactly at the right position in time. If the PMOS transistor switches slightly before the NMOS transistor again a form of aperture time difference will occur. The switch resistance is more constant than of a single transistor switch, but some modulation with the signal amplitude is present.

At extreme low voltages where $V_{T,N} + |V_{T,P}| < V_{DD}$, an input region exists where none of the two transistors will conduct. Crols and Steyaert [81, 82] presented the switched operational amplifier technique as a solution for this case. The functionality of the switch is implemented in the driver transistors of the opamp output stage. In case the opamp output stage switches to a high impedance, an additional switch is needed to connect the capacitor to a stable voltage. The switching operation in the opamp itself causes additional disturbances. An equivalent or better

Fig. 4.10 The channel charge in a MOS device acts as a transmission line when the transistor is switched off. If the impedances on either side differ, the charge splitting is asymmetrical
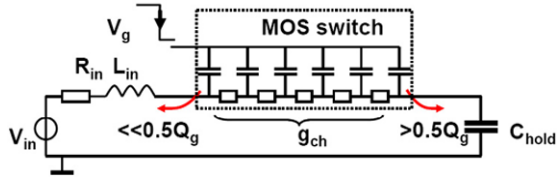
performance to the standard techniques is only reached in cases with extreme low-power supply. In many designs a better performance is reached with a bootstrapping technique.

The conduction of the switch requires a charge layer that connects the source and drain diffusions. The channel charge is the product of the gate capacitance and the net gate-source voltage. In a simple T&H structure the effective gate-source voltage consists of: $V_{DD} - V_T - V_{in}$, a fixed gate drive voltage minus the input voltage. In this case the channel charge is signal dependent. If the transistor switches off the charge in the hold capacitor increases by a amount of signal charge and a constant charge term:

$$V_{hold} = V_{in} + \frac{Q_g}{2C_{hold}} = V_{in} + \frac{WLC_{ox}(V_{in} - V_{DC})}{2C_{hold}}$$

$$= V_{in}\left(1 + \frac{WLC_{ox}}{2C_{hold}}\right) - \frac{WLC_{ox}V_{DC}}{2C_{hold}} \tag{4.5}$$

In fact there is a slight amplification at the moment of sampling. This effect can jeopardize pipeline and algorithmic based converters, see Sects. 8.4 and 8.6.2. In advanced technologies the channel contain much less charge and therefore charge splitting is less an issue.

When the switch turns to its isolating stage, the switch charge must be removed and creates the pedestal step. In first order the residual charge will split evenly between signal source and capacitor, see Fig. 4.10. The typical rise/fall time of a switching pulse is 50 ps in 65-nm CMOS. At these speeds the channel cannot be considered a single element, but must be analyzed as a transmission line along which the charge moves to the terminals of the switch. Depending on the impedances on both sides of the switch and the switching speed of the sample pulse, significant deviations can occur in the splitting accuracy. This in turn leads to undetermined signal components in the voltage over the hold capacitor. In some cases the impedance can be made more equal by adding a capacitor on the input side of the T&H circuit.

Compensation of the pedestal by means of half-sized transistors is possible. In a practical design Fig. 4.11 the dummy switches are single devices while the pass transistor consists of two parallel transistors for optimum cancellation. The control of the sample pulse edges requires careful balancing. If the "CLK" pulse switches off while the inverse pulse is not yet active, there will be a lot of switching glitches. If the inverse pulses become active while the "CLK" pulse is still active, the compensation charge will be supplied by the source and the whole method is ineffective. The occurrence of these two situations depends on the level of the input signal.

Fig. 4.11 The inherent
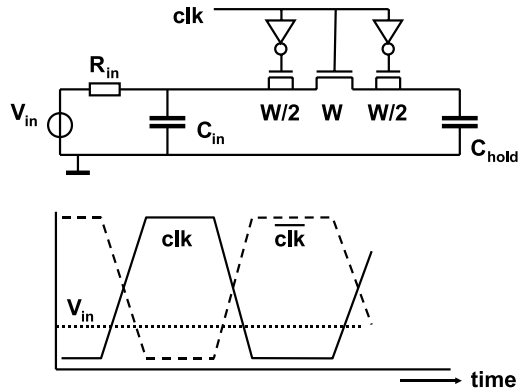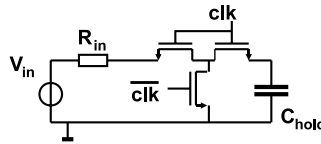charge splitting can be
compensated by dummy
switches

Fig. 4.12 The T-switch
reduces the signal
feed-through

The dummy-switches remove the constant part of the pedestal step. Mismatch between the transistors creates a random charge component. This random part is increased as the random variation in the channel charge of the switch-transistor adds to the random variation of the compensating charge.

In case some capacitive coupling remains between the signal and the hold capacitor a T-switch can be applied, see Fig. 4.12. A T-switch is a series connection of two MOS devices both clocked with the sample pulse. A third device connects the common source-drain terminal of the series switches to ground. This device is clocked in anti-phase.

4.3.3 Bottom Plate Sampling

In the previous circuits a signal dependent channel charge influences the sampling process. "Bottom-plate" sampling introduces a second transistor switch on the grounded side of the capacitor, see Fig. 4.13. This switch uses the full power supply range for a maximum gate drive and is not impaired by input signal variations. The bottom-plate switch can be turned off before the original sampling switch, thereby isolating the capacitor. A capacitor isolated on one side cannot exchange charge on the other side. As long as the top-side switch turns off before the charging of the parasitic capacitors connected to the bottom plate becomes a problem, this bottom-plate sampling will partly reduce the modulation problems with the sample switch.
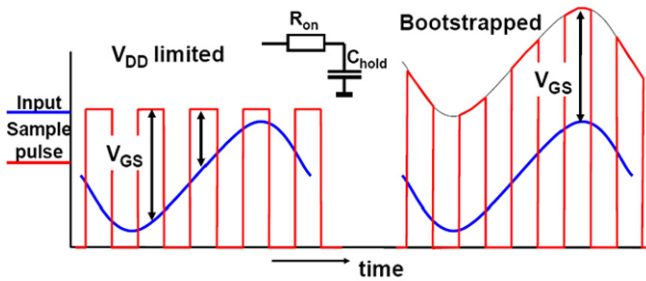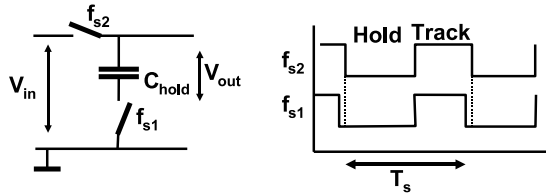
Fig. 4.13 Bottom plate sampling [83]

Fig. 4.14 The drive voltage of the switch varies if the power supply voltage is the limiting factor (*left*). If the switched is bootstrapped, ideally the drive voltage is constant over the input range
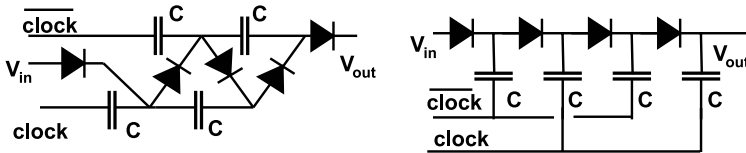

Fig. 4.15 Two techniques. *Left*: the series connected Cockcroft-Walton voltage multiplier (1932). *Right*: the parallel connected Dickson multiplier [84]

4.3.4 The CMOS Bootstrap Technique

A popular technique to circumvent most of the above problems is the class of bootstrapping circuits, Fig. 4.14. These circuits aim to solve two problems associated with the one-transistor switch: the limited input range due to the threshold voltage and the switch resistance variation. In these schemes the effective drive voltage is increased beyond the power supply limits.

Various techniques exist to multiply voltages. These techniques are either based on inductive (buck and boost converters) or capacitive multiplication schemes. In the context of sampling the required energies are limited and only capacitive schemes are selected that use the available components in a CMOS technology. Figure 4.15 shows two basic voltage multiplication techniques. On the left side the capacitors are connected in series and driven by two complementary pulses. The diodes charge the intermediate nodes to the swing of the driving pulses. If all parasitic elements are under control this scheme can generate output voltages in excess of the individual capacitor break-down voltage. A limitation of this scheme is that the effectiveness
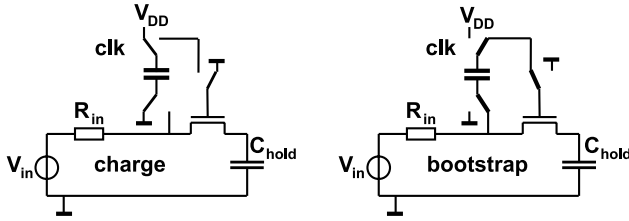
Fig. 4.16 The bootstrapping principle

depends on the ratio between the series connection of capacitors and the parasitic capacitors.

Dickson [84] proposed a parallel technique, which allows a more effective voltage multiplication in the presence of parasitic capacitive load on the nodes. Again the circuit is driven by complementary pulses and every node is charged to a voltage of roughly:

$$V_n = n \left(\frac{C}{C + C_{\mathrm{par}}} V_{\mathrm{clock}} - V_{\mathrm{diode}} \right) \tag{4.6}$$

This formula points to the main efficiency limiting effects in many voltage multiplication circuits: the ratio of the parasitic capacitance to the driving capacitance and the voltage loss over the rectifier element.

The above principle has been implemented in various forms, mostly with only one or two capacitors. Often this technique is referred to as bootstrapping.[2] Knepper [85] used in the era of enhancement/depletion NMOS technology a bootstrap circuit to solve the problem of driving an NMOS transistor with gate voltages higher than the power supply.

Figure 4.16 shows the general idea of bootstrapping the voltage of a sampling switch. A capacitor with a value of typically ten times the gate capacitance to drive is charged to the power supply during the hold phase of the T&H circuit. In the track phase this capacitor is connected between the input voltage and the gate of the transistor. This gate is now biased at a more or less constant voltage overdrive with respect to its source and drain.

Many practical circuits use in the track-and-hold circuit the bootstrapping principle [86]. In Fig. 4.17 transistors T_1 and T_2 implement a clock pulse bootstrap. If the clock is high, the source of T_1 is pulled low, while the gate is pushed up. This will (partially) charge the capacitor connected to T_1. In the next clock cycle the same will happen with the capacitor connected to T_2. In the following clock cycles the extra charge on one capacitor will help to enhance the charging of the other until both are charged to the full power-supply voltage. Parallel to T_1 and T_2, T_3 will fully charge the bootstrap capacitor C_B. After some clock cycles, the maximum gate voltages of T_1, T_2 and T_3 will reach voltage levels of 1.8 to 1.9 × V_{DD}. The requirements for

---

[2]A relation seems likely with the tales on Baron von Mŭnchhausen, who pulled himself up by his bootstraps, and "booting" of computers.
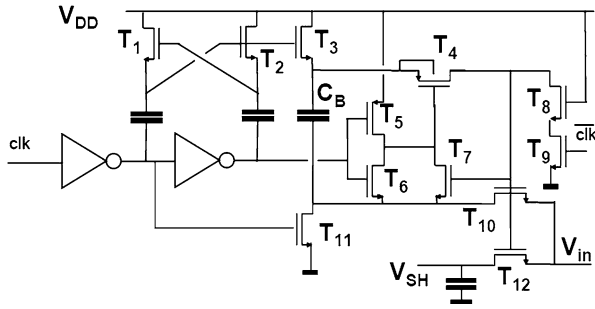
Fig. 4.17 The bootstrapping circuit according to Abbo [86]

process reliability demand that special measures prevent any two nodes (D, G, S) of a transistor to see a voltage drop of more than V_{DD}. T_8 serves as a protection device: in this way the high voltage on C_B is divided over both transistors T_8 and T_9. The drain-source voltage of T_9 can not exceed the voltage limits. Nevertheless reliability people require to prove that all devices are protected from over-voltage under any circumstance.

At the onset of the tracking phase (the clock goes high) T_3 and T_{11} switch off. T_6 connects the gate of T_4 to the low side of C_B. The charge in C_B will pass via T_4 to the central bootstrap node that connects to T_{10} and the T&H switch transistor T_{12}. Via T_{10} the input voltage is applied to the bottom plate of C_B and the bootstrap action can start. When the clock goes low for the hold phase T_8 and T_9 rapidly discharge the central bootstrap node.

4.3.5 Buffer

The voltage on the hold capacitor must be buffered before any succeeding operation can be performed. For low-resolution and high-speed operation open-loop amplifiers such as source followers [87] and degenerated differential pairs [50] are used. The inherent speed advantage must be traded off versus the non-linearity. Still an 8–10 bit performance level is possible.

Higher linearity is obtained with feedback operational (transconductance) amplifiers. These opamps buffer the capacitor voltage and are often mixed with special functions for the specific analog-to-digital architecture. This results in specific demands within each conversion topology. Some general remarks:

- The signal range of the switch, the input range of the buffer and the output range of the buffer are often linked together. Mostly the switch range must equal the output range and sometimes also the input range, in case a feedback topology is used. Input and output ranges of the buffer depend on the opamp configuration. An NMOS switch combined with an NMOS input stage of a buffer will loose signal range on both sides and is not an optimum combination. An optimum choice for the conversion architecture will allow a larger signal and make the circuit less

Fig. 4.18 An off-chip buffer drives the switch which is directly connected to the bondpad
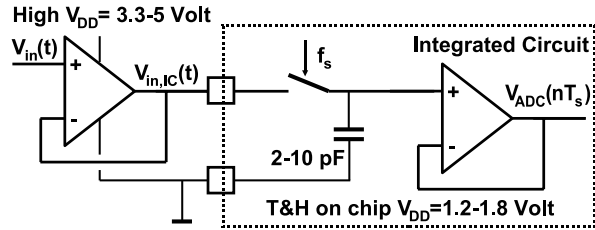
sensitive to parameter and power supply variations. The design of a track-and-hold circuit is started by identifying the range requirements.

- If the buffer is constructed as a unity feedback opamp, the minimum DC amplification of the opamp is mostly given by:

$$A_{DC} > \frac{1}{\epsilon} = 2^{-N}$$

where ϵ is the remaining settling error. The implicit assumption is that ϵ is unknown and is better kept a low level. In fact, in most opamp topologies the larger part of ϵ is linearly dependent on the signal and would result in a minor overall gain error. In most analog-to-digital converters this is quite acceptable, however in e.g. a 1-bit pipeline converter an exact gain is essential for accuracy.

- The speed of settling depends on the first moment in time when an accurate output signal of the T&H is needed. In order to reach a settling level of 2π time constants within one sample period ($e^{2\pi} = 527 \approx 2^9$), the unity gain frequency of the T&H buffer must be equal to the sample frequency. If only a fraction of the sample pulse is available for settling: UGBW $> 3f_s$.

- The buffer will load the T&H capacitor. In a lot of topologies this loading is constant and limited. If a source follower circuit is applied the hold capacitance/impedance becomes a complex function. An interesting analysis is found in [87, Appendix].

4.4 Track-and-Hold Circuit Topologies

4.4.1 Basic Configurations

The buffers that drive a track-and-hold switch and capacitor have to fulfill the maximum specifications in terms of bandwidth, distortion and timing accuracy. With the low supply voltages in modern IC processes these demands are difficult to realize. In some system partitioning solutions this problem is circumvented by using an off-chip driver, see Fig. 4.18. The off-chip driver uses higher supply voltages to meet the specifications, moreover the inherent S&H power consumption does not add to power in the data sheet.

A disadvantage of this solution is that the bondpad normally is connected to a protection circuit. A series resistor and some diffusion capacitor load the input
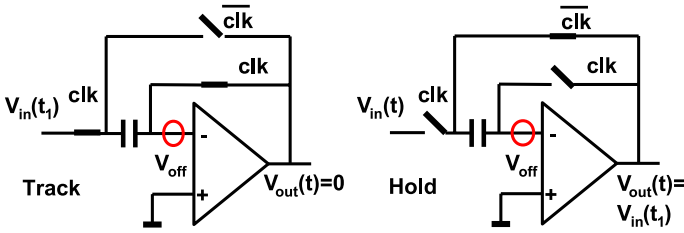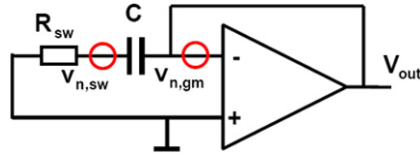
Fig. 4.19 A track-and-hold circuit with offset cancellation in hold mode

Fig. 4.20 The equivalent T&H circuit during track mode [88]

terminal and may limit the achievable bandwidth and distortion. The input signal swing can modulate the sample switch resistance, mostly a bootstrapped switch is needed.

Also the on-chip handling of the signal has a number of disadvantages. Buffering the capacitor voltage requires a sufficiently large input range. The buffer needs to be designed with a sufficiently high common mode rejection ratio and any offset or $1/f$ noise in the buffer adds up.

The switched-capacitor circuit in Fig. 4.19 circumvents the problem of a wide common mode range at the input of the buffer. The feedback topology is switched in a unity gain mode during tracking (clk closed). Any offset of the buffer or low-frequency noise will appear at the negative input terminal and is stored on the capacitor. In hold-mode only the inverse clock switch is conductive. The operational amplifier is now fed back via the capacitor and the output voltage will copy the level of the input signal in order to have the same difference voltage at its input terminals. Consequently the output is not affected by offset or low-frequency noise at the input. However the high frequency noise generated by the opamp is sampled into the signal (on top of the kT/C noise). The transfer of the offset source to the output signal is:

$$V_{out,error} = V_{off}(z)(1 - z^{-0.5})$$

The exponent -0.5 assumes a 50% duty cycle of the switching signal. The transfer is characterized in the frequency domain by a term

$$H(f) = 2\sin(\pi f/2f_s) \tag{4.7}$$

This transfer function suppresses a DC offset voltage effectively. However this scheme amplifies any (undesired) signal near $f_s/2$ by a factor of 2.

The behavior of noise in switched capacitor circuits such as track-and-hold circuits has to be considered carefully. During the initial track phase, time-continuous noise from the switch resistance is present on the capacitor. Next to this noise a noise

Fig. 4.21 A track-and-hold
circuit based on a
transconductance in
differential design [89]
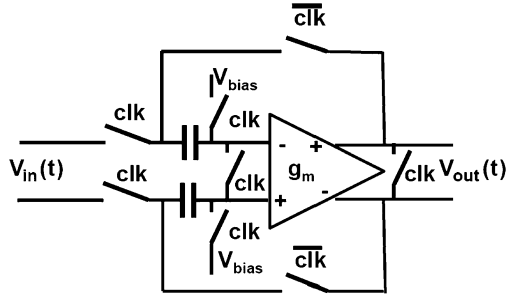
contribution on the negative terminal of the opamp is sampled. The noise generated inside the opamp is dominated by the noise of the transconductance of the input pair, see Sect. 2.5.8. As $1/g_m \gg R_{sw}$ and as the noise of the switch and sample capacitor spreads out over a much wider bandwidth, the opamp contribution is relevant. $v_{gm,n}$ models this contribution in Fig. 4.20. The resulting output noise $v_{out,n}$ experiences the (first-order) transfer function of the opamp. The feedback path returns this noise to the input terminal. Combining the spectral density of the input referred noise for the two transistors in the input pair and the unity-gain transfer gives:

$$v_{\text{out},n}^2 = 2 \times \frac{4kT}{g_m} \int_{f=0}^{f=\infty} \frac{1}{1 + (f/f_{\text{UGBW}})^2} \, df = \frac{4kT}{g_m} \frac{f_{\text{UGBW}}\pi}{2} \qquad (4.8)$$

with the transconductance noise correction factor $\alpha = 1$.

The factor $f_{\text{UGBW}}\pi/2 \approx 1.57 f_{\text{UGBW}}$ accounts for the energy up to f_{UGBW} and the energy in the roll-off part of an ideal first order transfer curve.

The unity gain frequency of the opamp is determined by the input transconductance of the input pair and the output load capacitance in case of a single stage opamp ($f_{\text{UGBW}} = C_{\text{load}}/2\pi g_m$), or by the transconductance and the Miller capacitance in case of a two-stage opamp. The resulting sampled noise is:

$$v_{\text{out},n}^2 = \frac{2kT}{C_{\text{load}} \text{ or } C_{\text{Miller}}}$$

If the bandwidth of the two-stage opamp is comparable to the single stage version the Miller capacitance should be of the same order of magnitude as the load capacitance. The load capacitance of the opamp certainly includes the sample capacitor. In total this means that the noise energy on the sample capacitance is about $2kT/C$.

After the inverse clock takes over, the noise on the capacitor consists of two components: the previously sampled noise and new time-continuous noise from the opamp. This noise will be sampled during the succeeding processing. A full cycle of switching activity in a switched capacitor circuit will result in several separate and independent contributions of noise, whose energies are summed.

Figure 4.21 shows a variant of the standard scheme. The operational amplifier is replaced by a transconductance stage. During sampling the capacitors are directly connected to the input signal. In feed back the capacitors are connected to the output of the transconductance. This variant does not cancel the input offset, but it avoids sampling the high-frequency components of the noise of the transconductance.
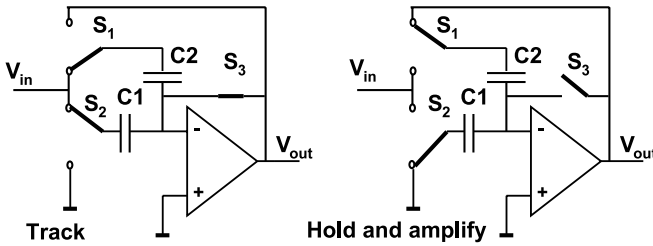
Fig. 4.22 A track-and-hold with multiplying capacitors as is used for pipeline converters

4.4.2 Amplifying Track-and-Hold Circuit

A more complex track-and-hold uses the capacitors to multiply the signal. The circuit is shown in Fig. 4.22. During the track phase (left) the switch S_3 connects the operational amplifier in unity gain feedback. the capacitors C_1 and C_2 are connected in parallel to the input signal. When the transition to the hold-phase occurs (right), switch S_1 creates via C_2 the feed back path for the amplifier. S_2 switches capacitor C_1 to ground. The original charge corresponding to the input signal on C_1 is transferred to C_2. The overall transfer of this circuit is:

$$V_{out} = \frac{V_{in}(C_1 + C_2)}{C_2} \tag{4.9}$$

The initial sampling of the signal gives a noise energy contribution proportional to $C_1 + C_2$. In the amplification phase C_1 is connected to the ground level. Input referred noise of the opamp is therefore also multiplied by 2 and an additional contribution of C_1 is added.

Although the exact specifications for the operational amplifier depend on the T&H architecture used, some general remarks apply. For sufficiently low distortion numbers the DC-gain of the operational amplifier should exceed the distortion number: for a 60 dB distortion the amplifier should be designed at 60–70 dB gain. As most amplifiers are used in (near) unity gain configuration the speed of the amplifier should allow settling to the gain error within the hold period. The settling time constant of the amplifier equals $1/2\pi f_{UGBW}$. If the unity gain frequency would be equal to the sampling frequency, only 2π settling time constants fit in a full sampling period ($e^{-2\pi} = 0.002$). Therefore the unity-gain frequency is mostly chosen at 1.5–2 times the sampling rate. In the topology of Fig. 4.22 the amplifier is in the hold phase in a $(C_1 + C_2)/C_2$ amplification mode. For an equivalent settling behavior the unity gain frequency must be increased with this factor.

4.4.3 Correlated Double Sampling

In most systems the track-and-hold circuit creates time-discrete samples from a time-continuous input. In some cases, the input is already time-discrete or is at least during a part of the sample period time-discrete. An example is the output of an
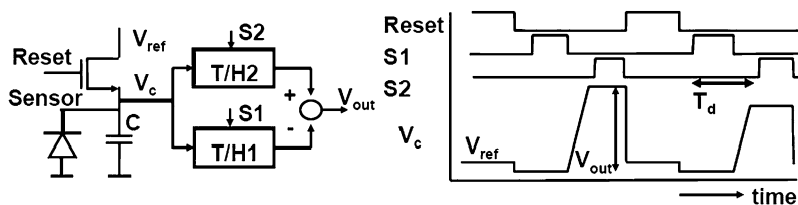
Fig. 4.23 The correlated double sampling technique

image sensor. In the sample period of such an output various phases can be distinguished: a reset phase which sets the sensitive node to a predetermined value, a "zero-signal" phase and a signal phase, Fig. 4.23. The resetting of the sensitive node introduces a number of unwanted components: noise from the reset transistor (kT/C), uncertainty of the charge splitting, $1/f$ noise components, etc. The correlated double sampling technique eliminates a number of these effects by first sampling the "zero-signal" level, including the before mentioned components, then sampling the signal period, and subtracting both. Figure 4.23 shows the basic functionality. A single capacitor can be used for sampling both the "zero-signal" and the wanted signal. The transfer function for the unwanted components is:

$$\begin{aligned} |H(s)| &= |1 - e^{-sT_d}| = |e^{-sT_d/2} 2\sin(sT_d/2)| \\ |H(\omega)| &= |2\sin(\omega T_d/2)| = |2\sin(\pi f T_d)| \end{aligned}$$

(4.10)

with T_d as the delay between the two sample moments. The main disadvantage also becomes clear from this transfer function: the noise near odd multiples of half of the sampling frequency is amplified.

4.4.4 A Bipolar Example

Bipolar design is still frequently used for analog realizations. In contrast to complementary MOS technology (CMOS), commercial bipolar processes are not always equipped with complementary devices of more or less similar performance level. Vorenkamp and Verdaasdonk [90] published in 1992 a often copied high-performance track-and-hold circuit in npn-only bipolar technology. The circuit of which a single-ended version is shown in Fig. 4.24 uses a pre-stage, the actual switch and capacitor and an output buffer. T_1 and T_2 form an one-time input amplifier, buffering the signal towards the base of T_6. In the "Track" mode "T" is high and "H" is low. T_3 is switched off and the current of that branch passes via T_4 and biases T_6 as an emitter follower. The signal is applied to the hold capacitor C_H and T_7, T_8 and T_9 form an output buffer. In order to check T_5 the base-emitter voltages are counted: using the input or the base of T_1 as a starting point, the emitter of T_5 is at the same DC level. The voltage on the hold-capacitor is at minus one V_{be} and again the output at the collector of T_9 reaches the same DC-level as the input node.
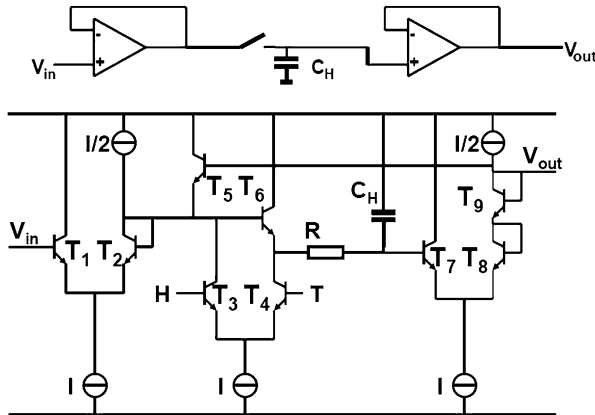
Fig. 4.24 A track-and-hold circuit in bipolar technology [90]

This voltage is applied to the base of T_5 which consequently has a zero base-emitter voltage and is not conducting.

In the "Hold" mode the terminals "H" and "T" have opposite voltages. The current passes via T_3 to T_5. Counting again the V_{be} voltages, starting from the hold capacitor node, it becomes clear that now T_6 has a zero base emitter voltage and is not conducting. The sample-and-hold switch is non-conductive. The positive rail is used as a signal reference. The power supply rejection for the negative power supply variations is high because only current sources connect to the negative power supply rail. So all internal voltages have a constant voltage difference with the positive power supply rail. Consequently the hold capacitor is referenced to that rail as well. If this circuit is designed as a differential track-and-hold the droop-effect due to the base current of T_7 becomes a common effect. As long as the single-ended stages remain correctly biased the droop will only marginally affect the signal. Another improvement of the differential circuit that is indicated in [90], is a pair of compensation capacitors for hold-mode feed-through that are cross-connected from the collectors of T_2 to the hold capacitors.

4.4.5 Distortion and Noise

The previous circuit example allows to illustrate a design choice that comes up in many track-and-hold circuits: the trade-off between distortion and noise. The dependence of the noise voltage on the capacitance is known from kT/C (3.16). The distortion component can be calculated from the ideal circuit in Fig. 4.25. If the bipolar emitter follower circuit[3] copies the input signal to the capacitor and the current source provides an ideal and constant current, the capacitive current

---

[3] A bipolar circuit is simple to analyze, of course the same holds for a MOS circuit.
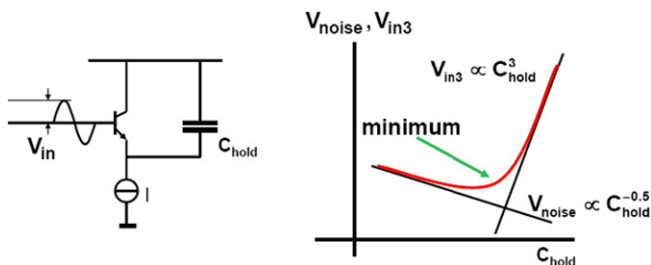
Fig. 4.25 This simple example shows that the hold capacitor cannot optimize the noise and distortion simultaneously

$i_C = j\omega C V_a \sin(\omega t)$ will be taken away from the current in the bipolar transistor. This variable current will modulate the base-emitter voltage:

$$I - i_C = I_0 e^{\frac{q(V_{BE} - \Delta V_{BE})}{kT}}$$

Reversing the exponent to a logarithm and Taylor expansion for the first three terms gives:

$$\Delta V_{BE} = \frac{kT}{q}\left[\frac{i_C}{I} - \frac{1}{2}\left(\frac{i_C}{I}\right)^2 + \frac{1}{3}\left(\frac{i_C}{I}\right)^3\right]$$

If the input signal to the emitter follower is a perfect sine wave, the voltage that is rendered to the capacitor will differ ΔV_{BE} containing second and third order terms. The terms in the equation are a function of the signal current over the total current. This ratio, sometimes called "the modulation depth", determines the relative magnitude of the higher order terms with respect to the fundamental. The second order and other even order harmonic terms can be eliminated by differential design, the third order and odd harmonics will however remain.

Substitution of the sine function leads to an estimate[4] of the fundamental and third order distortion component:

$$v_{C,1} = V_a - \frac{kT}{q}\left(\frac{V_a j\omega C}{I}\right)$$

$$v_{C,3} = \frac{1}{12}\frac{kT}{q}\left(\frac{V_a \omega C}{I}\right)^3$$

(4.11)

More current or less signal swing, lower frequency or lower capacitance will reduce the relative impact of the third order term. On the other hand less signal swing and a smaller hold capacitor lead to a lower signal-to-noise ratio. In the right hand side of Fig. 4.25 the distortion term is compared to the noise contribution. For a given set of frequency, signal amplitude and bias current parameters there will be a optimum capacitor value for a minimum combination of noise and distortion.

---

[4]Contributions from higher order terms are not included.

Chapter 5
Quantization

Digital signal samples consist of a group of bits. In binary format the bits take two values $(0, 1)$. The number of bits needed to form a digital word is called the word width. The binary representation of an analog signal is limited by the width of the digital word and the digital processing. Consequently the analog-to-digital conversion must round the analog signal to the closest digital representation. This process is called quantization.

In analog-to-digital conversion mostly the sampling of a time-continuous signal precedes the quantization. It is of course possible to start the signal processing with quantization: this (small) class of analog-to-digital converters is known under the name of level-crossing converters, see Sect. 8.10.1. Quantization was, as many other techniques, first developed for telephony transmission [91]. In this application the amplitude of a speech signal is represented by a number of pulses. The technique is called Pulse Code Modulation (PCM). This term is today generalized and describes any digital format of a quantized analog sample.

During quantization, the signal level is compared to a frame of reference values, see Fig. 5.1. The frame of reference values spans a limited number of levels, the amplitude of the continuous signal is rounded to the nearest level. Each analog-to-digital conversion produces therefore rounding errors. The step from the analog to the digital domain is consequently impaired with an error signal: the quantization error. The power associated with this quantization error is a fundamental limit to the quality of the process of analog-to-digital conversion.

Most amplitude discrete representations use a binary coding base. The number of levels is a power of the base number "2". The power N is called the resolution of a conversion, and defines the number of levels to which an amplitude continuous signal can be rounded as 2^N, see Fig. 5.2. The IEEE has standardized a number of conversion terms in standards IEEE 1057 [92] and IEEE 1241 [93, 94]. The accuracy of the conversion depends on the quality of the quantization levels. "Accuracy" and "resolution" are often confused. If the range between 0 and 0.8 V is theoretically quantized in 8 sections by 7 decision or trip levels: $0.1, 0.2, \ldots, 0.7$ V the resolution is $N = 3$ bit. In practice these levels are disturbed by offsets, gain errors and random errors. If the decision levels are shifted to $0.04, 0.14, 0.24, \ldots, 0.74$ V,

Fig. 5.1 An analog-to-digital
converter rounds the
amplitude of the continuous
signal on the time-discrete
sampling moments to the
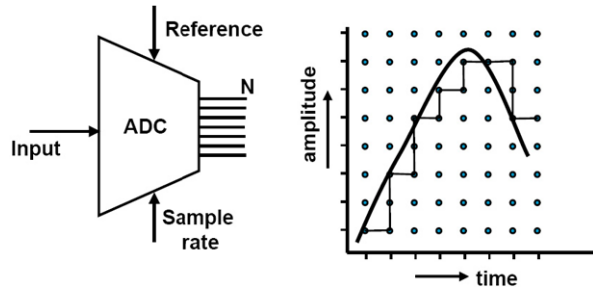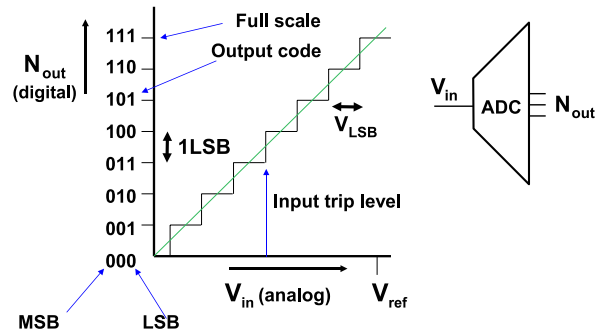closest value of the discrete
amplitude scale

Fig. 5.2 Definition of
analog-to-digital conversion
parameters

the relative error between the levels is 0%. The resolution is still perfect. But the
absolute accuracy is shifted by 0.04 V with respect to an absolute external voltage
level.

One possible signal representation in a binary system is the "straight binary representation"[1]:

$$B_s = \sum_{i=0}^{i=N-1} b_i 2^i = b_0 + b_1 2^1 + b_2 2^2 + \cdots + b_{N-1} 2^{N-1} \tag{5.1}$$

The coefficient of the 2^{N-1} term is called: "Most Significant Bit" or MSB. The
coefficient of the 2^0 term and the step between to successive quantization levels are
called "Least Significant Bit" or LSB. The term LSB is restricted to the numerical
domain. The physical equivalent of an LSB corresponds to A_{LSB} where A stands
for voltages, currents, charges or other physical quantities:

$$\frac{\text{full scale}}{2^N} = \text{LSB} \quad \Leftrightarrow \quad A_{LSB} = \frac{\text{physical reference}}{2^N} \tag{5.2}$$

where the reference is the range of the physical quantity where the analog signal
is expected. Of course there is no conversion available outside the range defined
by the full-scale range. Sometimes industrial converters are equipped with signals
indicating an "overflow" or "underflow".

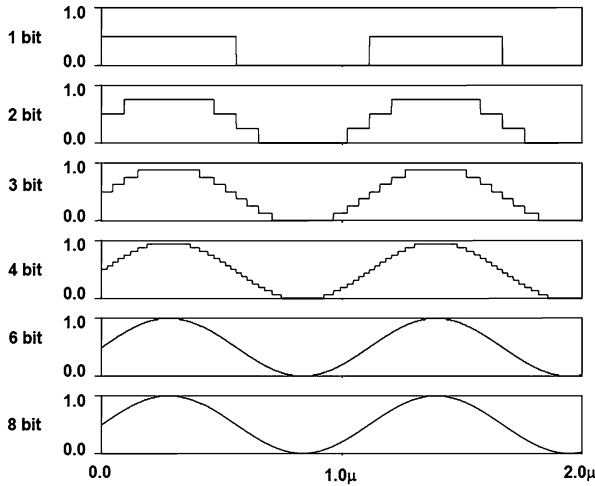[1] Other digital code formats are discussed in Sect. 7.1.1.

Fig. 5.3 Representation of a sine wave quantized with 1, 2, 3, 4, 6, and 8 bit resolution

Figure 5.3 shows the quantization of a 900 kHz sine wave at an increasing resolution.

5.1 Linearity

An ideal analog-to-digital converter will only show the imperfection of the quantization error. In practice there will be many additional errors, depending on the method of conversion, the available technology and circuit, the required resolution, the signal and sampling frequency and the skills of the designer.

5.1.1 Integral Linearity

In Fig. 5.4 $A(i)$ is the analog value where the digital code trips from code i to $i + 1$. The "trip levels" or "decision levels" of an ideal converter $A(i)$ are given by $A(i) = i \times A_{LSB}$. The integral linearity of a converter shows its deviation from the ideal conversion function.

$$\mathrm{INL} = \frac{A(i) - i \times A_{LSB}}{A_{LSB}}, \quad \forall i = 0 \dots (2^N - 1) \tag{5.3}$$

Often the INL is given as a curve, however also INL can be reduced to single number corresponding to the maximum deviation over the entire range:

$$\mathrm{INL} = \max \left| \frac{A(i) - i \times A_{LSB}}{A_{LSB}} \right|, \quad \forall i = 0 \dots (2^N - 1) \tag{5.4}$$

Fig. 5.4 Definition of the integral linearity error
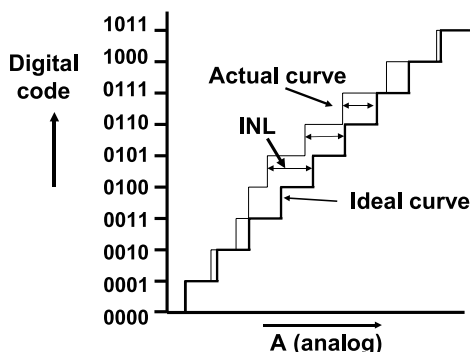

Figure 5.4 shows the ideal (bold line) conversion of the input signal to a digital code and a practical line. The converter in this example does not follow the ideal line, and in the middle a deviation of roughly one LSB is visible. The maximum number for the Integral Non-Linearity (INL) of this converter is therefore 1 LSB.

The above definition implies that the conversion process starts at an input signal at "0" level and ends at the full-scale value. This absolute requirement is important in some industrial and measurement systems. In many other systems offsets in the absolute value are acceptable, e.g. in ac-coupled circuits. Deviations in the slope of the transfer curve can also be handled by many systems and result in an amplification error. In those cases a more loose definition of the integral linearity is sufficient: the performance is then measured against the best fitting straight line. In the example of Fig. 5.4 this will mean a shift of roughly half an LSB, which will shift the INL specification from $+1/0$ to $+0.5/-0.5$. The maximum INL deviation is no reduced to 0.5 LSB.

The integral linearity is directly related to the harmonic distortion. The specific shape of the transfer curve of a converter as it is given by the INL will determine the magnitude of the specific harmonic components. The Total Harmonic Distortion (THD) is a well-known term from Sect. 2.1.3 to describe the linearity deviations in the frequency domain:

$$\text{THD} = 10^{10} \log\left(\frac{\text{Power of 2nd and higher harmonics}}{\text{Power of 1st harmonic}} \right) \tag{5.5}$$

THD is the power ratio between the signal and its harmonics. Usually the first 5 or 10 harmonics are counted as THD, while higher-order components and folded products are counted as SINAD contributions.

5.1.2 Differential Linearity

Next to the integral linearity the differential linearity is important in characterizing the DC-transfer curve of analog-to-digital and digital-to-analog converters. The Dif-
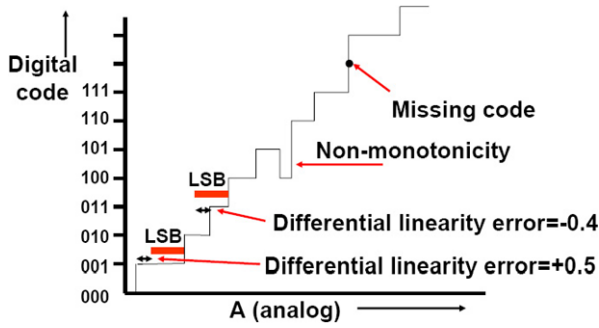
Fig. 5.5 Definition of the differential linearity

ferential Non Linearity (DNL) is the deviation of each step with respect to the ideal
LSB size. The mathematical formulation is:

$$\text{DNL} = \frac{A(i+1) - A(i)}{A_{\text{LSB}}} - 1, \quad \forall i = 0 \ldots (2^N - 2) \tag{5.6}$$

or as a single maximum number:

$$\text{DNL} = \max \left| \frac{A(i+1) - A(i)}{A_{\text{LSB}}} - 1 \right|, \quad \forall i = 0 \ldots (2^N - 2) \tag{5.7}$$

Figure 5.5 shows in the lower part a limited DNL error. For clarity small bars are
added to show the size of "1 A_{LSB}". Higher up in the curve two extreme situations
of DNL errors are illustrated. In certain constructions of converters, such as binary
coded analog-to-digital converters, an increasing input signal may result in a short
step towards a lower digital output code: the converter is non-monotonic.[2] This be-
havior can result in catastrophical problems if this converter is part of a control loop.
Therefore some systems explicitly demand for monotonic behavior: an increase in
input must lead to a zero or positive increase in output code.

In the upper part of Fig. 5.5 an increase of input signal leads to an increase of
two LSBs, one digital code is skipped: this error is called "missing code". Note that
a missing code is equivalent with a DNL $= -1$ and that is the lowest DNL value.
Figure 5.6 shows an example of an INL and DNL curve for an eight-bit converter.
Both curves give information on the performance but also on the architecture of the
converter. The saw-tooth shaped pattern in the INL curve indicates a sub-ranging
architecture, while the overall "sine" pattern is an indication that a third order distor-
tion is dominant. Obviously a rigorous differential design has eliminated the second
order component, which would have resulted in a "half sine wave".

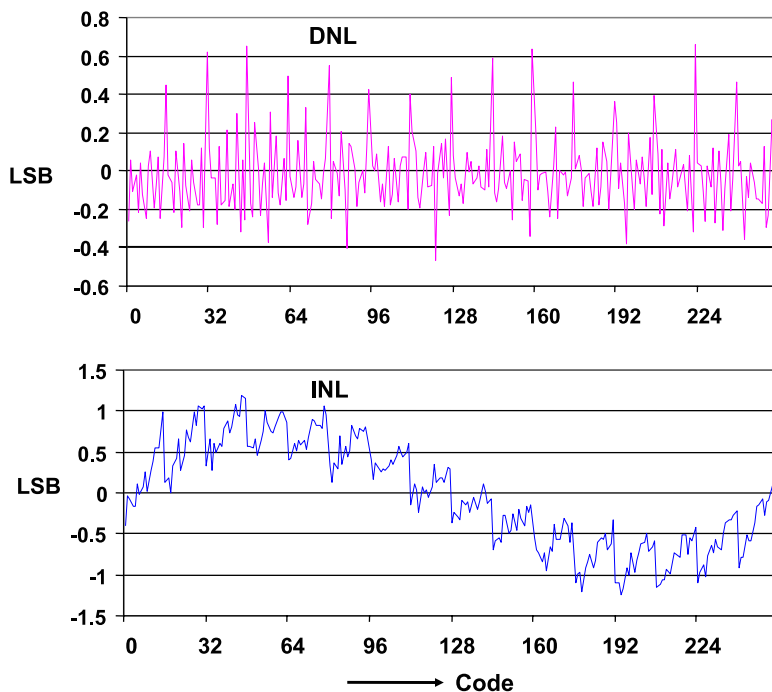[2]Dutchmen often confuse monotonic with monotonous which means "vervelend" (=boring).

Fig. 5.6 An example of a DNL curve (*upper*) and an INL curve (*lower*) of an eight bit converter

5.2 The Quantization Error

5.2.1 One-bit Quantization

The quantization energy limits the analog-to-digital converter's performance. The analysis of quantization errors is the subject of many mathematical treatises [95–97].[3] Figure 5.7 shows the spectrum of a 900 kHz signal quantized at 100 Ms/s with resolutions of $N = 1$, to $N = 8$. Quantization and sampling are mutually independent processes. The effect of sampling can be regarded as an independent pre- or post-processing of a quantized signal.

In the case of $N = 1$ the analog-to-digital converter is in fact a simple one-level comparator and reshapes the input sine wave into a simple block-wave. The quantization error equals the higher terms in the Fourier expansion of a sine wave, which were calculated in Sect. 2.1.2:

$$f(t) = \frac{4}{\pi} \sin(2\pi t/T) + \frac{4}{3\pi} \sin(6\pi t/T) + \frac{4}{5\pi} \sin(10\pi t/T) + \cdots \quad (5.8)$$

---

[3]N. Blachman has mathematically analyzed many processes around quantization. His publications from 1960–1985 are a good starting point.
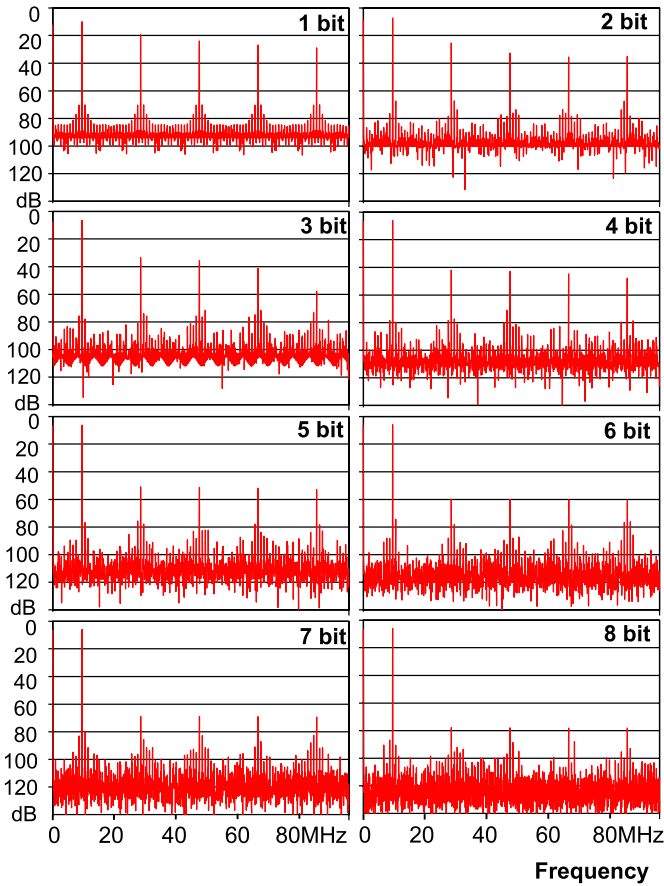
Fig. 5.7 The frequency spectrum of a 900 kHz signal quantized at 100 Ms/s with resolutions of $N = 1$, to $N = 8$

and the total energy ratio between fundamental and harmonic components amounted to a theoretical value of 6.31 dB.

5.2.2 2–6 bit Quantization

With a resolution of more than one bit the harmonic components will contain less power, as the approximation of the sine wave by the multilevel discrete signal will improve. Quantization is a distortion process, where the power is now in higher order components.[4] Blachman derives an expression for the p-th harmonic of quan-

---

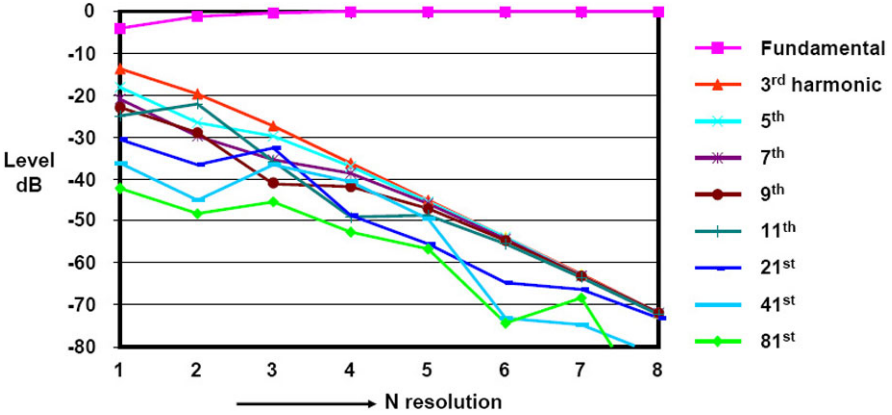[4] As the simulator for Fig. 5.7 requires to sample the signal, also alias components are visible.

Fig. 5.8 The frequency components of a 900 kHz signal quantized at 100 Ms/s with resolutions of $N = 1$ to $N = 8$. The fundamental and the harmonic at 3, 5, 7, 9, 11, 21, 41, and 81 times the fundamental frequency are shown as a function of the quantization resolution. A decay of 8–9 dB/bit is visible

tized signal $\hat{A}\sin(2\pi ft)$, where A_{LSB} is equal to 1 [96]:

$$
\begin{aligned}
y(t) &= \sum_{p=1,3,5,\ldots}^{\infty} A_p \sin(2\pi pft) \\
A_p &= \hat{A} \quad \text{for } p = 1, \\
A_p &= \sum_{n=1}^{\infty} \frac{2}{n\pi} J_p(2n\pi\hat{A}) \quad \text{for } p = 3, 5, \ldots,
\end{aligned}
\tag{5.9}
$$

A_p are the coefficients of the harmonic terms. J_p is a first-order Bessel function. For large amplitudes of $\hat{A}$, the last equation can be approximated by:

$$
\begin{aligned}
A_p &= \hat{A} \quad \text{for } p = 1, \\
A_p &= (-1)^{(p-1)/2} \frac{h(\hat{A})}{\sqrt{\hat{A}}} \quad \text{for } p = 3, 5, \ldots
\end{aligned}
$$

For $\hat{A} = 2^3, 2^4, 2^5, \ldots$, the value of $h(\hat{A})$ is more or less constant [96]. The ratio between the fundamental component A_1 and the odd distortion components equals $\hat{A}^{3/2} = 2^{3N/2}$. In dBs: the distortion amplitude of the odd harmonics reduces 9 dB per added bit. OudeAlink [97] finds a value closer to 8 dB/bit.

Figure 5.8 shows a simulation result of a 900 kHz fundamental frequency and some of its harmonic frequencies as a function of resolution. When the resolution is increased, a reduction of the third harmonic by some 8 dB per bit is seen. Also the amplitude of other harmonic components will reduce, be it at less regular steps due to the fact that at some frequencies multiple components interfere in this simulation.
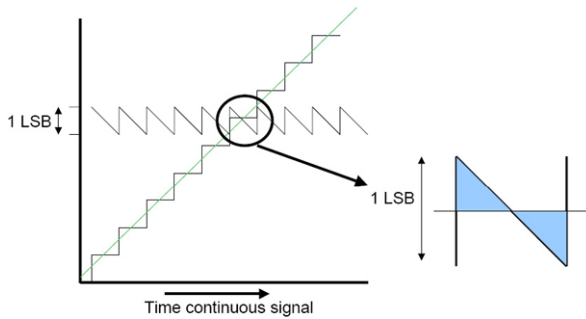
Fig. 5.9 As the analog-to-digital converter rounds the signal to the closest level, a triangular shaped error signal will arise. This saw-tooth signal is the fundamental quantization error

5.2.3 7-bit and Higher Quantization

The quantization error is a non-linear phenomena and an approximation can be used for quantization higher than 6 bit to handle this phenomena on a system level. The first order approximation for quantization of the signal supposes that the time-continuous signal has a uniform probability density of the occurrence of the signal within the range of conversion. This assumption does not take any signal specific properties into account, neither will the result be sensitive to small changes in the properties of the signal. Figure 5.9 shows the error signal that is generated while the input signal of an analog-to-digital converter steadily increases. The error signal is at its extremes at the trip levels, and varies linearly from +0.5 LSB to −0.5 LSB in between. The point at the arithmetic average between two trip points is the value used for reconstructing the signal. Optimum quantization [98, 99] or the lowest error power is reached if the quantization levels are uniformly spaced. The power contained in the error signal plays a dominant role in the analysis and application of analog-to-digital converters. For a small signal or for an analog-to-digital converter of low resolution the quantization error strongly depends on the input signal and will show up as distortion, see Fig. 5.7. However for a sufficiently large signal with a frequency that is not linked to the sample rate, the multitude of distortion products that fold with all multiples of the sample rate allow a statistical approximation of the error signal. This deterministic error signal after sampling is approximated as white noise in the band from 0 to $f_s/2$ and mirrored to the higher bands. Some authors call this quantization energy: "noise", however it shows only in a limited sense noise properties. Table 5.1 compares a few characteristics of thermal noise and quantization energy.

The probability density of the error signal between −0.5 LSB and +0.5 LSB is assumed constant and uniformly distributed. The power in that range, see Fig. 5.9, is determined by calculating the estimation value for the variance. The integration

Table 5.1 A comparison of thermal noise and quantization energy

Thermal noise	Quantization energy
Physical phenomenon caused by Brownian motion of charges	Mathematically determined distortion products
A white noise density	Modeled as white noise density
Temperature dependent	No physics involved
Resistor value dependent	Resolution dependent
Requires power to reduce	Requires more resolution to reduce
Amplitude is Gaussian distributed $\mu = 0, \sigma = v_{\text{noise}}$	Amplitude is uniform distributed $[-A_{\text{LSB}}/2, A_{\text{LSB}}/2]$
$v_{\text{noise}} = \sqrt{4kTR\Delta BW} = 4$ nV @$R = 1k\Omega, \Delta BW = 1$ Hz	
$v_{\text{noise}} = \sqrt{kT/C} = 0.22$ mV @$C = 1pF, BW = 0\ldots f_s/2$	$A_{\text{noise}} = \dfrac{A_{\text{LSB}}}{\sqrt{12}}, \quad BW = 0\ldots f_s/2$

of the amplitude squared times the probability density function yields the equivalent power for the quantization. The quantization power now equals[5] [95]

$$Q^2 = \frac{1}{A_{\text{LSB}}} \int_{\varepsilon=-0.5A_{\text{LSB}}}^{\varepsilon=0.5A_{\text{LSB}}} A_{\text{error}}^2(\varepsilon)\, d\varepsilon = \frac{1}{A_{\text{LSB}}} \int_{\varepsilon=-0.5A_{\text{LSB}}}^{\varepsilon=0.5A_{\text{LSB}}} \varepsilon^2 d\varepsilon$$

$$= \frac{A_{\text{LSB}}^2}{12} \tag{5.10}$$

Despite the constraints this formula is sufficiently accurate for most applications. The term A_{LSB} refers to the physical size of the least significant bit. Referred to the full-scale:

$$\text{Quantization energy} = \frac{A_{\text{LSB}}^2}{12} = \left(\frac{\text{full-scale}}{2^N \sqrt{12}} \right)^2 \tag{5.11}$$

where "full-scale" is the analog range of the conversion and N the resolution.

5.3 Signal-to-Noise

At a resolution of $N = 7$ the spectrum of an error signal is mostly flat enough for an approximation as "white noise". Yet the level of the odd harmonics is at -54 dB. In specific systems still special attention must be paid to the distortion side of the

---

[5]Note that in a formal sense just voltage-squared is calculated, which lacks the impedance level and the time span before reaching the dimension of power. In all applications this "voltage-squared" power is used to compare to another "voltage-squared" power, assuming that both relate to the same impedance level and the same time frame.

quantization error. If the assumptions are met where the white-noise approximation holds, the "noise" due to quantization dominates other noise sources up to 14–16 bit resolution.

Many systems use or are based on sine related specifications. A sine wave can be easily generated with high quality. Therefore in analog-to-digital converters sinusoidal signals are used to characterize the performance. The quantization error is related to the maximum sinusoidal signal that the analog-to-digital converter can handle. In fact this comparison violates the above assumption of a uniformly distributed signal over the conversion range, especially when the sine wave is at its peak values.

The signal-to-noise ratio (SNR) compares the noise power to the power in a full sine wave as in Sect. 2.2.5:

$$\text{SNR} = 10^{10}\log\left(\frac{\text{Signal power}}{\text{noise power}}\right) = 20^{10}\log\left(\frac{V_{\text{signal,rms}}}{V_{\text{noise,rms}}}\right) \qquad (5.12)$$

The quantization errors of the analog-to-digital converter are considered uncorrelated. The resulting noisy spectrum is modeled as a white noise spectrum, whose power is contained between $f = 0$ and $f = f_s/2$ and mirrored around the multiples of the sample rate. The quantization error power can be used to calculate an equivalent signal-to-noise ratio:

$$\text{Quantization power} = \frac{A_{\text{LSB}}^2}{12},$$

$$\text{Signal power} = \frac{1}{T}\int_{t=0}^{t=T} \hat{A}^2 \sin^2(\omega t)\, dt = \frac{\hat{A}^2}{2} = \frac{2^{2N} A_{\text{LSB}}^2}{8}$$

$$\text{Signal-to-Noise Ratio} = \frac{\text{Signal power}}{\text{Quantisation power}} = \frac{3}{2}2^{2N}$$

$$(5.13)$$

$$\text{SNR} = 10^{10}\log\frac{3}{2}2^{2N} = 1.76 + N \times 6.02 \text{ dB}$$

This last formula is an approximation that is often used in designing an analog-to-digital converter:

The maximum signal-to-noise ratio (SNR) represented by a digital word of N bits in a bandwidth of $f_s/2$ is $1.76 + 6.02 \times N$ dB.

As an example: a converter of 10 bit resolution can never reach a better signal-to-noise ratio in the full conversion band than 62 dB. And a 8-bit digital bus is limited to 50 dB. In video applications the top-top signal value of the video content is used as a reference. This description results normally in a 10 dB higher value for the SNR than in the sinusoidal case.

All simple descriptions have their limits, so the question arises: when is the white noise model a correct approximation for a phenomena, which in fact is a complex distortion spectrum. Table 5.2 compares the simulated signal to quantization

Table 5.2 Simulated signal-to-quantization power compared to the approximation formula

Resolution	Simulated SNR	$6.02N + 1.76$ dB	Theory
1	6.31 dB	7.78 dB	6.31 dB
2	13.30 dB	13.80 dB	
3	19.52 dB	19.82 dB	
4	25.60 dB	25.84 dB	
5	31.66 dB	31.86 dB	
6	37.71 dB	37.88 dB	
7	43.76 dB	43.90 dB	
8	49.80 dB	49.92 dB	

power ratio to the simple approximation of $6.02N + 1.76$ dB. For $N = 1$ the simulated value corresponds perfectly with the mathematical analysis. With increasing N the approximation gains accuracy, although a minor overestimation of the noise remains. A cause for this overestimation is that uniformity of the signal over the entire range was assumed, which is not the case for a sine wave approaching the top and bottom values. Note that if the signal amplitude spans 3% of the full range of a ten-bit analog-to-digital converter, this converter functionally is equivalent to a five bit analog-to-digital converter.

Looking at the quantization energy density in a fixed bandwidth, there are now two ways to reduce this density: increase the resolution N or spread out the noise thinner over more bandwidth by increasing f_s. Doubling the sampling rate halves the quantization power per Hertz and results for a fixed bandwidth in a higher SNR. These two directions correspond to the classification also used in this book:

- Nyquist converters use most of the available bandwidth between 0 and $f_s/2$. The bandwidth of interest can be high or system requirements force to use these converters. Often a form of track-and-hold circuit is used to have a stable copy of the signal. These converters require some array or string of passive or active components to subdivide the reference value. Connecting these converters to the digital world is trivial. Chapters 7 and 8 discuss Nyquist converters.

- Oversampled converters such as sigma-delta converters use the large difference between the bandwidth of interest and the technologically available sampling speed. If more than a factor of 10 can be reached oversampled converters are a worthwhile option. Now the accuracy comes from the time domain. Some form of glue circuitry is required to connect these converters to the digital processor. Chapter 9 discusses these converters.

5.3.1 Related Definitions

The Signal-to-Noise-And-Distortion (SINAD) stands for the ratio of the signal power to all the unwanted components: quantization errors, thermal noise, distortion, etc.:

$$\text{SINAD} = 10^{10}\log\left(\frac{\text{Power of 1st harmonic}}{\text{Power of all unwanted components}}\right) \qquad (5.14)$$

The Spurious Free Dynamic Range (SFDR) is the distance between the signal and the largest single unwanted component, the spurious signal.

The Dynamic Range (DR) is sometimes equivalent to SNR or SINAD as it represents the ratio of the full-scale input signal and the noise floor at a small signal input. The difference between DR and SNR is clearly present in e.g. range switching configurations.

In order to characterize the converter in a simple manner, the effective number of bits (ENOB) is calculated by reversing (5.13):

$$\text{ENOB} = \frac{\text{SINAD} - 1.76}{6.02} \qquad (5.15)$$

The ENOB allows an easy comparison of the real performance of a converter. Suppose a nominal 8 bit resolution is in the data sheet. If the measurements show only 6.8 ENOB, the converter loses a lot of performance. At 8-bit level no more than 0.5 ENOB can be lost. At 12-bit a loss of 1 ENOB is tolerable.

5.3.2 Non-uniform Quantization

Many signals exist with a non-uniformly distributed amplitude function. Video-camera signals show most detail close to the black level, while in full-white signals only little information is present. Speech signals are symmetrically distributed around the zero value and need much finer quantization around the mid-point than in the extreme excursions. A similar argument holds for OFDM modulated digital-communication signals. These signals consist of a large number of individually modulated carriers. On average the power of these carriers will add up in a root-mean-square sense: $v_{\text{rms}} = \sqrt{v_1^2 + v_2^2 + \cdots}$, however every now and than a time moment may occur where a significant number of carriers are in phase and their amplitudes add up: $v_{\text{top}} = v_1 + v_2 + \cdots$. The ratio between v_{top} and v_{rms} is called the crest factor or the peak-to-average ratio. Sine waves have a crest factor of $\sqrt{2}$ or 3 dB. In practical communication systems crest factors in excess of 10 dB will occur, requiring more than a full bit of extra resolution.

In communication literature various attempts have been made to come up with specific non-uniform quantization schemes, to improve the conversion quality at lower hardware costs, e.g. for normal amplitude distribution [99]. However in practice this means to design a specific analog-to-digital converter for each signal type.
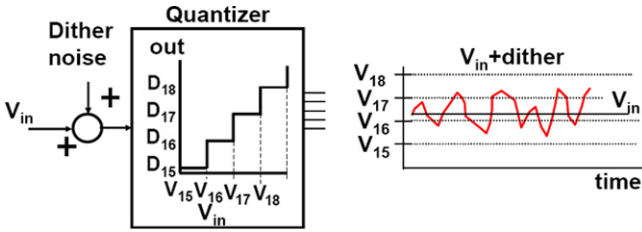
Fig. 5.10 The addition of a random signal allows to determine the value of a DC-signal at greater accuracy than the quantization process allows

A more common approach is to precede a uniform quantizer with an appropriate compression circuit. A starting point is the μ-law compression of function $f(t)$ in a compressed $g(t)$:

$$g(t) = \text{sign}(f(t)) \frac{\ln(1 + \mu|f(t)|)}{\ln(1 + \mu)} \qquad (5.16)$$

where μ determines the amount of compression. It is obvious that after compression and analog-to-digital conversion a de-compression step is required with the exact inverse function. Any intermediate signal processing must take the complex structure of the compressed signal into account.

5.3.3 Dither

Slow variations of a signal that occur between two trip levels, will be lost after rounding to the closest representation level. This quantization distortion can be annoying in low-resolution conversion systems. A helper signal or "dither" can be added to the original signal [100], see Fig. 5.10. This dither signal has typically the magnitude of 1 A_{LSB} and consists of a well-chosen random signal.[6] Due to the dither signal, the signal + dither in between the two trip levels will be pushed over and under these trip levels, more or less proportional to the distance of the original signal level to the trip levels. Additional signal processing like averaging, can now lead to resolution improvement for low-frequency signals (compare also Sect. 9.1). In spectrum terms: the dither signal turns the quantization distortion into high-frequency random noise. Conceptually the most simple form of dither is thermal noise which is however not so easily controlled. An alternative is a uniform distribution ranging from 0 to V_{LSB}. An input signal that is quantized:

$$V_{\text{in}} = (B_{\text{integer}} + Q_{\text{fraction}}) \times V_{\text{LSB}} \qquad (5.17)$$

is composed of the quantized digital signal $B_{\text{integer}} \times V_{\text{LSB}}$ and the (lost) quantization fraction of the signal $Q_{\text{fraction}} \times V_{\text{LSB}}$. The addition of the uniform noise dither will

---

[6]There has been an extensive search for optimum dither signals in the 1960–1970s. After that the interest for dither has reduced. The concept however still provides insight.

also activate the quantization trip level at $B_{\text{integer}} + 1$. For a total of M samples, the amount of hits on this level is M_{fraction} and proportional to Q_{fraction}. The average digital output over M samples is now:

$$\frac{(M - M_{\text{fraction}})B_{\text{integer}} + (M_{\text{fraction}})(B_{\text{integer}} + 1)}{M}$$

$$= B_{\text{integer}} + \frac{M_{\text{fraction}}}{M} \approx B_{\text{integer}} + Q_{\text{fraction}}\Big|_{M\to\infty} \tag{5.18}$$

This example of uniformly distributed dither with an amplitude of V_{LSB} may seem constructed and difficult to realize in practice, however note that the noise-shaper in Sect. 9.2 uses its own delayed quantization error as a dither signal.

5.3.4 DNL and SNR

The above analysis also allows to include the influence of a DNL error on the quantization noise performance. Differential Non-Linearity represents the non-uniformity of the trip levels as can be caused by various random processes (e.g. comparator offsets in full-flash converters). Often it is acceptable to characterize these deviations by a Gaussian distribution. Every trip level is modified by an instance of this distribution:

$$A(i) = i \times A_{\text{LSB}} + A_G(i) \tag{5.19}$$

where $A_G(i)$ is a sample from a Gaussian distribution with $\mu = 0$, and the variance σ_A. The addition of this random process has two consequences: on one hand the random fluctuation of the trip levels can be seen as a random noise source in series between the input signal and an ideal quantizer. This assumption is valid if the signal varies sufficiently and passes a large number of trip levels. On the other hand the variation in trip levels is measured in the DNL curve and summarized by its maximum value "DNL". The maximum value for $2^N - 1$ trip levels is then the threshold value n in the Gaussian probability function, for which the probability that $2^N - 1$ instances will remain within the interval $(-\text{DNL}, +\text{DNL})$ is acceptable, see Table 2.10. In a normal distribution n is in the range $n = 3.0\text{--}3.4$ for $N = 8$ to 10.

If a full-swing sine-wave input is assumed, the signal to noise due to quantization and DNL errors for N-bit resolution is:

$$\text{SNR}_{\text{Q+DNL}} = 10^{10}\log 2^{(2N-3)} - 10^{10}\log\left(\frac{1}{12} + \frac{\sigma_A^2}{V_{\text{LSB}}^2}\right)$$

$$= 6.02N - 9.03 - 10^{10}\log\left(\frac{1}{12} + \frac{\text{DNL}^2}{2n^2}\right) \tag{5.20}$$

For $n = 0$, $\text{DNL} = 0$ the formula results in the well-known $\text{SNR} = 6.02N + 1.76$ dB. In Fig. 5.11 the SNR has been plotted in dB versus the DNL for $N = 8$ and $n = 3.0$. The squares indicate the results of Monte-Carlo computer simulations,

Fig. 5.11 Signal-to-noise ratio for quantization with Gaussian-distributed random offset errors versus the expected value of the DNL. The squares indicate Monte-Carlo simulations, the triangles refer to measurements obtained using the A/D converter described in Sect. 8.9
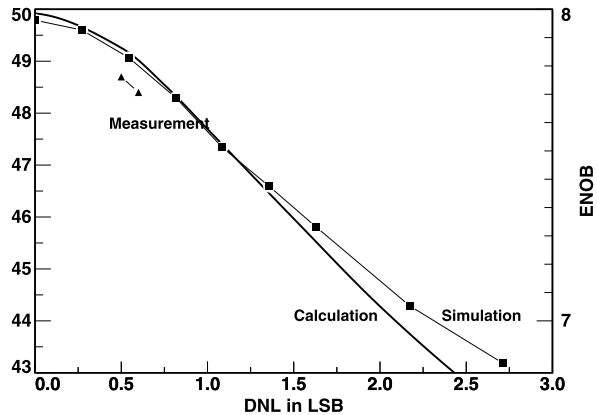
for both SNR and DNL. At higher DNL the assumption of independence of "DNL-noise" and quantization noise is less valid. Moreover, the simple analysis ignores the fact that sine-wave signals more often involve the quantization errors at the top and bottom than errors around mid-range. The triangles indicate measurement points.

Figure 5.11 shows that a DNL = 0.5 LSB yields a −49.2 dB noise level and a poor DNL = 1 LSB results in a SNR = 47.7 dB, corresponding to a loss of 0.37 ENOB.

Chapter 6
Reference Circuits

6.1 General Requirements

In many practical systems a reference quantity is present to define the signal range. This quantity is mostly a voltage, current, charge or time period. In the field of analog-to-digital conversion the reference value, and consequently the maximum input signal, of an N-bit converter is subdivided into 2^N LSBs, least significant bits, where V_{LSB} is the physical value corresponding to one LSB. During operation of an analog-to-digital converter or a digital-to-analog converter, this unit is subdivided, copied or multiplied, which causes various deviations, see Chap. 11.

The number range in a digital system with a fixed word width is limited. That limitation can be imposed by the bus width or due to the capabilities of the digital signal processing. The digital signal coming from an analog-to-digital converter is then the fraction of the maximum signal that belongs to that word width (mostly 2^N, where N represents the word width). On the analog side of an analog-to-digital converter in essence the same convention is used. The reference value is now the measure for the analog signal. The ratio between the analog signal and the reference value corresponds to the ratio of the digital signal to the maximum digital signal, see Fig. 1.2.

The reference quantity in the analog domain is a key component for the quality and performance of analog-to-digital and digital-to-analog conversion. Any disturbance or error in the reference value will cause a deviation of the converted signal of the same relative magnitude. A good reference quantity meets the following criteria:

- Stable for temperature changes. Systems for industrial and consumer use are subject to temperature variations from -20 to $+70\,^{\circ}\mathrm{C}$ on the printed circuit board. In special cases these variations can be much higher (space, oil drilling) or much lower (pace-maker).
- Limited drift as a function of time. Aging of components is often related to temperature and over-voltage. Stressing a part by means of extreme temperatures and/or higher voltages can be used to extrapolate the expected life time and aging effects. Often the stress mechanisms show a typical bathtub behavior during the product life cycle: a lot of defects in the beginning and at the end, and a rather

stable behavior during the long period in between. In order to enter this stable period, fresh products are subjected to a short stress period: the burn-in period.

- Not sensitive for low-frequency and high-frequency variations in the supplies. Next to the power supply also the effect of changes in surrounding potentials can affect the operation. E.g. in case of an integrated circuit, voltage variations in the substrate must be considered.
- Reproducible in systems that are mass-manufactured. Statistical variations must be minimized.
- Low energy usage. The amount of energy that signal manipulations require, depends on the complexity of the operation, the bandwidths and accuracy. A reference is normally not listed in the high energy categories.
- A low source impedance. Some system related mal-functioning or loss of performance is due to the sharing of references for too many purposes. A reference with a relatively high output impedance will be unable to adequately react on variations in the current pattern, caused by the connected circuits. Consequently the reference voltage will distribute spurious signals to all connected circuits.
- The reference should not introduce interference or noise into the surrounding circuitry. Most reference circuits use feed-back circuitry. If no attention is paid to stability issues, or if a reference is loaded with an unexpected combination of capacitive and inductive elements, the reference circuit can superimpose on the reference quantity an oscillation.[1]

Note that reference sources always are equipped with two connections: the reference voltage is defined between two terminals. And the current from the output of the reference circuit will flow from that output via the load back to the reference ground. Both terminals contribute equally to the performance and should therefore be considered as being of equal importance.

6.2 Bandgap Reference Circuits

A reference circuit is based on some physical quantity. There are many physical quantities to chose from, see Table 6.1.

The determining factor for the choice is the degree in which a reference can fulfill the above requirements. In the voltage domain a voltage division of a relative quiet power supply may serve as a reference voltage. Zener and threshold related potentials can serve in simple applications. Accurate systems require special reference sources, and only a limited number of physical phenomena are sufficiently stable. Accurate time reference is derived from phenomena on atomic scale. Voltage references often operate on the potential difference between the conduction and the valence band in pure materials (the bandgap). Industrial references can reach an additional stability by using temperature compensation or even by stabilizing their

---

[1]Don't say: "This will not happen to me".

Table 6.1 Various domains for choosing a reference

Domain	Example	Remarks
Voltage	Bandgap voltage, threshold voltage, Zener-voltage power supply	Physics related quantity
Current	Via external reference resistor	Every current reference is a voltage reference in disguise
Charge	Electron, voltage on a capacitor	
Time	Crystal clock	For low frequency application

own temperature in a control loop. Yearly calibration further improves the performance.

In integrated circuits for consumer applications a simple bandgap reference presents a suitable solution for modest demands: an absolute accuracy of around 1% and another 1% variation during a temperature trajectory from 0 to 70 °C. The bandgap reference idea was originally developed by Hilbiber [101] in 1964. Widlar [102] and Kuijk [103] refined the idea in the early seventies.

The reference quantity in a bandgap reference is the energy difference that an electron has to overcome for moving from the valence band in silicon to the conduction band, see Sect. 2.4.1. The corresponding potential difference is called the "bandgap voltage" and varies from 1.17 V close to 0 K to 1.12 V at room temperature. The bandgap voltage curve can be back-extrapolated from room temperature to 0 K and results in a value of 1.205 V. This value of the bandgap voltage can be accessed via a diode structure, see (2.80). The current over a pn-junction is analyzed in detail in Sects. 2.4.8 and 2.4.9 and is written as:

$$I_{pn} = C_0 T^{\eta} e^{\frac{q(V_{pn}-V_{bg})}{kT}} \tag{6.1}$$

All terms in (2.79) that have no temperature dependence have been collected in the term C_0. Note that $V_{pn} < V_{bg}$ and that an increase of temperature requires a decrease in V_{pn} to keep the exponential term and the current at the same value. The forward voltage of a pn-junction V_{pn} has therefore a negative temperature coefficient of roughly -2 mV/°C. The slope is mainly determined by the temperature dependence of the intrinsic carrier concentration n_i, see (2.60) and slightly curved due to the temperature dependence of the mobility. Mathematically the temperature power term is expressed via the term T^{η}. Practical values of η range between 2 and 4.

Figure 6.1 shows the principle of a bandgap reference [101]: the voltage over the pn-junction, V_{pn} with its negative temperature coefficient reaches after back-extrapolation to 0 K the bandgap voltage 1.205 V. In that situation the Fermi level of the n-region aligns with the valence band and the Fermi level of the p-region aligns with the conduction band, so $V_{pn} = V_{bg}$.

The task of a bandgap circuit is add to the pn-junction voltage a component with an equal magnitude but opposite sign temperature dependence. This component can be obtained by using the difference between two pn-junction voltages with
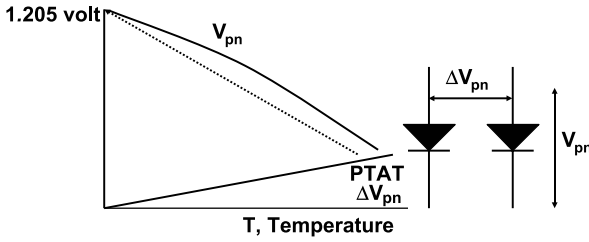
Fig. 6.1 The diode voltage V_{pn} decreases with increasing temperature at roughly 2 mV/°C. The differential voltage between two diode voltages ΔV_{pn} (e.g. created by different areas of the diodes) is proportional to the absolute temperature. By amplifying this last slope to a level where the differential voltage temperature coefficient is equal but opposite to the diode temperature coefficient, a temperature stable voltage can be obtained
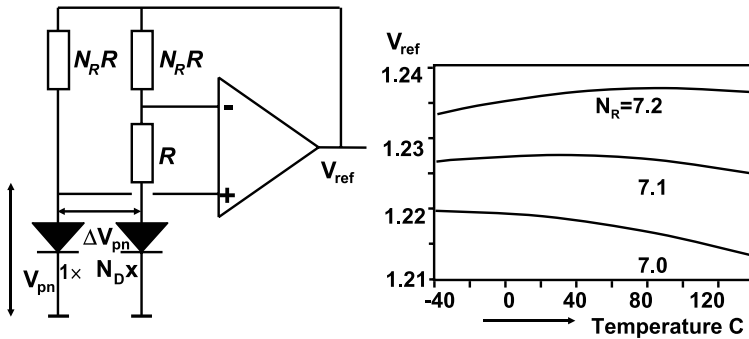

Fig. 6.2 The bandgap reference circuit [103]. The graph shows a simulation of this circuit with $N_D = 24$ and $N_R = 7.0, 7.1, 7.2$

unequal current densities. Suppose that the current density is varied by a factor N_D. This can be realized by feeding a current through a single pn-junction and a current with the same magnitude through a parallel connection of N_D pn-junctions.

$$\Delta V_{pn} = V_{pn,1} - V_{pn,ND} = \frac{kT}{q} \ln(N_D) \tag{6.2}$$

So ΔV_{pn} shows a positive temperature coefficient, often referred to as "proportional to absolute temperature", PTAT. This property of two pn-junctions is the basis for many temperature measurement schemes. In Fig. 6.2 the operational amplifier will force its input terminals to a zero potential difference. Consequently there is an equal voltage drop over the two resistors $N_R R$, and the currents in the two branches are equal. As one branch has one pn-junction and the other branch N_D equally sized pn-junctions, the desired PTAT voltage difference will exist between both pn-junctions. As the inputs of the amplifier are at the same potential, this voltage difference of ΔV_{pn} is equal to the voltage drop over the resistor R. This voltage is multiplied to $N_R \Delta V_{pn}$ over the top resistors. Because this voltage drop over R is PTAT and

the resistors are assumed to be temperature independent,[2] this will result in a PTAT behavior of the current in the branches. The output voltage of the circuit V_{ref} is now:

$$V_{\text{ref}} = V_{pn} + N_R \Delta V_{pn} \tag{6.3}$$

Of course is the choice for N_R essential for a correct compensation of the negative temperature coefficient of the pn-junction and the PTAT term.

For a more accurate calculation the above current equation of a pn-junction is substituted:

$$V_{\text{ref}}(T) = V_{pn} + N_R \Delta V_{pn} = V_{bg} + \frac{kT}{q} \ln\left(\frac{I_{pn}}{C_0 T^\eta}\right) + N_R \frac{kT}{q} \ln(N_D)$$

After differentiation for the temperature T and setting the result to zero, this maximum value is found at a temperature T_0:

$$\ln\left(\frac{I_{pn}}{C_0 T_0^\eta}\right) - \eta + N_R \ln(N_D) = 0$$

Substituting this result in the original formula gives:

$$V_{\text{ref}}(T) = V_{bg} + \frac{kT}{q} \eta \left(1 - \ln\left(\frac{T}{T_0}\right)\right)$$

The maximum point in the curvature is found at T_0 with a value:

$$V_{\text{ref}}(T = T_0) = V_{bg} + \frac{kT_0}{q} \eta$$

The choice of N_R determines T_0 and the value of $V_{\text{ref}}(T = T_0)$, see Fig. 6.2. Series expansion of the ln function with an argument close to "1" requires to use: $\ln(1 + \alpha) = \alpha$, and yields:

$$V_{\text{ref}}(T) = V_{bg} + \frac{kT_0}{q} \eta \left(1 - \left(\frac{T - T_0}{T_0}\right)^2\right) \tag{6.4}$$

The output voltage of a reference voltage circuit has consequently a parabolic shape. The temperature T_0 at which the maximum occurs, can be chosen by the resistor ratio N_R. If the diode ratio is $N_D = 8$, N_R will be around 11. The maximum output voltage is the bandgap voltage V_{bg} plus a few ($\eta \approx 2$–4) times the thermal voltage $kT/q \approx 26$ mV. The typical output voltage is 1.23–1.28 V.

The circuits in Figs. 6.2 and 6.4 have a second stable point: at 0 V output voltage there is no current and no drive to start-up the circuit. Consequently these reference circuits needs an additional start-up circuit to avoid this 0 V condition. A pull-up resistor on the reference voltage is often sufficient. More elaborate schemes compare the output to e.g. a threshold voltage and after sufficient output voltage is detected switch-off the start-up.

Most of the problems of this circuit are related to the roughly tenfold amplification in this circuit. All noise and offset components related to the input of the

---

[2]Temperature dependence of the resistors or other components is canceled during the fine tuning of the simulation.
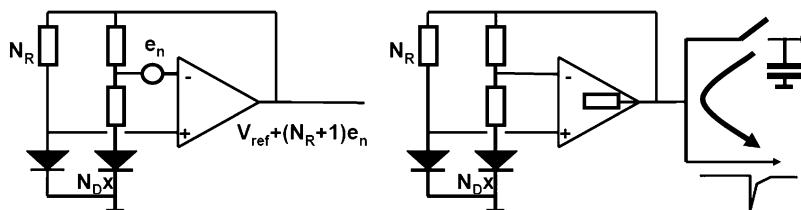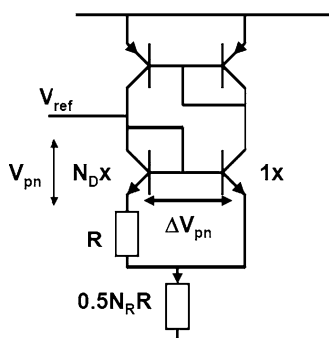
Fig. 6.3 A bandgap reference circuit amplifies unwanted input signals. If the output impedance is too high unwanted coupling will occur between connected circuits

Fig. 6.4 A bandgap reference circuit with bipolar transistors

operational amplifier (input related circuit effects) will be multiplied with this factor, see Fig. 6.3 (left). The bandwidth of the operational amplifier can be reduced to suppress high frequency components. This would also limit the response of the circuit to any load variations, or suppression of any interference on the reference output terminal. Without sufficient bandwidth the reference circuit will not be able to respond these changes at the output. In other words: for those frequencies the output impedance is too high. A proper choice for the bandwidth requires considering the loading variations in the succeeding circuits. In some cases separate reference circuits for each separate block are preferred.

6.2.1 Bipolar Bandgap Circuit

Only four bipolar transistors are needed to design a bandgap reference circuit [102], see Fig. 6.4. The base-emitter junctions of the two npn-transistors create the temperature-sensitive pn-junctions. The pnp-mirror keeps the currents in both branches equal. The npn-transistors in combination with the pnp-transistors form the operational amplifier. All currents in the circuit are PTAT. This circuit can be modified to act as a temperature sensor, for a detailed analysis see e.g. [104].
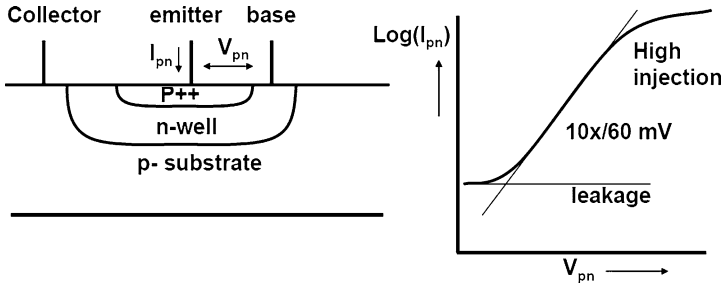
Fig. 6.5   The pn-diode or pnp transistor in a CMOS process and its $I-V$ curve

6.2.2 CMOS Bandgap Circuit

In a CMOS process the pn-diode is available as a parasitic structure created from a standard p-drain/source, an n-well and the p-type substrate [105], see Fig. 6.5. In fact this structure is a parasitic pnp transistor where the substrate serves as a collector, see Sect. 2.4.9. This structure also poses a danger: the hole current in the substrate may trigger latch-up. Latch-up in CMOS processes is possible if an intended or parasitic pnp structure feeds its collector current in an npn structure, see Sect. 2.5.9. A potential candidate for an npn structure is every n-MOS transistor, where the source, substrate and drain form the npn structure. If this npn transistor starts conducting, the base current of the pnp will increase. Now an uncontrollable process is started that can easily lead to excessive currents and damage to the circuits. The turn-on of the parasitic bipolar npn transistor depends on a sufficiently high base voltage. This base voltage can become high if there is a large resistor from the base to ground or the supply. It is important to surround the npn and pnp-devices with sufficient substrate contacts to lower this resistance and to avoid the pnp collector current to build up a high base voltage for the npn devices.

Typical emitter sizes of the pnps are $2 \times 2\ \mu m^2$ to $5 \times 5\ \mu m^2$. Due to the rather low-doped base (n-well) the current density can not be large in this structure. High-injection will cause deviation from the ideal $I-V$ curve. On the lower side leakage limits the usable range. Typical designs aim at a current range between $1\ nA/\mu m^2$ and $1\ \mu A/\mu m^2$.

The base width of this parasitic pnp transistor (0.5–1 μm) is large compared to advanced bipolar RF technology. High-speed bipolar devices normally use a base width of less than 0.1 μm and this relatively small base charge allows to operate them at high frequencies. For bandgap reference circuits this argument is irrelevant. There is even an advantage to the wide base of the parasitic pnp-transistor: a low V_{be} mismatch in the order of $\sigma = 0.2\ mV/\sqrt{area}$ where the area is measured in μm^2, see Sect. 11.4.5.

Resistors are formed from diffused p- or n-type material. In order to prevent large structures with lots of parasitics, unsalicidized material is used. This type of material will show resistivity values in the order of 50–500 $\Omega/\square$, mostly in combination with rather high temperature coefficients, see Table 2.11. These temperature coefficients need to be part of the resistor model in order to achieve a correct result on silicon.
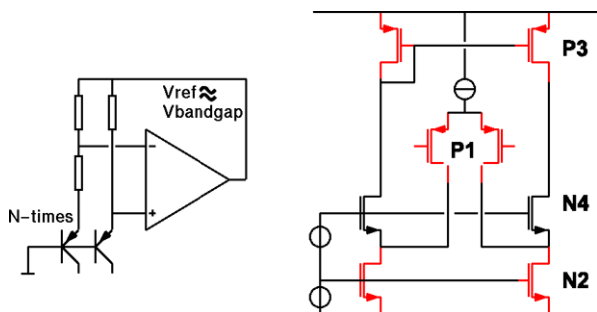
Fig. 6.6 The input stage of an operational amplifier for a bandgap reference circuit

The operational amplifier needs a sufficiently high amplification factor in order to reduce the input offset and to provide a low-ohmic output. Simple two-stage Miller operational amplifiers will serve the purpose, see Sect. 2.7.11.

Figure 6.6 shows the input stage for the operational amplifier of the reference circuit. The input transistor pair consists of PMOS devices. NMOS devices can be used as long as the NMOS threshold is not too high. The voltage over the pn-junction can reach values of 0.4 V at high temperatures, and although the temperature coefficient of the threshold voltage of NMOS devices is negative, the margins may become too small to keep high-threshold devices in inversion.

The input stage of Fig. 6.6 is build up as a folded cascode stage. As indicated in Fig. 6.3 the input offset is multiplied by the resistive ratio to the output voltage. The random offset component will cause variations from one circuit to another. The random offset in a bandgap circuit is mainly caused by the operational amplifier. The transistor pairs P1, N2 and P3 contribute to the input referred mismatch. The cascode pair N4 is not relevant for the mismatch budget as in normal operation this pair does not modify the current in the circuit.

The input-referred mismatch of pair P1 is easily found as it equals the threshold-voltage mismatch of P1 using (2.95). The threshold-voltage mismatches of pairs N2 and P3 are found in the same way, however, they need to be recalculated towards the input. Considering the threshold-voltage mismatch of N2 it is important to realize that this mismatch translates in a current mismatch. If the contribution of the other components is ignored, the input stage P1 will have to deliver an equal, but opposite current to cancel the mismatch current from N2. In order to create this current a gate voltage difference on P1 is needed. This is the input-referred mismatch voltage from N2. Mathematically this voltage is found from equating the mismatch currents from N2 to the input-referred mismatch current in P1, via the transconductances of both devices:

$$i_{\text{mismatch},N2} = g_{m,N2} \times \sigma_{VT,N2} = g_{m,P1} \times \sigma_{V\text{in},N2} \tag{6.5}$$

In a similar manner the contribution of P3 is referred back to the input. The total input referred mismatch is found in Table 6.2. The resulting variation in the output voltage of more than 20 mV is rather large. For that reason the circuit is modified to yield a lower spread of the reference voltage. The first option is to change the

Table 6.2 Calculation of input referred mismatch of the bandgap design, with $A_{VTn} = 14.2$ mV μm, $A_{VTp} = 20.5$ mV μm, $\mu_n = 2.5 \times \mu_p$

First design			Second design		
Transistor dimensions	σ_{VT} (mV)	σ_{Vin} (mV)	Transistor dimensions	σ_{VT} (mV)	σ_{Vin} (mV)
P1: 90/5	0.95	0.95	P1: 128/7	0.68	0.68
N2: 28/7	1.00	1.08	N2: 40/9	0.75	0.84
P3: 60/4	1.40	1.34	P3: 100/6	0.84	0.81
Total input referred:		1.97 mV	Total input referred:		1.35 mV
After multiplication (11×):		21.7 mV	After multiplication (7.8×):		10.5 mV
Measured reference s.d.:		24–27 mV	Measured reference s.d.:		9–10 mV

Fig. 6.7 Measured output voltage histogram of a bandgap reference circuit compared to simulation
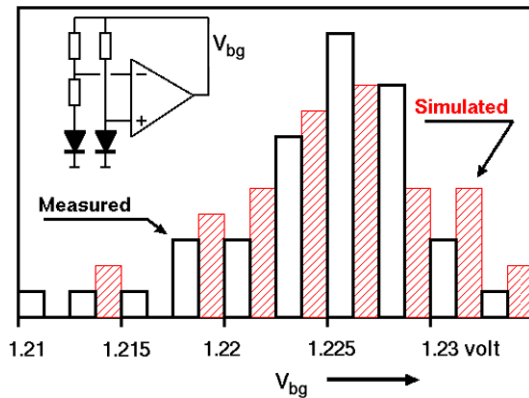
dimensions of the transistors. Larger transistors will reduce the inherent mismatch. A gate-length reduction for P1 and longer transistors for N2 and P3 lead to a lower contribution of N2 and P3 due to a more favorable g_m ratio. Next to that the multiplication factor of the opamp configuration must be addressed. A lower resistor ratio is possible with a larger ΔV_{pn} value. This requires to increase the ratio of the pnp-transistors from 1:8 to 1:24 and reduces N_R from 11 to 7.8. The second design is summarized on the right hand side of Table 6.2.

Next to these hand calculations it is well possible to use circuit simulation for estimating the input-referred mismatch voltage. Figure 6.7 shows a comparison of a typical simulation and a measured batch of reference circuits. Both the mean value and the standard deviation agree reasonably well for this sample size.

Fig. 6.8 A bandgap reference circuit for operation below 1.2 V [106]
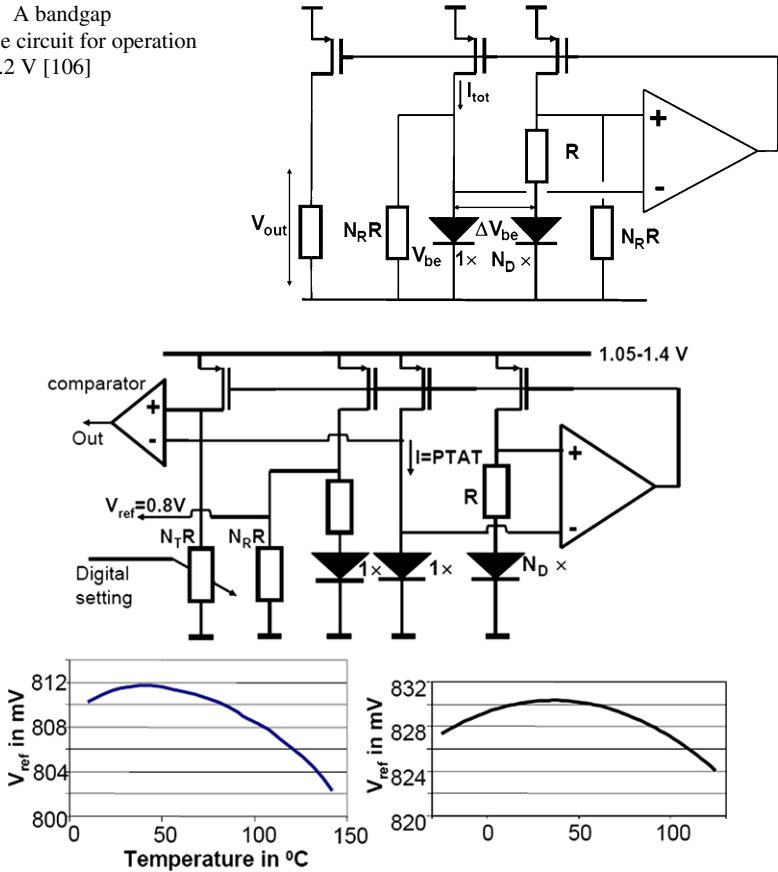
Fig. 6.9 A bandgap reference circuit of 0.8 V extended with temperature monitoring [107, 108]

6.2.3 Low-voltage Bandgap Circuits

Stacking a diode voltage and a temperature compensating voltage drop over a resistor requires a minimum power supply voltage of 1.4–1.5 V. In order to operate at lower supply voltages Banba [106] presented an elegant solution, Fig. 6.8. In this circuit it is not the voltage that is made constant but the current. The PTAT current is still generated as shown in the previous section. Given a PTAT current the question is what to add to get a temperature stable current. By connecting two resistors over the diodes an additional current is drawn which contributes in first order a term: $V_{be}/N_R R$ which has the negative temperature coefficient of the base-emitter junction. After having generated a stable current, a simple mirror operation allows to define any voltage level required. Again the absolute value of the resistor will show up in T_0: a shift of the maximum value over the temperature.

Another realization of a reference voltage is achieved by feeding the PTAT current into a network formed by two resistors and a third diode Fig. 6.9. Proper tuning
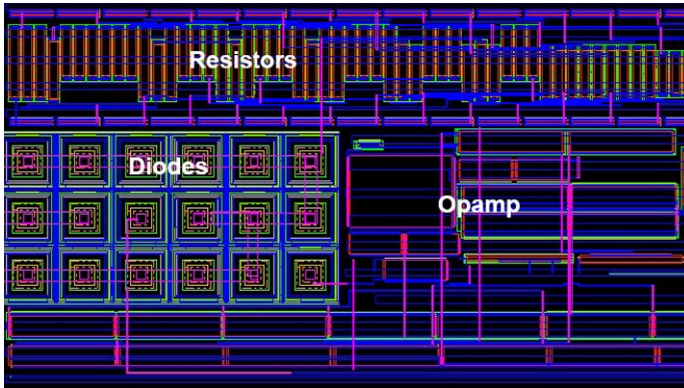
Fig. 6.10   The lay-out of a bandgap reference circuit in a 65-nm CMOS process

Fig. 6.11   A bandgap circuit with stacked PTAT sources [110]
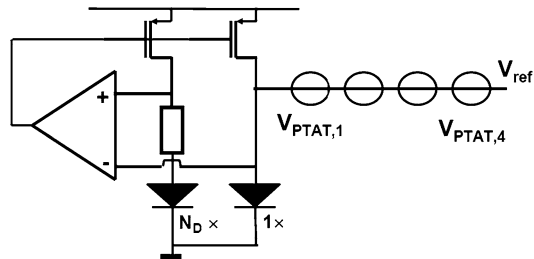
of the resistor values cancels the negative temperature coefficient of the diode at a level of e.g. 800 mV. This circuit was also used to compare a PTAT voltage (controllable via a programmable resistor string) with the diode voltage. In this way a temperature sensor can be constructed. The measurements of the reference voltage as a function of the temperature show the characteristic second order curvature.

Figure 6.10 shows a portion of the lay-out of a bandgap reference circuit. The characteristical array of pnp-transistors and the large resistors and operational amplifier components are visible.

A real avalanche of bandgap circuits for below 1 V operation is found in [109].

6.3   Alternative References

The bandgap of silicon is a stable and reliable physical quantity to serve as the base for a reference voltage. Some other quantities have been proposed to serve as a reference.

- The dominant problem in the conventional bandgap circuit is the linear multiplication of all input referred errors in the generated PTAT voltage. An alternative to this multiplication is to build up the total PTAT voltage from a series connection of several PTAT sources, Fig. 6.11. In stead of a linear multiplication of the error

Fig. 6.12 Two ways of
generating the difference
between two threshold
voltages T_1 and T_2
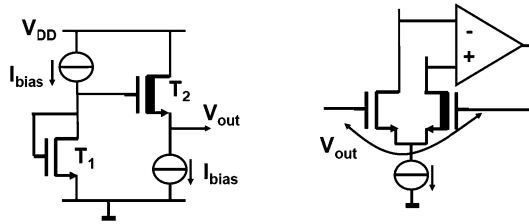
in one PTAT source now the uncorrelated contributions of N sources will add
up to a multiplication factor of $\sqrt{N}$ times. In order to realize the PTAT sources,
mostly transistors in the weak inversion regime are used, but also circuits based
on back-gate connected MOS devices (DTMOS transistor) are reported [110].

- The threshold voltage of the MOS transistor is a potential candidate for designing
  a reference source. Running a current in a diode connected transistor gives:

$$V_{\text{ref}} = V_T + \sqrt{\frac{2I}{\mu C_{ox} W/L}} \qquad (6.6)$$

Both the threshold of the MOS transistor and the mobility show a considerably
temperature coefficient [32]. The NMOS threshold decreases with 1–3 mV/°C,
while the square-root term increases in value due to the decreasing mobility.
There is a sweet spot where both effects compensate. Industrially this is con-
sidered to be not reliable. Therefore some form of controlled compensation is
needed, which is normally applied by making the current temperature dependent
[111].

- The difference between two threshold voltages can be exploited in various ways.
  In the circuit of Fig. 6.12 (left) the W/L ratios and currents are equal. Now the
  output voltage is easily found by subtracting (6.6) for transistor T_2 from T_1. In
  first order the current dependent part will cancel and $V_{\text{out}} = V_{T,T1} - V_{T,T2}$. If T_1
  and T_2 have threshold voltages of the same polarity V_{out} is small and will suffer a
  lot from variations. This circuit is more interesting if opposite threshold voltages
  are available [112].

- The threshold voltage of an EEPROM device can be trimmed during the pro-
  gramming process. This method and derived ideas require however some form of
  factory trimming.

Despite some interesting attempts these principles have never gained much accep-
tance compared to bandgap based references.

Chapter 7
Digital-to-Analog Conversion

Digital-to-analog converters fulfill two important roles in the data conversion chain. A digital-to-analog converter is needed at the end of the chain for converting the digital signal back into the physical domain. Next to that every analog-to-digital converter needs some form of digital-to-analog conversion for its operation. These two functions directly influence the requirements posed on the digital-to-analog conversion. A digital-to-analog converter that has to deliver a signal to the physical world needs to act in the continuous time domain and the signal has to show a high quality at every time moment. Moreover the signal must be delivered at some power level to a load impedance.

In an analog-to-digital converter the value delivered by the digital-to-analog converter is relevant only at a few (perhaps only one) time moments. The performance on other time moments is not critical. Together with a minimum demand on the drive capabilities, the demands of this application on the converter are much less challenging.

These application constraints limit the freedom of the choice of the architecture and the physical domain. In the next section of this chapter, the architectural and physical domain options are analyzed. Then some realizations of a digital-to-analog converter per domain illustrate the combination of these aspects.

7.1 Unary and Binary Representation

A reference quantity forms the basis for the digital to analog conversion. The next step is the subdivision of this reference in fractions that allow the generation of LSBs. Of course it is also possible to generate directly a reference of the size of an LSB and multiply this value. However, this will also multiply the noise and interference. A full-scale reference value reduces these sensitivities and results in the best performance.

The two most commonly used techniques for combining elementary units created by a reference are the unary and binary representation. Figure 7.1 shows both techniques.

Fig. 7.1 Two basic techniques for digital-analog conversion: unary and binary representation. For both representations two signals are shown corresponding to a value of 15 and 16 LSBs

Unary representation uses a series of 2^N identical elements. A unary numerical value is created as:

$$B_u = \sum_{i=0}^{i=2^N-1} b_i = b_0 + b_1 + b_2 + \cdots + b_{2^N-1} \tag{7.1}$$

where each coefficient b_i equals either "0" or "1". The analog equivalent is formed by summing copies of the physical equivalent A_{LSB} of an LSB:

$$A = \sum_{i=0}^{i=2^N-1} b_i A_{\text{LSB},i} = b_0 A_{\text{LSB},0} + b_1 A_{\text{LSB},1} + b_2 A_{\text{LSB},2}$$
$$+ \cdots + b_{2^N-1} A_{\text{LSB},2^N-1} \tag{7.2}$$

Each next higher value consists out of all the previous elements plus one new element. The obvious advantage of this method is that is provides an absolute guarantee on monotonicity (see Fig. 5.5). A practical implementation will consists of 2^N elements (resistors or current sources) attached to an extensive switching matrix. A converter based on unary coding will grow exponentially with N. Till $N = 10 \ldots 12$ unary coded converters will result in a good, and economically usable solution. This technique can be applied to resistor strings, current source arrays, capacitor arrays and in timing (counting conversion).

In order to avoid the exponential growth of components in a unary architecture, the exponential behavior must be included in the representation itself. In a binary structure the elements are chosen such that the resulting voltages or currents form an exponential series.

$$B_b = \sum_{i=0}^{i=N-1} b_i 2^i = b_0 + b_1 2^1 + b_2 2^2 + \cdots + b_{N-1} 2^{N-1} \tag{7.3}$$

with the analog series:

$$A = \sum_{i=0}^{i=N-1} b_i A_{\text{LSB}+i} = b_0 A_{\text{LSB}} + b_1 A_{\text{LSB}+1} + b_2 A_{\text{LSB}+2}$$
$$+ \cdots + b_{N-1} A_{\text{MSB}} \tag{7.4}$$

As a switch has two positions, it is practical (but not necessary) to choose 2 as a base. The example from Fig. 7.1 shows a series of elements with the values:

$A_{LSB}, A_{LSB+1}, \ldots, A_{MSB} = 1, 2, 4, 8$, and 16. In the first binary coded situation, the dark colored elements add up to the value of 15. After an increment, all elements chosen up to then must be switched off and the element of value 16 is switched on (as shown in the second part of Fig. 7.1). The implicit disadvantage of this method is the transition from the sum of elements forming one value to another value, at which many bits flip (e.g. $01111 \rightarrow 10000$). Although most transitions will result in a controlled LSB change, the mid-range transition will cause the highest valued element to switch on and all other elements to switch off. Both values should differ one LSB, however mismatch between the physical quantities A_{LSB}, A_{LSB+1} can easily create deviations. If a higher exponent element is smaller than the sum of the lower exponential elements, a non-monotonicity in the transfer curve will arise. In this case an increment on the digital input code will result in a decrementing analog output. In a control loop instability around such a non-monotonicity may occur and upset the system. Another potential problem with binary coded circuits may arise from the non-linear relation between the number of switched blocks and the output (e.g. for "15" four units are switched on, for "16" one unit, for "17" and "18" two units). Dynamic errors may occur as a result. On the other hand binary coded converters can be designed with low area and power consumption. Several implementations have been reported in literature [113–118].

Both techniques, unary and binary coding, are applied in practical design. In the case of converters with a high resolution the problem with the above schemes is the large number of units involved in a unary design or the wide range of binary values for a binary coded DA converter. Segmentation is generally applied to circumvent these problems: a converter of N-bit resolution is subdivided into a cascade of 2 sub-converters of M and $N - M$ bits. Also partitioning in more than two segments is possible.

For converters with high resolution, segmentation allows combining unary and binary techniques: a 16-bit converter can be build effectively by designing the 6 MSBs in a 64-element unary array each with a value of $2^{10} A_{LSB}$. A 10-bit binary array is coupled to the unary array and will code for the lower 10 LSBs. This set-up assumes that a 10-bit accuracy in a binary architecture can be reached. Examples of segmentation are found in Sects. 7.2.3 and 7.5.1.

Most digital-to-analog (sub)schemes can be classified along the above lines, though there are a few deviating forms such as ternary coding $(+1, 0, -1)$, which is sometimes used in combination with sign/magnitude representation.

7.1.1 Digital Representation

The unary and binary structures from the previous descriptions assume a positive signal. Of course most systems use signals with negative values as well. There are several ways to represent negative signals in a conversion process. The choice how to represent the signal will influence several aspects of the conversion and of the

Table 7.1 Various forms of digital representation

Straight binary		Two's complement		Sign+magnitude		Gray coded	
15	1111	7	0111	7	0111	15	1000
14	1110	6	0110	6	0110	14	1001
13	1101	5	0101	5	0101	13	1011
12	1100	4	0100	4	0100	12	1010
11	1011	3	0011	3	0011	11	1110
10	1010	2	0010	2	0010	10	1111
9	1001	1	0001	1	0001	9	1101
8	1000	0	0000	0	0000	8	1100
7	0111	−1	1111	0	1000	7	0100
6	0110	−2	1110	−1	1001	6	0101
5	0101	−3	1101	−2	1010	5	0111
4	0100	−4	1100	−3	1011	4	0110
3	0011	−5	1011	−4	1100	3	0010
2	0010	−6	1010	−5	1101	2	0011
1	0001	−7	1001	−6	1110	1	0001
0	0000	−8	1000	−7	1111	0	0000

analog and digital processing. Table 7.1 shows in a table various representations of digital signals.[1]

The straight binary code in the first column is well suited for positive signals. Negative signals can be used if the entire scale is shifted by half of the full amplitude. In this "two's-complement" code "1000" is the virtual zero value. In the digital domain now the positive and negative signals can be properly handled. Addition and subtraction can be executed without prior knowledge of the signs of the operands. Multiplication requires the digital numbers to be extended to fit a format corresponding to the result. Positive numbers are extended by zero's and negative values are extended with one's. A direct translation of a "two's complement" code in the analog domain requires two analog power supplies. This is mostly not an economical solution, therefore the code "0000" corresponds in the analog domain with half of the reference value. Now small signals in the analog domain move around half of the reference value. This costs power and noise and other artifacts associated with half of the reference value deteriorate the signal-to-noise ratio. E.g. in current digital-to-analog converters, half of the reference value corresponds to half of the maximum current. Thereby a zero-valued signal will show thermal and $1/f$ noise which is imperative to this current. Obtaining a good signal-to-noise ratio for small signal is made difficult by the choice for two's complement representation.

The "sign and magnitude" code is linked to the analog hardware. The MSB of the code is the sign and the remaining part of the code is the amplitude of the signal.

---

[1]Many more representations exist, this table only lists the ones used in this book.
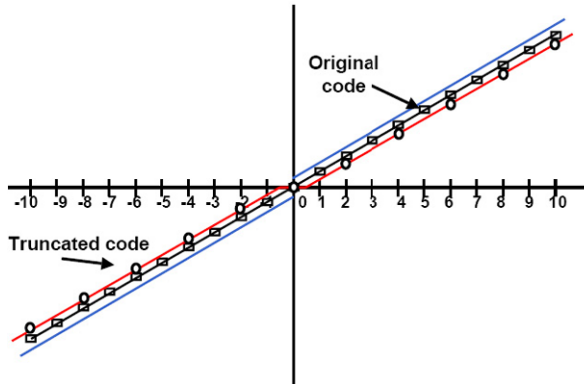
Fig. 7.2   Basic truncation or rounding creates distortion near the zero-code

From a circuit design point of view the MSB signal can directly be used to switch e.g. a polarity switch in a digital-to-analog converter. The amplitude component of a "sign and magnitude" code is straight binary. Consequently this kind of code allows implementations that avoid problems with noisy half-reference values. In the digital domain this code is less pleasant: a circuit will be needed to decode this signal in a format that can be processed.

If a "sign-and-magnitude" signal is rounded or truncated in the digital domain with simple rounding or truncation rules an error will occur, Fig. 7.2. In the case of rounding or truncation for a "straight binary" or "two's-complement" signal both truncations of positive and negative numbers will result in a shift in the same direction. For "sign-and-magnitude" signals the positive and the negative part of the signal will reduce in amplitude and shift towards zero. Straight forward rounding or truncation will have a cross-over problem near zero and a distortion component.

7.1.2 Physical Domain

In the physical domain the output value of a digital-to-analog converter can be formed using voltages, currents, charges or time. In each of these physical or analog domains both unary and binary architectures can be used, see Table 7.2.

Voltages can be subdivided by means of resistors. The upper left scheme of Fig. 7.3 shows the concept: a digital decoder selects one of the switches that determines the output voltage. In a similar manner a row of current sources and switches implement a unary current source digital-to-analog converter. Converters operating in the charge domain use capacitor banks and unary implementations in the time domain use pulse trains, that switch on or off a physical unit.

Creating exponential sequences of physical quantities is less simple. In the voltage domain "R-2R" structures are applied (see Sect. 7.2.2), while in the current domain currents can be split by means of transistor pairs. Capacitor arrays can use

Table 7.2 Various forms of analog representation and physical domains. In italic the main application area is indicated

	Unary	Binary
Voltage	Resistor string	R-2R
	Flash ADC	*Low-performance DAC*
Current	Current matrix	Current splitting
	High bandwidth DAC	
Charge/capacitor	Capacitor bank	Capacitor bank
	Low power DAC	
Time	PWM, SD mod	Limited by distortion
	Low bandwidth DAC	
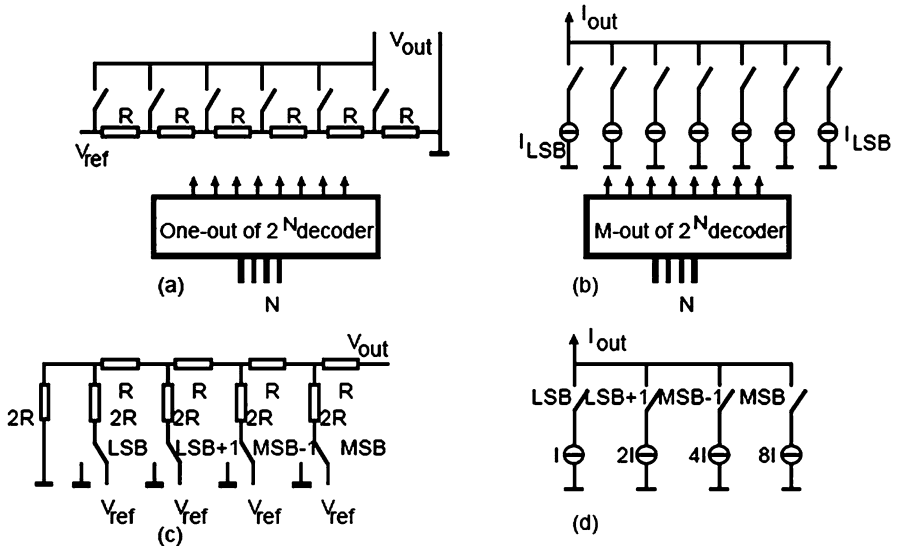
Fig. 7.3 Unary (*upper schemes*) and binary forms of resistor string and current source digital-to-analog conversion.

the R-2R principle as well. Using pulses of exponentially increasing lengths in the time domain is realizable, however, it is unclear what advantage that can bring.

Next to the combination of signal representation and physical domains Table 7.2 shows the major application area. Except for binary weighted timing all principles find usage.

Fig. 7.4   A digital-to analog
converter based on a resistive
divider
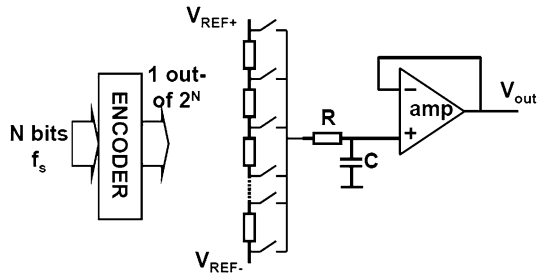
Fig. 7.5   The impedance of a
resistor string various with
the position
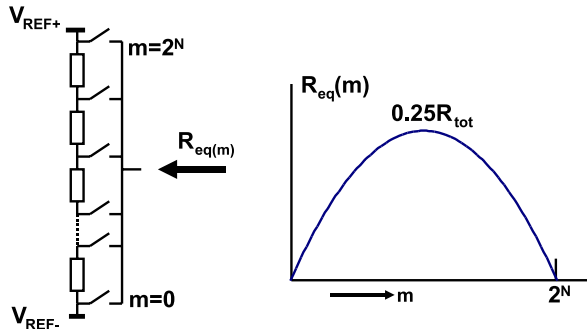
7.2 Digital-to-Analog Conversion Schemes

7.2.1 DA Conversion in the Voltage Domain

A practical example of a unary digital-to-analog converter in the voltage domain
consists of a series connection of resistors between two reference voltages see
Fig. 7.4. These structures are called resistor ladders or resistor strings. A resistor
ladder can be turned into a digital-to-analog converter if a selection and switching
network is connected to the resistor ladder taps. A buffer is needed to lower the
output impedance.

   An important problem in this resistor string is the variation of the impedance of
the resistor ladder: on both ends the impedance is equal to the impedance of the
reference source and close to zero. At a code close to the middle of the ladder the
impedance is equal to the parallel connection of two half-ladders. If $m = 0 \ldots M$ is
the position of the nodes in a ladder with $M = 2^N$ resistors, the impedance on each
node is:

$$R_{eq}(m) = \frac{\frac{m}{M}R_{tot} \times \frac{M-m}{M}R_{tot}}{\frac{m}{M}R_{tot} + \frac{M-m}{M}R_{tot}} = \frac{m(M-m)}{M}R_{tot} \tag{7.5}$$

Figure 7.5 shows the parabolic behavior of the effective impedance as a function of
the position, with a maximum in the middle. The time constant formed by the ladder
impedance and the loading capacitor will be position and signal-value dependent.
High frequency signals will show distortion.

The impedance variation requires buffering the output of the resistor string. The remaining time constant of the resistor string and the capacitance at the input of the buffer must be kept as low as possible to avoid the code-dependent distortion.

The resistor string itself will show capacitive loading. This capacitive loading is distributed over the ladder. Exitations by spurious voltages or charges will cause a settling behavior. In order to facilitate an analysis the resistor string is modeled as a continuous resistive strip, where the capacitance and resistance per unit length are defined by r in Ω/m and c in F/m. The voltage over this strip with length L is described by the diffusion equation. This equation is also referred to as "heat equation"[2] and describes in classical thermodynamic theory the evolution of the temperature as a function of time and position $\text{Temperature}(x, t)$:

$$\frac{\partial \text{Temperature}(x, t)}{\partial t} = D \frac{\partial^2 \text{Temperature}(x, t)}{\partial x^2} \tag{7.6}$$

where D is the thermal diffusion constant. Transferring this equation to the voltage domain with the function $v(x, t)$ describing the voltage in time and position over the resistive structure:

$$rc\frac{\partial v(x, t)}{\partial t} = \frac{\partial^2 v(x, t)}{\partial x^2} \tag{7.7}$$

With the boundary conditions at $v(x, 0) = v_{\text{start}}(x)$ and at $v(0, t) = v(L, t) = 0$ an exact solution can be obtained of the form:

$$v(x, t) = \sum_{i=1}^{\infty} e^{\frac{-i^2\pi^2}{rcL^2}t} a_i \sin\left(\frac{i\pi x}{L}\right) \tag{7.8}$$

The solution is orthogonal for time and position, both are described by separate functions. The start condition is brought into the equation by solving the equation for $t = 0$:

$$v(x, 0) = v_{\text{start}}(x) = \sum_{i=1}^{\infty} a_i \sin\left(\frac{i\pi x}{L}\right) \tag{7.9}$$

The terms a_i with the sin function are a Fourier description in a one-dimensional direction.

The initial condition that is defined by $v_{\text{start}}(x)$, will exponentially decay. The decay behavior is dominated by the first term:

$$v(x_m, t) \approx v_{\text{start}}(x_m) e^{\frac{-\pi^2}{rcL^2}t} \tag{7.10}$$

with a time constant: $\tau = rcL^2/\pi^2$.

Looking at this problem from an engineering point of view, the starting observation is that the largest errors due to the distributed delay line will occur in the middle of the structure. From this point the resistive impedance towards the ends is

---

[2]It is convenient to look in literature for solutions of the "Heat equation" problem with your specific boundary conditions and rewrite them to voltage equations.
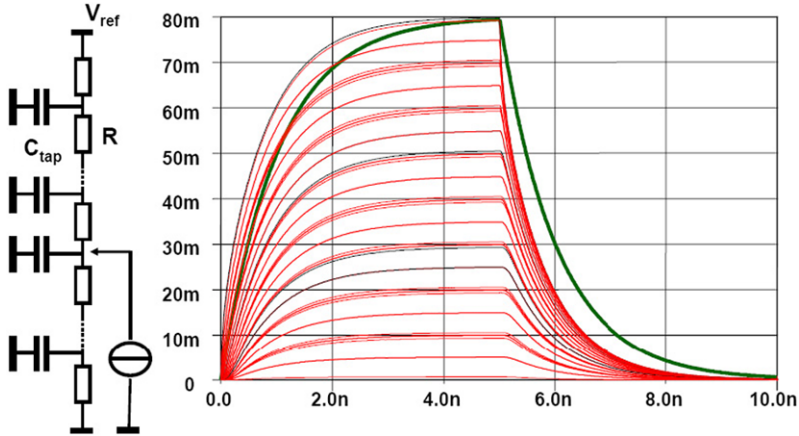
Fig. 7.6 In a ladder string with 256 resistors of 1.25 Ω and $C_{tap} = 0.1$ pF each, a current is injected at the middle tap and switched off after 5 nanoseconds. The *thin-line* plots show the time behavior at various nodes in the ladder. The *bold line* is the lumped RC approximation with $\tau = R_{tot}C_{tot}/8 = 80 \ \Omega \times 12.8$ pF
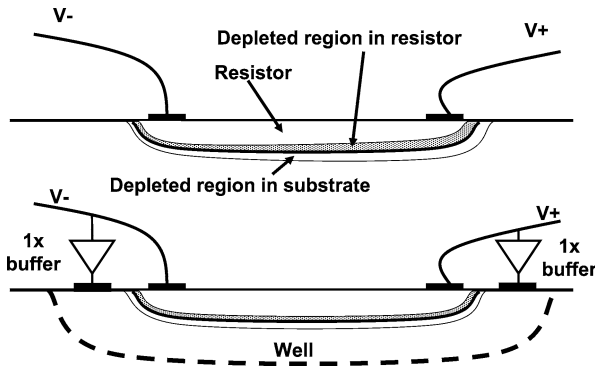
Fig. 7.7 Due to the voltage difference over a diffused resistor, the depth of the active resistor is less at higher voltages. Placing the resistor in a well of opposite dope and biasing the well with the same voltage difference avoids the voltage dependence

$R_{tot}/4 = rL/4$. The capacitors at the intermediate nodes need to be charged. The capacitors close to the middle tap must be fully charged, and capacitors closer to the ends will be charged proportional to their position. On average the total capacitance is charged to half the value of the middle tap. That totals to a time constant on the middle tap of: $\tau = R_{tot}C_{tot}/8 = rcL^2/8$, which is unequal to the solution of the heat equation, but close enough for a first order guess. Figure 7.6 compares the approximation with the distributed line solution.

The actual design in the lay-out of a resistor ladder requires to consider various aspects. The available resistive material must show low random variation, zero-voltage coefficient and low parasitic capacitance. For high-speed conversion
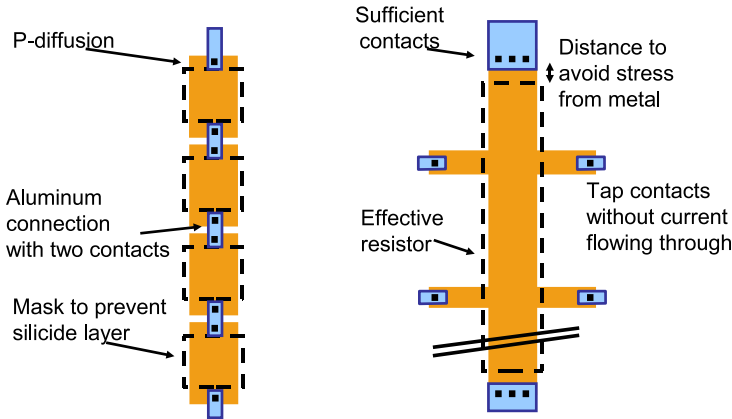
Fig. 7.8 On the *left-hand side* is a ladder build up with discrete resistors. On the *right-hand side* a ladder structure is shown for high accuracy, see Sect. 11.3

(>100 Ms/s) the total parasitic capacitance forces to choose extremely low values for the tap resistances. Realizing tap resistances in the order of magnitude of 1 Ω poses considerable problems and requires to use special material layers on top of the chip.

In sub micron CMOS processes a polysilicon layer has less parasitic capacitance and a lower temperature coefficient, see Table 2.15. A diffused layer gives a better matching performance, see Table 11.8. The voltage dependency of the diffused resistor in Fig. 7.7 is canceled by placing the resistor in a well of opposite doping and biasing this well with the same voltage difference.

Figure 7.8 shows two lay-out examples. The left-hand construction uses a fixed resistor lay-out connected by wiring. This is not an optimum construction. The current has to pass through contacts and contact areas show a lot of additional variation in resistance. On the right-hand side a construction is shown where the tap voltage connections do not carry current. Any variation in position, resistivity of the contact etc is not relevant. The main connections to the ladder are designed with sufficient contacts and are placed at some distance to reduce the material stress caused by the presence of aluminum wires, see Fig. 11.14. Preferably a few dummy taps are inserted between the main connections and the first relevant taps.

This construction cancels a number of effects, such as contact hole resistance. For large ladder structures also attention must be paid to gradients. Due to processing or heat sources the resistivity of the ladder material is not constant with distance. Cross-coupling, as shown in Fig. 7.9 eliminates gradients. In this structure the second ladder is connected upside-down to the first ladder. Only the two extreme connections carry (large) currents. The currents in the intermediate connections are ideally zero, but will never be large. See Sects. 11.3 and 11.4 for some more background.

Resistor ladders are crucial as building blocks in flash analog-to-digital converters. When a resistor ladder converter is applied as an output device for driving an

Fig. 7.9 A cross-coupled ladder structure
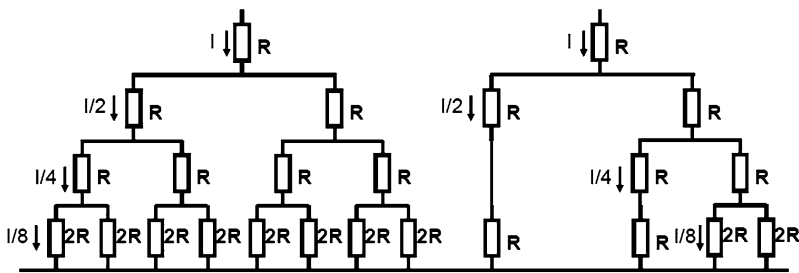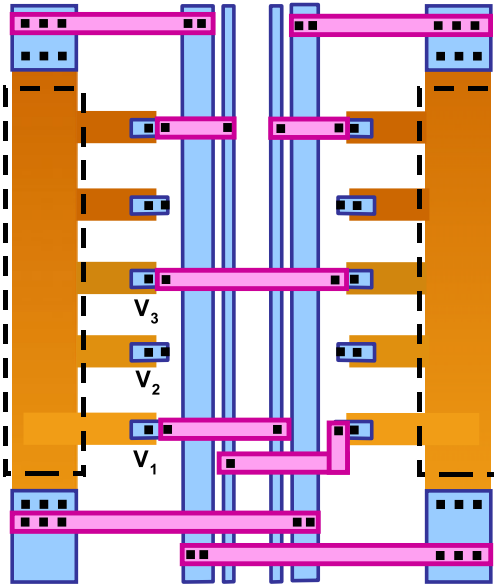
Fig. 7.10 Currents can be split in binary weighted portions (*left*). The constant impedance in each branch allows to reduce the branches of the structure

application, a careful design of the buffer is required, see Sect. 7.5.1. This topology allows an excellent static performance at a limited power level. The speed is limited by the performance of the buffer.

7.2.2 R-2R Ladders

Figure 7.10 shows on the left side a binary tree structure of resistors with value R. This tree is terminated with resistors values $2R$. In every layer twice the number of currents flow of half of the value of the layer above. Moreover from each node looking downwards the impedance equals R. This property allows to replace in Fig. 7.10 (right hand side) the branches by resistors R and thereby constructing the

Fig. 7.11 R-2R
digital-analog converter with
digital encoding
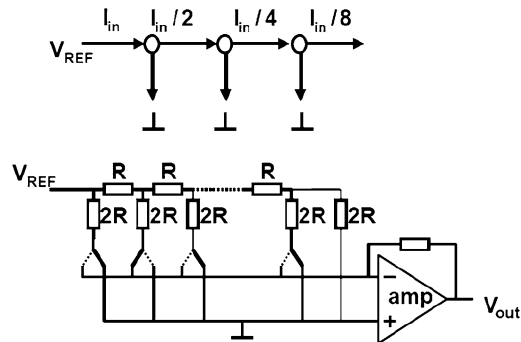
R-2R structure.[3] Figure 7.11 shows a more abstract binary coded digital-to-analog converter based on the "R-2R" principle. Current entering the R-2R resistor circuit, splits at every node in two equal parts, because the impedance in both directions equals 2R. The combination of branches therefore generate a power-of-two series of currents, that can be combined via switches into an output current.

An analysis of the "R-2R" ladder in Fig. 7.11 is simple, provided the analysis starts from the LSB side. The resistor network is terminated with a resistance 2R. This resistance is put in parallel to an equally sized value to generate the LSB current. The combined value of the two resistors is R. If the resistor in series is added, the total impedance of the section is back to the original 2R. This impedance is connected in parallel with the 2R resistor for the LSB-1 and can again be combined to yield an R value. In this example a buffer is used to convert the current from the R-2R ladder into a suited output format. The bandwidth of the buffer is limiting the overall bandwidth of this converter. This technique allows to design digital-to-analog converters of a reasonable quality (8 to 10 bits) at low power consumption and low area. Moreover the digital coding can be directly taken from a straight binary representation.

The main accuracy problem in a R-2R digital-to-analog converter is the inequality in the splitting process. If the switch in the first vertical branch has a 1% larger resistance, the current will split in 0.495 and 0.505 portions, limiting the achievable resolution (with DNL < 1) to 7 bit. The application is limited to low-resolution low-cost applications such as offset correction.

7.2.3 Digital-to-Analog Conversion in the Current Domain

Figure 7.12 shows the block diagram of unary and binary coded digital-to-analog converters in the current domain.

The buffer in Fig. 7.12 provides a low-impedance load for the current sources, avoiding modulation of the currents due to their finite output impedance. Major

---

[3]Explanation from Colin Lyden (ADI).

Fig. 7.12 Digital-analog converter with current sources: unary and binary forms
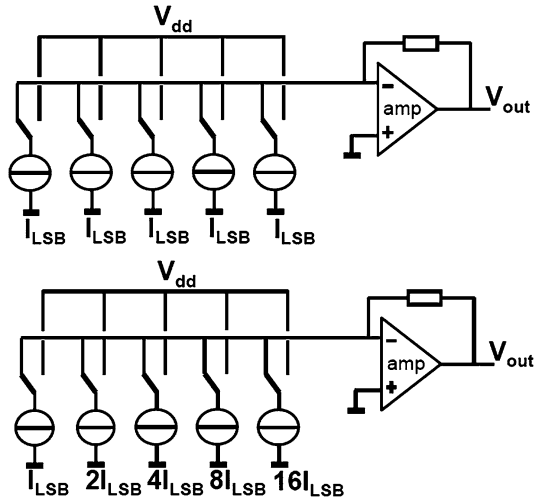
disadvantage of this arrangement is the feedback stabilization of the opamp used for the buffer. If the buffer has to drive a load consisting of resistive and capacitive elements, the unity gain of the buffer has to be designed at a frequency lower than the dominant pole of the output load in order to avoid output ringing. Consequently the overall bandwidth is degraded. When the choice for a buffer is made, a capacitor array or resistor ladder solution is preferred.

The block diagram in Fig. 7.13 shows a digital-to-analog converter based on current sources without a bandwidth-limiting buffer. The current sources are directly feeding the load impedance. This arrangement is only speed limited by the pole of the output impedance. This topology is suited for delivering a high-performance time-continuous signal into a 50 to 75 Ω loads. This architecture is mostly called a "current-steering digital-to-analog converter".

These converters combine unary current sources for the MSB values, thereby ensuring good DNL and binary coded current sources for the lower bits. The total area is kept rather small.

Figure 7.14 shows the basic current source schematic. The current source transistor is DC-biased and its size is chosen in a way to reduce mismatch effects and noise contributions (e.g. long transistor length). The switch transistors are optimized for fast switching and are controlled by differential pulses.

The switches in current-source digital-to-analog converters complicate the design. In contrast to the resistor ladder, where only one switch is active, in a current matrix a number of switches must be toggled if the signal changes from one value to another value. Each switch (MOS or bipolar) needs some charge to create a conducting path. This charge is taken from the current source and the signal during switching on, and is released into the output signal during switching off. If a current-source digital-to-analog converter is used as an output device for driving an application this charge disturbance distorts the output signal and must be minimized. This disturbance is called a "glitch" and is harmful as this converter is used
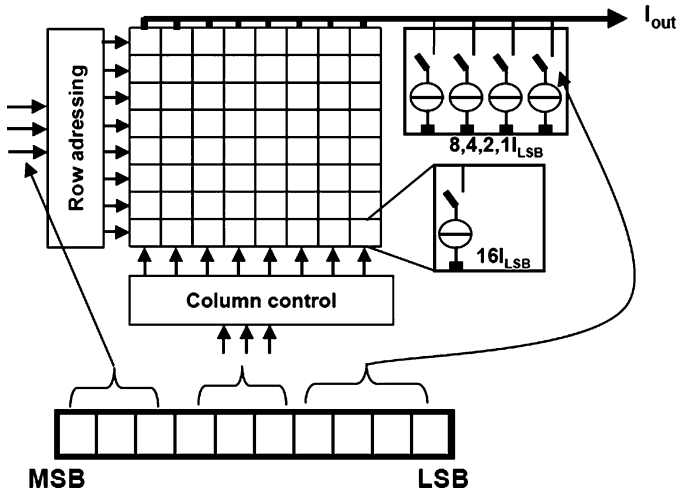
Fig. 7.13 10-bit digital-analog converter: the 6 MSBs are implemented as 64 unary current sources in a matrix configuration, while the four LSBs are designed in a binary series
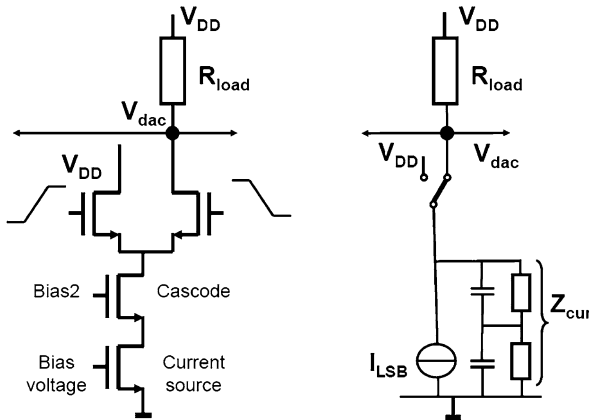
Fig. 7.14 The basic digital-analog converter current cell and its small-signal equivalent circuit

to produce a time-continuous output signal. The glitch produces a voltage excursion over a certain time period. The associated area under the glitch is called the "glitch energy" and often expressed in picoseconds-Volt (psV). Careful design of the switches is necessary. The voltage swing on the switches must be kept at a minimum and the clock edges must be strictly timed in order to limit the glitch energy. Other methods try to keep the glitches constant for all codes.

One of the main issues in this architecture is the modulation of the current sources by the output voltage. Depending on the effective internal impedance, this voltage variation will result in a current modulation. When the digital signal or the fraction

of current sources that is switched on, is $0 \leq \alpha \leq 1$, the ideal output voltage at a sample moment becomes:

$$V_{DD} - V_{dac} = \alpha 2^N I_{LSB} R_{load} \tag{7.11}$$

With N as the resolution of the unary part of the digital-to-analog converter. A finite output impedance of each active current source Z_{cur}, as shown in the right-hand side of Fig. 7.14, causes an additional current that is added to the digital-to-analog converter's output current. The total error current flowing into the finite output impedance is

$$I_{err} = \alpha 2^N V_{dac}/Z_{cur} \tag{7.12}$$

Both the fraction of active current sources α and the output voltage V_{dac} are proportional to the signal. This results in a second-order distortion term in the current:

$$V_{DD} - V_{dac} = R_{load}\left(\alpha 2^N I_{LSB} - \frac{\alpha 2^N V_{dac}}{Z_{cur}}\right) \tag{7.13}$$

Substitution of a full-swing output signal $\alpha = 0.5 + 0.5\sin(\omega t)$ results after some manipulation in:

$$HD2 = \frac{\text{second order component}}{\text{first order component}} = \frac{2^N R_{load}}{4Z_{cur}} \tag{7.14}$$

The second order distortion is directly related to the ratio of the output impedance of the current sources to the load impedance. In case of $N = 10$, $R_{load} = 50\ \Omega$ and a desired -60 dB level of the second order distortion HD2, the effective impedance of a single current source must be better than 12.8 MΩ.

Next to the resistive modulation Z_{cur} also can contain capacitors, such as the parasitic capacitor on the drain of the current source transistor. The capacitive current will increase with higher frequencies and so will the second order distortion. This effect and the time deviations of the switching pulses are the root cause for performance loss at higher signal frequencies.

If the switches are used in their conductive mode as cascode stages, the output voltage modulation has less effect. Also dedicated cascode stages can be used for this purpose. High performance devices use a differential output scheme.

A 10-bit converter can be built from 64 unary current sources and 4 binary current sources, which requires considerably less area than 1024 sources for a full unary implementation. The choice between unary and binary block size depends on technology and required DNL performance [119–121]. In high-resolution converters the advantage of using binary coding becomes marginal. If the accuracy of each current source is limited by random effects that are inversely related to the area, equation (2.95), the total amount of gate area for a certain resolution will be the same for both architectures. Secondary arguments such as wiring overhead will decide.

The digital decoding of the input word into current source control signals requires a lot of attention. A popular arrangement is a column-row addressing scheme resembling a Random-Access Memory, as indicated in Fig. 7.13. The current sources that
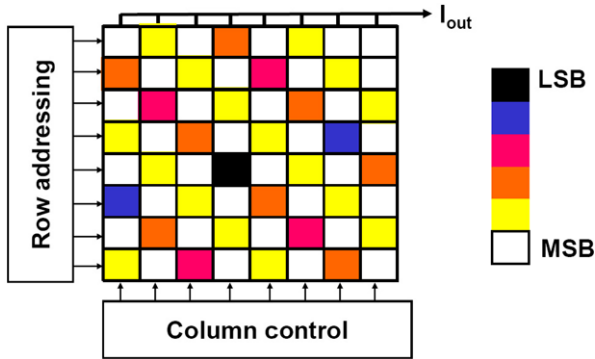
Fig. 7.15 Placement of unary weighted current sources in a 64-element array. Similar colored squares represent current sources that form together one bit level [122]

form the unary part of the digital-to-analog converter are placed in the lay-out in a way that cancels gradients [116, 122]. These gradients can occur due to technological deviations such as doping or oxide thickness variations, power supply drop due to voltage drop over current carrying lines, clock timing gradients, temperature gradients. In practical design the technological gradients are rather limited in magnitude. Especially the voltage, time [123] and temperature gradients can become rather severe. Figure 7.15 shows a simple solution to the first order gradient problem in an array where the current sources are connected in a binary way: the common centroid topology. Each group of current sources, forming the LSB up to the MSB, is arranged symmetrically around the center in both lateral dimensions. This will cancel any linear gradient. More advanced schemes are known as Q^2-walk schemes [124]. These schemes use further subdivision of the current sources. Groups of sub-current sources are arranged to compensated for second order components. Another approach is to randomize the current sources from one sample to the next [125]. Other points to consider with matrices of current sources are:

- The lines that are routed over the current-source cells can interfere with these cells due to capacitive coupling, but also in a technological way. The metal covers the current source transistors and affects the annealing of the underlying gate structure which causes mismatch, Sect. 11.3.
- The decoding network consumes power and increases cross-talk.
- Close to the final switches the selection signals must be re-timed by the main clock signal. Delays between the switching-on and off of various current sources create unequal $\int i\,dt$ contributions to the output signal. Proper distribution of the clock signal with equal traveling distances for each cell (1 mm wire equals 10 ps) is needed.

These converters do not need a speed-limiting buffer and can reach high operating frequencies. The speed constraint at the output of these converters is limited by the RC time constant of the output load, which also serves as a first alias filter. Next to that the speed of the digital processing and the timing inaccuracy limit the

Fig. 7.16 Die photograph of a 16-bit current-steering digital-to-analog converter in 180-nm CMOS. *Courtesy*: J. Briaire

Fig. 7.17 Spurious-free dynamic range versus bandwidth for current-steering digital-to-analog converters published from 2000–2007. *Courtesy*: J. Briaire

performance. The penalty for these advantages is power. When a current is not contributing to the output current, it cannot be switched off. This would inevitably lead to a discharge of the inversion layers in the (rather large) current source transistors and the parasitic capacitors. Building up this charge after reconnection takes a lot of time and will lead to linear and non-linear distortion, therefore the unused current must be drained in a power rail. Consequently these converters always consume the maximum current needed. For a single terminated 50 Ω cable this leads to 20 mA per converter. Next to that clock and decoding schemes require a significant amount of power. Alternatively the drained current is used to implement a differential output that will cancel the second order distortion component. An advantage of the constant current consumption is that the impact of power wiring impedances and bond-wire inductance is less. Figure 7.16 shows a die photograph of a current-steering digital-to-analog converter.

Figure 7.17 compares various data from publications [120, 123, 125–127] and data sheets on current-steering digital-to-analog converters in the time frame 2000–2009. The plot suggests a first-order relation between the bandwidth and the level of spurious and harmonics.

Fig. 7.18 A semi-digital filter and digital-analog converter [128]

Current-steering digital-to-analog converters are the industry's primary choice for output devices in demanding applications. Their dynamic performance is unmatched and the required high power is accepted.

7.2.4 Semi-digital Filter/Converters

After the conversion from the digital to the analog domain with a standard zero-order hold function, the alias bands contain a lot of energy. This energy must be removed before the signal can be used in applications.[4] Su and Wooley [128] have proposed a digital-to-analog structure which converts the digital signal back into its analog form and (at the same time) filters the result. This semi-digital filter in Fig. 7.18 uses a digital shift register were every position of the shift register controls a digital-to-analog converter. The outputs of these converters are weighted and summed. A simple implementation uses a one-bit code that will switch on or off a weighted current source. With the help of Sect. 3.2.1 the result is written as:

$$I_{\text{out}}(nT_s) = \sum_{k=0}^{k=K-1} a_k D_{\text{in}}((n-k)T_s) \tag{7.15}$$

The coefficients are chosen using similar constraints as for a normal FIR filter.

The advantage of this structure is the suppression of the alias components. Moreover, a voltage signal will only appear after the current summation. So higher order signals components can be effectively suppressed in the current domain and do not generate distortion when they appear over non-linear components. It is important to realize that errors in the coefficients in first order will affect the filter characteristics and not the overall linearity.

This structure is useful in the reconstruction of delta-modulated signals [129].

---

[4]Think of all the energy your tweeter loudspeakers would have to consume.

Fig. 7.19 Digital-analog converter based on unary capacitors
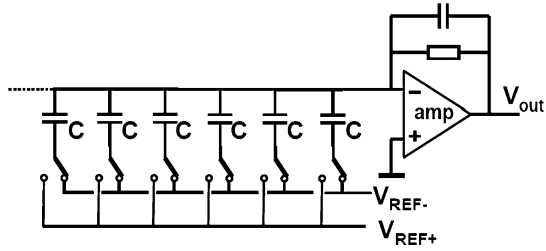
Fig. 7.20 Binary-weighted digital-to-analog converter in a stray-insensitive switched capacitor configuration
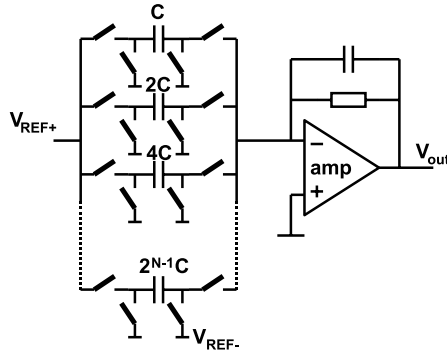
7.2.5 DA Conversion in the Charge Domain

Charge-domain converters differ from voltage-domain converters because the charge is now the information carrying quantity. Even in the presence of perfectly linear capacitors, this distinction is still relevant, as e.g. offset-voltages can have different effects in the charge or voltage domains.

In a first order approach in a digital-to-analog converter the resistor circuit is replaced by switched capacitors [130], see Fig. 7.19. Both unary and binary conversions can be realized, depending on the choice of capacitor values. An example of a full-differential binary weighted implementation is found in [131, Fig. 9]. A unary decoding scheme with equal capacitors and switches is used to implement the conversion.

The configuration in Fig. 7.19 suffers from parasitics that are attached to the top plates of the capacitors, e.g. the capacitance of the switching transistor. A stray-insensitive switching topology uses a standard switched capacitor technique to move charges. The parasitic capacitances connected to the switched capacitors in Fig. 7.20 are either charged from voltage sources or do not see any charge change because they are connected to the (virtual) ground. In Fig. 7.20 the unit capacitors are grouped in a binary ascending scheme: $2^0C, 2^1C, 2^2C, \ldots, 2^{N-1}C$. Only one switch per group is necessary to implement a binary coded converter. This topology principally suffers from the inequality of the capacitor banks at higher resolutions. However with capacitors that can achieve matching performance that allows 12–14 bit accuracy, the more practical limitation is in the exponential growth of the capacitor banks.
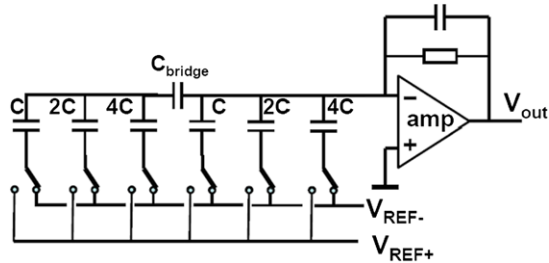
The span of capacitance values in digital-to-analog converters based on binary capacitor arrays requires some trade-off. In some situations the choice for the unit capacitor and the lowest capacitor value is determined by the kT/C noise. However also the technological realization of small capacitors can be an issue: it is difficult to fabricate accurately a 1 fF capacitor. In that case a bridging scheme can be used, see Fig. 7.21. In this example the three MSB bits are formed in a conventional binary fashion. The three LSB bits use a bank with the same size capacitors, but are coupled via a bridging capacitor C_{bridge} to the MSB side. If the capacitance of a unit cell equals C and on the LSB side a total amount of $i(=0,\ldots,2^k-1)$ unit cells are switched from negative reference to positive reference and $i2^k-1-i$ unit cells remain connected to the negative reference, the voltage change on the left hand side of C_{bridge} is

$$\frac{iC(V_{\text{ref}+}-V_{\text{ref}-})}{C(2^k-1)+C_{\text{bridge}}} \tag{7.16}$$

This charge injection in the integration capacitor is compared to the desired charge injection in the summation node of the opamp:

$$\frac{C_{\text{bridge}}}{C_{\text{bridge}}+(2^k-1)C}iC(V_{\text{ref}+}-V_{\text{ref}-})=\frac{i}{2^k}C(V_{\text{ref}+}-V_{\text{ref}-}) \tag{7.17}$$

From this equation $C_{\text{bridge}}=C$ is found.

In first instance the capacitive schemes seem a one-to-one copy of the resistor schemes in Fig. 7.4. Yet there are some important differences. In the charge domain no constant current flow is required, except for the buffer. So a good power efficiency is possible. A second difference is in the way jitter influences the output. If time uncertainty occurs in the switching moments, the transfer of charge will be a bit early or late, however the total magnitude of the packet remains intact. In a voltage or current domain digital-to-analog converter, the overall packet consists of the time period multiplied by the current or voltage amplitude. The jitter thereby changes the signal and jitter is directly translated into noise.

Translation of current and voltage domain converters to the charge domain is possible. The charm of the charge domain is, however, in the observation that storage is for free. A large number of interesting algorithms is possible. A basic charge redistribution digital-to-analog converter is shown in Fig. 7.22. This converter operates on a sequential binary principle: every bit is evaluated successively. After the reset switch has discharged capacitor C_2, the sequence can start. First the LSB value
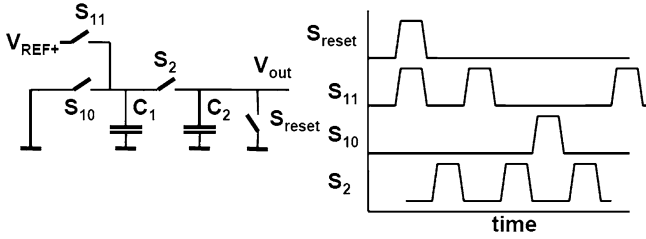
Fig. 7.22 A charge redistribution converter [130]

decides whether C_1 is charged to the positive or negative reference. Then switch S_2 connects the capacitors in parallel, redistributing the charge from the reference over both capacitors. If these capacitors are equal, half of the charge is in either. Now the LSB+1 bit is used to charge C_1, thereby destroying the remaining charge. The sequence continues with a charge redistribution, where the new charge is added to the previous charge and halved.

$$V_{\text{out}}(i) = b_i \frac{C_1}{C_1 + C_2} V_{\text{ref}} + \frac{C_2}{C_1 + C_2} V_{\text{out}}(i-1) \tag{7.18}$$

where $b_i \in \{0, 1\}$, with $i = 0 \dots (N-1)$. If both capacitors are equal this principle is only limited by the accumulated kT/C noise and will result after N sequences in an output value of:

$$V_{\text{out}}(N) = \sum_{i=0}^{i=N-1} \frac{b_i 2^i}{2^N} V_{\text{ref}} \tag{7.19}$$

Using $\delta = (C_2 - C_1)/(C_1 + C_2)$ to describe the error between both capacitors, an estimate of the overall error in V_{out} can be made:

$$\delta V_{\text{out}}(N) = \sum_{i=0}^{i=N-1} \frac{b_i 2^i \delta^{N-1-i}}{2^N} V_{\text{ref}} < \delta V_{\text{ref}} \tag{7.20}$$

Assuming that the error at the MSB transition must remain under the value of an LSB, sets a limit for the capacitor deviation:

$$\delta = \frac{C_2 - C_1}{C_1 + C_2} < 2^{-N} \tag{7.21}$$

An error of 0.1% in the capacitors limits the achievable resolution to $N = -^2\log(0.001) \approx 10$ bits.

The reproducibility of the capacitor ratio is of crucial importance in all switched-capacitor designs. Special care has to be taken to create capacitor lay-outs with optimum reproducibility properties. Figure 7.23 shows a few different design styles. The first implementation is used in processes with two poly-silicon layers separated by a high quality dielectric layer of silicon oxide (double-poly processes). In this case a top contact connects the second layer that does not cross the underlying first layer. Some process flows do not allow the second layer to extend over the first
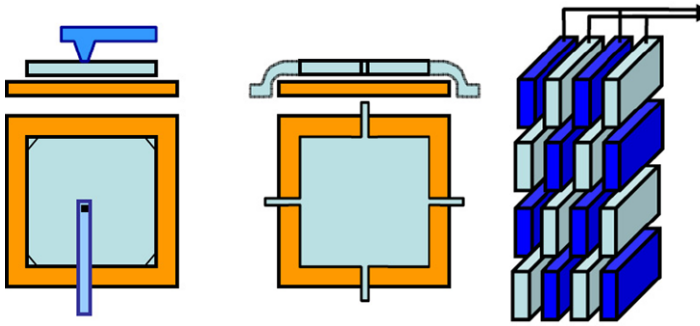
Fig. 7.23 Three capacitor implementations. The *left-side* configuration uses the fact that a top contact is allowed in some processes. The *middle* lay-out connects on the sides whenever top–contacts are not permitted. The *right* fringe-capacitor configuration is used in high-metal density configurations

layer. The process first deposits the two poly-silicon layers and after that patterns the second and the first layer. This procedure creates a high quality capacitor, but reduces the lay-out freedom. The corners of the structures are slanted to prevent etching issues of the corners.

Often a top-contact is not allowed. The aluminum in the contact hole can potentially penetrate into the polysilicon and even through the dielectric layer. In the middle lay-out the second layer is allowed to extend over the first and is used to connect with minimum width stripes. The 2-layer capacitors are not symmetrical with respect to their terminals. It is crucial to consider where to connect the lower plate of the capacitor. The parasitic capacitor on this node can disturb any capacitor ratio, but can also create an unwanted coupling path to a noisy substrate terminal.

In advanced processes with many interconnect layers, the sides of the wires are so high and closely spaced that the lateral capacitance can be used for designing high quality "fringe capacitors". The combination of flat surfaces due to the chemical-mechanical polishing (CMP) process steps and the 8–10 layers of metal lead to a capacitor with excellent properties.

For large arrays of capacitors gradient reduction techniques similar to Fig. 7.15 are applied, see Table 2.18 for some data.

These switched-capacitor digital-to-analog converters are mostly applied in larger system chips, where they perform low-power conversions.

7.2.6 DA Conversion in the Time Domain

Besides subdivision in the voltage, current or charge domain, digital-to-analog conversion can also be realized by means of time division. The signal information is contained in the succession of switching moments, see Fig. 7.24. One of the first techniques to digitize information was based on the succession of pulses [91] and
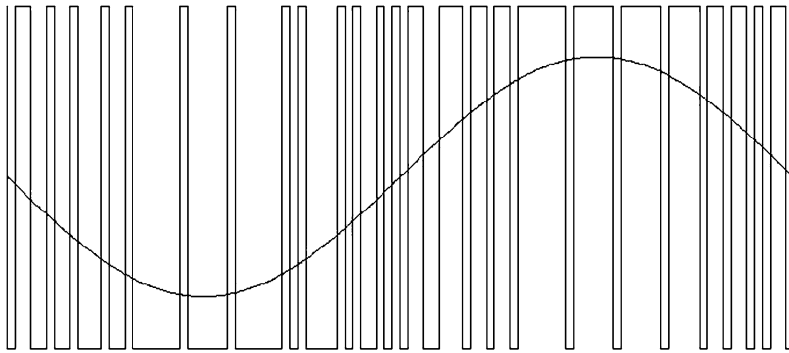
Fig. 7.24   The low-frequency content of a pulse sequence contains the signal
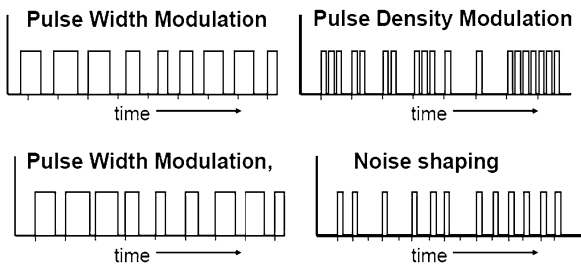
Fig. 7.25   Several forms of time-domain digital-to-analog conversion

was originally called Pulse Code Modulation (PCM). This specific implementation is today referred to as Pulse Density Modulation.

All time division schemes have in common that the output switches between a few (two) levels of a physical medium (voltage or current). Any DC-variation on these levels will manifest itself as a gain factor and will not affect the conversion quality. The linearity problems due to component inaccuracies are circumvented. Pulse Width Modulation or Pulse Density Modulation only uses two levels to create a fast switching pulse sequence. The high and low time of the digital pulse train is ordered in such a way that the low-frequency component of the signal is correctly represented. In a simple form a pulse-width or pulse density modulation (PWM and PDM) can be converted into the analog domain by filtering the pulses in a suitable low-pass filter.

Figure 7.25 shows four differently coded sequences. On the left side are the Pulse Width Modulation formats. The amplitude of the signal is proportional to the width of the pulse. These pulses can be synchronized to the sample rate or free-running. PWM signals are not quantized in amplitude: the pulse width is proportional to the original amplitude. On the right-hand side pulse-density modulated signals are shown. In PDM the amplitude is quantized. Pulse density avoids issues with distortion due to asymmetries in falling and rising edges by using just one type of

Fig. 7.26 Comparison of jitter in normal sampling (*dotted line*) and in one-bit signals where a moderate oversampling ratio of 5 is used
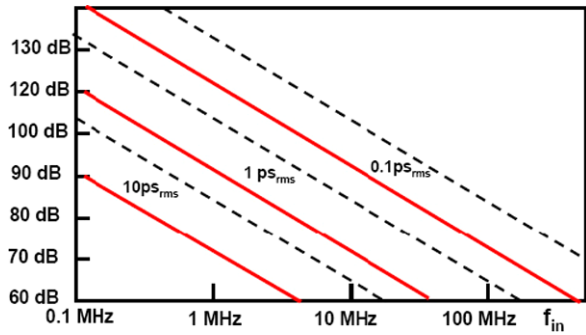
pulse. Sigma-delta modulation and noise shaping are forms of pulse density modulation.

Whatever form of time-domain conversion is used, it is important to realize that a large portion of the available energy in the time-domain representation is unwanted:

$$\text{Energy in PWM signal} \propto (+/-A)^2 = A^2$$

$$\text{Energy in maximum sine wave} \propto \int (A\sin(\omega t))^2\, dt = A^2/2$$

Even when the time-domain signal is representing a full-scale sine wave half of its energy consists of spurious components and needs to be removed.[5] These unwanted signal components are normally located at higher frequencies, however if these frequency components are applied to non-linear elements, down mixing of harmonics of the fundamental signal can occur.

The accuracy of generating time moments is limited as well. Jitter affects the actual position in time of the transitions. As jitter normally is a signal independent process, the accuracy of the conversion improves at low signal frequencies.

Any misplacement of a switching edge is directly multiplied with the maximum signal swing. The effect of jitter on a pulse-density coded signal is therefore considerably larger than in the sampling of analog signals, compare Sect. 3.1.4. The maximum sinusoidal signal power that is contained by a pulse train switching between $+A$ and $-A$ is the power of $A\sin(\omega t)$. The time jitter is σ_{jit} which results in $2A\alpha\sigma_{\text{jit}}f_s$ after multiplication with the amplitude and normalization with respect to the sample rate.[6] $0 < \alpha < 1$ is an activity factor indicating the fraction of transitions compared to the sample rate. This gives a signal-to-noise ratio of:

$$\text{SNR} = \frac{\frac{1}{T_a}\int_{t=0}^{T_a}(A\sin(\omega t))^2\, dt}{\frac{1}{T_a}\int_{t=0}^{T_a}(2A\sigma_{\text{jit}}\alpha f_s)^2\, dt} = \frac{1}{8(\alpha f_s\sigma_{\text{jit}})^2} = \left(\frac{1}{4\sqrt{2}\text{OSR}\alpha f_{\text{sig}}\sigma_{\text{jit}}}\right)^2 \quad (7.22)$$

---

[5]Compare this result to the power in the harmonics of a block wave in (2.6).

[6]Assuming only one edge is jittering.

with OSR $= f_s/2f_{\text{sig}} \gg 1$ as the oversampling ratio. Comparison with the signal-to-noise ratio of sampled signals in (3.22):

$$\text{SNR} = \left(\frac{1}{2\pi f_{\text{sig}}\sigma_{\text{jit}}}\right)^2 \tag{7.23}$$

shows that the impact of jitter on PWM signals is related to the sample rate. Therefore jitter in time-domain digital-to-analog conversion is an order of magnitude higher than in voltage or current-domain converters, see Fig. 7.26.

7.2.7 Class-D Amplifiers

A well-known form of pulse-width modulation is in audio class-D amplifiers as shown in Fig. 7.27. The incoming analog signal is compared to a triangular signal (as can be generated from the integration of a block wave). At the crossings the PWM signal flips polarity. This PWM signal is used to drive the switches that connect to the positive and negative power. This power-PWM signal is low-pass filtered and applied to the loudspeaker. Potentially this method can result in 85–90% efficient output stages. More advanced schemes are reported in e.g. [132].

7.3 Accuracy

7.3.1 Limits to Accuracy

The accuracy in which a quantity can be subdivided limits the obtainable performance of a digital-to-analog converter. In Fig. 7.28 a ladder of $M = 256$ equally designed resistors is connected between a reference voltage and ground. Ideally the voltage at the mth position is:

$$V(m) = \frac{m}{M}V_{\text{ref}} = \frac{mR}{mR + (M-m)R}V_{\text{ref}} = \frac{R_1}{R_1 + R_2}V_{\text{ref}} \tag{7.24}$$

Although devices can be equally designed, some random variation is unavoidable. If all resistor values are subject to a mutually independent random process with normal distributions with a mean value R and a variance σ_R^2, the variance on the voltage at the mth position can be calculated.

The resistance R_1 is a string of m resistors and shows an expectation value and variance:

$$E(R_1(m)) = E(mR) = mE(R) = mR, \qquad \sigma_{R1}^2 = m\sigma_R^2 \tag{7.25}$$

The other quantity involved $R_2(M-m)$ is the string of $(M-m)$ resistors which is independent of the complementary string R_1. For the string R_2 the mean and the

Fig. 7.27 Class-D amplifier with PWM modulation
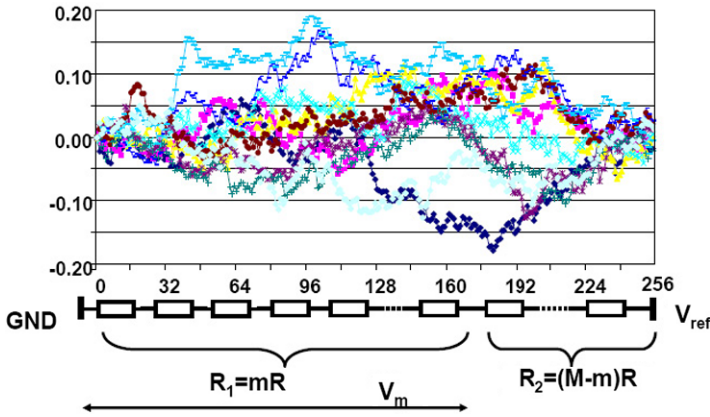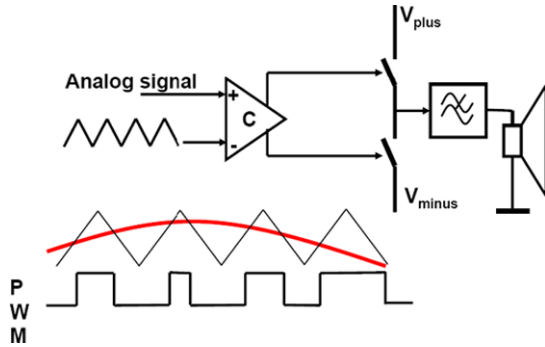
Fig. 7.28 A resistor string is connected between a reference voltage and ground. The simulation shows 10 instances of the effect of $\sigma_R/R = 1\%$ mismatch. In this 256 devices ladder the maximum deviation is $\pm 0.2R$

variance are found in a similar way.[7] The variance of the voltage $V(m)$ is now found by applying the statistics formula for multiple stochastic variables, (2.18):

$$
\sigma_V^2(m) = \left(\frac{\partial V(m)}{\partial R_1}\right)^2 \sigma_{R1}^2 + \left(\frac{\partial V(m)}{\partial R_2}\right)^2 \sigma_{R2}^2
$$

$$
= \left(\frac{R_2}{(R_1 + R_2)^2}\right)^2 \sigma_{R1}^2 V_{\text{ref}}^2 + \left(\frac{-R_1}{(R_1 + R_2)^2}\right)^2 \sigma_{R2}^2 V_{\text{ref}}^2
$$

$$
= \frac{m(M - m)}{M^3} \frac{\sigma_R^2}{R^2} V_{\text{ref}}^2
$$

---

[7]It seems that a shortcut is possible by using the string of M resistors, however this string shares m resistors with R_1 and the covariance has to be included, which is a possible route, but not pleasant.

Compare to (7.5). The maximum value of the variance occurs at $m = M/2$ for which the variance is found as:

$$\sigma_V^2(m = M/2) = \frac{1}{4M} \frac{\sigma_R^2}{R^2} V_{\text{ref}}^2 = \frac{M}{4} \frac{\sigma_R^2}{R^2} V_{\text{LSB}}^2 \tag{7.26}$$

where $V_{\text{ref}} = M \times V_{\text{LSB}}$.

The position of the maximum value of the variance is not equivalent to the position of the maximum voltage deviation of one particular string. Nevertheless (7.26) can serve to estimate the INL. With the help of Table 2.10 the sigma margin can be chosen.

In the example of Fig. 7.28 the ladder contains 256 resistors with a relative resistor mismatch of 1%. The relative sigma value in the middle of this ladder is therefore 8% and in the 10 random simulations excursions up to 20% of an LSB are seen. These values directly impact the integral linearity.

The differential linearity is given by the variation in the step size itself and equal to the expected maximum deviation of one resistor. The DNL of a resistor string is completely determined by the single resistor variance and the number of resistors. The DNL equals the maximum value of M random varying resistors with a spread σ_R/R. In order to guarantee the DNL for a production lot the estimation will have to be extended to P converters times M resistors per converter.

As an example the yield for a DNL of 4% in the previous example is calculated. The probability p that one resistor R_a with $\sigma_R/R = 1\%$ deviates more than 4% from R is:

$$p = P\left(\left|\frac{R_a - R}{R}\right| > \frac{4\sigma_R}{R}\right) \tag{7.27}$$

with Table 2.10 $p = 6.3 \times 10^{-5}$. The yield per resistor is $(1 - p)$. The yield per resistor string is $(1 - p)^M = 98.4\%$ (for $M = 256$). In a similar manner a unary current source array or a capacitor array can be calculated. The variance in current formed by the sum of m independent sources is $\sigma_{mI}^2 = m\sigma_I^2$. The maximum error occurs if all M current sources are used. This scenario implies the strict definition of linearity. If the "best-fitting" straight line algorithm is used, the overall variation is better referred to the total current. If the current of m sources is evaluated relative to the total current:

$$I(m) = \frac{mI}{mI + (M - m)I} I_{\text{tot}} \tag{7.28}$$

the mathematical description is identical to the resistor string problem and yields:

$$\sigma_{MI/2}^2 = \frac{1}{4M} \frac{\sigma_I^2}{I^2} I_{\text{tot}}^2 \tag{7.29}$$

This equation allows to estimate the INL while the DNL is directly coupled to the variance of the individual current sources. The DNL is therefore small, which emphasizes the architectural advantage of the unary architecture. Figure 7.29 shows the simulation of a unary array of 256 current sources with a relative current mismatch of 1%. The coding of the current sources is in a thermometer manner: each sample is
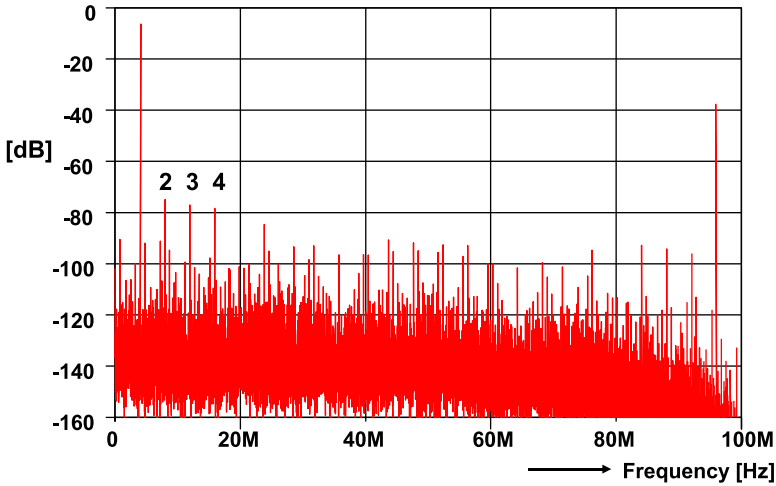
Fig. 7.29 Spectrum of a digital-to-analog converter with random mismatch. An array of 256 current sources is simulated with a normally distributed mismatch of $\sigma_I/I = 1\%$. The signal frequency is 3.97 MHz at 100 Ms/s sample rate, the quantization error is removed. The second, third and fourth order distortion components are labeled

built up starting with source number 1 until the number of sources corresponding to the digital code, is active. If the random process creates an error in a current source then this will result in an error at a fixed value of the signal. This form of coding transforms random errors in the current sources partially into distortion products of the signal and partially in a fixed-pattern noise.

Digital-to-analog converters based on a binary architecture require some modification of the above scheme. The random variation will affect the transition between the exponentially weighted portions of the unit. In an N-bit current source configuration the currents are ordered along powers of 2:

$$I_{N-1} = 2^{N-1} I_{\text{LSB}}, \qquad I_{N-2} = 2^{N-2} I_{\text{LSB}}, \qquad \cdots,$$
$$I_1 = 2^1 I_{\text{LSB}}, \qquad I_0 = I_{\text{LSB}} \tag{7.30}$$

The major transition is at the code $011\ldots11$ and $100\ldots00$ where in absence of random variation, the expected change in current is equal to one I_{LSB}:

$$\Delta I = I_{N-1} - I_{N-2} - \cdots - I_1 - I_0 = 2^{N-1} I_{\text{LSB}} - 2^{N-2} I_{\text{LSB}}$$
$$- \cdots - 2^1 I_{\text{LSB}} - 2^0 I_{\text{LSB}} = I_{\text{LSB}} \tag{7.31}$$

If the currents are subject to random variation, the value of ΔI will show significant variations. These variations result in a (potentially) large DNL, the architectural weak point of the binary coding. Depending on the implementation of the architecture several situations are possible.

- Every current branch is composed of a parallel connection of 2^k basic current sources as in Fig. 7.12 (lower). The current branch of weight k in the array is

described by $I_k = 2^k I_{LSB}$. Let each basic current source I_{LSB} suffer from a variation mechanism characterized by a normal distribution with mean value I_{LSBm} and a variance σ_I^2. Then each branch will be characterized by a mean $2^k I_{LSBm}$ and a variance $\sigma_{Ik}^2 = 2^k \sigma_I^2$, see (2.19). With the same equation the variance for the current step on the MSB transition is:

$$\sigma_{\Delta I}^2 = (2^{N-1})\sigma_I^2 + (2^{N-2} + \cdots + 2^1 + 2^0)\sigma_I^2 = (2^N - 1)\sigma_I^2 \quad (7.32)$$

Monotonicity requires that the value of ΔI remains positive. If a 3σ probability (99.7%) is considered an acceptable yield loss, then $3 \times \sqrt{(2^N - 1)}\sigma_I < I_{LSB}$. For an 8-bit converter this requirement results in $\sigma_I < 0.02 I_{LSB}$.

• In an R-2R ladder as in Fig. 7.11, the current splits at the first node in an MSB current I_{MSB} and a similar current for the remaining network. The impedances R_1 and R_2 are nominally equal to $2R$, but the individual resistors suffer from random variation characterized by a normal distribution with a variance $\sigma_{R1,2}^2$. With the help of (2.18)

$$I_{MSB} = \frac{R_2}{R_1 + R_2} 2^N I_{LSB}$$

$$\frac{\sigma_{IMSB}^2}{(2^N I_{LSB})^2} = \left(\frac{\partial I_{MSB}}{\partial R_1}\right)^2 \sigma_{R1}^2 + \left(\frac{\partial I_{MSB}}{\partial R_2}\right)^2 \sigma_{R2}^2$$

$$= \frac{R_1^2 \sigma_{R2}^2 + R_2^2 \sigma_{R1}^2}{(R_1 + R_2)^4} \approx \frac{2\sigma_{R1}^2}{4(R_1)^2} \quad (7.33)$$

Requiring for monotonicity that $3 \times \sigma_{IMSB} < I_{LSB}$, results in $\sigma_{R1}/R_1 = 0.002$. The ten-times higher precision that is needed for the resistor split, is due to the fact that the error is determined by the combination of just two resistors, whereas the MSB in the current architecture was built from 2^{N-1} current sources.

Van Den Bosch et al. [121] give some design optimization based on the above analysis.

7.4 Methods to Improve Accuracy

Small variations in parameters as discussed in the previous section lead to additive and multiplicative errors in analog circuits. In an amplifier additive errors can be regarded as a DC-offset, while multiplicative errors affect the overall gain. The digital-to-analog converters from the previous paragraphs consist of a division mechanism based on multiple copies of components (e.g. a resistor string or a set of current sources) and a switching mechanism that chooses the appropriate settings. Additive and multiplicative errors in analog-to-digital and digital-to-analog converters where signals use different paths in the circuit, affect only a certain range of the signal, see Fig. 7.30. These errors will generate complex and signal-dependent patterns that result in distortions, spurious tones or noise-like behavior. With the design guidelines of Table 11.6 the systematic errors in components can largely be eliminated. And large devices reduce the random error to a level that allows a 10 to 12 bit accuracy. If more accuracy is needed additional circuitry is required in the design.
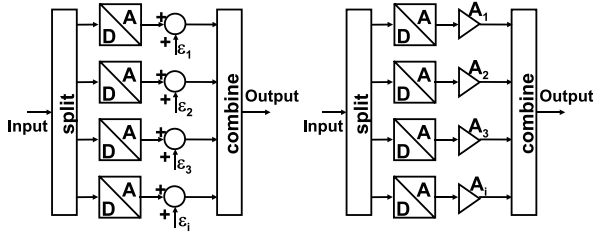
Fig. 7.30 Additive and multiplicative errors in conversion
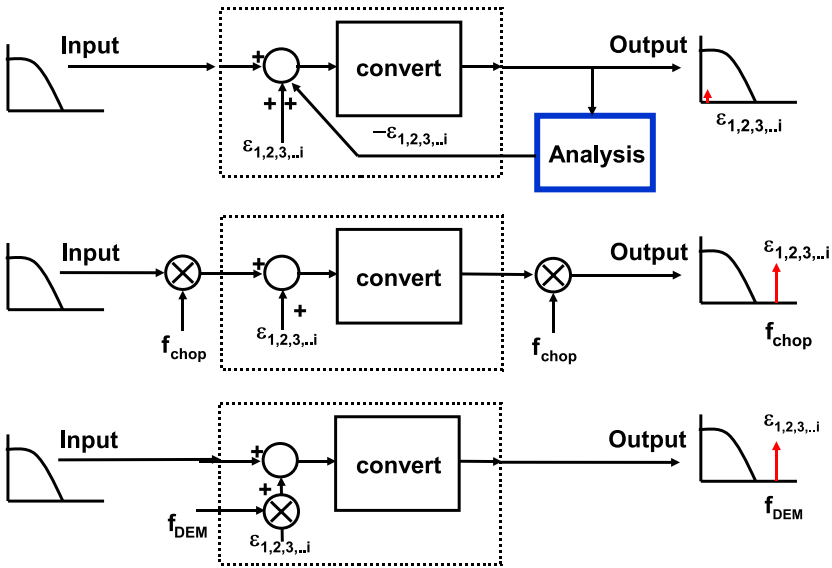
Fig. 7.31 Three methods to mitigate additive errors. *Top*: calibration, *middle*: chopping, *bottom*: dynamic element matching

In Fig. 7.31 the errors $\epsilon_{1,2,3,i}$ affect the different paths through the analog-to-digital or digital-to-analog converter. Three classes of mitigation of additive errors are shown. On top is the calibration principle. This principle assumes that the additive errors are localized and that a feedback path is available to compensate these errors at the node or close to the node where they originate. The feedback is generated by means of a form of measurement of the errors. The feedback loop can take many forms. E.g. a single measurement can be done during the production test and the feedback consists of a laser trimming some resistor value. A second option is a form of on-chip measurement, which runs continuously during the life time of the device. The advantage of calibration is that once in place the remaining circuit acts as an error-free circuit. The error is removed, not only moved. The electronic feedback path in such a system is continuously present and may introduce noise into the
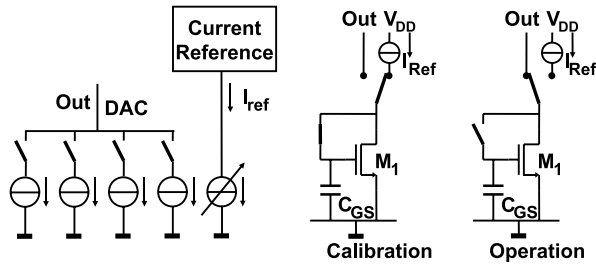
Fig. 7.32   The current calibration technique [117]

signal path. An example of the calibration of a matrix of current sources is given in Sect. 7.4.1.

The second method to mitigate additive errors was 50 years ago already in use for precision instrumentation. The chopping technique, Fig. 7.31 (middle), first modulates the input signal to a higher frequency band which is processed by the converter. The error is located at DC or a low-frequency band and will not interfere with the signal. Now the processed and composite signal is modulated back with the same frequency as used at the input. The signal is restored to its frequency range and the additive error is modulated to the chopping frequency. This technique relies on good quality mixers/modulators and of course requires the conversion function to process the signal at that frequency. Although this technique is rarely used in a pure converter [133], it is often implicitly present in digital radio systems. In contrast to calibration the error is still present and its energy may come back into the signal path after e.g. distortion.

The last principle is addresses the error at its source, just as in calibration. The error is not removed but moved to a less critical frequency range. Examples of this technique are dynamic element matching and data-weighted averaging, Sects. 7.4.2 and 7.4.3.

The discussion in this section is limited to mitigation of additive errors. The impact and avoidance of multiplicative and timing errors is discussed in Sect. 8.8.

7.4.1 Current Calibration

With the change from bipolar devices to CMOS techniques new forms of calibration became possible. Figure 7.32 shows an NMOS transistor that is used as a current source for a unary array of current sources. In order to mitigate the inherent mismatch in these sources the current sources are calibrated. For this purpose in Fig. 7.33 the unary row of current sources is extended by one, from 64 to 65. The spare current source is connected as a sort of track-and-hold circuit by closing the switch between gate and drain and feeding this arrangement with a reference current. The gate voltage will settle at the level needed to balance the reference current. After opening the switches the gate capacitor will hold the current and this current
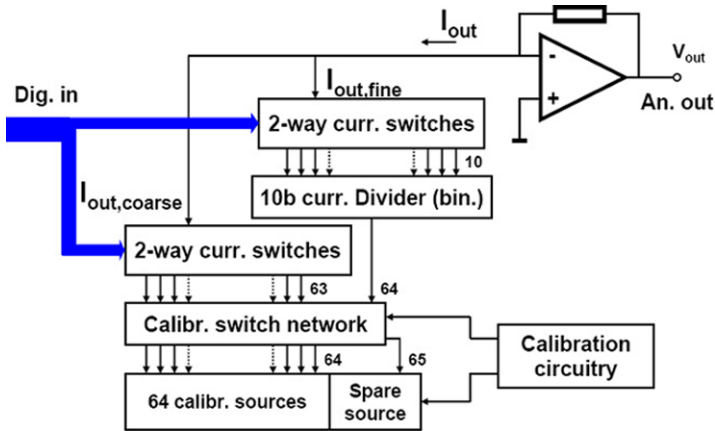
Fig. 7.33 An example of a digital-to-analog converter applying current calibration [117]

source is put back into the array. A new current source is taken out of the array and will be calibrated. In this way the calibration mechanism rotates through the array tuning every 65 cycles all current sources. The unary array of calibrated current sources in the digital-to-analog converter of Fig. 7.33 is completed by a 10-bit passive current divider to yield a 16-bit digital-to-analog converter. The main problem in this arrangement is the sampling of noise in the calibration cycle. Unwanted low-frequency noise as $1/f$ noise is suppressed up to half of the calibration frequency, but the remaining contributions influence the current considerably. This is mainly due to the relatively low calibration frequency which is experienced by the individual current sources. For that reason an approach where most of the current is generated by a standard current source and only a few percent is calibrated in the above way is favored.

7.4.2 Dynamic Element Matching

An often used technique is swapping components, currents or voltages. In Fig. 7.34 the two currents I_1 and I_2, have been created by dividing a main current, however these currents show a small offset with respect to each other. A switching box alternates the connection to the outputs. The resulting currents I_3 and I_4 are composed of both I_1 and I_2 and average out the difference over time. The resulting error in I_3 and I_4 is the product of the original inequality times the duty cycle of the switching pulse.

$$\frac{I_3 - I_4}{I_3 + I_4} = \frac{I_1 - I_2}{I_1 + I_2} \times \frac{t_1 - t_2}{t_1 + t_2} \tag{7.34}$$

This technique is known as "Dynamic Element Matching" [113] and is applied in many types of converters. Figure 7.35 shows an extension of the technique to generate three currents in the ratio $1 : 1 : 2$. Cascoding a number of these stages leads to a
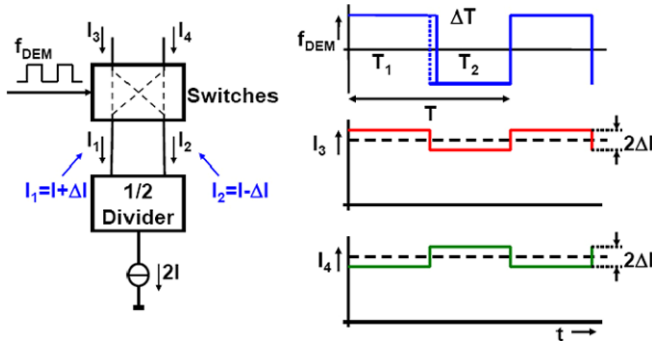
Fig. 7.34 Dynamic element matching [113]
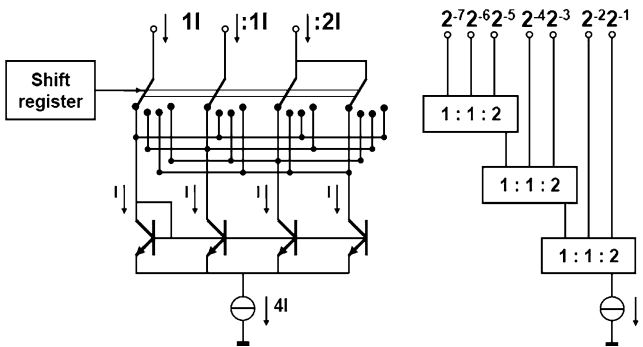
Fig. 7.35 An example of a digital-to-analog converters applying dynamic element matching [116]

16-bit current digital-to-analog converter. Unfortunately the cascoding costs voltage head room, therefore the dynamic element technique is today more used in order to tune a small set of devices. Note that the switching frequency will not interfere with signal path, the signal switching can be performed after the D.E.M. procedure and spurious frequency components can be filtered out before the signal switching takes place.

7.4.3 Data-weighted Averaging

The elements in the current source array in Fig. 7.12 or the resistor string are always used in the same order as depicted in Fig. 7.36 (left). The signal in this digital-to-analog converter is being built up starting with the left most element and the somewhat larger third element will show up at a fixed position in the signal. This will result in harmonic distortion, as is shown in Fig. 7.29.

   The dynamic element matching technique rotates the various components in a pre-determined order and under control of an independent frequency. The idea is to

Fig. 7.36 Data weighted averaging [134, 135]
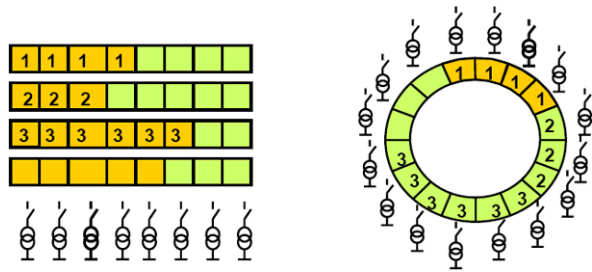
transform a fixed DC-error into a modulated component at a high frequency that is either filtered out or moved to a frequency band where it causes no problem.

A different solution to the problem is to change the starting point for building up the signal in the array. E.g. the next sample starts at an offset of one position. This offset is increased for every next sample. The offset can even be chosen randomly [136]. These techniques transform the error at the third position in this example into a sort of noise as the correlation with a fixed position in the signal amplitude is avoided. Mathematically this principle can be understood by assuming that the elements x_i, $i = 0, \ldots, \infty$ form an infinitely long line of elements. A signal sample formed with p of these elements is disturbed with an error $\sigma_x \sqrt{p}$. If the average value of p is around mid range $E(p) = 2^N/2$, the long term average noise per sample is $\sigma_x \sqrt{2^{N-1}}$. In the frequency domain this results in a white noise spectrum.

This idea is taken one step further in the data-weighted average principle.[8] The elements are now functionally arranged in a circle, see Fig. 7.36 (right). Every next sample starts after the last element of the previous sample. If an array is composed of 2^N elements with random varying values x_i, $i = 0, \ldots, 2^N - 1$, the cyclic operation causes that every element is used an equal amount of times in the long term. Therefore it is useful to specify the average or mean value of every element by means of the expected value:

$$m_x = E(x_i) = \frac{1}{2^N} \sum_{i=0}^{i=2^N-1} x_i \tag{7.35}$$

A sample at time $t = nT$ contains k elements and its output value will be:

$$y(nT) = \sum_{i=0}^{i=k-1} x_i = \sum_{i=0}^{i=k-1} m_x + \sum_{i=0}^{i=k-1} (x_i - m_x) \tag{7.36}$$

The first term is the desired signal value of the digital to analog conversion and the last term in the summation represents the unwanted or noisy part of the array of elements. This formulation modifies the ideal amplitude of the actual element x_i into its mean value m_x. This may (slightly) change the amplitude of the signal if the random component has a non-zero mean value, however in practice the result will be a minor change in gain.

---

[8]Two sources are available as the originator of DWA: Michael Story [137] and A. Maloberti [138].

Fig. 7.37 A single mismatching current source will appear at a low frequency if the average signal value is close to the extremes of the range
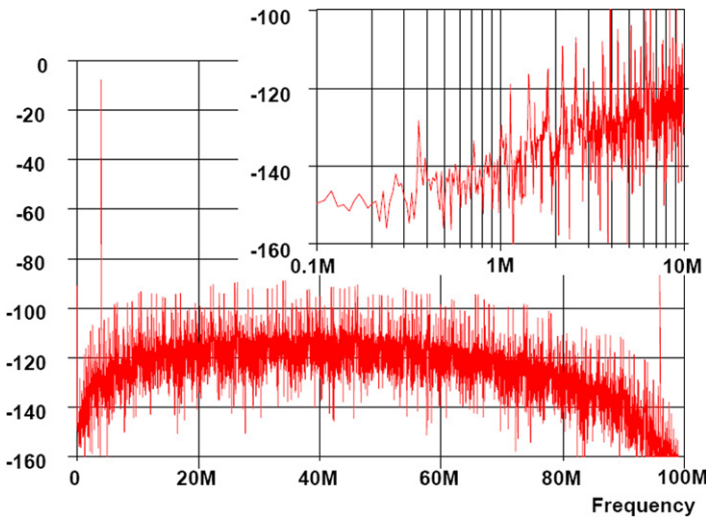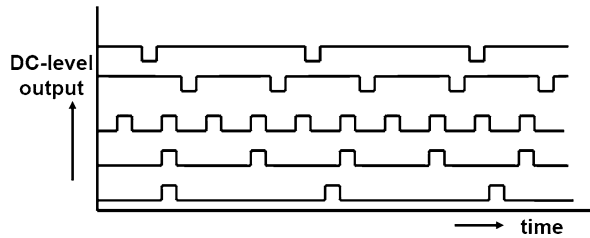
Fig. 7.38 Spectrum of a data weighted averaging digital-to-analog converter. 256 current sources are simulated with a normally distributed mismatch of $\sigma_I/I = 1\%$ (no quantization error). The signal frequency is 3.97 MHz at 100 Ms/s. In the *insert* a small fraction of the range is visible. The spectrum contains a lot of tones

The error signal contains spectral components. If there would be only one deviating element x_p with an error ϵ_p, Fig. 7.37, the error signal due to ϵ_p will appear in the output depending on the number of times the element x_p is requested. This results in an unwanted spurious contribution f_p where the frequency depends on the level of the output signal. E.g. if the average signal level is a one-quarter of the full-scale, the error element will appear one on every four sample pulses, so the error component will appear around $f_s/4$. With an average signal around mid-range the error will produce tones around $f_s/2$ and generate maximum error power. The error frequency and the error power associated to it are proportional:

$$P_p \propto f_p \epsilon_p^2 \tag{7.37}$$

For a single error the energy of spurious component is proportional to the frequency. In a spectrum this leads to a first-order behavior of the error tones versus frequency. For a full error pattern the same conclusion holds [134]. Unfortunately a full error

pattern will not completely randomize the error signal. The possibility that a fair
number of samples of a DWA digital-to-analog converter produces tones is likely.[9]

If there is no dominant error source, the spectrum of the random components will
show a first-order frequency shaping, as is seen in Fig. 7.38.

A formal proof requires the introduction of some stochastic mathematics [134].
The following reasoning summarizes the line of thought without pretending to be a
solid mathematical proof. The second term of (7.36) defines the error for a sample
at time $t = nT$. As the total number of sources is bounded to $2^N - 1$, the error of
the remaining sources is given by:

$$\sum_{i=k}^{i=2^N-1} (x_i - m_x) = \sum_{i=0}^{i=2^N-1} (x_i - m_x) - \sum_{i=0}^{i=k-1} (x_i - m_x)$$

$$= - \sum_{i=0}^{i=k-1} (x_i - m_x) \tag{7.38}$$

The remaining error has the magnitude of the error at time $t = nT$ with an opposite
sign. The DWA algorithm by its definition uses for the next samples the remaining
sources. So the remaining error will show up at the next sample or at the next few
samples, depending on the values of the succeeding signal samples. If the delay of
the remaining error term with respect to the originating error term is estimated at
αT_s, with $\alpha \geq 1$, the transfer function of the error becomes:

$$E(z) = \sum_{i=k}^{i=2^N-1} (x_i - m_x)(1 - z^{-\alpha}) \tag{7.39}$$

With a crude approximation the frequency behavior of the error is:

$$E(\omega) \propto \sum_{i=k}^{i=2^N-1} (x_i - m_x)(2 \sin(\omega T_s/2)) \tag{7.40}$$

This result for a single error can be extended to all errors that will occur during
the conversion of a complex signal resulting in a first order shaping of the ran-
dom errors. Figure 7.38 shows the result of the data-weighted averaging operation.
Firstly the harmonic distortion products have been considerably reduced due to the
randomizing effect of the selection. Secondly the DWA algorithm has cleaned the
spectrum around DC, see insert. The noise level at higher frequencies has increased
as is inevitable in "noise-shaping". The consequence is that DWA must be applied
in oversampling applications: the sample rate must largely exceed the signal band-
width in order to allow some frequency band for the excess noise. Other switching
sequences of the DWA algorithm allow to choose the frequency of maximum sup-
pression at arbitrary points in the spectrum [129]. Miller and Petric [139] show a
high speed implementation of the digital decoding.

---

[9]In Chap. 9 "idle patterns" are discussed. The patterns in data-weighted averaging bear a lot of
resemblance but come from a completely different origin.

7.5   Digital-to-Analog Conversion: Implementation Examples

7.5.1   Resistor Ladder Digital-to-Analog Converter

In a system chip with analog-to-digital and digital-to-analog converters, it is advantageous to have similar references for the analog-to-digital and digital-to-analog converters. The tracking of input and output ranges for processing variations, temperature, etc. is then guaranteed and the overall gain of analog-to-digital and digital-to-analog converters is better controlled. The voltage dependence and the mutual matching of large-area polysilicon resistors[10] allows the design of a converter with high integral and differential linearity. Basically, the variation in the polysilicon resistance value is determined by its geometry variations: the variations in length and width result in local mismatches and the variation in thickness results in gradients.

The design of a digital-to-analog converter with a single 1024-tap resistor ladder and sufficiently fast output settling requires tap resistors in the order of 6–10 Ω. The size of such resistors in conventional polysilicon technology is such that accurate resistor matching, and hence linearity, becomes a problem. The solution to this problem is to use a dual ladder [140] with a matrix organization [141].

Figure 7.39 shows the ladder structure: the coarse ladder consists of two ladders, each with 16 large-area resistors of 250 Ω which are connected anti-parallel to eliminate the first-order resistivity gradient. The coarse ladder determines 16 accurate tap voltages and is responsible for the integral linearity. A 1024-resistor fine ladder is arranged in a 32-by-32 matrix, in which every 64th tap is connected to the coarse ladder taps. This arrangement makes it possible to avoid the need for small tap resistors (see Fig. 7.6). The fine ladder tap resistance is chosen at 75 Ω without loss of speed. The wiring resistances can be neglected compared to the 75 Ω tap resistors. There are only currents in the connections between the ladders in the case of ladder inequalities; this reduces the effect of contact resistance variance. The current density in the polysilicon is kept constant to avoid field-dependent non-linearities. The coarse ladder is designed with polysilicon resistors in order to avoid voltage dependence of diffused resistors. The fine ladder is designed in either polysilicon or diffusion, depending on secondary effects in the process implementation. The double ladder structure is also used in all of the three A/D converters discussed in Sect. 8.9.

In a basic ladder design consisting of one string of 1024 resistors, the output impedance of the structure varies with the selected position on the ladder and therefore with the applied code. The varying output impedance in combination with the load capacitance results in unequal output charging time and consequently signal distortion of high-frequency output signals. This source of varying impedance has been eliminated by means of a resistive output rail. The insert in Fig. 7.40 shows part of two rows of the matrix.

The second source of the varying output impedance is the switch transistor; usually its gate voltage equals the positive power supply, but the voltage on its source

---

[10]Diffused resistors are a preferred alternative in more advanced processes.
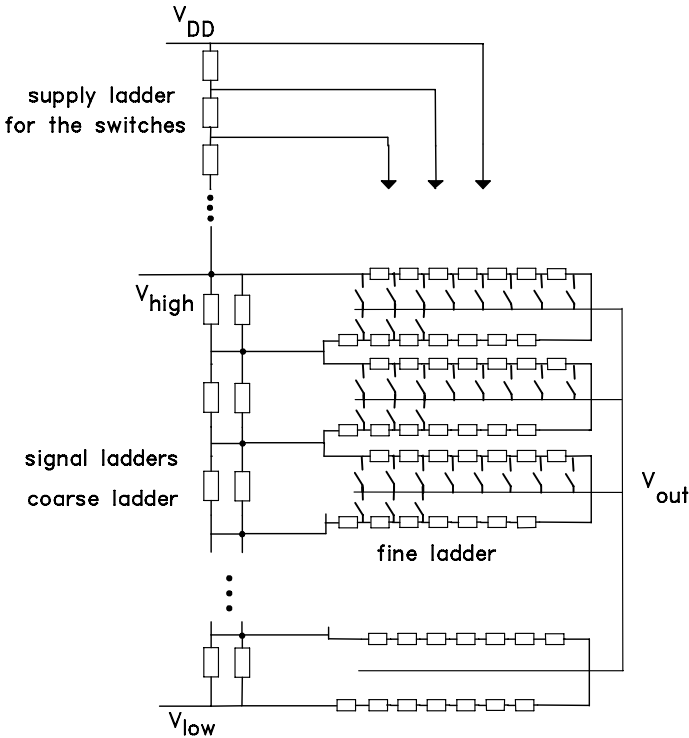
Fig. 7.39 Resistor network for a video digital-to-analog converter [142]

terminal is position dependent. The turn-on voltage doubles from one end of the ladder to the other. In this design an additional supply ladder is placed on top of the signal ladders to keep the turn-on voltage of the switches more constant. The turn-on voltage of each switch transistor is effectively made to correspond to the lowest turn-on voltage of the ladder digital-to-analog structure.

The total ladder configuration can still be fed from the 3.3 V analog power supply; the signal ladders are in the range between ground level and 40% of the power supply, the supply ladder goes from 60% to V_{DD}.

The core of the digital-to-analog converter is formed by the 32-by-32 fine-resistor matrix. The two decoders are placed on two sides of the matrix. The two sets of 32 decoded lines are latched by the main clock before running horizontally and vertically over the matrix. In the matrix, the 1024 AND gates perform the final decoding from the 32 horizontal MSB lines and the 32 vertical LSB lines.

A voltage-domain digital-to-analog converter generates only the necessary momentary current and is hence more efficient in power consumption. However, a voltage-domain digital-to-analog converter requires an on-chip buffer, which introduces two drawbacks: the output always needs some offset from the power supply rail and the opamp is inherently slower. The output buffer is a folded-cascode opamp. The p-channel input stage operates on input voltages ranging from 0 to
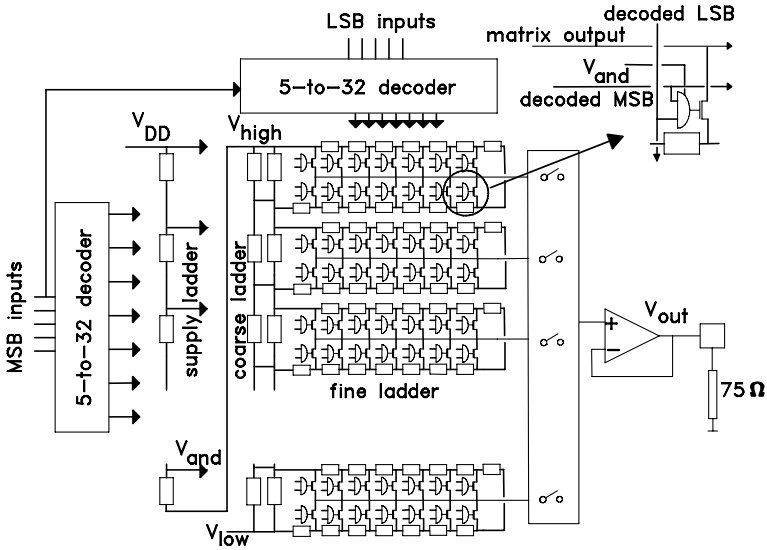
Fig. 7.40 Block diagram of the digital-to-analog converter [142]
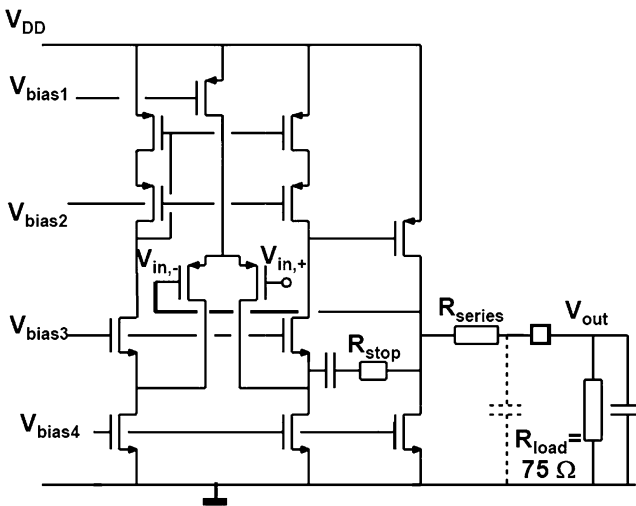
Fig. 7.41 Folded cascode opamp circuit used for the buffer

2.2 V, see Fig. 7.41. The main current path through the output stage goes from the PMOS driver ($W/L = 1400$) down into the output load. A resistive load is consequently needed for optimum performance.

The on-chip stop resistor R_{stop} is on the order of 25–75 Ω; it keeps up a feedback path even at frequencies at which the bond pad capacitance shorts the circuit output. It also serves as a line termination. The swing of the output load resistor is
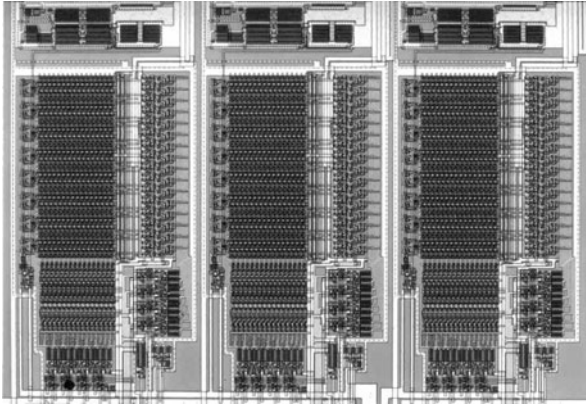
Fig. 7.42 (Color online) This 10-bit triple digital-analog converter based on voltage dividers is used for applications in the video domain for the basic video colors: *red*, *green* and *blue*

consequently half of the buffer input voltage. The actual value of the stop-resistor can be controlled to within 10%; the resulting gain error is no serious drawback in video equipment, as there is always a total gain adjustment.

The power distribution over the 10-bit digital-to-analog converter is dominated by the output stage: with a full-swing sine wave (0.1 V to 1.1 V on 75 Ω), the average current through the driver transistor is 7.3 mA. The remaining part of the driver requires 1 mA. The ladder current is 1 mA while the digital part running at 50 MHz is limited to 0.7 mA, resulting in a total power supply current of 10 mA.

Table 7.3 in Sect. 7.5.3 summarizes the performance that is more extensively reported for a 1 μm technology in [142]. Figure 7.42 shows a photograph of a triple 10-bit converter.

7.5.2 Current Domain Digital-to-Analog Conversion

High-speed CMOS digital-to-analog converters designed with a current cell matrix allow fast and accurate settling [114, 143].

A 100 Ms/s 10-bit digital-to-analog converter was designed with an array of 64 current sources for the 6 MSBs. This structure was completed with a 4-bit binary-coded LSB section. Figure 7.13 shows the block diagram of a 10-bit design. The major design issue in this circuit is the switching of the 64 current sources. Decoding delays in the 64 current sources are reduced by an additional latching stage just before the current switch. The current switch itself is designed with a low-swing (1 V) differential pair. This measure results in a low clock feed-through on the output line, while in this case the switch transistors also act as a cascode stage, thus reducing the modulation of the output current by the output voltage. Figure 7.43 shows the chip photograph of this design.
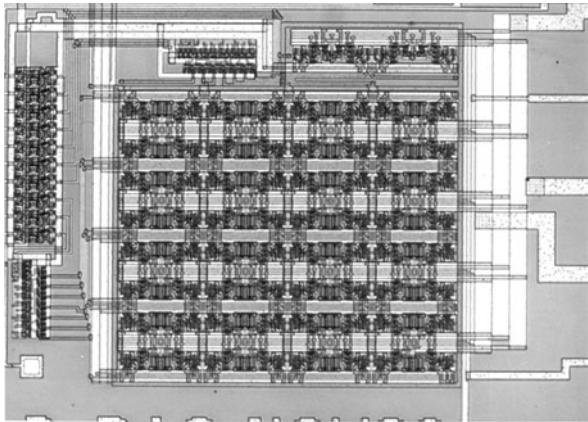
Fig. 7.43 Chip photograph of a 10-bit current digital-to-analog converter

Table 7.3 Comparison of measured specifications of a ladder and a current digital-to-analog converter, both loaded with 75 Ω and 25 pF

Digital-to-analog converter type	Ladder	Current
Process	1.0 μm CMOS	
DC resolution	10 bit	
Sample frequency	>100 MHz	
Area 1 μm CMOS	1.05 mm^2	0.7 mm^2
Differential linearity error	<0.1 LSB	<0.6 LSB
Integral linearity error	<0.35 LSB	<1 LSB
Glitch energy	100 psV	100 psV
Rise/fall time (10%–90%)	4 ns	1 ns
Settling time (1 LSB)	20 ns	5 ns
Signal bandwidth (-1dB)	20 MHz	>20 MHz
Minimum power supply (THD < -40 dB)	3 V	3 V
Output in 75 Ω	1 V	1 V
Output at min. code	100 mV	<1 mV
Av. current (50 MHz, 75 Ω)	10 mA	15 mA
Av. current (50 MHz, 2×75 Ω)	10 mA	28 mA
THD $f_{signal} = 1$ MHz, $f_{clock} = 27$ Ms/s	-58 dB	-60 dB
THD $f_{signal} = 5$ MHz, $f_{clock} = 100$ Ms/s	-50 dB	-44 dB

7.5.3 A Comparison

Table 7.3 compares the specifications of the two 10-bit video digital-to-analog converters. Remarkable differences are the differential linearity error and the distortion. The DNL error in the current digital-to-analog is a direct effect of the current source

mismatch: especially between the coarse and fine sections. In the voltage digital-to-analog this problem is circumvented by a fully unary approach: 1024 resistors in series are used. The consequences are of course seen in a larger area. The harmonic distortion has different origins in the two converters: in current digital-to-analog converters the non-linear behavior of the large output diffusion node and the output transconductance of the current sources is important. The distortion in the voltage digital-to-analog converter is caused by the limited performance of the driver stage in the output buffer. The modulation of the switch resistance in the resistive digital-to-analog is effectively canceled by the ladder organization, while the reduced swing scheme of the current source switching limits the switch distortion in the current digital-to-analog converter.

The dynamic behavior of the digital-to-analog converter is determined by the output pole: in the current digital-to-analog converter this is the dominant pole with a 25 pF/75 Ω load. The opamp that implements the buffer of the resistive digital-to-analog converter sees this pole as its second pole because the internal Miller-compensation is the dominant pole. The buffered output is consequently slower than the current output, which is seen in differences in the rise/fall time. In most systems the values reported here will be sufficient.

The minimum value for the output code is significantly larger than 0 V for a buffered output because a minimum saturation voltage is needed in the output stage.

The difference in power dissipation is less pronounced because there is more overhead for ladders and biasing in a voltage domain. Even so, a factor of 1.5 to 2.8 remains.

Voltage domain digital-to-analog converters route about 75% of their average current into the 75 Ω output load; with those (system-determined) output loads the potential for further power reduction seems to be low on an implementation level.

A high-speed digital-to-analog converter can be chosen if system requirements are mapped onto these specifications.

7.5.4 An Algorithmic Charge-based Digital-to-Analog Converter

Algorithmic principles [41, 144, 145] allow to realize digital-to-analog converters for high-resolution application in CMOS switched-capacitor technology. The major problem which has to be solved in many high-resolution converters based on switched-capacitor techniques, is performance loss due to opamp offset and mismatching in capacitor values. This digital-to-analog converter is based on an algorithm which ensures monotonicity despite offset and capacitor mismatch.

For the implementation of an N-bits converter the basic digital-to-analog conversion formula is rewritten:

$$V_{\text{out}} = V_{\text{LSB}}\left(a_0 2^0 + a_1 2^1 + \cdots + a_p 2^p + \cdots + a_{N-1} 2^{N-1}\right)$$
$$= V_{\text{LSB}}\left(W_{\text{LSB}} + W_{\text{MSB}} 2^p\right) \tag{7.41}$$

Fig. 7.44 Basic switched capacitor network with the timing diagram for converting "1011010" [144]
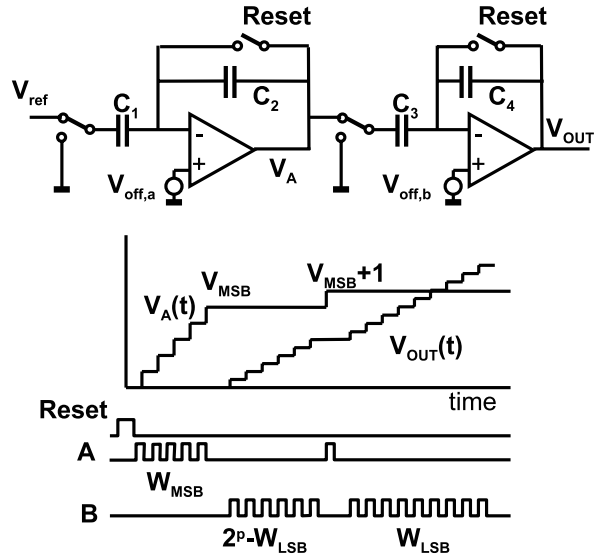
where W_{LSB} denotes the digital word formed by the p least significant bits: $(a_0 \ldots a_{p-1})$ and W_{MSB} is the digital word formed by the $(N - p)$ most significant bits $(a_p \ldots a_{N-1})$. Under the constraint that coefficients a_i, W_{LSB} and W_{MSB} are non-negative integers this equation belongs to the class of "diophantic equations". Diophantic equations are composed of polynomials with only integer coefficients and integer variables. Direct implementation in CMOS switched capacitor technique could be done in a two stage schematic as in Fig. 7.44 by generating the analog value of an LSB at the output of section A. This V_{LSB} is W_{LSB} times transferred to section B: $V_{out} = V_{LSB} \times W_{LSB}$. Now the output of section A is raised to $2^p V_{LSB}$ and W_{MSB} transfers take place to section B. Resulting in: $V_{out} = (W_{LSB} + 2^p W_{MSB})V_{LSB}$. This implementation leads however to a signal-dependent offset at the output. The offset of section B $V_{off,B}$ is transferred $(W_{MSB} + W_{LSB})$ times into the output signal:

$$V_{out} = (W_{LSB} + 2^p W_{MSB})V_{LSB} + (W_{MSB} + W_{LSB})V_{off,B} \qquad (7.42)$$

thereby creating a non-linear dependence and signal distortion.

The contribution of the offset of section B at the output is made constant by keeping the number of charge transfers of section B constant for the conversion of any code. This can be obtained by rewriting the last part of (7.41) in:

$$V_{out} = V_{ref}\frac{C_1 C_3}{C_2 C_4}(W_{MSB}(2^p - W_{LSB}) + (W_{MSB} + 1)W_{LSB}) \qquad (7.43)$$

If M_1 through M_4 are the sequential numbers of transfers, where M_1 and M_3 denote the number of transfers of section A and M_2 and M_4 are the transfers of section B, then the choice:

$$M_1 = W_{\text{MSB}},$$
$$M_2 = 2^p - W_{\text{LSB}},$$
$$M_3 = 1,$$
$$M_4 = W_{\text{LSB}}$$

results in $M_2 + M_4 = 2^p$ transfers of the offset of section B, which is no longer signal dependent. In the lower part of Fig. 7.44 the transfers for $N = 7$ and $p = 4$ have been indicated. Equation (7.43) can be interpreted as an interpolation algorithm: section A forms voltages which are proportional to W_{MSB} and $W_{\text{MSB}} + 1$, while section B interpolates in 2^p cycles between these values according to the value of W_{LSB}. The maximum number of charge transfers is $2^p + 2^{N-p}$. The fixed offset of section A is added to V_{LSB}, resulting in a gain error, and the offset of section B is multiplied by $2^p C_3/C_4$ which is signal independent and added to the output voltage. The capacitor ratios form a multiplication factor for the complete conversion, thereby influencing only the absolute gain. With a capacitor ratio of 2^{-p} a maximum swing at all opamp outputs is obtained.

The principle described by (7.43) has been the basis of the implementation of a 15-bit CMOS digital-to-analog converter with three sections of 5 bits. During the first half of the conversion period, (7.43) is used to form the value represented by the 10 most significant bits, then (7.43) is once more used to obtain the resolution for the 5 least significant bits.

Two important phenomena that influence the performance of the device are the limited DC gain and the settling of the charge transfer. In switched capacitor filter applications the finite gain is merely a constant factor for the transfer characteristics from the input voltage to the output. In this device the input is the number of transfers for a section. Now non-linearity occurs because the size of the transferred charge packet changes with the gain-error voltage between the opamp inputs. This error voltage is again proportional to the number of previous charge transfers. The non-linear transfer is mainly generated in the first section while the error magnitude in the other sections is reduced by 2^{-p}, because the number of transfers in these sections has been made constant:

$$V_{\text{out}} \approx V_{\text{LSB}}(W_{\text{LSB}} + 2^p W_{\text{MSB}}) \frac{C_1 C_3}{C_2 C_4} \left(1 - \frac{W_{\text{MSB}} C_1}{2 A_{\text{DC}} C_2} \right) \qquad (7.44)$$

With 80 dB opamp gain this effect is sufficiently reduced.

The settling of the charge transfer is important as it transforms clock jitter into output noise: the magnitude of the transferred charge from C_1 to C_2 is determined by the moment where the discharging of C_1 stops. With a limited unity gain of 6 MHz and the second pole of the opamp at 30 MHz the charge fraction which is not transferred, is 5×10^{-4} after 100 ns. With an average of 80 transfers per conversion, the contribution to the noise of clock jitter to the total S/N ratio is in the order of -90 dB if the clock jitter is below 1 ns. The gain and settling requirements of the opamp have been realized by means of a folded cascode configuration followed by a Miller stage. The input stage and the current source transistors contribute to the noise, and have to be designed carefully.

Table 7.4 Measured performance of the algorithmic digital-to-analog converter

DC resolution (monotonicity)	15 bit
S/(N+THD)	74 dB
Dynamic range	87 dB
Sample frequency	44 kHz
Clock frequency	5.6 MHz
Power consumption (2.5 μm CMOS)	22 mW

The results of the implementation are summarized in Table 7.4. Measurements have been performed with an external sample-and-hold circuit. The DC measurements show the inherent monotonicity of the 15-bit digital-to-analog converter. The noise and distortion figures of Table 7.4 include the sample-and-hold contributions.

The main drawback of this approach is that after the converter an additional sample-and-hold has to be used in order to create a full-time signal.

Chapter 8
Analog-to-Digital Conversion

The analog-to-digital converter compares the input signal to a value derived from a reference by means of a digital-to-analog converter, Fig. 8.1. The basic functions of an analog-to-digital converter are the time and amplitude discretization. The circuit ingredients that implement these functions are the sample- or track-and-hold circuit where the sampling takes place, the digital-to-analog converter and a level-comparison mechanism. The comparator circuit is the location where the signal changes from its pre-processed analog form into a digital decision. After the comparator a digital circuit processes the decisions into a usable digital signal representation.

  Various subdivisions of analog-to-digital converters can be made. Based on the timing of the conversion four categories can be identified, see Fig. 8.2:

- The "flash" converters or "parallel search" converters use one moment in time for the conversion. The input signal and all the required reference levels must be present at that time moment. From a topological point of view no sample-and-hold on the analog side is required. The latches in the comparators form a sort of digital hold structure. Fast converters are possible, however the requirement to provide decisions on 2^N reference levels leads to an exponential growth of the area and power of the converter. An elegant variant is the "folding" converter. Application that require the highest speed and modest accuracy are served by these converters.
- The second category is the "sequential search" converter. This class builds up the conversion by choosing at every new clock strobe a new set of reference levels based on the information processed up to that moment. For each step a suitable resolution can be chosen, mostly based on a power of 2: 2^1, 2^2, 2^3, etc. A fundamental choice is to use the same hardware for all processing steps of one sample, or to use dedicated hardware for each next resolution step. In the last approach the total processing time for a sample is still the same, but more samples can be processed at the same time. Principles in this category are: successive approximation conversion, pipe-line conversion, multi-step conversion. The combination of high accuracy and rather high speed makes these converters suitable for many industrial and communication application.

Fig. 8.1 Components in
analog-to-digital conversion
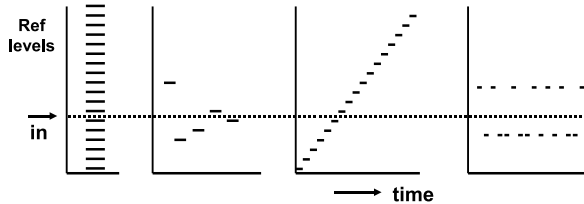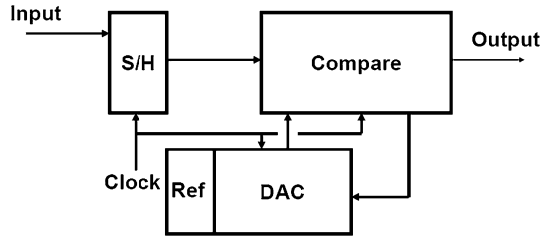
Fig. 8.2 Classification of analog to digital conversion principles. In the vertical dimension the available levels are depicted. *Left*: in direct conversion all levels are provided. The sequential search algorithm chooses the next level based on the information of the previous comparisons. In linear search each level is available at consecutive clock periods. Finally the oversampled conversion switches between a few levels

- On the opposite side of the spectrum is the "linear search" converter. All potential conversion levels are in increasing or decreasing order generated and compared to the input signal. The result is an extremely slow conversion, built with a minimum amount of hardware and tolerant to component variations. An example is the dual-slope converter. The robustness of these converters makes them popular for slow-speed harsh environments, such as sensor interfaces.
- The last category of conversion principles is mentioned here for completeness. The oversampled converters use mostly a few reference levels and the output switches frequently between those reference levels to create a time average approximation of the input. The accuracy comes from the time domain. These feedback based delta-converters do not provide the conversion result at a determined point in time, but are accurate over a larger number of samples. The special techniques and analysis tools for these delta modulators are discussed in Chap. 9.

The first three categories are called "Nyquist-converters". These circuits convert a bandwidth[1] close to $f_s/2$ and often operate at the speed limit of the architecture and process. This chapter discusses the flash, subranging, successive approximation and linear converters after a discussion of the comparator.

---

[1]This bandwidth can start at DC, but also far above f_s using sub-sampling in Sect. 12.1.4.

Fig. 8.3 Three basic comparator designs: straight-forward amplification, amplifier with hysteresis and a latched or regenerative amplifier
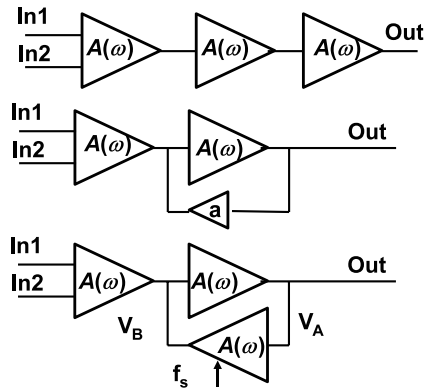
8.1 The Comparator

Every analog-to-digital converter contains at least one comparator. There are as many comparator circuits as there are analog-to-digital converter designers. Many aspects need consideration when designing a comparator. These aspects may vary for every different application, for specific class of signals, a technology, etc. No universal "one design fits all converters" comparator exists. Yet some general remarks on the design of the comparator can be made. Accuracy and speed are the global design parameters that have to be balanced versus the power consumption.

The fundamental task of a comparator is to amplify a big or small difference between its input terminals into a digital decision. In other words to extract the sign of the differential input signal. A number of requirements can be specified for a comparator:

- A large amplification is imperative.
- A wide bandwidth for operating on high-frequency signals.
- Accuracy of various forms is required. In practical terms this means a low input offset, a low noise figure, for $1/f$ noise as well as for thermal noise. Clocked comparators should not add uncertainty to the decision moment: a low timing jitter is required.
- A low power consumption, especially in analog-to-digital converters employing a lot of comparators.
- A wide input common mode voltage range, with a high common mode rejection.
- The previous comparator decision should not affect the following, no memory effect.

Figure 8.3 shows three topologies for comparators. The straight-forward limiting amplifier (top) consists of a cascade of amplification stages to obtain as much gain as possible. For a given current consumption the Unity Gain Bandwidth per stage determines the overall speed of this chain of amplifiers, compare Sect. 2.7.2.

If a single amplifier stage has a Unity gain Bandwidth ω_{UGBW} and an amplification A, then the dominant pole is at $\tau = A/\omega_{\text{UGBW}}$. The response of a cascade of M comparators is:

$$H(\omega) = \left(\frac{A}{1+s\tau}\right)^M = \left(\frac{\omega_{\text{UGBW}}}{s+\omega_{\text{UGBW}}}\right)^M$$

An excitation with a step function results in:

$$V_{\text{out}}(t) = A^M v_{\text{in}}(t=0)\left[1 - e^{-t/\tau}\sum_{i=1}^{M}\frac{(t/\tau)^{i-1}}{(i-1)!}\right] \tag{8.1}$$

For $t < \tau$ the cascade of comparators behaves as a cascade of integrators, therefore the response is proportional to $(t/\tau)^M$.

A second problem concerning the speed of this comparator is the recovery from the previous decision. Unless limiter circuits like diodes are used, the last stages of the amplifier will go into saturation. This will cause the inversion charge of MOS devices to be lost. The time required to resupply this charge leads to a delay during the next comparison.

Yet, the straight-forward limiting amplifier is popular as it requires no activation by a clock pulse. Often a string of inverters is used or some simple differential stages.

The second topology of Fig. 8.3 uses a small amplifier as a positive feedback element. As long as this amplifier is much weaker than the feed-forward path it will only marginally contribute to the amplification. What it can do, is add a threshold in the decision process. This so-called hysteresis is discussed in Sect. 8.1.2.

In the last topology the feedback path of the amplifier is of comparable strength to the forward path. The idea of this regenerative stage or latch is that the already build up difference in the forward stages feeds the positive feedback in order to reach a fast decision. This mode of amplification must be reset by a clock phase that clears the data after the decision.

The analysis of two positively fed back amplifiers or a latch is done in the Laplace domain. The nodes are labeled $v_A(t)$ and $v_B(t)$ or in the Laplace domain: $V_A(s)$ and $V_B(s)$. It will be assumed that there is an initial condition $v_B(t=0) \neq 0$.

$$V_A(s) = \frac{A}{1+s\tau}\left(V_B(s) - \frac{v_B(t=0)}{s}\right) = \left(\frac{A}{1+s\tau}\right)^2 V_A(s) - \frac{A}{1+s\tau}\frac{v_B(t=0)}{s}$$

where τ and A have the same meaning as above, but A is negative.

$$V_A(s) = v_B(t=0)\frac{-A(1+s\tau)}{s((1+s\tau)^2 - A^2)}$$

$$= v_B(t=0)\left(\frac{-A/(1-A^2)}{s} + \frac{A\tau/(2(1+A))}{1+s\tau+A} + \frac{A\tau/(2(1-A))}{1+s\tau-A}\right)$$

The inverse Laplace transform gives:

$$v_A(t) = v_B(t=0)$$
$$\times \left(\frac{-A}{(1-A^2)} + \frac{A}{2(1+A)}e^{-(A+1)t/\tau} + \frac{A}{2(1-A)}e^{(A-1)t/\tau}\right) \tag{8.2}$$

Fig. 8.4 A simple regenerative comparator circuit
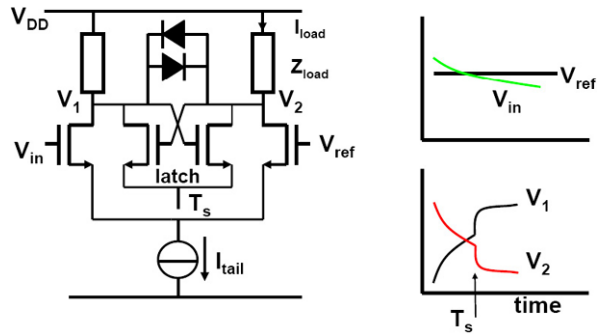
After a few time constants have elapsed only the middle term in brackets is relevant:

$$v_A(t) \approx \frac{-v_B(t=0)}{2} e^{-(A+1)t/\tau} \approx \frac{-v_B(t=0)}{2} e^{\omega_{\mathrm{UGBW}}t} \tag{8.3}$$

The nodes of the latch show an exponential increase determined by the start value and the unity-gain bandwidth. In order to reach a gain comparable to a cascade of M amplifier with a gain A time of $M\ln(A)/\omega_{\mathrm{UGBW}}$ is needed. Ultimately the node voltage will be limited by the circuit and its power supplies. The exponential signal growth makes a latch a fast decision element in a regenerative comparator.

8.1.1 The Dynamics of Transistor Comparator

Most comparators use an input differential pair where the difference is formed between the input value and the reference value. The differential pair allows to create some tolerance for the common mode level of the input signal. This differential current is applied to a positive feedback latch. A clock pulse will activate the latch and amplify the small input voltage difference to a large signal. A simple example is shown in Fig. 8.4.

The gain of the comparator is determined by the transconductance of the input differential pair and the load: $g_{m,\mathrm{in}} \times Z_{\mathrm{load}}$. This gain must be sufficient to suppress any secondary effects in the circuit. On the other hand there is no reason to boost the DC-gain to high values for high speed operation as the unity-gain bandwidth determines the achievable speed performance, see the previous section.

The bandwidth of the comparator must be analyzed both in small-signal mode as well as in large-signal mode. In the first case the input transconductance and the capacitive load of the nodes V_1 and V_2 determine the bandwidth:

$$\omega = \frac{g_{m,\mathrm{input}}}{C_{\mathrm{load}}} \tag{8.4}$$

A large input transconductance is beneficial for a large small-signal bandwidth. Choosing a wide input transistor also helps in reducing the input referred mismatch, but creates a larger capacitive load. Also the parasitic coupling between the drain voltages and the inputs ("kick-back") becomes stronger.

In order to achieve the overall performance also the large-signal bandwidth of the circuit must be considered. Two main issues are crucial for the large-signal bandwidth:

- The fastest change of the signal that is amplified without distortion is limited by the slew-rate:

$$\text{slew-rate} = \frac{dV}{dt} = \frac{I_{\text{tail}}}{C_{\text{load}}}$$

$$C_{\text{load}} \frac{dV_1(t)}{dt} \leq I_{\text{tail}} \tag{8.5}$$

$$2\pi f_{\text{in}} V_{1,\text{max}} C_{\text{load}} \leq I_{\text{tail}}$$

where $V_{1,\text{max}}$ is the amplitude of an equivalent sine wave. If the charging current during the transient of the signal exceeds the tail current, the charging of the capacitors in the circuit will be limited and distortion can follow.

- A large-signal effect in this comparator is saturation.[2] The large signals that drive a comparator will force the input transistors and the internal components in a saturated "on" or "off" regime. Saturation of the input transistors creates significant currents in the gate connection, see Sect. 8.1.5. In order to revitalize the comparator all saturated components will have to be brought back into their linear operation regimes. This process requires current and the signal processing will experience a delay time. This will result in signal distortion. To prevent saturation effects, the comparator of Fig. 8.4 uses two diodes to limit the signal swing on the internal nodes from saturation. Various other circuit techniques can be used to reduce the internal voltage swings (non-linear loads, cascode stages, etc.).

8.1.2 Hysteresis

A comparator is designed to discriminate between positive and negative levels at its input terminal, irrespective how small these levels are and what the previous decision was. In many comparator designs this ideal situation is not achieved. Either intentionally or as an unwanted consequence of the topology, the comparator remembers its previous state. Figure 8.5 shows an elementary comparator consisting of two inverters. The output of the second inverter is partly fed back in positive phase to the input. The comparator input has now a preference to keep its present state and a small input value around the trip level of the inverter will not cause the output to change. Depending on the resistor ratio, the input will see a trip level that depends on the present state of the output. Going from negative input to positive or vice-versa results in different switching characteristics, see Fig. 8.5 (right). This effect is in analogy with magnetic materials called "hysteresis". In this example the hysteresis creates an additional threshold for change. The example circuit in Fig. 8.5

---

[2]The term saturation is used to indicate that the circuit is far out of its operating point. Saturation of circuits has no relation with the operating regime of a transistor.
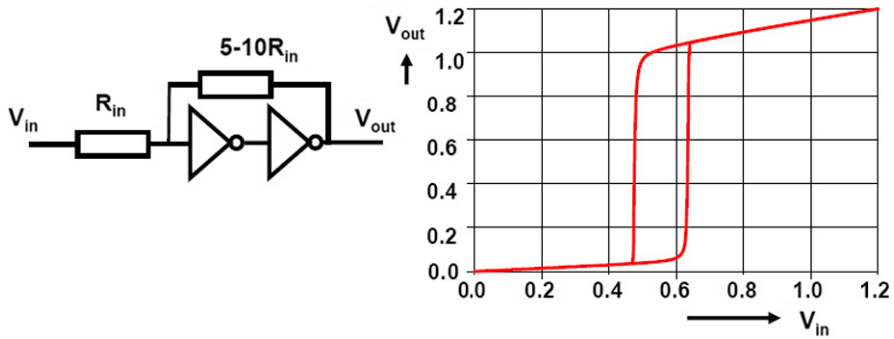
Fig. 8.5 Hysteresis created by positive feedback in a two-stage buffer
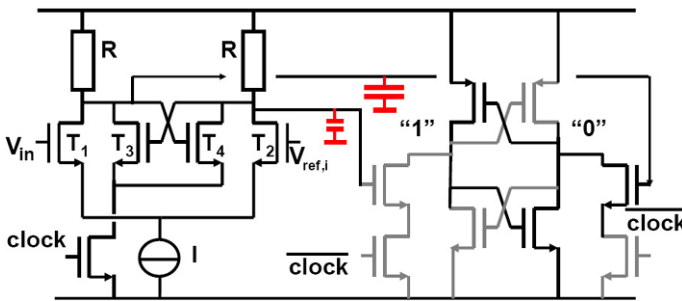
Fig. 8.6 Hysteresis created by the signal-dependent loading of the input latch by the second latch

belongs to the class of "Schmitt-trigger" circuits. These comparators are designed to allow a single decision in case of noise signals. In other words: a Schmitt-trigger circuit decides only if a clear input signal is present and the hysteresis avoids that noise generates false outputs.

A practical example of unwanted hysteresis is a pre-latch stage that is connected to a second stage consisting of a latch activated by the inverse clock. In Fig. 8.6 the signal is fed from the amplifier to the latch via two transistors that are activated via two clocked transistors. The left-hand transistors will be out of inversion when that side of the latch stores a "1" signal. The right-hand transistor sees a "0" signal and is in inversion. Thereby the two coupling transistors create an unequal capacitive loading of the first stage. This loading favors a similar signal in the first latch when the latch sees only a small input signal. The consequence is a comparator with hysteresis.

In other configurations the opposite effect may be the case: the hysteresis favors change and the comparator toggles at every clock cycle for an input signal close to the trip level.

The previous state can be removed by resetting the latch in the comparator. Some designs simply disconnect the latch from the supply. This method assumes that there is sufficient time to discharge the latch nodes. A more active approach
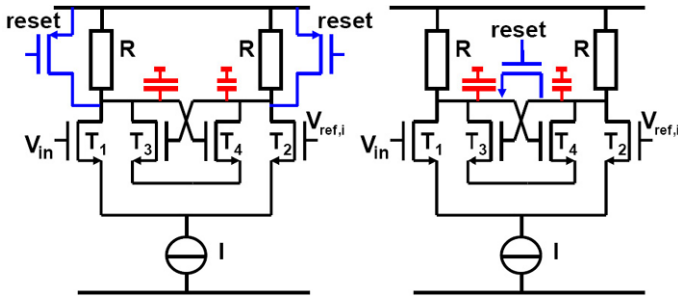
Fig. 8.7 Two methods to clear the state in comparator circuit

consists of adding a specific clear or reset mechanism to the latch. Figure 8.7 shows two configurations. The left-hand circuit connects the latch nodes to the power supply and actively removes all state information. The low-ohmic switches change the DC-levels in the amplification branch. The DC-level must be restored and during that process any inequality in capacitive loading will show up as inaccuracy.

The construction on the right-hand side circumvents the restoration problem by pinching the two latch nodes together. This will keep the DC-level of the nodes in tact and sets the latch to its trip point. A proper choice of the impedance of the switch allows to use it as a differential load for the current of the input pair. The reset transistor must be bootstrapped in case of low power supply.

8.1.3 Accuracy

The required accuracy of a comparator depends on the required specification and the architecture of the analog-to-digital converter. In dual-slope, successive approximation and pipeline designs, the static random offset is a second order effect and generates a (random) DC-shift of the entire signal. In a full-flash converter the input referred random offset is crucial because it will shift the individual reference levels and it impacts both integral and differential non-linearities. Also in multiplexed analog converters the random offset returns as a spurious tone. The input referred random offset $\sigma_{V\text{in}}$ must be designed to a value a factor 5 to 10 lower than the size of the LSB at the comparator as an initial target. Fine tuning of this requirement depends on e.g. the tolerance to conversion errors on system level.

Accuracy has static and dynamic components. Just as in every analog circuit the static accuracy is (in a carefully laid-out circuit) determined by the random mismatch of a transistor pair, see Sect. 11.4. Of course the input differential pair is not the only contributor, also the mismatch of the load and the latch must be taken into account.

In the schematic diagram of Fig. 8.8 current differences caused by the input pair, the load and the latch all come together on the drains of the transistors. All

Fig. 8.8 Mismatch sources in a comparator
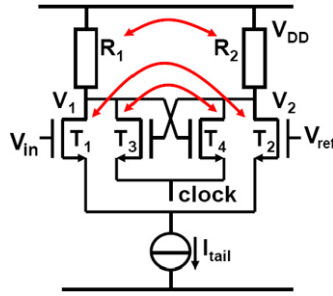
Fig. 8.9 Thermal noise will create a probabilistic behavior of the comparator decision (*dotted line*)
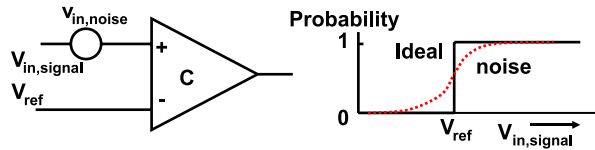
contributions can be referred back to an equivalent input-referred random mismatch $\sigma_{V\text{in}}$. With the help of (2.18), the input referred random offset is described as:

$$\sigma_{V\text{in}}^2 = \left(\frac{\partial V_{\text{in}}}{\partial V_{T,12}}\right)^2 \sigma_{VT,12}^2 + \left(\frac{\partial V_{\text{in}}}{\partial V_{T,34}}\right)^2 \sigma_{VT,34}^2 + \left(\frac{\partial V_{\text{in}}}{\partial R_{1,2}}\right)^2 \sigma_R^2$$

$$\sigma_{V\text{in}}^2 = \sigma_{VT,12}^2 + \frac{g_{m,34}^2}{g_{m,12}^2}\sigma_{VT,34}^2 + \frac{I_{\text{load}}^2}{g_{m,12}^2}\frac{\sigma_R^2}{R_{1,2}^2}$$

(8.6)

The equation assumes that the random mismatch can be calculated in a linearized model of the circuit. Unfortunately in this equation the transconductance of the latch does not behave in a linear fashion. In one extreme there is no current running in the latch and its contribution to the input-referred random offset is zero. Directly after the clock is activated and a current flows in the latch transistors, a step function on the nodes V_1 and V_2 will occur. If the gates are not equal or in the presence of mismatches due to differences in capacitive loading, different charges are drawn from the local nodes and will create random differences in these voltage steps.

If, on the other hand, the latch is constantly fed with a small tail current that is not sufficient to activate the latch, the dynamic errors can largely settle before the clock is activated. However the non-zero transconductance will contribute an additional component to the input-referred random offset.

Thermal noise and $1/f$ noise create additional accuracy limitations. Because these contributions are time-varying, also offset compensated analog-to-digital converter architectures suffer from this limitation. The contributions to the noise of the individual components are referred back to an equivalent input noise source in a

similar manner as for mismatch. If the input transconductance dominates the noise, the effective value of the input referred noise is:

$$v_{\text{in,noise}} = \sqrt{4kT \frac{BW}{g_m}} \tag{8.7}$$

with a bandwidth of 10 GHz and an input transconductance of 1 mA/V the value for the input referred rms noise is $v_{\text{in,noise}} = 400$ µV. The thermal noise amplitude distribution creates a cumulative Gaussian probability distribution around the trip level, Fig. 8.9. In order not to exceed the accuracy specifications the bandwidth has to be reduced or more power has to be spent on improving the impedance levels.

8.1.4 Metastability and Bit-Error Rate

In the comparator a fundamental problem arises in the form of metastability. Metastability is associated with any form of comparison and is also well-known in e.g. synchronization circuits. The crucial observation is that there is a fraction α of the LSB size were the comparator has insufficient voltage difference on its nodes to reach a decision in a limited time T_s. Now the comparator will not generate a clear "zero" or "one" output level. The succeeding logic circuitry may generate a large error, e.g. if this digital signal is crucial for determining the MSB.

For small signals the latch can be viewed as two single-pole amplifiers in a positive (regenerative) feedback mode. The voltages on the nodes of the latch will develop exponentially in time with a time constant τ, see (8.3). This time constant is determined by the internal node capacitance and the transconductance of the latch transistors. Now α can be approximated by:

$$\alpha \approx \frac{V_{\text{latch}}}{V_{\text{LSB}}} e^{-T_s/\tau} \quad \Rightarrow \quad \text{BER} \approx 2^N e^{-T_s/\tau} \tag{8.8}$$

where the ratio between the maximum latch swing V_{latch} and V_{LSB} is estimated as 2^N. The bit error rate is an important parameter in the design of fast converters with a high accuracy. In CMOS the time constant τ is formed by the parasitic and gate capacitances and the achievable transconductance. A typical example with 8 bits is: 5 fF total capacitance for 1 µm gate width, with 5 µA/V transconductance for the same gate width and $T_s = 20$ ns, results in a BER of 10^{-7}. This BER can be improved to better than 10^{-8} by means of more current in the latch transistors. For converters with sample rates in excess of 100 Ms/s a BER of the order 10^{-8} means one error per second. Especially in industrial and medical equipment such an error rate can be unacceptable. A method to get an impression of the bit error rate is to apply a slow sine wave to the device with an amplitude that guarantees that no more than one LSB change occurs at the digital output. Now the bit error rate can be estimated by recording output codes that differ more than one bit from the preceding code.
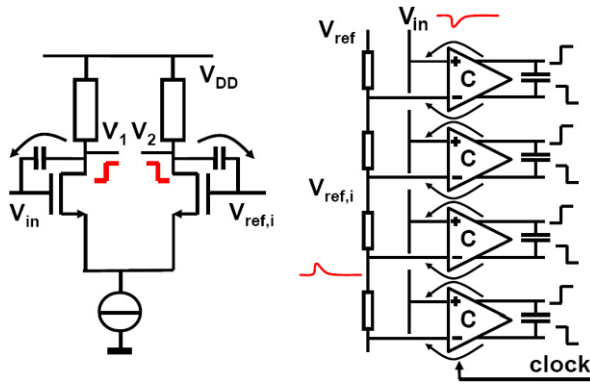
Fig. 8.10   Kick-back in a comparator circuit

From a fundamental point of view the BER can not be avoided completely, however decreasing the time constant (by lower capacitance and higher transconductance), a BER of 10^{-13} is possible. Measures to improve the BER include: improving the latch speed with more current and smaller capacitances, additional gain stages or even a second latch, and a special decoding scheme avoiding large code errors due to a metastable state.

8.1.5 Kick-back

The comparator is a non-linear element: the only relation between the input signal and the output is the sign. The decision process in the comparator is often implemented with transistors that pass through all operation modes of the devices. The charges controlling the devices will show considerable variation. These charges have to be provided and will be dumped preferably in the power supply lines, however also an exchange to the input and reference signals will occur.

Figure 8.10 (left) shows a differential input pair. The input voltage difference will lead in the succeeding circuits to a decision marked by sharp voltage transitions. Despite that the input differential pair remains in normal operating mode, the drain-gate capacitors of these MOS devices will pass these fast edges to the input and reference nodes thereby generating "kick-back". In some comparator designs these transitions can be as large as the power supply voltage. Figure 8.10 (right) shows a full-flash design. In this architecture the kick-back effect is strong as the contributions of many comparators add up. During the simulation of the comparator ideal voltage sources instantly drain the charges from saturation effects and therefore will result in an optimistic performance prediction. Realistic impedances at the input terminals of the comparator allow a better simulation prediction of the performance in the presence of kick-back.

Reducing the kick-back signal requires to decrease the dimensions of the input related circuit elements and lower the swing of the signals. Unfortunately these

Fig. 8.11 A comparator in CMOS based on an NMOS enhancement/depletion design [146]
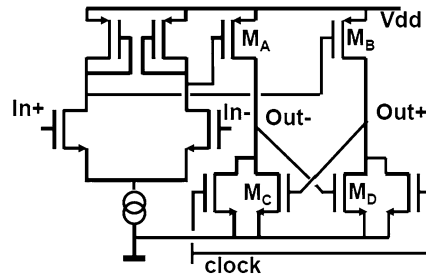
measures compromise the performance. Small input transistors cause high random offset. Lowering the swing of the transitions requires some form of separation between the full-swing digital output and the input. Separation stages, however, will create additional delay.

8.1.6 Comparator Schematics

Many comparator designs have been published in literature and there is obviously not a standard circuit topology. This is on one hand caused by the evolution of the technology. The drive capabilities of the transistors have improved, but the power supply voltage and the input signals have reduced. On the analog-to-digital converter architecture side also a lot of developments have taken place over the last 25 years. The rise in popularity of the pipeline converter has allowed to spend more area and power on a comparator than a flash converter can tolerate. The following examples of comparator circuits are a subset of the published work and are meant to educate and perhaps inspire a new design.

The comparator published in [146] was designed in a 7 μm NMOS metal-gate technology. Figure 8.11 shows a CMOS variant of that design. A similar topology is found in [131, Fig. 13]. The input stage serves to amplify the differential signal and allows some common mode rejection. The second stage is fed with the amplified signal to generate a digital decision. The intermediate amplification nodes separate the pre-amplifier from the latch. The problem with an intermediate node is that it creates an additional pole and slows down the circuit. In the design of Fig. 8.11 the poles are formed by the PMOS mirrors. These mirrors cannot easily be operated at a large gate-source voltage as the input range is reduced. This mirror also contributes to the input random offset. The strict separation of latch and input allows a small kick-back effect.

The original design (1981) resulted in a random offset at the input of 6.1 mV, 5 bit resolution, 4 MHz signal bandwidth at 20 Ms/s.

The comparator in Fig. 8.12 consists of a preamplifier (left) and a latch. The preamplifier does not use a differential pair with tail current, but the inputs are capacitively coupled into the comparator. In the cross-coupled loop formed by the source followers and the capacitors, capacitive voltage shifting is used. The residual offset is 2 mV and in a 2 μm process 40 Ms/s was reported.
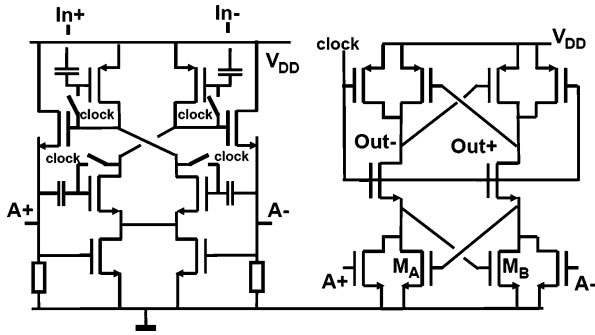
Fig. 8.12 A comparator based on the work in [147]

Fig. 8.13 A comparator
designed for a folding
analog-to-digital converter
[148]
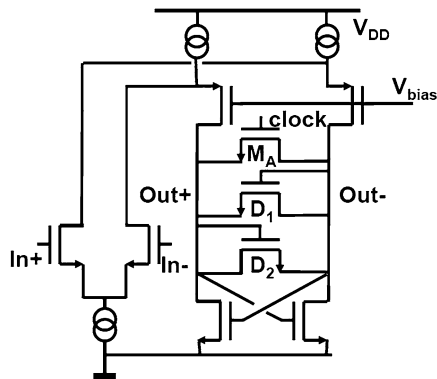
The comparator in Fig. 8.13 is designed for a folding analog-to-digital converter. It consists of an input stage that feeds its differential current in a cascode stage. This construction avoids an intermediate pole which would slow down the circuit. On the other hand a reasonable shielding for kick-back is obtained. The latch is equipped with two diode-connected transistors to have a current path available even if the latch is activated. Saturation of the input stage is thereby prevented. In a 0.8 μm CMOS process the 1-sigma input referred random offset is 2.5 mV. The circuit runs up to 70 Ms/s sample rate in this process. Its speed will certainly benefit from advanced processes.

The comparator in Fig. 8.14 is based on former bipolar designs. The comparator routes a constant current through either the amplification side (clock is high) or through the latch. The speed is limited by the latch pair and its load. A lot of imperfections add up in the current path, therefore some more input referred mismatch is expected. Also the kick-back noise can be significant. This effect will depend on how well the drains of the input pair are kept stable. The accuracy of the crossing of the clock and the inverse clock is crucial. The design was used in a 0.5 μm CMOS process with a clock speed of 80 Ms/s.

A comparator for a 4-bit 12 Gs/s analog-to-digital converter uses a two stage comparator, Fig. 8.15. The middle PMOS in the pre-amplifier serves as a load for

Fig. 8.14 A comparator designed for a folding analog-to-digital converter [149]
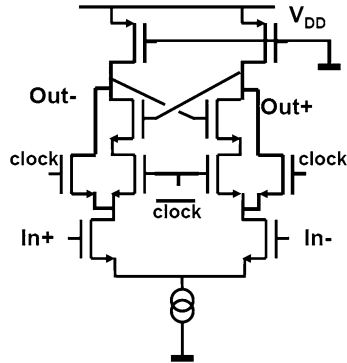
Fig. 8.15 A comparator for high speed operation [150]. The latch (*right*) is based on a design for the "StrongArm" processor [151]
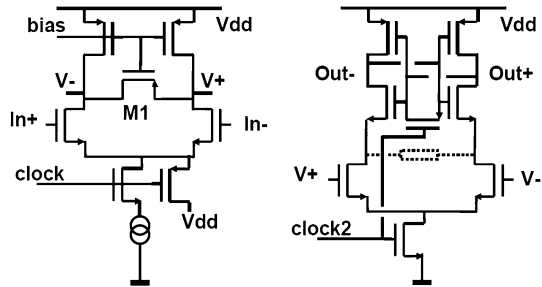
the input pair. The sources of the input pair are switched to the positive power supply at the end of the amplification, which produces kick-back noise via the coupling of the gate. The latch (right) is also used in the "StrongArm" design [151] and is not intended to resolve small voltage differences as it serves as an edge-triggered latch. If the clock is low the cross-coupled pair is reset and no current flows. A high-ohmic resistor (dotted) prevents leakage currents to charge nodes asymmetrically. At rising clock the input signal is amplified and regenerated in the latch to a full digital signal. After full settling the latch is consuming no current. The 1-sigma input referred random offset in [150] was in 0.25 μm CMOS 60 mV and the bandwidth exceeded 2.5 GHz at 12 Gs/s. Another example of an implementation of this comparator is in a flash converter [152] where a good energy efficiency is achieved.

8.1.7 Auto-zero Comparators

Input referred random offset is one of the key problems to battle in comparator design. Early on attempts have been made to cancel [154] or "auto-zero" [153] this offset. The basic schematic of Fig. 8.16 is similar to the track-and-hold circuit in Fig. 4.19. During the auto-zero mode the switches are in the dashed positions and the inverter is essentially in unity gain mode. The capacitor is charged on its left hand side to V_{ref} and on the right-hand side to the input offset. When

Fig. 8.16 An auto-zero comparator as used in [153]
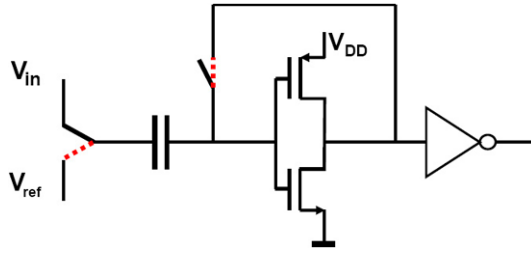
the switches toggle the change on the input of the comparator equals $V_{in} - V_{ref}$. This voltage is amplified with the small-signal gain of the inverter and its successors. Some residual offset is due to limited gain and charge variation in the switches. This scheme was used in a 1.4 μm CMOS process [153] and allowed an analog-to-digital converter running at 40 Ms/s with 10 MHz bandwidth. It is clear that the charging and discharging of the input capacitor puts extreme demands on the input and the reference impedances. The implementation as depicted in Fig. 8.16 has the additional disadvantage that the power-supply rejection ratio (PSRR) is low. A large part of the power supply variation is translated in a signal contribution on the input. In the case of an inverter the PSRR is around 0.5. So the power supply has to be kept clean to the level of a few V_{LSB}. Also the timing of the offset-cancelation is a point of consideration. Offset compensation per clock cycle costs 30–50% of the available sample period and reduces the maximum speed. Offset compensation at a fraction of the sample rate is possible, but can introduce spurious components at those lower frequencies.

The auto-zero cancellation, as shown in Fig. 8.16, acts as a transfer function for all (unwanted) signals $e(z)$ that originate at the input of the inverter. The error component in the output signal is found in a similar manner to the analysis in Sect. 4.4. The error component in the output signal is:

$$V_{out,error} = e(z)(1 - z^{-0.5}) \qquad (8.9)$$

The exponent -0.5 assumes a 50% duty cycle of the switching signal. This transfer is characterized by a term $2\sin(\pi f/2f_s)$. Low frequency components are suppressed, but higher frequencies can even experience some amplification. Moreover the switching sequence acts as a sampling mechanism for unwanted high-frequency signals thereby shifting these high frequencies into the signal band. The auto-zero capacitor acts as a sampling capacitor for which the kT/C noise analysis applies. Running this offset compensation scheme at low frequencies results in stacking of noise and down-sampling of spurious signals (e.g. from the substrate). See also the current-calibration technique in Sect. 7.4.

A more advanced offset canceling comparator is shown in Fig. 8.17. The design is differential, thereby eliminating the PSRR problems. The inherent comparator is estimated at 36 mV 1-sigma random input offset. Still some time is needed for the auto-zeroing in this 0.18 μm CMOS design. Interleaving of two input stages reduces in this converter this speed penalty.

Fig. 8.17 A comparator
design for an interleaved
pipeline converter [87]
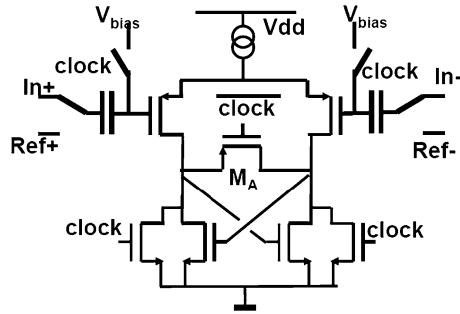
Fig. 8.18 A "full-flash"
analog-digital converter
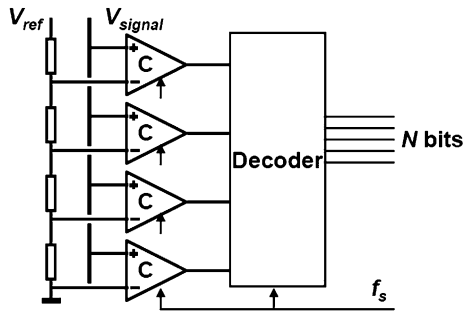
8.2 Full-flash Converters

Full-flash[3] converters are used for two main purposes. As a stand-alone device, this
converter can achieve the highest conversion speeds for low (6-bit) resolutions. An-
other important field is the application in lots of other analog-to-digital conversion
architectures, such as subrange converters. A full-flash converter is comprised of a
resistor ladder structure whose nodes are connected to a set of $2^N - 1$ compara-
tors, see Fig. 8.18. A decoder combines the comparator decisions to a digital output
word.

After the sampling clock becomes active, a part of the comparator circuits with an
input signal lower than the local ladder voltages will generate a logical "zero", while
the other comparators will show a logical "one". The digital code on the outputs of
the comparators is called a "thermometer code". A digital decoder circuit converts
the thermometer code in an N-bits output format (mostly in straight binary format).

The input signal for a full-flash converter is only needed at the moment the
latches are activated. The sample-and-hold function is inherently present in the dig-
ital latches of the comparators. For most applications an external sample-and-hold
circuit is not needed. Full-flash converters are the fastest analog-to-digital convert-
ers, however their complexity and their power consumption grows with the number

---

[3]The origin of the addition "full" in full-flash can refer to the conversion of the full range. "Partial-
flash" converters can refer to a sub-ranging architecture.
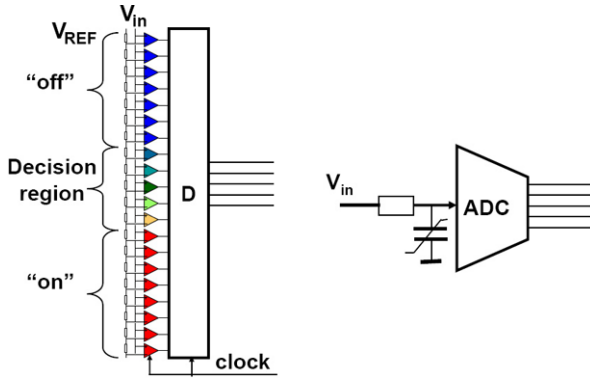
Fig. 8.19 A "full-flash" converter presents a non-linear input impedance

of comparator levels 2^N. Also the area and the input capacitance grow exponentially with N.

The capacitance at the input of a full-flash converter is largely determined by the input capacitance of the comparator. If that input capacitance is formed by e.g. a differential pair, the effective input capacitance of the comparator will depend on the input signal. Low input signals with respect to the local reference voltages of the comparators will result in a current starved input branch of the differential pair and a high signal will keep the input transistor in inversion. An input signal lower than the local reference voltage will see a low input capacitance as the MOS input transistor is out of inversion. A high input signal creates an inversion charge, a considerable input gate-source capacitance and a high input capacitance. The group of comparators below the decision level with a high input capacitance and the group above the decision level with a low capacitance, Fig. 8.19, make the input impedance of a full-flash converter signal dependent. Second-order distortion will arise if the converter is fed from a source with source impedance. If the input capacitance is defined as:

$$C(V_C(t)) = C_0 + \Delta C \frac{V_C(t)}{V_{\text{ref}}} \tag{8.10}$$

with V_C as the voltage over the non-linear capacitance, the capacitive the capacitive current is found using (2.52):

$$
\begin{aligned}
I_C &= C(V_C(t)) \frac{dV_C(t)}{dt} + V_C(t) \frac{dC(V_C(t))}{dt} \\
&= C_0 \frac{dV_C(t)}{dt} + \frac{2V_C(t)\Delta C}{V_{\text{ref}}} \frac{dV_C(t)}{dt}
\end{aligned} \tag{8.11}
$$

A correct analysis requires to equate this current to the current through the resistor, resulting in a non-linear differential equation. The substitution $V_C(t) = 0.5V_{\text{ref}} + 0.5V_{\text{ref}} \sin(\omega t)$ assumes that the non-linear term is relatively small and will not substantially change $V_C(t)$. Evaluation of the last part of the equation leads
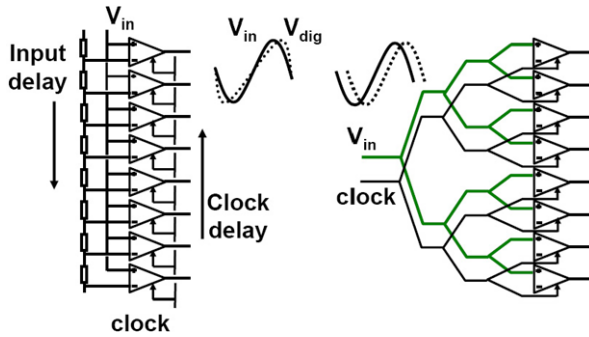
Fig. 8.20 A simple distribution strategy for input signal and clock can easily lead to performance loss at high frequencies (*left*). A tree-like lay-out is necessary to avoid delay differences

to a second harmonic current, which is multiplied with the input resistor R and compared to the first harmonic term in $V_C(t)$:

$$\text{HD2} = \frac{\omega R \Delta C}{2} \tag{8.12}$$

This estimate for HD2 is proportional to the signal frequency, the input impedance and the amount of capacitive variation. For $f = 1$ GHz, $R = 50\ \Omega$ and a capacitive variation of 100 fF a distortion level of -37 dB results.

Full-flash converters are predominantly used in high-speed applications. These applications require tight control on the timing of the converter. Jitter control is crucial to high performance, see Sect. 3.1.4. However, also in the actual design of the converter attention has to be paid for balancing the timing accurately. The impedance of the wiring in combination with a string of load elements (e.g. inputs to the comparators) can easily lead to delay differences between individual comparators. In the example of Fig. 8.20 (left) the clock and the signal come from opposing sides in the structure. For signal levels close to the bottom of the structure the signal will be delayed with respect to the sample clock, while at levels close to the top the signal will be advanced with respect to the sample clock. In a structure where the signal conversion is linearly related to the position of the comparators, the relative delay (ΔT) of the signal versus the clock is proportional to the signal.

$$
\begin{aligned}
V_{\text{in}}(t + \Delta T(t)) &= V_a \sin(\omega(t + \Delta T \sin(\omega t))) \\
&\approx V_a \sin(\omega t) + V_a \omega \Delta T \sin(\omega t)\cos(\omega t) \\
&= V_a \sin(\omega t) + 0.5 V_a \omega \Delta T (1 - \cos(2\omega t)) \tag{8.13}
\end{aligned}
$$

This timing error translates in a second order distortion component with a relative magnitude of $0.5\omega\Delta T$. Even with modest signal frequencies of e.g. 10 MHz and a delay of 100 ps this results in a -50 dB distortion component. This example shows that the relative timing of the signal and sample must be accurate. Therefore a tree-like structure as in Fig. 8.20 (right) is applied in high performance converters.

8.2.1 Ladder Implementation

In Sect. 7.2.1 the dynamic behavior of a resistor string was analyzed. This theory is applicable to the design of the ladder structure in the full-flash converter. The time-constant for settling was found to be:

$$\tau = rcL^2/\pi^2 \tag{8.14}$$

with r and c the resistivity and capacitance per unit length. In a full-flash converter with N bit resolution this equation is rewritten as:

$$\tau = R_{\text{tap}}C_{\text{tap}}2^{2N}/\pi^2 \tag{8.15}$$

R_{tap} is the resistance between two tap positions on the resistor ladder and C_{tap} represents the total capacitance on a tap and is composed of the input capacitance of the comparator, the bottom-plate stray capacitance of the resistor and all wiring parasitics. With $C_{\text{tap}} = 0.1$ pF, $R_{\text{tap}} = 1\ \Omega$ and $N = 7$ a time constant $\tau = 0.16$ ns will result, which would allow a sampling speed of around 1 Gs/s. The ladder impedance is 128 Ω. A 1 V reference voltage over the ladder already requires 8 mA.

In a similar manner as the time constant, the DC-current that e.g. bipolar transistors or some auto-zeroing schemes can draw from the ladder can be estimated:

$$\Delta V_{\text{middle}} = R_{\text{tap}}I_{\text{tap}}2^{2N}/\pi^2 \tag{8.16}$$

With the same parameters and $I_{\text{tap}} = 10^{-6}$ A, the voltage deviation in the middle of the ladder equals 1.6 mV. This loading effect of the ladder by the comparators is more pronounced in bipolar design. Darlington stages in the comparators are used to reduce the loading of the ladder by base currents and at the same time reduce the kick-back effects [155].

This example shows that even modest specifications in a standard full-flash converter require a low-ohmic ladder. Often the ladder is constructed from the metal layers or special materials on top of the device. These materials can easily exhibit gradients and although no more than 6–7 bit of resolution is required, some precautions have to be taken to mitigate the gradients. An anti-parallel connection of two ladders reduces the gradient, Fig. 7.9.

8.2.2 Comparator Yield

The key requirements for the choice of a comparator are a low capacitive load, no DC input current, low random offset, low kick-back, high switching speed and bandwidth and low power. With the exception of low random offset, most of these requirements can be met by using small size transistors.
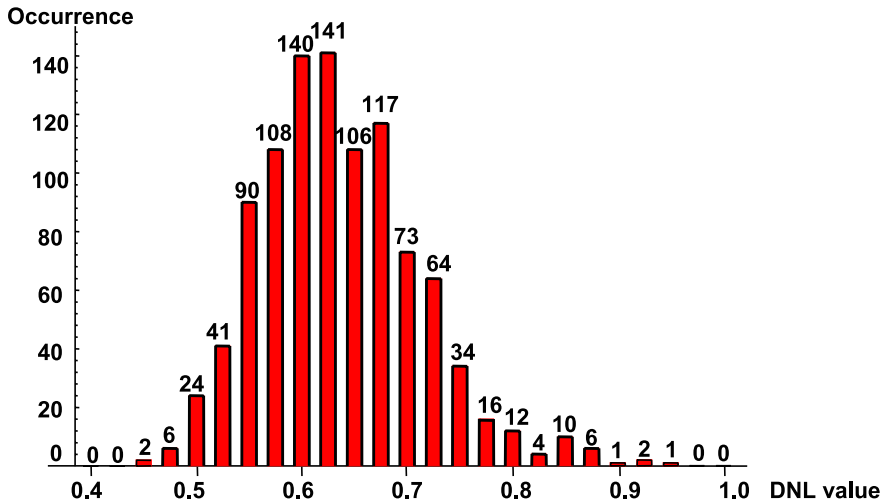
Fig. 8.21 Typical histogram of DNL values for an 8-bit full-flash architecture, with $\sigma = 0.15 V_{\text{LSB}}$. The mean DNL in this simulation is 0.64 LSB

One of the major differences in the design of CMOS and bipolar analog-to-digital converters is the lack of accuracy in CMOS comparators. A bipolar differential pair has a random offset on the base-emitter voltage V_{be} in the order of $\sigma_{\Delta Vbe} = 0.3$ mV. A pair of NMOS transistors with small gate lengths show random offsets in the order of $\sigma_{\Delta VT} = 2$–6 mV, see Sect. 11.4. Because of the low CMOS gain, the offsets of the entire comparator accumulate, which may lead to $\sigma = 5$–20 mV as a total input-referred random offset.

The random comparator offset in an N-bit full-flash analog-to-digital converter (with $2^N - 1$ comparators) affects the DNL or non-monotonicity (DNL $= -1$ LSB). The DNL for a converter with 2^N conversion levels is specified as:

$$\text{DNL} = \frac{V_{j+1} - V_j}{V_{\text{LSB}}} - 1, \quad \forall j = 0, \ldots, 2^N - 2 \tag{8.17}$$

V_j is the value of the input signal which causes comparator j to flip and consequently contains the comparator random input referred offset. V_{LSB} is the physical value corresponding to an LSB.

The DNL is often used as an elimination criterium and determines the yield. The actual value of the DNL, being the maximum of $2^N - 1$ differences of stochastic variables, is a random function itself. Figure 8.21 shows the distribution of the actual DNL of 1000 8-bit full-flash converters, generated by means of Monte-Carlo simulation. The shape of the distribution is characteristic for production measurements of various types of converters. The distribution is not Gaussian as the DNL is a non-linear function. The comparator random offset was in this example chosen to be $\sigma = 0.15 V_{\text{LSB}}$, corresponding to $\sigma = 0.586 \times 10^{-3} V_{\text{ref}}$. There are almost no trials with an actual DNL < 0.5 LSB, although all trials result in monotonicity. For

Fig. 8.22 Yield on monotonicity versus the standard deviation of the comparator random offset
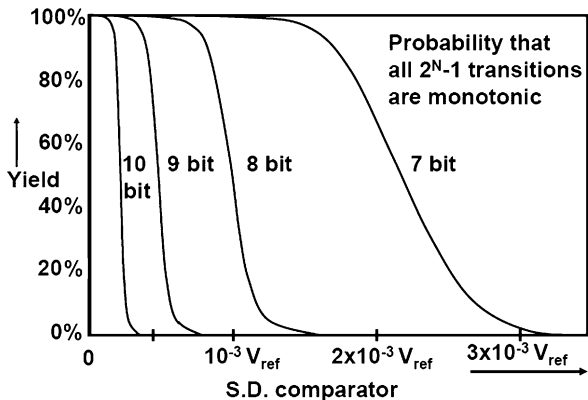
half of the trials the actual value of the DNL is lower than 0.64 LSB, which is also found in the following first-order calculation:

$$\text{Yield} = (1 - p)^{2^N - 2} < 0.5 \quad \rightarrow \quad p > 0.9973$$

$$\text{from Table 2.10} \quad p = 0.9973 \quad \leftrightarrow \quad \alpha = 3.0 \quad \rightarrow \quad \frac{\text{DNL} \times V_{\text{LSB}}}{\sigma \sqrt{2}} = 3.0 \quad (8.18)$$

$$\text{with } \sigma = 0.15 V_{\text{LSB}} \quad \rightarrow \quad \text{DNL} = 0.64$$

The $\sqrt{2}$ factor in the second line stems from the fact that the DNL is due to the difference between two comparator trip levels.

A more formal analysis results in a yield prediction. If the input random offset of every comparator is given by a Gaussian distribution with zero-mean and a standard deviation σ, then the probability of all the comparators being within the monotonicity limit can be calculated. This probability is an estimation of the yield of the analog-to-digital converter. Ideally, $(V_{j+1} - V_j) - V_{\text{LSB}} = 0$. Non-monotonicity occurs when comparator $j + 1$ (fed with a rising input signal) switches before the adjacent comparator j with a lower reference voltage does. In mathematical formulation the probability that non-monotonicity occurs, is: $p = P(V_{j+1} < V_j)$. Then $(1 - p)$ is the probability of comparators j and $j + 1$ switching in the correct order; this condition must hold for all $2^N - 2$ pairs of comparators, so:

$$\text{Yield} = (1 - p)^{(2^N - 2)} \quad \text{with}$$

$$p = P\left(\frac{V_{j+1} - V_j - V_{\text{LSB}}}{\sigma \sqrt{2}} < \frac{-V_{\text{LSB}}}{\sigma \sqrt{2}}\right) \quad (8.19)$$

The mean value of the argument in the probability function for p is 0, while the standard deviation of the argument corresponds to $\sqrt{2} \times$ the standard deviation of a single comparator σ in millivolts. The probability function is normalized to a standard normal curve $N(0, 1)$.

Figure 8.22 shows the curves that relate the yield to the standard deviation σ of the comparator random offset for an input voltage range of V_{ref}. A 10 bit converter requires a $\sigma < 0.25 \times 10^{-3} V_{\text{ref}}$. This is still achievable in bipolar technology [156].

Fig. 8.23 Yield on DNL < 0.5 LSB, on monotonicity for 8-bit full-flash and on monotonicity for an 8-bit 2-step subrange architecture
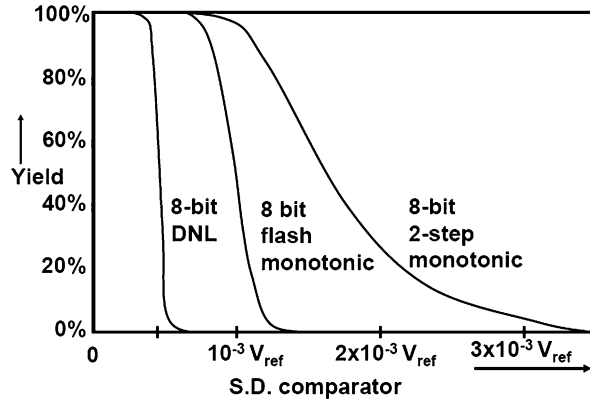
For higher accuracies trimming, higher input voltages or other forms of offset correction are needed. The 8-bit converter from Fig. 8.21 with $\sigma = 0.586 \times 10^{-3} V_{\text{ref}}$ is indeed monotonic with close to 100% yield.

At first sight one bit more resolution requires that the comparator random offset should reduce by a factor of two. A 7-bit full-flash converter has to reach an input standard deviation of $1.8 \times 10^{-3} V_{\text{ref}}$ to achieve an acceptable yield. At the 8 bit level comparators with random offset better than $0.8 \times 10^{-3} V_{\text{ref}}$ are needed for the same yield. One bit more resolution translates in a factor two reduction of the random offset standard deviation. Moreover, the same yield must be reached with twice the number of comparators resulting in a random offset reduction factor of 1.8/0.8, slightly higher than 2.

In Fig. 8.23 the requirements are more severe: now the probability of the converter achieving a DNL of less than 0.5 LSB has been calculated. The probability p of two adjacent comparators exceeding the DNL limits is:

$$\text{Yield} = (1 - p)^{(2^N - 2)} \quad \text{with}$$
$$p = P\left(\left|\frac{V_j - V_{j-1} - V_{\text{LSB}}}{\sigma\sqrt{2}}\right| > \frac{\text{DNL} \times V_{\text{LSB}}}{\sigma\sqrt{2}}\right) \tag{8.20}$$

As expected, the σ required for a DNL value of 0.5 LSB has been more than halved with respect to the non-monotonicity requirement.

The demands for a two-stage subranging architecture are also shown for (a theoretical minimum of) 15 coarse and 15 fine comparators. Owing to the steep nature of the Gaussian distribution the advantage of having only 15 critical comparators results in marginally more tolerance on the input random offset voltage at a 95% to 99% yield level. This example indicates that these yield considerations for full-flash architectures give a good first-order approximation for more complex architectures.

Table 8.1 shows the effect of increasing the resolution with one bit in a full-flash analog-to-digital converter. In this table it is assumed that the comparator's design comprises three transistor pairs that contribute to the random offset: the input pair, a current source pair and a latch pair.

Table 8.1 Comparison of N and $N + 1$ bit full-flash analog-to-digital converters

		N bit	$N + 1$ bit
1	Input range	V_{ref}	V_{ref}
2	LSB size	$2^{-N} V_{\text{ref}}$	$2^{-(N+1)} V_{\text{ref}}$
3	Number of comparators	$2^N - 1$	$2^{(N+1)} - 1$
4	Probability per comparator pair for 95% ADC yield	$p_N = \sqrt[2^N-1]{0.95}$	$p_{N+1} = \sqrt[2^{(N+1)}-1]{0.95} \approx \sqrt{p_N}$
5	Required σ's in $N(0, \sigma)$	$S_N \approx 3 \ldots 4$	$S_{N+1} \approx S_N + 0.3$
6	Input referred random error	$2^{-N} V_{\text{ref}}/S_N \sqrt{2}$	$2^{-(N+1)} V_{\text{ref}}/S_{N+1} \sqrt{2}$
7	MOS pairs in comparator	3	3
8	Random error per pair	$\sigma_N = 2^{-N} V_{\text{ref}}/S_N \sqrt{6}$	$\sigma_{N+1} = 2^{-(N+1)} V_{\text{ref}}/S_{N+1} \sqrt{6}$
9	Area per MOS	$WL = A_{VT}^2/\sigma_N^2$	$WL = A_{VT}^2/\sigma_{N+1}^2$
10	Capacitance of all gates	$3 \times 2^N A_{VT}^2 C_{ox}/\sigma_N^2$ $= 18 S_N^2 2^{3N} A_{VT}^2 C_{ox}/V_{\text{ref}}^2$	$3 \times 2^{N+1} A_{VT}^2 C_{ox}/\sigma_{N+1}^2$ $= 18 S_{N+1}^2 2^{3(N+1)} A_{VT}^2 C_{ox}/V_{\text{ref}}^2$

The first three lines specify the range and number of comparators in a full-flash converter. In line 4 the required overall yield of e.g. 95% is recalculated to the required probability that a pair of adjacent comparators remains monotonic. This number is close to 1. For one bit more and double the number of comparators, this probability is even closer to 1. Now in line 5 a table for a normal distribution is used to find the number of sigma's needed to reach this probability outcome. In line 6 the input referred mismatch is found by dividing the value of one bit by this number of sigma's. Another division by $\sqrt{2}$ accounts for the step from a difference between two comparators to a single comparator. In lines 7 and 8 this mismatch budget is divided over 3 relevant pairs of transistors in a comparator. In line 9 the required gate area is shown and finally line 10 gives the total capacitance per analog-to-digital converter. The input capacitance is dominated by the resolution 2^{3N}, yet it remains important to come to a high ratio between reference voltage (equals the signal swing) and mismatch coefficient. The difference between the input capacitance of an N bit and an $N + 1$ bit converter is:

$$\frac{C_{\text{in},N+1}}{C_{\text{in},N}} = \frac{8 S_{N+1}^2}{S_N^2} \approx 10 \quad \text{for } N = 5, \ldots, 8 \tag{8.21}$$
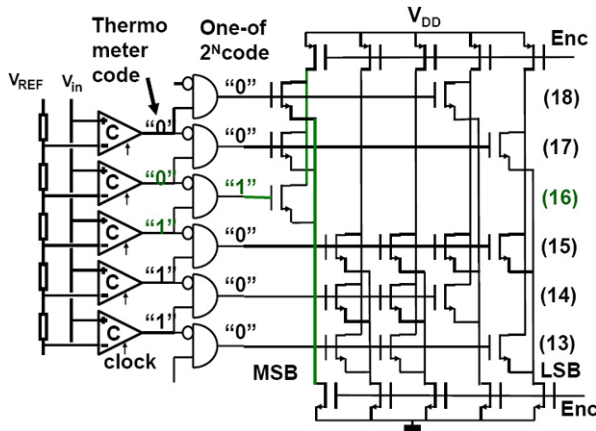
Fig. 8.24 A wired NOR based decoder scheme

8.2.3 Decoder

In many analog-to-digital converters the decoding of the comparator decisions into a digital output word is not a problem, only the algorithm behind the decoding is interesting. Figure 8.24 shows a basic wired NOR decoding scheme for a full-flash converter. The decoding starts by a simple gate that converts the thermometer code at the output of the comparators in a 1-of-N code. This is the digital form of the mathematical derivative function. The function will indicate the pair of comparators where the input signal is between the reference values. The corresponding decode line is connected to a matrix of transistors laid-out in a straight binary code. The above sketched operation can be disturbed by various mechanisms. If mismatch creates a situation where a lower or higher comparator switches too, more than one wired-NOR input lines will be active. These errors are called "sparkles" or "bubbles". This situation can also occur in the presence of high slew rate signals or meta-stability errors. Most of these sparkles will create a deviating code, however, if the sparkle affects a decode region around a major bit, full-scale errors can be the result. There are numerous ways to avoid that these sparkles upset the decoder. Figure 8.25 simply extends the thermometer decoding with an additional input. Many other solutions exist where different trade-offs with speed are made.

In many medium performance applications this scheme will work fine. However if the operating frequency is increased to the limits of the technology, performance problems arise. At high operating frequencies the inherent capacitive load of a wired NOR structure becomes a burden. Buffering is required at the cost of power. In advanced processes the parasitic wiring capacitance fortunately reduces significantly thereby allowing fast processing of the decoding signals.

The above decoding schemes show a disadvantage in case of meta-stability. A meta-stable comparator output in Fig. 8.24 will affect two decode gates and their associated decode lines. If the meta-stable condition continues, opposing signals may appear in the two decode gates and a major decoding error will happen.

Fig. 8.25 A wired NOR based decoder scheme with bubble correction
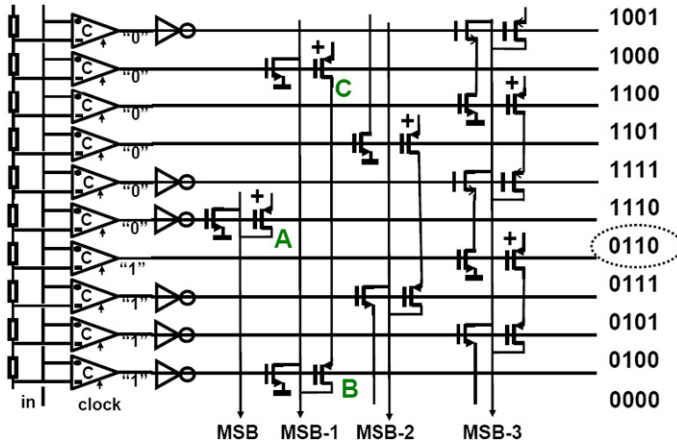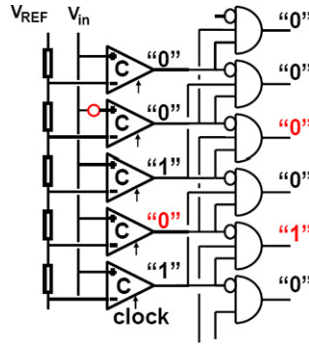
Fig. 8.26 A wired Gray-decoder scheme

The MSB line in the Gray-decoding scheme of Fig. 8.26 is controlled via a comparator and two inverters. The last inverter, labeled "A", drives the output line. The MSB-1 uses two gates "B" and "C". For a signal on the input terminal lower than the reference voltage of this scheme, all comparators will signal a logical zero. Gate "B" receives a logical one and the NMOS will pull the line to a logical zero state. After the input signal increases to the level corresponding to the encircled code in Fig. 8.26 the input of gate "B" will go to logical zero just as gate "C". Via both PMOS devices the MSB-1 line is set at a logical one state. If the input signal exceeds the reference levels in this drawing, the input to gate "C" will be a logical one, pulling the MSB-1 line to zero.

Each comparator in a Gray decoder sets one single output line. A potential metastable condition does not spread out over more than one cell and can still be resolved if more time is allowed. The Gray-code is an often used strategy in the first part of the decoder. For large decoders the remaining decoding can be done in conventional style.
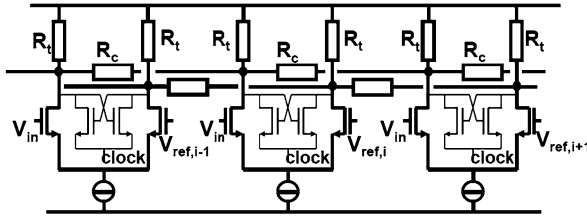
Fig. 8.27 The averaging scheme after [157]

Fig. 8.28 Starting point for the analysis is the equivalent resistor network *on the left*. Equation (8.23) is shown *to the right* [158]
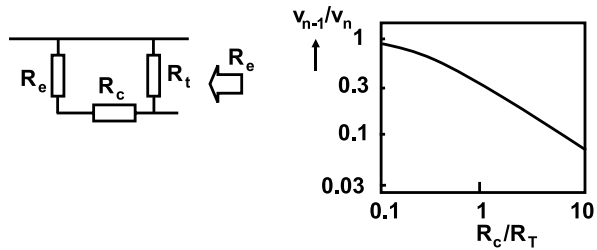
8.2.4 Averaging and Interpolation

The main problem for improving the accuracy of a full-flash converter is the mismatch of the comparator stages. This mismatch directly translates in INL and DNL performance loss. In order to alleviate this problem the ratio between signal amplitude and input referred random offset of the comparators must be improved. Since the early nineties two techniques are applied for this purpose: averaging and gain stages with interpolation.   The random mismatch problem in traditional full-flash converters is based on the ratio of the input signal and random offset in one comparator stage. The fundamental observation by Kattmann and Barrow [157], is to combine multiple input stages. The signals of these stages are added up linearly, while their mismatch adds up in a root-mean-square manner. Consequently their ratio will improve if more input stages are combined. Figure 8.27 shows the topology. The input stages of the comparators generate a differential current that form a differential voltage over the resistors R_t. These currents are combined through coupling resistors R_c.

Figure 8.28 (left) shows a part of the resistor network. R_e is the equivalent impedance seen to the left and R_t and R_c add another stage to this network. Now the equivalent resistance after the addition of these two elements should again be equivalent to R_e. Some arithmetic gives [157, 158]:

$$\frac{R_e}{R_t} = -\frac{1}{2}\alpha + \frac{1}{2}\sqrt{\alpha^2 + 4\alpha} \tag{8.22}$$

where $\alpha = R_c/R_t$. Based on this equivalent impedance the effect of the coupling resistor R_c on various performance aspects can be derived. If one input pair creates

Fig. 8.29 A gain stage is applied to increase the voltage swing on the comparators
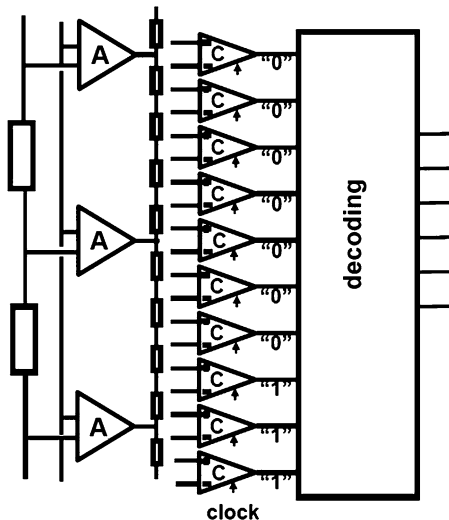
an offset current resulting in a voltage v_n over resistor R_t, then the impact of v_n on the neighboring voltage v_{n-1} is:

$$\frac{v_{n-1}}{v_n} = \frac{-\alpha + \sqrt{\alpha^2 + 4\alpha}}{\alpha + \sqrt{\alpha^2 + 4\alpha}} \tag{8.23}$$

With $\alpha = 1$ this factor is 0.38. Equation (8.23) is depicted in Fig. 8.28 (right). A small α (low R_c) increases the coupling between the node voltages and improves the ratio between the linear signal component and the stochastic variation. Also the range of comparators that contribute is increased. This range cannot be made infinitely long because only comparators with input stages operating in the linear regime can effectively contribute. As an example Bult [159] used an averaging over 5 stages.

   An issue with the averaging technique is the termination of the range. A low α factor means that a larger number of neighboring comparator stages are combined and a better random offset reduction is achieved. Near the limits of the range this means that a relatively large number of additional comparators are needed for achieving also an offset reduction at the extremes of the converter range. Scholtens [158] has proposed to use a dedicated termination cell at the end of each side of the comparator structure, which reduces this problem. Other techniques involve folding or are based on the Möbius band.

   A second form of improving the ratio between the signal and random mismatch is shown in Fig. 8.29. In this scheme a group of additional amplifiers is placed before the comparators. The amplifiers locally boost the difference voltage between the signal and the adjacent reference voltages. The result is applied to an interpolation ladder. A typical gain stage will serve 4–8 comparators. The trade-off in this scheme is between the additional power needed for the high-performance gain stages and the reduction on the side of the comparators.
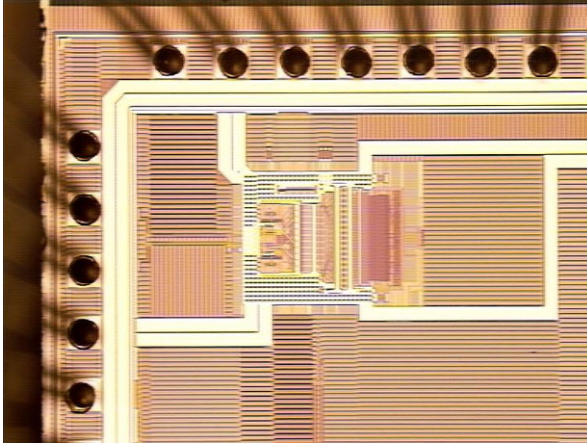
Fig. 8.30 A picture of a 6-bit full-flash ADC converter [158]

Fig. 8.31 Random offset reduction by means of a gain stage
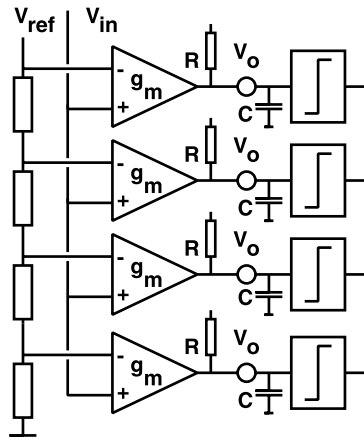
Figure 8.30 shows a full-flash converter placed in a system-on-chip design. Frequently averaging and interpolation are combined into the same design [158, 160].

So far, no frequency dependencies of the input gain stages have been considered. Figure 8.31 shows an example of comparators during offset reduction. The offset reduction is accomplished by means of a gain stage before the latch stage. These elements can be identified in most offset reduction schemes. The latch stage is considered ideal except for a load capacitor C and a random offset source V_o. Both are related to the latch transistor dimensions by $C = WLC_{ox}$ and $\sigma_{Vo} = A_{VT}\sqrt{N_T/WL}$. A_{VT} is the process constant for threshold matching (Sect. 11.4) and N_T is the number of latch transistor pairs that contribute to the random offset. For low frequency operation the input referred mismatch due to the latch mismatch is calculated by dividing the latch mismatch by the effective DC-gain. For high-frequency operation the input-referred random offset due to the latch is given by σ:

$$\sigma = \frac{\sigma_{Vo}(1 + j\omega RC)}{g_m R}$$

$$\approx \frac{A_{VT}\omega C_{ox}\sqrt{N_T WL}}{g_m} \quad \omega RC > 1 \tag{8.24}$$

This is only a rough approximation, which must be adapted for the sample-and-hold operation, timing and different architectures. It indicates that random-offset reduction is frequency limited and depends on design (N_T, g_m), technology (A_{Vt}, W, L) and power (g_m, number of comparators). Example: in a 1 μm CMOS technology $\sigma_{Vo} \approx 20$ mV, $gm/2\pi C \approx 250$ MHz, so the desired input random offset of 0.8 mV can be achieved up to 10 MHz bandwidth.

8.2.5 Technology Scaling for Full-flash Converters

In 90-nm, 65-nm and 45-nm various analog design constraints appear that affect the performance of converters. A few of these effects are:

- The power supply drops to 1 V. As a consequence all signal levels will be lower. Differential design is imperative, yet it will be difficult to handle more than 0.3 V of single sided signal swing.
- Thresholds go down in order to keep the current drive at an acceptable level. Low threshold voltages are not convenient for cascode structures, because the input and output DC-levels start to differ too much.
- The mismatch for a fixed area device is less predictable due to additional implants. At 45 nm the same transistor size may show comparable random variation as in 90 nm.
- The beta factor is low for minimum size devices, it can be improved by backing-off on the designed gate length, see Fig. 2.37.
- The construction of the transistor with high additional pocket dopes near channel to control the electric fields (halo implant), leads to a low intrinsic amplification factor of the device.
- Deep n-well constructions reduces substrate noise and allow floating NMOS devices next to the floating PMOS devices.
- Gate leakage of transistors with less than 1.5-nm effective gate-oxide becomes a issue as the gate current (1–10 nA/μm gate width for 1.2 nm gate oxide) will load the ladder.

8.2.6 Folding Converter

A variant of a full-flash converter is the folding analog-to-digital converter [2, 148, 161, 162] in which a pre-processing stage "folds" the input signal. Figure 8.32

Fig. 8.32 A folding
analog-to-digital converter
folds the input signal into a
smaller signal range
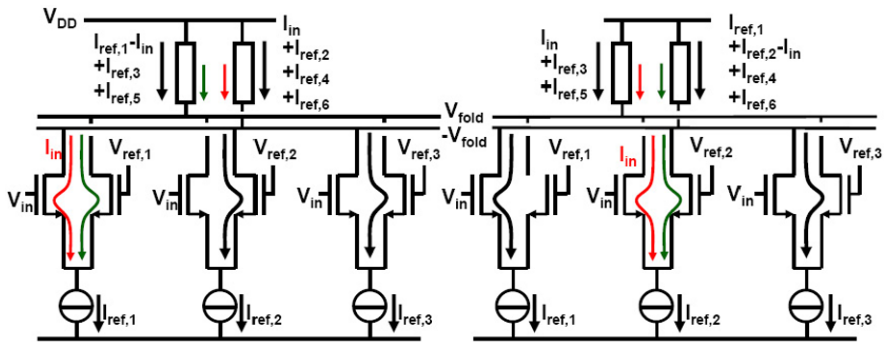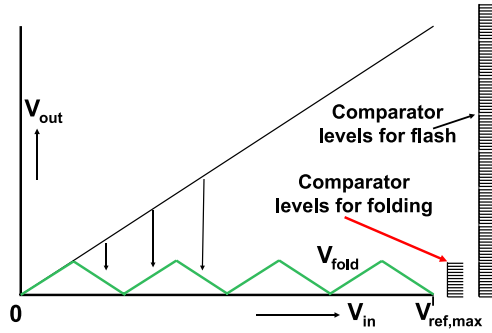
Fig. 8.33 The basic folding circuit: the input pairs are designed to reverse the signal current in the
top resistors for each folding section

shows the general idea: the input signal is folded into (in this example) 8 sec-
tions. The resulting folded range is applied to the succeeding analog-to-digital con-
verter and is reduced to 1/8 of the original range. Next to this analog-to-digital
converter, this method requires a circuit that performs the folding operation and a
coarse analog-to-digital converter to keep track of the section number. This parti-
tioning reduces the total number of comparators. If an 8-bit flash converter is split
in 8 folding sections (3-bit) and a remaining 5-bit fine flash converter, a total of
$2^3 + 2^5 = 40$ comparators are needed. The folding principle was originally devel-
oped in bipolar technology [161, 163]. Later on the folding principle was applied
in CMOS analog-to-digital conversion [148, 149]. In both technologies the basic
topology is a parallel arrangement of differential pairs. These pairs are driven by the
input signal and have each a local reference voltage of 1/16, 3/16, 5/16, ... of the
full-scale reference voltage. In Fig. 8.33 the first three sections of a folding stage
are depicted. In this example the reference voltages are chosen for $F = 8$ folding
sections spaced at $1/F$ fractions of the overall reference voltage. The transconduc-
tance of each stage is designed to cover an input voltage range of $\pm 1/2F$ of the
reference voltage. In this setting the F stages cover the entire input range. The sig-
nal itself will select the input stage. The output of the folding stage on an input ramp
is a triangle shaped signal with a periodicity of half of the number of fold sections.

Fig. 8.34 The basic folding circuit suffers from distortions near the switch points. In a practical design two folding circuits with a mutual offset, allow an interpolation scheme to determine the switching levels
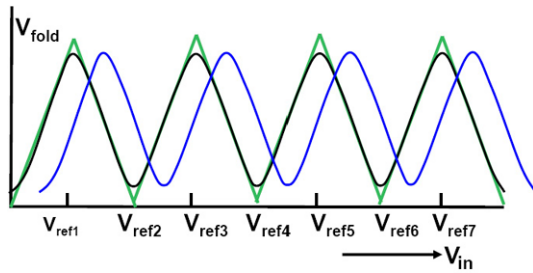
Fig. 8.35 A folding analog-to-digital converter architecture incorporating a sample-and hold [148]
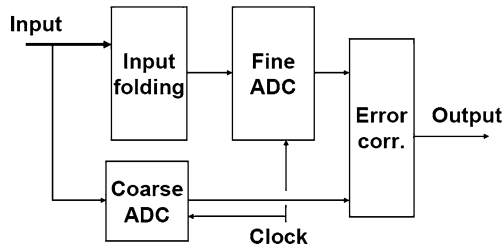
A full-range sinusoidal input signals will be multiplied with this triangular function and generate frequency folding signals at $(F + 1)$ and $(F - 1)$ times its own frequency. A full-range input signal results in a folding signal with $(F + 1)$th and $(F - 1)$th harmonics. Smaller signals use less folding stages and result in lower frequency harmonics. This multiplication effect is a potential cause for distortion products in a folding analog-to-digital converter when this higher order signals proceed to the output. Moreover the bandwidth of the output stage needs to accommodate these frequencies. After the folding stage some gain can help to reduce the accuracy requirements for the comparators in the succeeding full-flash stage.

The cross-over point between the stages in the basic concept is a weak point in the architecture of Fig. 8.33. In order to mitigate the problems with the cross-over points, mostly a dual folding architecture as in Fig. 8.34 is used. The second folding stage is shifted by half of the range of a single folding section $1/2F$, thereby creating a linear transfer at the cross-over points of the first folding stage. The coarse sub analog-to-digital converter selects the linear part of the transfer curves. More recent designs use a resistive interpolation technique between the outputs of these voltage-shifted folding stages. This results in bundles of transfer curves where the comparators trigger at the cross-over of the appropriate combination of (interpolated) folding signals [148, 159]. Proper design of the pre-processing stage where high-speed and high yield are combined, is the critical issue.

The architecture of Fig. 8.35 requires a sample-and-hold circuit. The first observation is that the input signal uses two paths through this converter: the path via the coarse converter and the path through the folding stage. The outcome of these paths must remain synchronous within one clock period. The advantage of using a S/H circuit is that signal propagation errors in the analog pre-processing are reduced and that architectures can be used that make multiple use of the input signal. The second
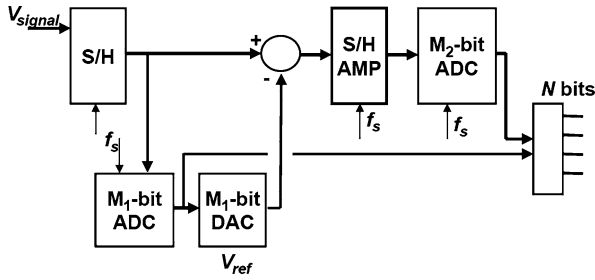
Fig. 8.36 Block diagram of a two-step converter

reason for using a sample-and hold lies in the observation that folding stages generate frequency multiples of the input. A sample-and-hold will reduce this problem to a settling issue. Design experience has shown that a high-speed S&H running at full-signal and bandwidth requires 30% of the total analog-to-digital converter's power budget.

The remaining problem with folding analog-to-digital converters is random offset in the folding input pairs. Any deviation in one of the folding stages will translate in performance loss. An attempt has been made to address this issue [159] by using an interpolation method at the cost of a lot of power. In bipolar technology a performance level of 10 bit has been reported [164], while in CMOS an 8–9 bit level is state-of-the art [148, 149, 159]. A form of calibration can further improve the resolution [165].

8.3 Sub-ranging Methods

For accuracies of 8 bits or more full-flash converters are no economical solution due to the exponential growth of input capacitance, area and power. Multi-step methods allow to achieve higher resolutions. In Fig. 8.36 the signal flow in a two-step converter is shown. The signal is sampled and held at the input. A limited resolution analog-to-digital converter (a small full-flash) with a resolution M_1 estimates the signal. This information is fed to a digital-to-analog converter and subtracted from the held input signal. The subtraction results in a residue signal, with a maximum amplitude fraction of 2^{-M_1} of the input range. This signal is amplified in a second S&H gain stage. A second converter with a resolution M_2 converts this residue signal. This simple approach results in a converter with a resolution of $M_1 + M_2 = N$. In this elementary approach of this converter only $2^{M1} + 2^{M2} \ll 2^N$ comparator circuits are needed. Converters based on this principle are known under names as: "sub-ranging analog-to-digital converters", or "coarse-fine analog-to-digital converters".

For this type of conversion additional components are present such as sampling circuits and an additional digital-to-analog converter. After the subtraction the amplitude of the remaining signal is small: an additional amplification step boosts the signal and consequently reduces the effects of errors in the succeeding processing.

Fig. 8.37 Coarse-fine conversion: on the *left-hand scale* the trip points of the first (coarse) converter are indicated. The remaining signal (*bold line*) is amplified and converted by the second (*fine*) converter. The *dashed arrow* indicates a situation where the coarse converter has decided for the wrong range. Over range in the fine converter corrects this decision
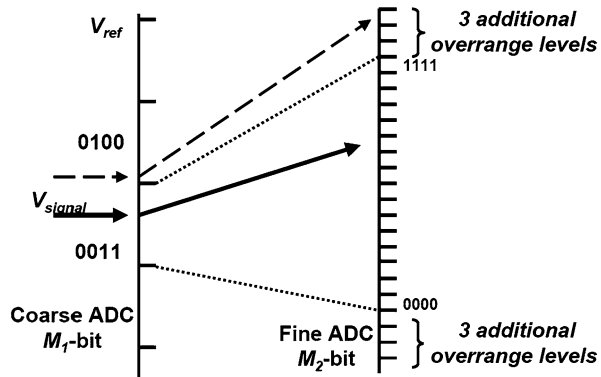
A disadvantage of the set-up in Fig. 8.36 with $M_1 + M_2 = N$ it the need for a perfect match between the ranges of the first and second converters. If the first (coarse) converter decides for the wrong range, there is no correction possible. However by adding additional levels on both sides of the second converter range, see Fig. 8.37, errors from the first converter can be corrected [166]. This "over-range" limits the accuracy requirements on the coarse analog-to-digital converter. In some designs the overrange doubles the total range of the fine converter.

Full accuracy (N-bit) is still needed in the digital-to-analog converter and in the subtraction circuit. Implicit in this method is that the subtracted portion of the signal is known with the accuracy of the full converter. The speed of this converter is in first instance limited to the time needed for the input track-and-hold, the first coarse conversion, the digital-to-analog settling, the subtraction and the second fine conversion. This processing can however be pipelined over two sample periods by inserting a track-and-hold circuit behind the subtraction and amplification point. Now two successive samples are processed in a pipe-lined fashion, and the sampling speed is limited to the processing speed of just a single section.

The over-range feature allows to reduce the time for the signal to settle in the chain from coarse converter, via digital-to-analog converter and subtraction node. The resulting error and the potential offsets should remain within the overrange section. This observation allows a considerable increase of the speed. The principle of coarse-fine conversion can be extended to three or more stages and resolutions of 14–15 bits [167].

The principle of sub-ranging is limited by the quality of the correspondence between the digital-to-analog conversion and the fine conversion range. There are various ways to link the output of the coarse conversion to the fine range. In Fig. 8.38 a elementary connection topology is depicted. The fine ladder is connected to the selected range of the coarse ladder via two buffers A and B. If these buffers suffer from opposite offsets the range defined by the coarse ladder will be stretched or shrunk. In Fig. 8.39 the buffers A and B move over the coarse ladder in a fixed mutual position: A is connected to the low-side of the coarse-ladder tap and B to the high side and this order remains for a second sample which is in the next higher coarse-ladder segment. This sequence causes an INL and a DNL error at the transition points and

Fig. 8.38 The effect of offset in a sub-ranging scheme
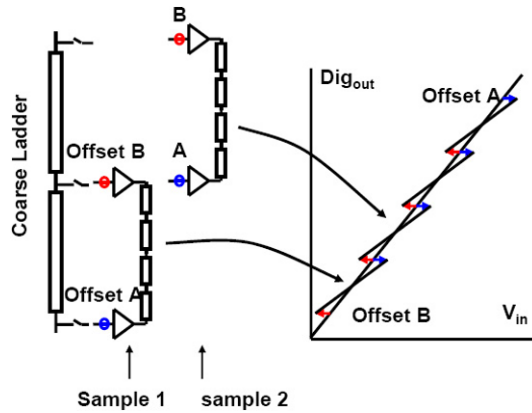
Fig. 8.39 The effect of offset can be reduced by appropriate switching schemes
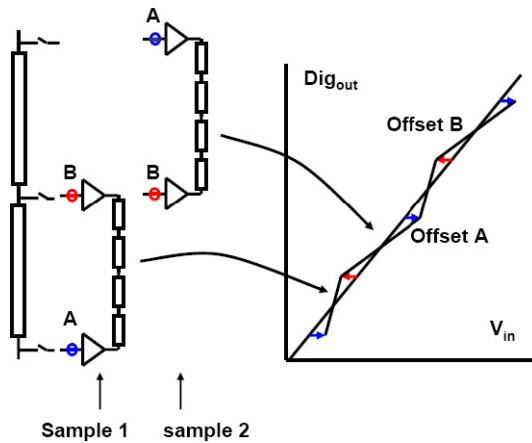
shows up in the integral and in the differential linearity plots. Especially the sharp transition at the coarse conversion transitions leads to large DNL errors. In the INL plot these errors are visible together with the errors in the fine converter as a repetitive pattern, see Fig. 8.40.

In Fig. 8.39 the control of the connections is different: if the converter decides to connect to a higher segment the lower buffer (A in this example) is disconnected and reconnected to the top side of the next higher segment, while buffer B remains connected to the same tap of the coarse ladder. When increasing the input voltage slowly the buffers will alternately connect. This method is often referred to as "monkey-switching".[4]

The set up of the timing is essential especially if additional time periods are allocated to offset cancelation. Typically the coarse converter is activated after some

---

[4]Some similarity exists with the way a monkey climbs a tree.
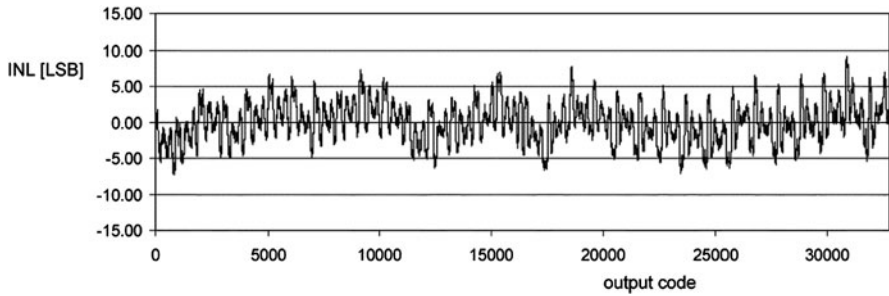
Fig. 8.40 Sub-ranging converters show a repetitive pattern in their transfer curve. This curve is from a 15-bit resolution converter [167]
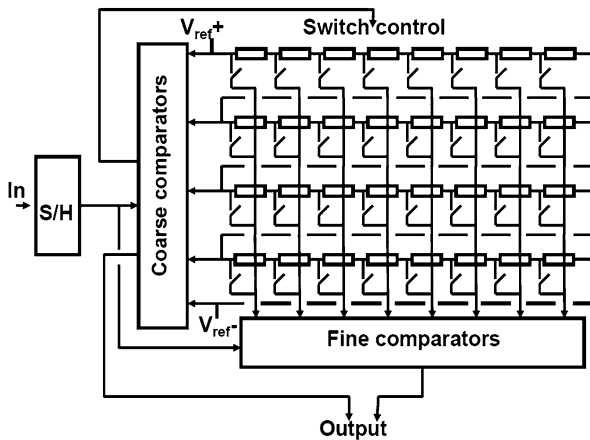
Fig. 8.41 A subranging analog-to-digital converter based on a switched resistor array [140, 153, 168]

10–20% of the hold period. This allows maximum time for the digital-to-analog converter, the substraction mechanism and the succeeding amplifier to settle.

An alternative to the standard sub-range scheme avoids the subtraction and the issues due to offsets between fine and coarse sections. A resistive ladder structure feeds both the coarse and the fine converter sections [140, 153, 168, 169] as shown in Fig. 8.41. The coarse comparators are connected to ladder taps spaced at 8 LSB positions apart. The decision of the coarse converter will select a row of switches which is then fed to the fine comparators. This method can of course easily be extended with overrange. Another extension [153] is to use two banks of fine comparators that will alternately digitize the signal, thereby allowing more time for the settling process. In [168] a capacitive interpolator is used to form the intermediate values for the fine conversion.

These methods use the main resistive ladder structure for both the coarse and fine conversion. This requires special measures to keep the spurious voltage excursions
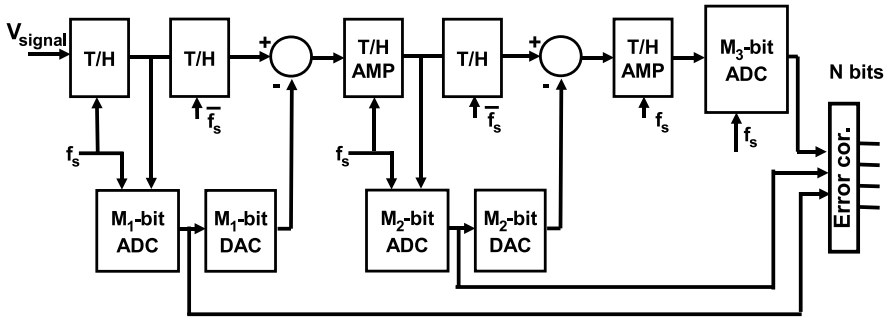
Fig. 8.42 A subranging analog-to-digital converter with three sections. A typical partitioning uses $M_1 = M_2 = 5$ bits and $M_3 = 4$ bit for a 14-bit converter

on the ladder under control. In [140] this is realize via a second low-ohmic ladder in parallel to the main ladder similar to Fig. 7.39.

The dominant implementation of stand-alone sub-range analog-to-digital converters for industrial applications[5] uses a three stage approach [170], see Fig. 8.42. The first stage typically converts 4–5 bits and is equipped with a high-performance track-and-hold. The reduction of the distortion in the track-and-hold is the dominant challenge in these converters. A low distortion results in an excellent spurious free dynamic range (SFDR), which is required for communication systems such as mobile phone base stations. The succeeding stages use 5 bit and 4 to 6 bit in the last stage. The different track-and-hold stages run on different clocks to allow the optimum usage of the sample period.

For a nominal resolution of 14 bit the ENOB ranges from 11.3–11.8 bit. Sample rates between 150 Ms/s and 400 Ms/s are available with power consumptions ranging from 400 mW at 80 Ms/s to 2 W for 400 Ms/s. The SFDR ranges from 80 to 95 dB. Special low-swing digital output stages (LVDS) are applied to suppress digital noise. For a more elaborate comparison see Sect. 12.2.

8.4 Pipeline Converters

A popular variant of the multi-stage analog-to-digital converter in CMOS technology is the "pipeline" converter, Fig. 8.43 [166, 171]. This converter consists of a pipeline of N more-or-less identical stages. For each bit of resolution there is one stage. Each stage comprises a track-and-hold, a comparator connected to a one-bit digital-to-analog converter, a subtraction mechanism and a multiplication circuit. The operation of the pipeline converter can be viewed from different angles. A pipeline converter can be seen as the extreme form of sub-ranging. The number

---

[5]See data sheets from ADI, TI and NXP. Some 2010 products: AD9640, AD6645, ADS5474 and ADC1410. In some data sheets these converters are called pipeline converters. In the terminology of this book they are classified as subrange converters.
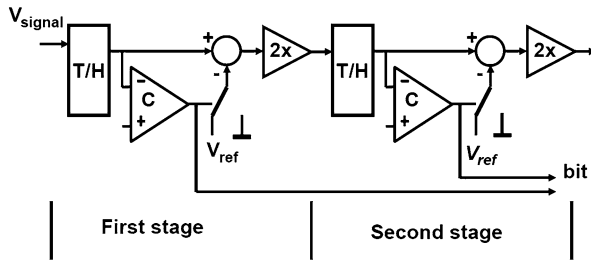
Fig. 8.43 Basic pipe-line analog-to-digital converter

Fig. 8.44 Transfer characteristics for a pipe-line analog-to-digital converter. *Left*: the first stage, *right*: threes stages before the end
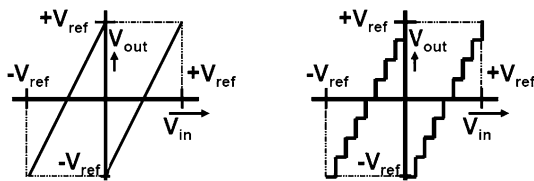
of bits per subrange is reduced to just one. The function of the analog-to-digital converter in each stage is reduced to a single bit decision.

A track-and-hold circuit stores the intermediate value of the signal. A comparator determines whether the input is higher or lower than the reference. The comparator decision leads to a subtraction or addition of the reference value. The result is multiplied with a factor 2 and passed to the next stage. For the i-th stage:

$$V_{\text{out},i} = 2 \times V_{\text{in},i} - D_i V_{\text{ref}} \tag{8.25}$$

where $D_i = 0, 1$ is the decision bit. This is essentially the kernel of the successive approximation algorithm from Fig. 8.58. The pipeline converter can therefore also be understood as a successive approximation converter, where every separate approximation step is implemented in dedicated hardware.

The transfer function of the input to the output of a single stage is shown in Fig. 8.44. For a stage at the input the transfer can be seen as a continuous function represented by a pair of straight lines. On the right hand side the discrete steps in low-resolution stage are visible.

The choice to operate the sections at one-bit resolution has a number of consequences. The analog-to-digital coarse converter becomes a comparator and similar to the subrange converter its inaccuracy can be overcome with redundancy as will be shown later. The digital-to-analog converter reduces in this scheme to a single bit converter deciding between a positive and a negative level. Two levels are by definition perfectly linear, solving an important distortion problem in the digital-to-analog converter.

The pipeline stages are separated by multiply-by-two track-and-hold stages. The factor two in the stage gain implies that the output range of a section $(2V_{\text{in}} - V_{\text{ref}})$ equals the input range. Errors in the second stage will affect the input by half of their magnitude. For the third stage the errors are reduced by 4 etc. Random errors,

Fig. 8.45 Two errors in a
pipeline stage. *Left*: the gain
is not equal to 2 and *right*: the
comparator suffers from a
positive or negative offset.
A converter with these errors
in the first stage has an
overall transfer characteristics
as shown in the lower plots
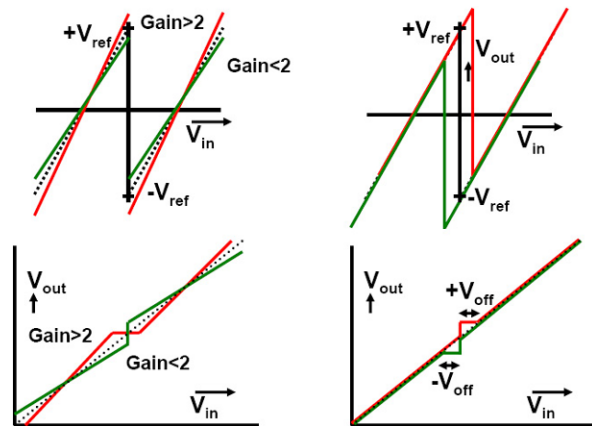
such as mismatch and noise(see Sect. 4.4), reduce in a root-mean-square sense. The
random errors of the second stage and next stage add up as:

$$v_{\text{in,random}} = \sqrt{v_{1,\text{random}}^2 + 2^{-2}v_{2,\text{random}}^2 + 2^{-4}v_{3,\text{random}}^2 + \cdots} \qquad (8.26)$$

The limited resolution reduction per stage and the low amplification per stage, make
that the effect of errors of many stages remains traceable at the input. Preferably a
designer would like to reduce the size of the capacitor for the second-stage track-
and-hold and to minimize the associated currents. However, the impact of the second
stage noise does not allow drastic reductions. In many designs a similar size capaci-
tor is used for the first three stages. For the successive stages the requirements on the
gain and accuracy drop by a factor two per stage. "Stage-scaling" implies reduction
of the currents in these stages resulting in power-efficient designs.

   With a conversion of one bit per stage, N stages are needed and N samples are
simultaneously present in the converter. Additional time is needed for the digital
reconstruction and error correction. Therefore the time between the first sampling
of the signal and the moment the full digital value is on the output is rather long
($N + 3$ clock periods). The resulting delay, called latency, will impair the system
performance if this converter is part of a feedback loop.

8.4.1 Error Sources in Pipeline Converters

Two important errors in this architecture are gain errors and comparator offset,
Fig. 8.45. At the moment the transfer curve exceeds V_{ref} either on the positive side
or on the negative side, the signal goes out of the input range of the succeeding stage.
If, on the other hand, the transfer curve does not reach the reference values, not all
of the digital codes will be used. Both types of errors result in loss of decision levels
or missing codes in the transfer characteristic. The comparator offset must remain
under 0.5 LSB to guarantee a correct transfer curve. With some offset compensation
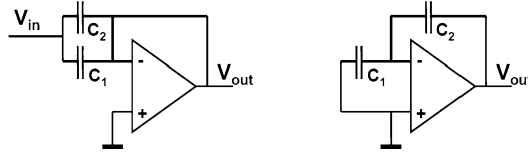
Fig. 8.46 The multiply-by-two operation in a pipeline converter is implemented by sampling the signal on two equal capacitors (*left*). In the multiply phase (*right*) all charge is transferred to C_2 [131]

this value can be reached for 10–12 bit accuracies. The comparator offset problem is effectively solved in 1.5 bit pipeline converters.

The multiply-by-two operation in a pipeline converter is implemented by sampling the signal on two equal capacitors, see Fig. 8.46. In the multiply phase all charge is transferred to C_2:

$$V_{\text{out}} = \frac{C_1 + C_2}{C_2} V_{\text{in}} \tag{8.27}$$

If $C_1 = C_2$ this operation results in an exact multiplication by two. A gain error is the result of insufficient opamp gain, capacitor mismatch or switch charge injection, see Sect. 4.3.2. The error affects the transfer of the residue signal and must remain within a fraction (0.1–0.2) of an LSB. The opamp DC-error after settling is given by feedback theory and must be smaller than the overall required accuracy:

$$\epsilon_{DC\text{gain}} = \frac{1}{A_0 C_2/(C_1 + C_2)} \ll 2^{-N+i} \quad \rightarrow \quad A_0 > 2^{N-(i-1)} \tag{8.28}$$

with resolution N and the feedback ratio is assumed to be equal to 2. For stages further on in the pipeline the demands are less critical. The stage number $i = 1, 2, \ldots, N$ illustrates the option for stage-scaling. The amplification of the first stage is the highest. The amplification requirement holds for the entire output range, so any saturation effects in the output range will cause loss of accuracy. Errors in the capacitor ratio can be the result of technological problems or deviations caused by the lay-out environment. A deviation in gain creates either missing range on the analog side or missing codes on the digital side. The overall DC-accuracy requirement is translated in a minimum mismatch error for the capacitor ratio of:

$$\epsilon_{\text{cap}} = \frac{C_1 - C_2}{C_1 + C_2} \ll 2^{-N+i} \tag{8.29}$$

The limited settling speed of the operational amplifiers determines the speed of the pipeline converter. Limited settling will cause errors comparable to static gain errors. The time constant corresponding to the unity-gain bandwidth of the operational amplifier $\tau_{\text{UGBW}} = 1/2\pi f_{\text{UGBW}}$ and the available time $T_{\text{settle}} = \alpha T_s$ define the settling error:

$$\epsilon_{\text{settle}} = e^{-2\pi f_{\text{UGBW}} T_{\text{settle}}} \ll 2^{-N+1} \quad \rightarrow \quad f_{\text{UGBW}} > \frac{f_s(N-i)}{2\pi\alpha} \tag{8.30}$$

$\alpha < 0.5$ is the fraction of the clock period T_s that is allocated to the settling process. If the converter is operated in a simple mode where the comparator is active in one

Fig. 8.47 The gain in the stage is deliberately made smaller than 2. The converter is calibrated starting at its back end and working towards the front-end [173]
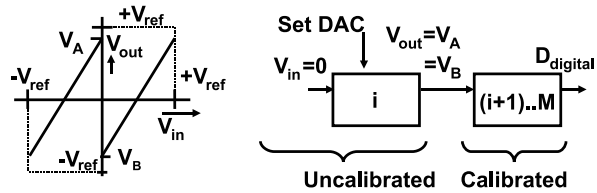
clock phase and the settling takes place in the other clock phase, α must be set according to the minimum duty cycle of the clock. The UGBW of the operational amplifier will have to be high requiring a lot of power. If a more complex timing circuit is used, the settling time for the operational amplifier and the time needed for the comparator can be better controlled allowing more time for the settling and saving of some power.

The above mentioned error sources must not exceed a total of $0.5V_{\rm LSB}$. The balance between the error sources depends on the required specifications and available technology: high accuracy requires high gain and low mismatch. High speed requires simple transconductance stages. A first estimate of the error allocation over the stages of the converter is according to the relative weight: 1/2, 1/4, 1/8, etc.

Next to minimizing the capacitor mismatch in a technological manner, several schemes try to mitigate this error in an algorithmic way. Li [172] proposed the ratio-independent technique. A signal is sampled in a capacitor and in a second cycle stored in an intermediate capacitor. In a third cycle the first capacitor takes another sample, and in a fourth cycle the first sample is retrieved and added to the second sample.

Song [88] presented a scheme where the multiply-by-two is executed in two phases: in the first phase the capacitors are arranged as shown in Fig. 8.46 and the second phase C_1 and C_2 are interchanged. The results of both are averaged yielding a reduced effect of the capacitor mismatch. Unfortunately these schemes cost hardware and several clock cycles.

8.4.2 Digital Calibration

In order to circumvent the accuracy problems Karanicolas [173] proposed to use a gain factor which is deliberately smaller than 2. The gain factor is chosen to avoid range excess either by offset or gain errors, see Fig. 8.47. With less than 1 bit resolution per stage, more stages M are needed than the resolution N.

In order to reconstruct the input from a given output value the following equations are used:

$$\begin{aligned}
V_{\rm in} < 0 \quad &\rightarrow \quad V_{\rm in} = V_{\rm out} - V_A \\
V_{\rm in} > 0 \quad &\rightarrow \quad V_{\rm in} = V_{\rm out} - V_B
\end{aligned} \tag{8.31}$$

The reconstruction requires that the intersection points with the vertical axis V_A and V_B are known. During the digital calibration cycle the values of V_A and V_B
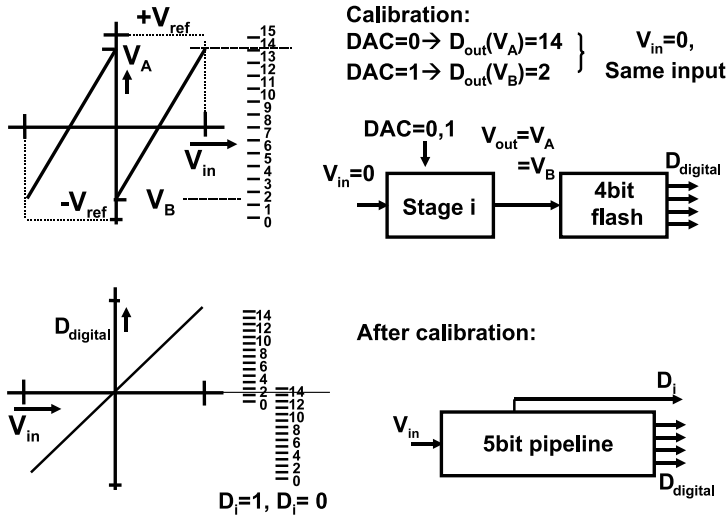
Fig. 8.48   Example of pipeline calibration

are measured as digital codes with the help of the calibrated part of the converter itself. In calibration mode the input of section i is grounded and the digital-to-analog converter is set to $+V_{ref}$. The output will be equal to the value V_A. Similarly V_B is obtained with the digital-to-analog converter set to $-V_{ref}$. The digital values of the intersection points correspond to a fraction of the reference voltage. Calibrating stage i requires that stages $i+1, \ldots, M$ have been measured before.

Figure 8.48 gives an example of the calibration. Before the pipeline converter is used, the calibration cycle has to determine the digital values for V_A and V_B. In this example the last converter stage is a 4-bit full-flash converter and V_A and V_B are measured as digital values of 14 and 2. In fact these two codes correspond to one and the same input voltage (in this example $V_{in} = 0$). Therefore the conversion curves corresponding to DAC = 0 and DAC = 1 can be aligned on these points. The resulting converter is close to a 5-bit range with 29 codes.

The local reference and the quality of the first subtraction or addition must still achieve the accuracy level of the converter. The reconstruction of signal samples is now a simple addition of the digital codes weighted with the decision of the comparator in each section. This calibration relaxes the accuracy requirements on the overall gain. Also comparator offset is not critical, as long as V_A and V_B stay within the reference range.

8.5 1.5 Bit Pipeline Analog-to-Digital Converter

The straight-forward pipeline converter of Fig. 8.43 requires full and accurate settling in all stages. Digital correction by is possible when the gain factor is reduced.

Fig. 8.49 The architecture of
a 1.5-bit pipe-line
analog-to-digital converter is
derived from the coarse-fine
architecture
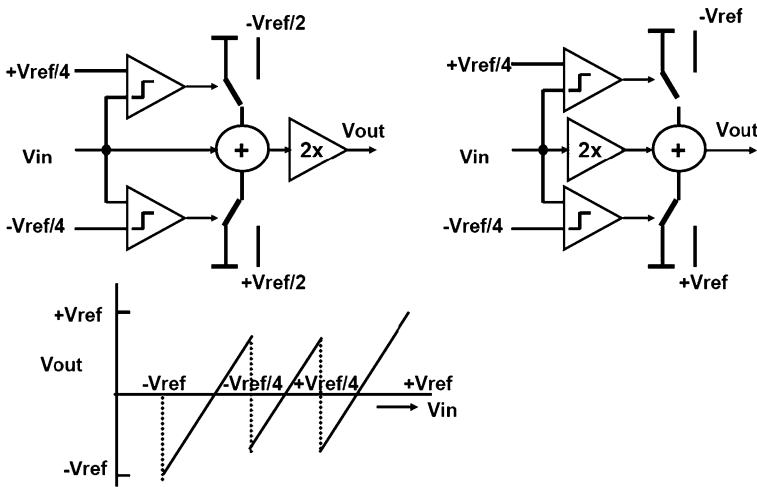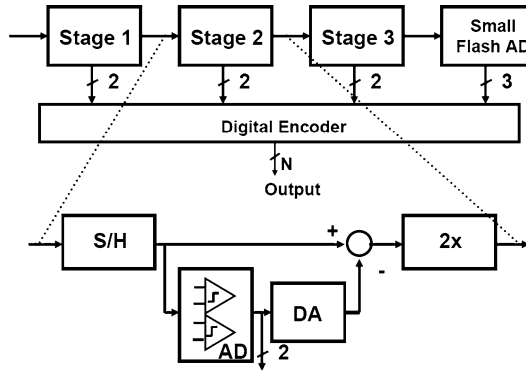
Fig. 8.50 1.5-bit pipe-line analog-to-digital converter

The 1.5 bit pipeline converter allows a more extensive error correction as in each
stage two comparators are used, see Fig. 8.49. The trip levels of these comparators
are typically located at 3/8 and 5/8 of a single sided reference or at $\pm 1/4$ of a
differential reference, see Fig. 8.50 [86, 174–178]. These decision levels split the
range in three more or less equivalent pieces, thereby optimally using the dynamic
range of the circuit. In case of a double positive decision of both comparators the
digital-to-analog converter will set the switches to subtract a reference value. In case
of two negative decisions the digital-to-analog converter will add a reference value
and in case of a positive and a negative decision, the signal will be passed on to the
amplifier. Every stage therefore is said to generate 1.5 bit.

$$V_{\text{out}} = 2V_{\text{in}} + \begin{pmatrix} +V_{\text{ref}} \\ 0 \\ -V_{\text{ref}} \end{pmatrix} \tag{8.32}$$

The digital-to-analog converter uses three levels. In a fully differential design these levels can be made without loss of accuracy: zero, plus and minus the reference voltage. The last two levels are generated by straight-forward passing the reference voltage or twisting the connections.

The typical transfer curve from Fig. 8.50, shows that negative signals are shifted upwards and positive signals are shifted downwards in the plot with respect to the values around zero.

8.5.1 Design of a Stage

Like in most CMOS analog-to-digital principles also 1.5 bit pipeline converters need good capacitors, high-performance operational amplifiers and low-leakage CMOS switches to implement high-quality track-and-hold stages and to create a pipeline of N stages. The first choice that has to be made is the signal swing. Operating the converter in differential mode is preferred as it doubles the amplitude with respect to a single-sided approach. Moreover it minimizes even order distortion, power supply and substrate noise influence. However, analog-to-digital converters that take their input directly from external sources will normally see a single-sided signal. Also many RF and filter circuits before a converter are designed single-sided to allow optimum and/or minimum use of components. In both cases it is desirable to have an as large as possible input signal swing.

For high-resolution converters the track-and-hold capacitor must fulfill the kT/C noise requirement. This capacitor choice is a determining factor in the design. The total noise accumulates as in (8.26). See also the noise discussion in Fig. 4.20. When all stages are 1.5-bit, the contribution of the second and third stages is still significant. Using the same capacitor value for these stages and scaling only after the third stage is a realistic option. The alternative is to use more levels in the first stage: e.g. 6 [179] or 7 [87].

With a large swing of the input signal ($2\,\mathrm{V}_{pp}$) the capacitors can remain relatively small, causing less problems with area, power, etc. The input switches can either be complementary or bootstrapped, see Sect. 4.3.4 and are of minimum length. A potential issue is formed by the bondpad that is connected to the input terminals. Its protection measures must be examined as too much parasitic capacitance will affect the performance. Special RF-bondpads are an alternative.

The processing stage of Fig. 8.51 (left) allows a large signal swing without any repercussion for the operational amplifier. Its inputs remain at virtual ground level thereby avoiding common mode problems (such as in telescopic amplifiers, see Sect. 2.7.10).

During the track phase in Fig. 8.51 (left). The switch S_3 puts the operational amplifier in unity gain feedback. The capacitors C_1 and C_2 are connected in parallel to the input signal. When the transition to the hold-phase occurs, switch S_1 creates via C_2 the feedback path for the amplifier. S_2 switches capacitor C_1 to one of three reference voltages, depending on the result of the two comparators. The original
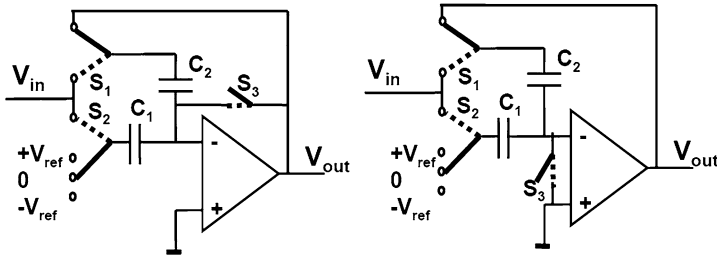
Fig. 8.51 A track-and-hold with multiply-and-subtract stage as used in 1.5-bit pipeline converters. Two alternative schemes are shown for creating a virtual ground node in pipeline converters

charge corresponding to the input signal on C_1 is transferred to C_2. In addition a charge packet corresponding to the chosen reference voltage is moved into C_1 and out of C_2. The overall transfer of this circuit is:

$$V_{\text{out}} = \frac{(C_1 + C_2)}{C_2} V_{\text{in}} \begin{pmatrix} + \\ 0 \\ - \end{pmatrix} \frac{C_1}{C_2} V_{\text{ref}} \tag{8.33}$$

Both signal and reference are multiplied by two. This principle has a lot of similarity with algorithmic analog-to-digital converters, see Sect. 8.6.2.

The amplification stage has to fulfill similar demands as in the 1-bit pipeline converter. The capacitor accuracy is given by (8.29). Plate or fringe capacitors normally allow to reach 12–14 bit capacitor matching. Moreover, the size of the capacitors is in this range more determined by the noise requirements. The topology of the opamp depends on the specifications. Low accuracy and high speed will push towards simple transconductance stages, while high accuracy may involve (folded) cascode amplifiers. The DC-gain is set by the overall resolution in (8.28), and the settling in (8.30). The implementation of the opamp starts with the current. The current must be able to generate sufficient transconductance to reach the UGBW of (8.30). Moreover the slew rate requirements must be met:

$$I_{\text{bias}} > I_{\text{slew}} = C_{\text{load}} \frac{dV}{dt} \tag{8.34}$$

where the load capacitance C_{load} includes the feedback and parasitic capacitances as well as the input capacitance of the next stage. A safe measure for the voltage slope is $V_{\text{max,swing}}/\tau_{\text{UGBW}}$. The first stage is most demanding, the next stages can be designed with relaxed specifications, following the stage-scaling principle.

In this example of Fig. 8.51 (left) switch S_3 sets the opamp in unity-gain mode in order to create a virtual ground during track mode. The consequence of this choice is that the opamp and the switches operate as described for auto-zeroing comparators in Sect. 8.1.7 with an increased noise level due to the sampling of the input referred noise of the opamp. The switch S_3 can be moved to a position where it shorts the input terminals of the amplifier Fig. 8.51 (right), see e.g. [179]. The real ground node for the sampling process removes the input-referred opamp noise. This configuration
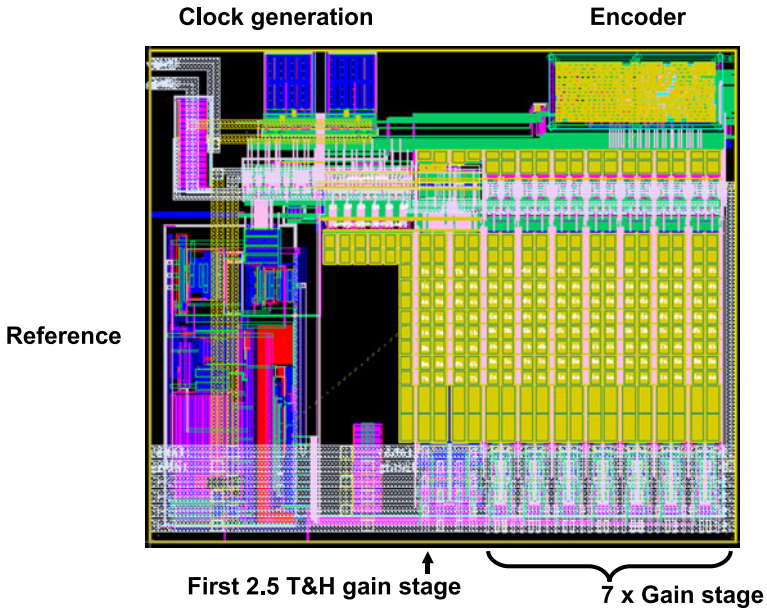
Fig. 8.52 A 10-bit pipeline layout. The first stage contains 2.5 bit, the seven remaining stages 1.5 bit. The depicted area in 90 nm CMOS is $500 \times 600 \ \mu m^2$. *Courtesy*: Govert Geelen NXP [179]

trades off the offset cancelation of the scheme in Fig. 8.51 (left) for reduced noise performance in Fig. 8.51 (right).

Figure 8.52 shows the lay-out of a 10-bit pipeline converter. Most of the area is used by the capacitors.

8.5.2 Redundancy

Figure 8.53 shows the redundancy mechanism of a 1.5 bit scheme. The upper curve shows seven stages and the voltage levels in between. An input voltage equal to $V_{in} = 0.6 V_{ref}$ clearly exceeds the upper comparator threshold at $0.25 V_{ref}$. Therefore the first coefficient is determined to be $a_{N-1} = 1$. The residue is formed as $2 \times V_{in} - V_{ref} = 0.2 V_{ref}$. This value generates in the second stage a $a_{N-1} = 0$ decision. In the following stages the rest of the conversion takes place. The result "77" equals $77/128 = 0.602 V_{ref}$.

In the lower trace of Fig. 8.53 the comparator level of the third section is moved to $0.45 V_{ref}$. It is easy to see how the redundancy scheme corrects for this error. See also the remarks for the RSD converter in Fig. 8.66. The redundancy eliminates the need for accurate comparators, but leaves full accuracy requirements on the switched-capacitor processing stages.

Figure 8.54 shows the lay-out of a 16-bit pipeline converter.

Fig. 8.53 1.5-bit pipe-line analog-to-digital converter uses the redundancy of two decisions per stage. The *upper sequence* shows the ideal behavior. In the *lower sequence* one comparator level is shifted, e.g. due to offset. The redundancy corrects the mistaken decisions
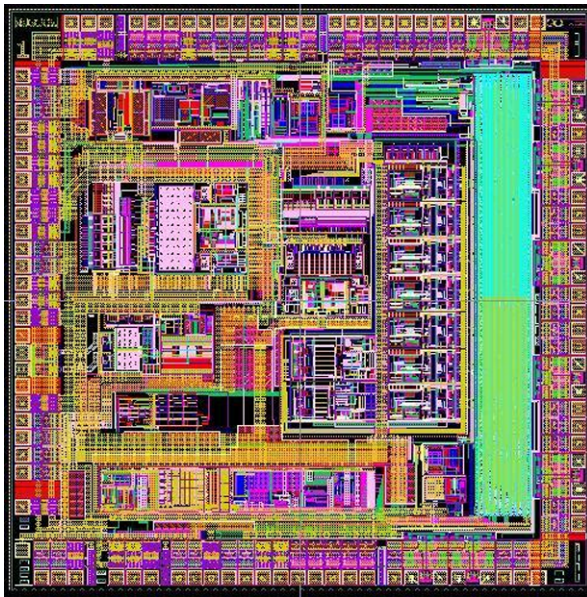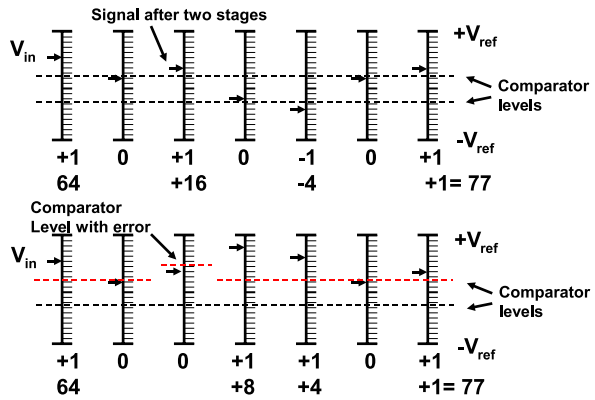
Fig. 8.54 Lay-out of a 16-bit 125 Ms/s pipeline analog-to-digital converter ENOB = 11.6. *Courtesy*: Ph. Gandy, NXP Caen

8.5.3 Pipeline Variants

Judging from the number of published papers, the pipeline converter is the most popular analog-to-digital converter. Most effort is spent on converters in the range of 10 to 14 bit of resolution and 50 to 500 Ms/s sampling rates. These high-performance converters aim at the same markets as high-end sub-ranging converters. Similar to the high-performance sub-ranging converters the main problems are found in the design of the first track-and-hold circuit: constant switch resistance, the settling behavior of the opamp, and parasitics. In order to avoid too many stages contributing
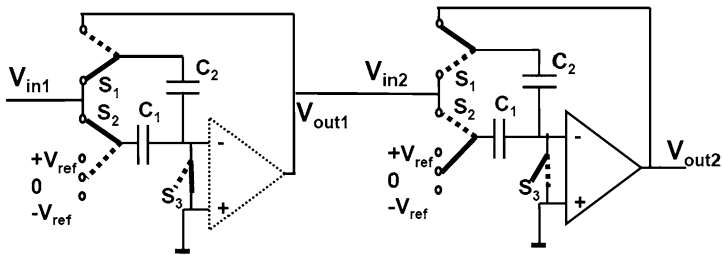
Fig. 8.55 The opamp is shared between two stages in a 1.5-bit pipeline converter. The *left stage* is sampling and the *right stage* is amplifying. During the next half phase the opamp is moved to the first stage which amplifies, while the second stage samples without an opamp. The switches for reconnecting the opamp have been left out for clarity

to the noise and distortion budget, most of these converters use input stages with 2.5 to 4 bit resolution [181–183]. The analysis in [175] indicates that this is an optimum range.

Power efficiency is a driving force in 1.5 bit pipeline converters. One possibility to improve power efficiency is found in a better use of the hardware. A feature of the configuration in Fig. 8.51 (right) is the fact that during sampling the opamp is redundant. Opamp sharing is a technique that effectively uses the opamp redundancy in the scheme of Fig. 8.55. The opamp now serves two sections of the pipeline converter and reduces the required number of opamps by a factor two, [174, 180]. During the sampling the input capacitor C_1 and the feedback capacitor C_2 do not need the opamp as long as a separate ground switch S_3 is present. The opamp can be used for the second stage, where the subtraction and multiplication by a factor two takes place. The odd and even sections now run half a sample phase delayed and one opamp can serve two sections. The opamp can show a "memory-effect" due to signal charges stored in the internal capacitors. This interference between one sample and the next is undesired. A fraction of the clock period or special switching schemes are used to avoid coupling between the samples.

The pipeline converter uses digital calibration techniques to overcome analog imperfections. Many authors propose to extend these calibration techniques to eliminate the energy consuming opamps. Mehr and Singer [184] propose to remove the dedicated front-end track-and-hold function. The pipeline stage as in Fig. 8.51 samples directly the time continuous input and performs the first processing. The timing mismatch between this stage and the first set of comparators must be solved.

Replacing the opamps with less performing building blocks or allowing incomplete settling [177, 178] requires more extensive calibration. Sensitivity to changing environmental conditions (power supply, bias, temperature) is unclear.

A radical idea to avoid the opamp in the processing of a pipeline converter is proposed by Sepke [185]. A comparator switches on and off two current sources in Fig. 8.56. I_1 is used for fast charging, however, some overshoot must be expected due to the delay between the crossing of the levels at the input of the comparator and the current switching. A second current source I_2 discharges at a slower rate and reaches the required output level.
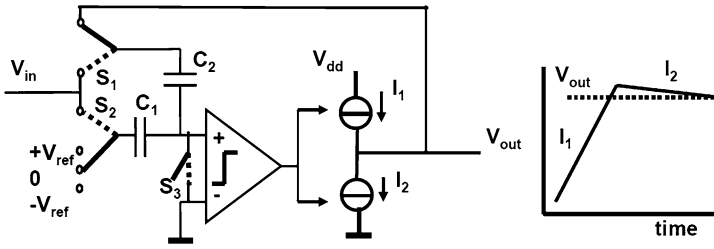
Fig. 8.56 The opamp is removed and replaced by a comparator and two current sources
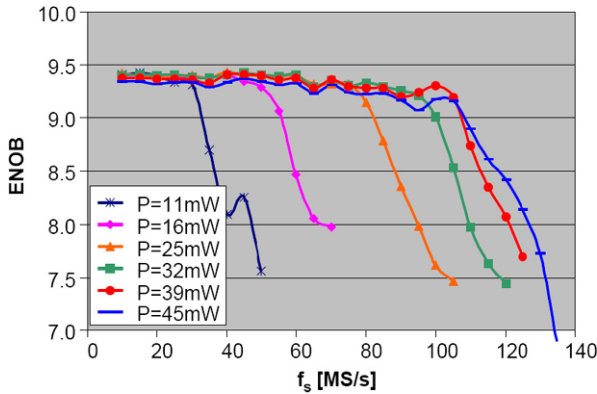
Fig. 8.57 Measured effective number of bits versus sample rate at different power levels. The signal frequency is each time just below $f_s/2$. *Courtesy*: Govert Geelen NXP [179]

Power efficiency is crucial for the application of converters in large system chips. Next to optimizing the power consumption of the individual blocks and techniques such as opamp sharing, also converters are published with adjustable power and performance specifications see Fig. 8.57 [179]. Another approach to calibrate the errors in a fast pipeline converter is in combination with a slow high-precision analog-to-digital converter. Synchronization of both converters is achieved by a track-and-hold circuit or post processing [186].

8.6 Successive Approximation Converters

Where a full-flash converter needs a single clock edge and a linear converter 2^N clock cycles for a linear approximation of the signal, the successive approximation converter (SAR stands for successive approximation register) will convert the signal in N cycles. Figure 8.58 shows an abstract flow diagram of a successive approximation algorithm. The output bits a_{N-1} to a_0 are set to 0. In the first cycle the coefficient corresponding to the highest power a_{N-1} is set to 1 and the digital word

Fig. 8.58 A flow diagram for successive approximation

$V_x = V_{in}(t)$
$i = N$
$a_{N-1} = a_{N-2} = \ldots a_2 = a_1 = a_0 = 0$

$i := i-1$
$a_i = 1$
$V_{DAC} = V_{ref} 2^{-N} (a_{N-1} 2^{N-1} + \ldots a_2 2^2 + a_1 2^1 + a_0 2^0)$

If $V_x < V_{DAC}$
then $a_i = 0$
else $a_i = 1$

no

$i = 0$?
yes
ready
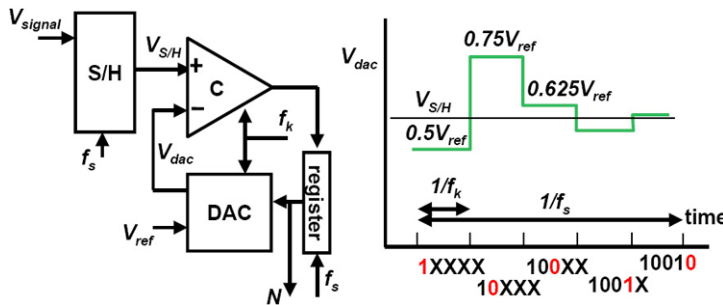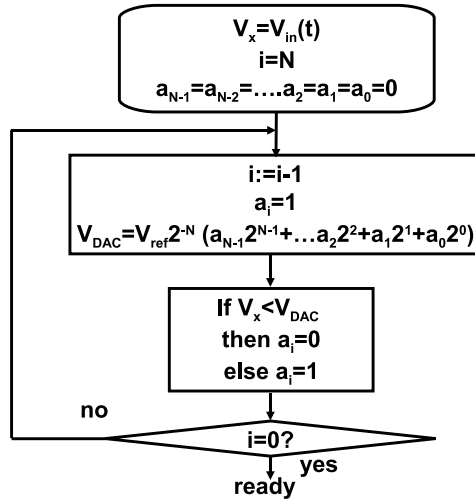
Fig. 8.59 A successive approximation analog-to-digital converter with approximation sequence

is converted to a value V_{DAC}. The input level is compared to this value and depending on the result the bit a_{N-1} is kept or set back to 0. This cycle is repeated for all required bits.

Figure 8.59 shows a circuit implementation. After the signal is stored in the sample-and-hold circuit the conversion cycle starts. In the register the MSB is set to 1 and the remaining bits to 0. The digital-to-analog converter will generate a value representing half of the reference voltage. Now the comparator determines whether the held signal value is over or under the output value of the digital-to-analog converter and keeps or resets the MSB. In the same fashion the next bits in the output register are determined. The internal clock runs much faster than the sample clock, for every sample a sample-and-hold cycle and N clock cycles are needed. In this scheme the digital-to-analog value approximates the input value. Another implementation reduces the input value by successive subtractions.

Offsets in the sample-and-hold circuit or the comparator will generate a shift of the conversion range, but this shift is identical for every code. This principle allows sample rates of tens of MegaHertz. The demands on the various constituent parts

Fig. 8.60 An early implementation of a successive approximation analog-to-digital converter is based on capacitor switching, after [187]
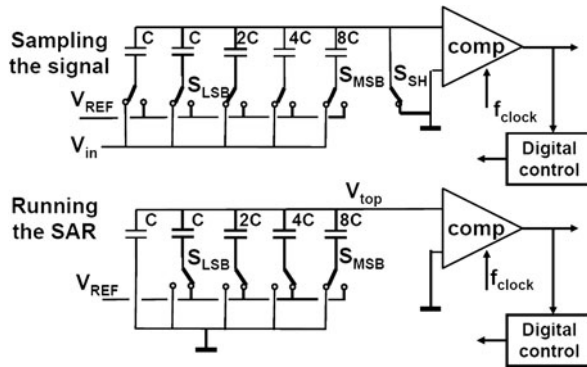
of this converter are limited. Main problem is a good sample-and-hold circuit that needs a good distortion specification for relatively low sample periods. Next to the sample and hold, the digital-to-analog converter determines the overall linearity and will take up most of the area. The conversion speed is determined by the settling of the digital-to-analog converter. Especially in larger structures with a lot of elements this settling time constant τ_{DAC} can be rather long. For reaching half-an-LSB accuracy in an N bit converter the settling time requirement is:

$$t_{settle} > \tau_{DAC} \ln(2^{N+1}) \tag{8.35}$$

For many applications that do not need the maximum conversion speed that is possible in a technology, successive approximation is a safe and robust conversion principle. In applications with a built-in sample-and-hold function (e.g. a sensor output) the combination with a successive approximation converter is appropriate. In combination with a micro-controller, the register function and the timing can be controlled with software. However, special attention must be paid to processor interrupts that can easily disturb the conversion process.

8.6.1 Charge-redistribution Conversion

One of the early fully integrated CMOS successive approximation analog-to-digital converters is known as "charge-redistribution" converters [187]. The principle is shown in Fig. 8.60 and utilizes optimally the properties of CMOS technology: good switches and capacitors. In the sampling phase the input signal is stored on a capacitor bank with a total capacitance value of 16C. "C" is the unit capacitor and is laid-out in a standardized manner. In the second phase the ground plates of the capacitors are switched one after the other from ground to the reference voltage. If the MSB switch is toggled the top plate voltage changes from ground to:

$$V_{top} = -V_{in} + \frac{8C}{C + C + 2C + 4C + 8C} V_{ref} \tag{8.36}$$

Fig. 8.61 A successive approximation analog-to-digital converter based on capacitor switching [188, 189]
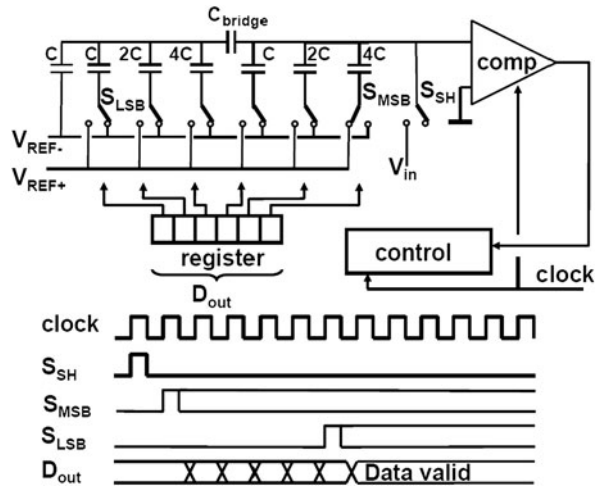
Depending on the original value of the input voltage, the comparator will decide to keep the MSB switch in this position or return to ground. Every bit is subsequently tested. In this implementation the result is a digital code and an output voltage of the digital-to-analog converter that approximates the original input signal.

In another implementation the sampling is performed on the upper plates of the capacitors and the algorithm will converge to the digital code that complements the input voltage to full-scale.

Figure 8.61 shows a second example of a successive approximation analog-to-digital converter in standard CMOS technology. The digital-to-analog converter is implemented as a bridged capacitor array, see Fig. 7.21. An important difference is that the summation node is not a virtual ground node as it is in the digital-to-analog converter. The voltage swing on the input node of the comparator depends on a correct charge sharing between the capacitors connected. The bridging capacitor and the left hand in series capacitors must sum up to a unit capacitor, so:

$$\frac{C_{\text{bridge}}2^kC}{C_{\text{bridge}} + 2^kC} = C \tag{8.37}$$

where k is the resolution of the left-hand side array. From this equation the bridge capacitor is found as:

$$C_{\text{bridge}} = \frac{2^k}{2^k - 1}C \tag{8.38}$$

The bridging capacitor can be rounded to unity if k is large [188], thereby avoiding the need for a fraction of a capacitor with various mismatching issues.

The conversion cycle starts after the sample is loaded on the capacitors through switch S_{SH}. In this example this action also resets the structure, however, also a separate reset clock cycle and switches can be used. The reference voltages are chosen with equal but opposite voltages with respect to the signal ground level. The purpose of this successive approximation is to make the signal on the input of the
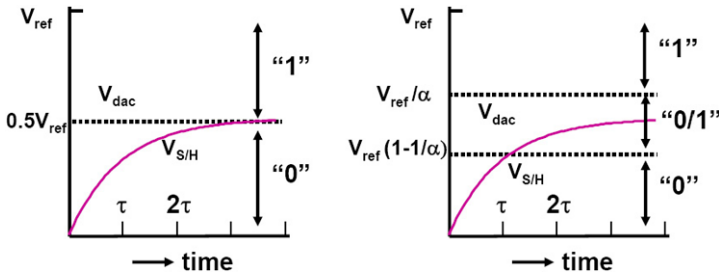
Fig. 8.62 In a base-2 successive approximation the settling of the input signal must reach the final accuracy level before the next step is taken. In non-binary search the base is smaller than 2 and an intermediate region of signal levels exists, where initial decision can be corrected

comparator equal to the ground level. To achieve that goal in this implementation the MSB-switch in the sampling phase is connected to the plus reference, while the other switches are connected to the negative reference. After the sampling the MSB can be tested and then sequentially all other bits.

These successive approximation converters use a limited amount of hardware and good energy efficiencies have been reported [188, 189].

Successive approximation converters can be easily equipped with forms of redundancy. In a successive approximation with a base of 2, the first decision is the final decision on the MSB. It is necessary to let the signal from the track-and-hold stage settle till sufficient accuracy is reached before the decision is taken, see Fig. 8.62. In [190] the base for the digital calculation is not 2 but e.g. $\alpha = 1.85$. After the MSB decision, the range for the remaining search is not the half of the original range but a fraction $1/\alpha$. This decision range for a "0", see the right side of Fig. 8.62, extends over the "1" decision level and therefore allows an intermediate range where the initial decision can be corrected. Of course some more clock cycles are needed for this redundancy, however as the signal needs far less settling time the frequency can be a factor 2–3 higher. This type of search is also called: "non-binary" search, compare Sect. 8.4.2.

8.6.2 Algorithmic Converters

In the previous examples of successive approximation converters the searching process is implemented by comparing the input value to a set of values from the digital-to-analog converter. Next to the design of the sample-and-hold, the accuracy and the area are determined by the digital-to-analog converter. Algorithmic or "cyclic" analog-to-digital converters keep the reference value constant and avoid a large digital-to-analog structure. By capacitive manipulation the signal is modified [171].

A flow diagram of a basic algorithm is shown in Fig. 8.63. The value V_x is set to the input value and compared to half of the reference. If V_x exceeds half of the

Fig. 8.63 A flow diagram for
a cyclic converter
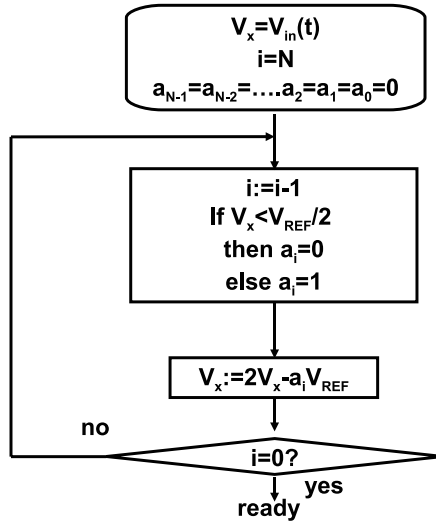
reference, this value is subtracted. The remainder is multiplied by 2 and treated as
the new input value of the process. This multiplication is an advantage over the el-
ementary successive approximation algorithm. Now errors in the smaller bits count
less. Obviously the result of N executions of this flow diagram is:

$$V_x = 2^N V_{in}(t) - V_{ref}(a_{N-1}2^{N-1} + \cdots + a_1 2^1 + a_0 2^0) \tag{8.39}$$

If the remainder V_x is set to zero (ideally it should be less than an LSB), V_{in} will
equal a binary-weighted fraction of the reference voltage. The critical factors in this
algorithmic converter are the offsets and the accuracy of the multiplication by 2. The
total multiplication error must remain within one LSB. If the amplification equals
$(2+\epsilon)$ the difference between the value at the MSB transition and the (MSB-1LSB)
transition (the DNL at that code) equals:

$$\begin{aligned}
D &= (2+\epsilon)^{N-1} - [(2+\epsilon)^{N-2} + (2+\epsilon)^{N-3} + \cdots + (2+\epsilon)^1 + (2+\epsilon)^0] \\
&= (2+\epsilon)^{N-1} - \frac{1-(2+\epsilon)^{N-1}}{1-(2+\epsilon)} = \frac{1-\epsilon(2+\epsilon)^{N-1}}{1-\epsilon} \\
&\approx 1 - \epsilon 2^{N-1} \text{ [in LSB]} \tag{8.40}
\end{aligned}$$

The error in the multiplication factor is itself multiplied by the term 2^{N-1}. In order
to keep the DNL sufficiently low $\epsilon < 2^{-N}$.

Figure 8.64 shows a basic circuit topology of the converter. After the sample-
and-hold circuit has acquired a sample and all capacitors have been discharged, the
signal is multiplied by two and compared with the reference voltage to generate
the MSB bit. Based on this bit zero or the reference voltage is subtracted from the
signal. This remainder signal is fed back to the sample and hold for the next run.

Offsets are critical. The comparator offset must remain under an LSB. In
switched-capacitor technique the offsets at the inputs of op-amps and comparators

Fig. 8.64 An example of a cyclic analog-to-digital converter, after [172]
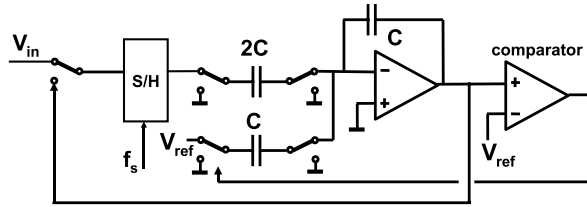
Fig. 8.65 The redundant signed digit algorithm [191]
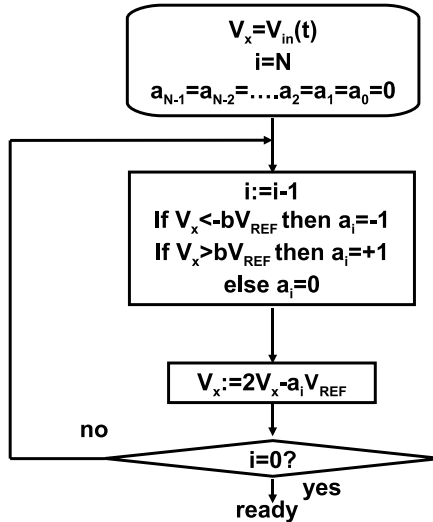
can be removed. The remaining problem is the required accuracy of the multiplication by 2. The minimum capacitor value is mostly determined by the accumulated noise. And the minimum gain of the operational amplifiers is given as in (8.28). The implicit mismatch of the capacitor structure may jeopardize accuracy. However, a careful lay-out where the capacitor "2C" is built from two parallel capacitors "C" and properly surrounded by dummy structures, will reduce this error source to the 10 to 12 bit level. The injection of charge by the switches is especially an issue in older technologies. The channel charge in the relatively large switch transistors varies with the signal and will affect the overall gain, see (4.5). Differential operation or special switching sequences help to reduce this effect [172].

Several accuracy issues can be removed if some redundancy is built in. The flow diagram in Fig. 8.65 shows that instead of a single decision now the signal is compared to values bV_{ref} and $-bV_{\text{ref}}$, with $b \approx 0.25$. The redundancy is due to the three values that each coefficient can take: $a_i = (-1, 0, +1)$. In simple terms: the algorithm assigns only a $+1$ or -1 value to a coefficient if the signal is unambiguously positive respectively negative. In case of doubt, the signal is left unaltered. In Fig. 8.66 the algorithm converts a signal $V_{\text{in}} = 0.6V_{\text{ref}}$. The first decision is therefore $a_{n-1} = +1$, and the remainder is formed as $(2 \times 0.6 - 1)V_{\text{ref}} = 0.2V_{\text{ref}}$. This new residue value leads to a $a_{n-2} = 0$ decision and the next remainder is simply
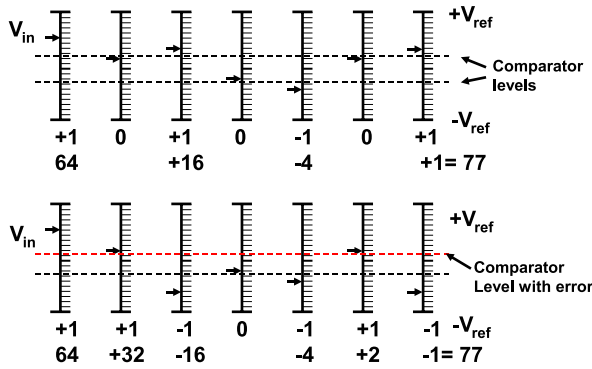
Fig. 8.66 *Upper sequence*: An input voltage of $0.6V_{\text{ref}}$ is converted in seven cycles with comparator levels at $\pm 0.25 V_{\text{ref}}$. *Lower sequence*: the positive comparator level is now at $0.15 V_{\text{ref}}$ e.g. due to offset, however with the help of redundancy is the same code is reached: $77/128 = 0.602$. Note that the exactly matching end result is partly a coincidence. The difference at the one-but-last stage is 2 LSB
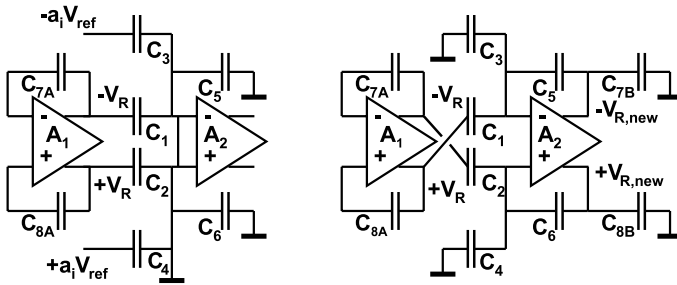
Fig. 8.67 The redundant signed digit (RSD) converter [191]

$2 \times 0.2 V_{\text{ref}}$. The process repeats until after seven cycles a value of 77 is found: $77/128 = 0.602$.

In the lower part of the figure, the upper comparator level is reduced to $0.15 V_{\text{ref}}$, e.g. due to offset. Still the algorithm converges to the same overall result. The RSD algorithm[6] creates in this way robustness for comparator inaccuracies. Figure 8.67 shows an implementation of the RSD principle [191]. During the first phase of the cycle the residue is stored in capacitors C_{7A}, C_{8A}. The result of the comparison results in a value for a_i. In the second phase of the cycle, the signal is multiplied by two by means of the cross-coupling of the differential signals on C_1, C_2. This trick allows that C_1, C_2, C_5, C_6 are all equal. Gain errors due to capacitive mismatches are reduced by interchanging the pair C_1, C_2 with C_5, C_6 for every odd-even cycle. The new values for V_R are stored on C_{7B}, C_{8B} which take the place of C_{7A}, C_{8A} in the

---

[6]The general form of this principle is in [191] identified as the Sweeney-Robertson-Tocher division principle.
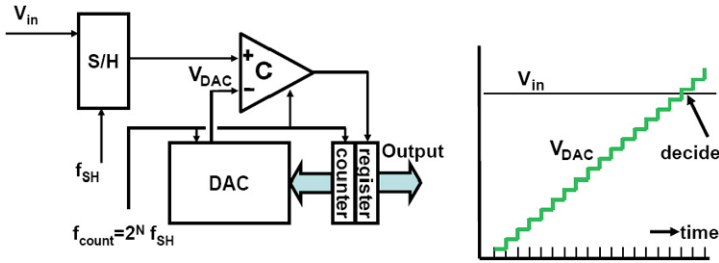
Fig. 8.68 A counting analog-to-digital converter
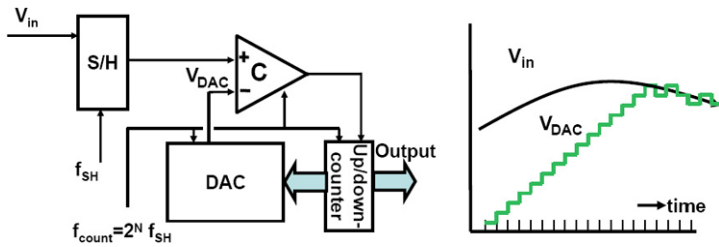
Fig. 8.69 A tracking analog-to-digital converter

next cycle. The typical performance of algorithmic converters is in the 12-bit range at relatively low power consumption. Applications are found in sensors, e.g. [192].

8.7 Linear Approximation Converters

Full-flash converters and successive approximation converters will yield at every clock cycle a result. In the case of a full-flash conversion that will be after a latency of 1 or 2 clock periods, and for multi-stage conversion that delay may last some $N + 3$ clock periods. Linear approximation methods and converters need for their operation a operating clock frequency that is at least 2^N times higher in frequency than the sample pulse.

Figure 8.68 shows a simple implementation of a counting analog-to-digital converter or a "digital-ramp" converter. After a start pulse an input sample is taken and the counter is reset. The counter will start incrementing the code that is applied to the digital-to-analog converter. If the level of the input sample is reached, the register will copy the counter value and deliver this value as the conversion result to the succeeding logic.

An example of a linear converter with clock subdivision is [193].

This counting converter can be turned into a tracking analog-to-digital converter [194] by changing the counter in an up-down counter and connecting the comparator output to the up-down switch, Fig. 8.69. The system will operate in a way that the counter and digital-to-analog converter output will follow the input signal. The
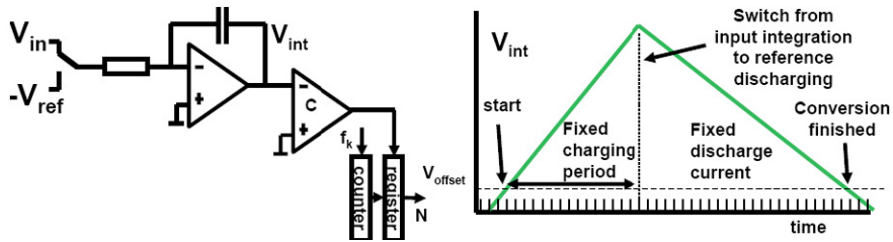
Fig. 8.70 A "dual-slope" analog-to-digital converter

speed is limited to the speed of the counter clock. The accuracy is determined by the digital-to-analog converter. Yet with simple means a reasonable analog-to-digital converter can be built, e.g. for microprocessor interfacing of slow signals. The same type of feedback loop can also be used in other domains, e.g. mechanical tracking systems. Some offset correction systems, e.g. [195], also shows similarity with a tracking analog-to-digital converter.

A "dual-slope" converter is suited for slowly varying input signals. In Fig. 8.70 an integrator circuit integrates the input signal during the fixed sample period. During a second time frame a reference current discharges the integrator, while a counter measures the number of clock periods. The maximum number of clock cycles is around 2×2^N. The input integration has a transfer characteristic comparable to a sample and hold circuit:

$$H(\omega) = \frac{\sin(\pi \omega T_{\text{int}})}{\pi \omega T_{\text{int}}} \tag{8.41}$$

The low-pass characteristic of this operation makes that this principle is tolerant to RF noise and interference.

A dual-slope converter is an example of a zero-point detecting method or zero-crossing method. The converter determines the value of the unknown signal by subtracting an equivalent signal from the digital-to-analog converter until the zero starting level is reached. The advantage of zero-crossing methods is that the system needs to be linear only around the zero level. Voltage dependency of the integration elements (non-linear capacitor) in Fig. 8.70 will not impair the conversion accuracy. Moreover an offset in the comparator or integrator is implicitly canceled by the operation as long as the offset in the crossing of the rising edge is still present when the signal returns to zero at the end of the cycle. Hysteresis around the zero crossing can not be tolerated.

Dual-slope converters find their application especially in the harsh industrial environments and multi-meters.

8.8 Time-interleaving Time-discrete Circuits

All converter types except for the full-flash converter, need several intermediate timing cycles to convert a signal. This reduces the overall throughput speed. De-

Fig. 8.71 Three different errors in a time-interleaved or multiplexed time-discrete structure
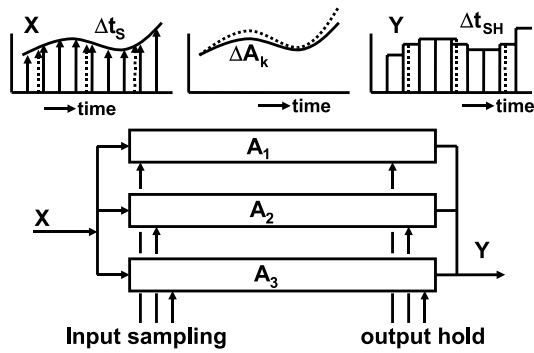
multiplexing or time-interleaving the signal over several identical structures allows to match the limited speed of the basic structure to the high-speed requirements. Time-interleaving can also help reduce the power consumption. Equivalent to the situation in digital circuits, demultiplexing will not reduce the number of steps to be taken, on the contrary the number of steps may increase. Power saving is possible because the steps can run at lower speed with lower (stand-by) currents and supply voltages.

Figure 8.71 analyzes multiplexing of analog time-discrete circuits. A chain of analog-to-digital and digital-to-analog converters is included. The disadvantage of multiplexing is of course the unwanted output signal which consists of a fixed pattern noise, due to DC inequalities, and uncanceled folding products around multiples of the sample frequency of the individual structures. The problem in designing a chip with analog multiplexing is that a two-dimensional wiring pattern has its limitations in achieving a fully identical structure. E.g. a one-layer metalization gives asymmetrical crossings. And the speed of signal propagation is roughly half of vacuum, so 1 picosecond corresponds to 150 μm. Unwanted components to the output signal are caused by random and systematic errors.

The systematic errors can be subdivided according to their origin:

- Technological fixed errors are reduced if all multiplexed structures have the same orientation on chip: no mirrored or rotated placements, see Table 11.6. Although it is important to keep identical circuits close together (e.g. in order to avoid gradients), a trade-off is necessary with respect to the proximity effects of other structures. Many technological fixed errors cause either patterns at the multiplex frequencies or gain errors.
- Electrical errors such as voltage-drop errors are reduced by star-connected signal, pulse and power wiring, see Fig. 8.20. RC time constants should be matched in value, but preferably consist of identical components. Digital and analog power supplies are separated, however the common substrate is difficult to avoid and precautions in the digital part have to be taken as well.
- Timing errors affect in first order only the input sampling and the output restoration. The relative position of the pulses has to be accurate. Moreover, fast edges are important to avoid that the moment of sampling or holding becomes signal-dependent.
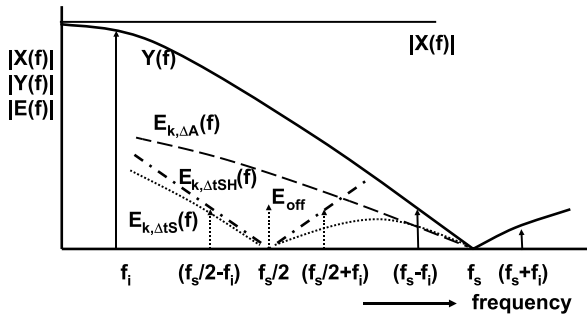
Fig. 8.72 The signal and the trajectory for different error components due to two-times multiplexing

In Sect. 7.4 several techniques are discussed to avoid additive errors. Here especially the multiplicative errors and timing errors are analyzed.

The error signal due to multiplexing of time-discrete structures can be analyzed by means of the spectrum of the output signal [196, 197], see Fig. 8.72 for $N = 2$.

The time continuous analog input signal $X(f)$ is in this analysis sampled with a frequency f_s and reconstructed to a signal $Y(f)$ with a zero-order hold function with a hold time T_s. The signal is multiplexed over N branches with N samplers in N multiplex phases, each running at $f_{sN} = f_s/N$ and spaced in time at kT_s with $k = 0 \ldots (N-1)$. At the end of the chain the N signals are combined after the sample-and-hold operation over a period T_s.

The spectral contribution of a signal in multiplex line k to the spectrum of the output signal is identical to a sampling system running at frequency $f_{sN} = f_s/N = 1/NT_s$ and is written as (compare (3.12) and (4.2)):

$$Y_k(f) = \sum_{r=-\infty}^{r=\infty} A_k X(j2\pi(f - rf_{sN}))e^{j2r\pi f kT_s}\frac{\sin(\pi f T_s)}{\pi f T_s}e^{-j\pi f T_s} \quad (8.42)$$

with:

$Y_k(f)$ = the contribution of line k to the output signal $Y(f)$,
r = index of spectrum repetition around f_{sN},
A_k = amplification of line k,
$X(f)$ = the spectrum of the input signal,
T_s = the period of the overall clock frequency.

The formula is composed of the amplification A_k, the sampled spectrum $X(f)$, the multiplexing phase shift and the sample-and-hold $\sin(\pi f T_s)/\pi f T_s$ function including the delay term. The analog input spectrum is repeated around multiples of the subsample frequency f_{sN}, but each structure k has a phase shift with respect to structure $(k-1)$ corresponding to the time delay between the sampling moments T_s.

Adding up the outputs of the N lines, ideally only gives non-zero contributions for those values of r that are multiples of N. All other spectra around multiples r or

f_{sN} (except those which are multiples of f_s itself) extinguish. The ideal total output of this sampled data system is therefore:

$$Y(f) = \sum_{m=-\infty}^{m=\infty} AX(j2\pi(f - mf_s)) \frac{\sin(\pi f T_s)}{\pi f T_s} e^{-j\pi f T_s} \tag{8.43}$$

m is the index of spectrum repetition around f_s. The errors that cause uncanceled components in the output spectrum, can be subdivided into four categories:

- The first error is a DC-offset between the N branches. This error leads to fixed frequency components at multiples of the branches:

$$E_{\text{off}}(t) = \sum_{r=0}^{r=\infty} V_{\text{off},r} \sin(j2\pi r f_{sN} t) \tag{8.44}$$

- A static time error Δt_s in the input sampling of line k will result in:

$$\left| \frac{E_{k,\Delta t}(j2\pi f_r)}{AX(j2\pi f_i)} \right| = \frac{2}{N} \sin(\pi f_i \Delta t_s) \frac{\sin(\pi f_r T_s)}{\pi f_r T_s} \tag{8.45}$$

- An amplification error ΔA_k in line k leads to:

$$\left| \frac{E_{k,\Delta A}(j2\pi f_r)}{AX(j2\pi f_i)} \right| = \frac{1}{N} \frac{\Delta A_k}{A} \frac{\sin(\pi f_r T_s)}{\pi f_r T_s} \tag{8.46}$$

- A time error Δt_{SH} in the sample-and-hold pulse in line k gives:

$$\left| \frac{E_{k,\Delta SH}(j2\pi f_r)}{AX(j2\pi f_i)} \right| = \frac{2}{N} \sin(\pi f_r \Delta t_{SH}) \frac{\sin(\pi f_i T_s)}{\pi f_r T_s} \tag{8.47}$$

The errors are expressed as the amplitude ratio of a spurious output signal $E_k(j2\pi w f_r)$ at signal frequency: $f_r = f_i + r f_{sN}$ with respect to the amplified input signal at frequency f_i. The three errors due to uncanceled folding components can easily be recognized: the timing errors have input-frequency dependent amplitudes, and the amplification and input timing errors are filtered by the sample/hold transfer function.

Figure 8.72 shows the error components in the case of $N = 2$. E_{off} is the fixed error at f_{sN}. At $f_s/2 \pm f_i$ two errors are indicated caused by time-interleave errors. The dotted and dashed lines $E_k(f)$ indicate the trajectory of each component. The timing related errors extinguish if the input frequency is low, because a timing error will not lead to a change in input/output signal. Consequently a test with a proper choice of f_i will allow to analyze the type of error in a practical multiplexed device.

8.9 An Implementation Example

Based on a single comparator architecture a full-flash, a successive approximation and a multi-step pipeline converter are described.
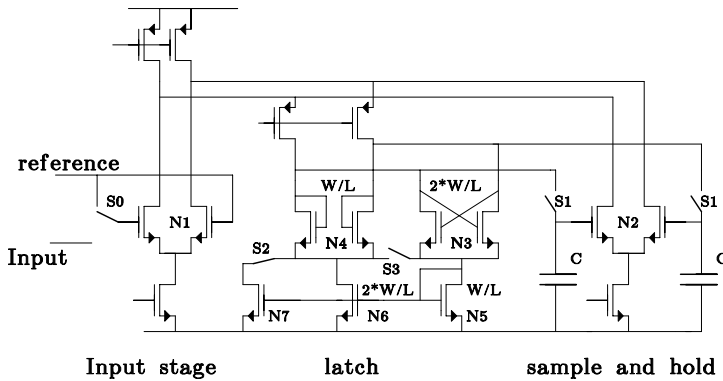
Fig. 8.73 The basic comparator schematic. Switches are shown in the position for the amplification phase. *Design*: J.v. Rens

8.9.1 An Auto-zero Comparator

Random-offset reduction is necessary to achieve 8-bit DNL performance. In most offset reduction schemes the offset+signal and the offset are determined at different points in time, so at least one must be stored in a capacitor. Authors of [153, 168] describe a well-accepted method to perform offset reduction with the aid of a capacitor at the input of the comparators which switches between input and reference, see Fig. 8.16. The disadvantage of this method is that the high capacitive input load requires a low-ohmic ladder (300 Ω) and poses high demands on the circuitry for driving the analog-to-digital converter. The single-sided input capacitor has a large parasitic capacitance to a bouncing substrate in a mixed signal chip. The inherent capacitive input voltage division attenuates the signal. The basic inverter-type comparator has a poor power supply rejection [169], which affects the performance if power supply bouncing occurs between coarse and fine cycles. These comparator design points have to be solved for embedded operation.

Figure 8.73 shows the comparator that is used in the full-flash analog-to-digital converter designs (see Sect. 8.9.2). In two other analog-to-digital converters discussed in Sects. 8.9.3 and 8.9.4, modified comparators that allow larger signals, are used.

With the exception of input and ladder terminals, the design is fully balanced and the PMOS current sources allow a good PSRR. The design consists of an input stage N1, from which the signal is fed into a sample-and-hold stage N2, S1, C. Comparison involves three cycles: sampling, amplification and latching. During the sampling phase switch S1 is conducting and $V_{in} - V_{ref} + V_{off}$ is stored on both capacitors. These capacitors are grounded on one side and do not suffer from parasitic coupling to substrate. V_{off} represents the sum of all the offsets in the comparator. During the sampling phase the negative conductance of latch stage N3 is balanced by the positive conductance of the load stage N4. Their combination acts as an almost infinite impedance, which is necessary for good signal+offset storage. The

latch stage has twice the W/L of the load stage, but its current is only half due to the current mirror ratio N5, N6. For large differential signals the effect of the factor 2 in W/L of N3 and N4 becomes important: the conductance of N3 reduces at a much higher rate that of N4, so the effective impedance of N3–N4 collapses. The large-signal response is consequently improved by the reduced time constant. The feedback of N2 via the switches S1 allows a 150 MHz bandwidth, resulting in a high-quality S/H action.

After the sampling phase, the switch S0 at the input connects to the reference voltage, effectively disconnecting the common input terminal from the comparator. As switches S1 are disconnected, the sample-and-hold stage will generate a current proportional to $V_{in} - V_{ref} + V_{off}$, while the input stage and the rest of the comparator generate only the part proportional to V_{off}. The (differential) excess current is almost free of offsets and will be forced into the load-latch stage N3–N4. Switch S2 is made conductive, which increases the conductance of N4 and decreases the conductance of N3; the gain of N2 on N3–N4 is now about 8 for small differential signals.

Finally, S3 is made conductive; the current now flows in a 2 : 1 ratio into N3–N4, thereby activating the latch operation. The latch decision is passed on to the decoding stage and a new sample can be acquired. Remaining random offsets (approx. 0.4 mV) are caused by limited gain during the sampling phase, charge dump of S1 and the difference in matching contributions of N3 and N4 in the sampling and amplification phases.

8.9.2 Full-flash Analog-to-Digital Converter

The comparator described in the previous section forms the basis for 3 types of converters: an 8-bit full-flash analog-to-digital converter, a 10-bit successive-approximation analog-to-digital converter and the so-called multistep analog-to-digital converter, which combines a 3-bit full flash with a 5-bit successive approximation in a multiplexed approach.

The full-flash analog-to-digital converter has been subdivided into 8 sections of 32 comparators; each section uses local 5-bit Gray-coding for the LSBs. This scheme allows fast and efficient (2 MOS/comparator) decoding. The 32nd, 64th, etc. comparator in a 3-bit MSB analog-to-digital converter decides which section output will be passed to the data rails.

The ladder structure is identical to that described in Sect. 7.5.1: a two-ladder structure provides a good quality reference voltage. Due to the fact that the comparator uses a differential input pair, the input load and reference ladder load are limited to parasitic charges of the input switch. The reference supply current and the decoupling required in this design are both small. Care has been taken to avoid delay skew between the on-chip-generated clock scheme and the input signal; the generation of bias voltages for 255 comparators and the power distribution also requires a careful design and layout.

The main characteristics of the analog-to-digital converter are summarized in Table 8.2.

Fig. 8.74 Signal-to-noise-and-distortion at 20 Ms/s and 25 Ms/s as a function of the input frequency
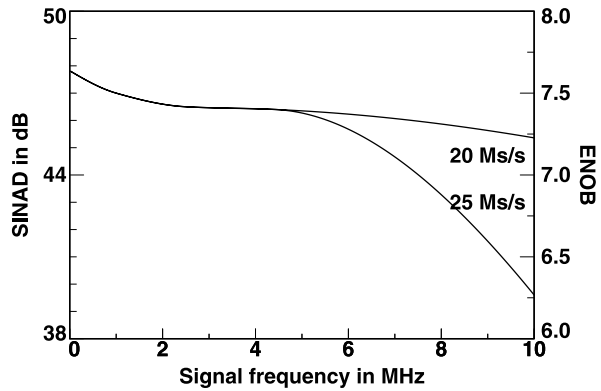
Fig. 8.75 Successive-approximation analog-to-digital converter
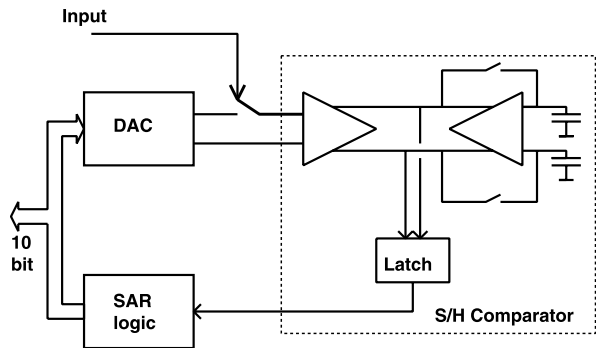
Figure 8.74 shows the measured SINAD in dB; the ENOB has been plotted on the right axis. The converter achieves an ENOB of 7.5 bits at low signal frequencies and 7.4 bits at 4.43 MHz and 20 Ms/s.

In order to the test substrate immunity, a 17.7 MHz 200 mV$_{pp}$ pulse wave was applied to the substrate of a 13.5 Ms/s running analog-to-digital converter digital-to-analog converter combination. The resulting 4.2 MHz (17.7–13.5) disturbance was 45 dB below the 4.4 MHz converted signal at the digital-to-analog converter output. The analog-to-digital converter in 1 μ CMOS technology was used on several prototype and production chips [198–200].

8.9.3 Successive-approximation Analog-to-Digital Converter

Basically, the successive-approximation technique consists of comparing the unknown input voltage with a number of precise voltages generated by a digital-to-analog converter, by means of a single comparator (see Fig. 8.75). As the input voltage may not change during the comparisons, there has to be a sample and hold circuit in front of the actual analog-to-digital. The offset-compensated comparator described in the previous section has a built-in sample-and-hold stage and can
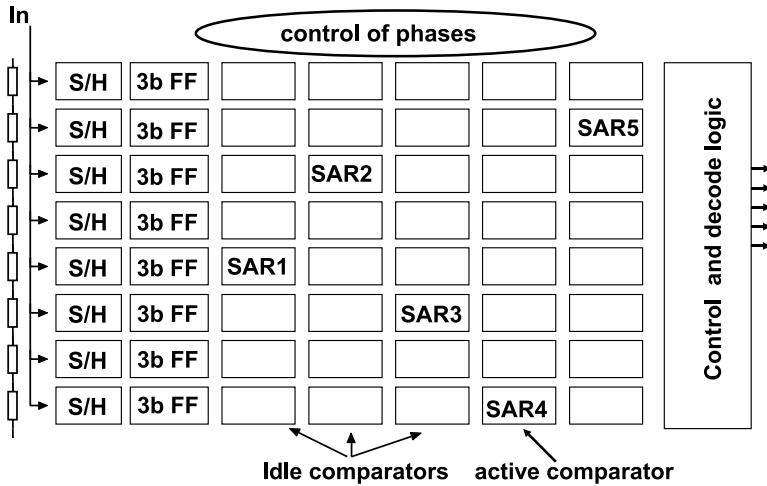
Fig. 8.76 Multi-step analog-to-digital converter. *Design*: J.v. Rens

therefore be used for comparison as well as for storage of the input signal. The SA analog-to-digital converter realized consists of one offset-compensated comparator, a 10 bit digital-to-analog converter and some control logic for controlling the digital-to-analog converter and storing the intermediate results. The converter requires 11 clock cycles for complete conversion. Therefore, the actual sampling frequency is limited to about 1–2 Ms/s. Inherent to the architecture, the DNL of the analog-to-digital converter is good. A disadvantage of the converter is the relatively small signal input range (or the risk of INL errors in the case of large signal swings) which is a consequence of using long-channel tail pairs at the input of the comparator to improve the common mode rejection.

The successive-approximation analog-to-digital converter design is used in servo applications and micro-controller inputs. Specifications are given in Table 8.2.

8.9.4 Multi-step Analog-to-Digital Converter

The multistep analog-to-digital converter[7] (Fig. 8.76) is an 8 bit converter based on a technique involving a combination of successive approximation, flash and time-interleaving. The 3 most significant bits of the conversion are determined by means of a flash conversion; the remaining 5 bits are realized through successive approximation. Multiple time-interleaved signal paths have been used to increase the maximum sampling frequency.

The hardware of the analog-to-digital converter consists of an array of 56 comparators with a built-in sample and hold stage. The array is grouped in 7 channels

---

[7]This section is based on a design by Jeannet v. Rens.
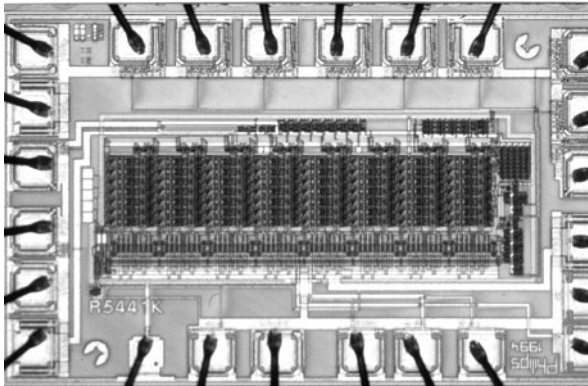
Fig. 8.77 Photograph of multi-step analog-to-digital converter TDA8792

of 8 comparators in a flash structure. The channels operate in a time-interleaved manner. The actual conversion takes place in 7 clock cycles. First, the input signal is sampled by, and stored in, one channel of 8 comparators (sample phase). A flash decision generates the three coarse bits and selects the comparator that stores the replica of the unknown input signal closest to a reference voltage. This comparator is used in a successive approximation loop to determine the remaining bits, while the other comparators are idle.

The input ranges of the eight comparators in that form the flash structure, determine the signal input range of the analog-to-digital converter. Note that use of parallel signal paths can be successful only if the different channels match well. Offset, gain and timing mismatches between multiple channels give rise to fixed patterns which manifest themselves as spurious harmonic distortion in the frequency domain. The effect of offset is minimized by the use of the previously described offset-compensated comparators. Gain mismatch is minimized by the use of a common resistor ladder digital-to-analog converter and timing mismatch by the use of a master clock which determines the sampling moments of all the channels. The production device is shown in Fig. 8.77.

This converter was produced as a stand-alone product TDA8792 and was the building block for many video integrated circuits.

8.9.5 A Comparison

Table 8.2 shows the main specifications of the three analog-to-digital converters. Basic to all converters is the comparator of Fig. 8.73, in which signal speed and accuracy have been traded-off versus power. The decisive factor for the power comparison is the comparator current. In the successive approximation design this current is higher because this comparator has to handle a larger signal span. The lower kickback of the single comparator in the successive approximation analog-to-digital converter also makes it possible to increase the ladder impedance. All the converters

Table 8.2 Specifications of the three analog-to-digital converters. All the converters are based on the comparator shown in Fig. 8.73

A/D converter	Full-flash	Successive app.		Multi-step
Resolution	8 bit	10 bit		8 bit
Sample rate	25 Ms/s	2 Ms/s		30 Ms/s
Differential linearity	0.6 LSB	0.5 LSB		0.5 LSB
Integral linearity	0.6 LSB	1 LSB		0.6 LSB
ENOB at input 4.43 MHz	7.4	8.5 (1 MHz)		7.4
SINAD (4.43 MHz)	46 dB			46 dB
SD (2–5 harm, 4.43 MHz)	>52 dB			>52 dB
Input bandwidth (1dB)	>70 MHz	20 MHz		70 MHz
Input signal swing	2 V	1.5 V		1.6 V
Ladder resistance	1200 Ω	4800 Ω		1200 Ω
Active area	2.8 mm^2	1.2 mm^2	0.4 mm^2 (8 bit)	1.1 mm^2
Technology		0.8–1 μ (1 PS, 2 Al)		
Current	55 mA	3 mA		13 mA
Current/comparator	200 μA	500 μA		200 μA
Number of comparators	256	1		56

have been extensively used in consumer ICs: digital video, picture-in-picture, instrumentation, etc. In 0.5 μm CMOS the 8-bit multi-step runs at 50 Ms/s, while the 9-bit version achieves 8.2 ENOB.

8.10 Other Conversion Ideas

Many other principles exist to convert signals from the physical domain to bits. Not all of them are relevant for a larger community, yet some of them may be considered in specific circumstances.

8.10.1 Level-crossing Analog-to-Digital Conversion

In the previous sections analog-to-digital converters were designed by sampling the signal and subsequently quantizing the signal to reference levels. In this process rounding errors occurred, that were labeled quantization errors. The sequence of sampling and quantizing can also be reversed. The level-crossing analog-to-digital converter in Fig. 8.78 [201, 202] generates a new digital output code at each time moment an amplitude quantization level is passed. With infinite time resolution this level-crossing algorithm will lead to a digital representation of the input signal with

Fig. 8.78 The principle of a level crossing analog-to-digital converter
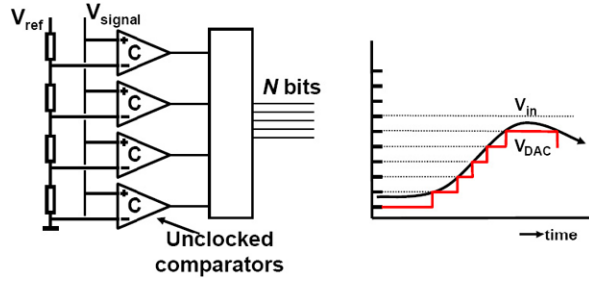
some harmonic distortion, depending on the density of the levels. There is no folding back of spectra and a rather high quality signal representation be obtained. In a conventional digital system, it is impractical to process this pulse-width modulation signal and therefore a rounding to a time grid is needed. That step introduces rounding errors and quantization problems. Suppose that the rounding will be towards a time grid specified by a sample frequency f_s with a time period T_s. If the level crossing occurs ΔT_s before a sample moment $n T_s$, the amplitude error is:

$$\Delta A(n T_s) = \frac{d V_{\text{in}}(t)}{dt} \Delta T_s \tag{8.48}$$

Assuming that the signal $V_{\text{in}}(t) = A \sin(\omega t)$ is so slow that only one level is passed during a time period T_s, that the probability of the occurrence of a level crossing moment is uniformly distributed, and that the level crossing is independent of the signal derivative, the expectation value of the error is:

$$E((\Delta A)^2) = \left(\frac{d V_{\text{in}}(t)}{dt}\right)^2 \times E\left(\Delta T_s^2\right) = \frac{A^2 \omega^2 T_s^2}{6} \tag{8.49}$$

The resulting signal-to-noise ratio between signal power and error power is:

$$\text{SNR} = 10 \log(3) - 20 \log(\omega T_s) \tag{8.50}$$

An increase in sample rate or a decrease in sample time of a factor two results in 6 dB of signal-to-noise ratio or the equivalent of 1 bit. When discussing oversampling in Sect. 9.1, the increase in sample frequency in a converter must be a factor of four higher to gain the same amount of signal-to-noise ratio. In a level-crossing device the time axis serves as the quantization axis and a frequency doubling doubles the accuracy. In the oversampled situation, the time axis only serves to spread out the quantization energy.

The requirement that the signal passes no more than one quantization level in one sample period makes this principle not suited for many application domains.

8.10.2 Asynchronous Conversion

Asynchronous converters do not wait for a level-crossing trigger, but constantly monitor the signal. Several low-resolution high-speed examples are reported [203,

Fig. 8.79 The asynchronous successive approximation analog-to-digital converter [203]
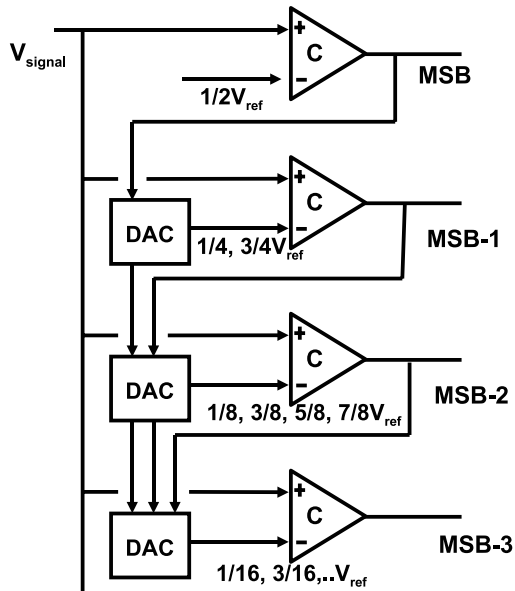
204]. An example is shown in Fig. 8.79. The unclocked MSB comparator continuously monitors the signal. The MSB bit is fed into a simple digital-to-analog converter that supplies either a $1/4V_{ref}$ or $3/4V_{ref}$. The second comparator uses this input reference to determine the MSB-1. If a signal passes through its range the converter will follow the digital code. A settling time of a few nanoseconds for 6 bits suffices.

Comparable to the level-crossing analog-to-digital converter the interfacing to the (clocked) digital world introduces quantization errors.

8.10.3 Time-related Conversion

In some systems, where a full analog-to-digital converter is not the optimum solution for area or power reasons, the conversion of a voltage to an intermediate quantity (such as a frequency) can be a solution. A voltage controlled oscillator (VCO) is by principle a voltage-to-frequency converter. The frequency deviation is in first order proportional to the voltage deviation. Both domains, input and output, are time-continuous and amplitude continuous, see Fig. 8.80. However a frequency is easily mapped on the digital domain by using a time window to count the number of frequency periods. The resolution is proportional to the time window and the sample rate is inversely proportional.

In some systems a time interval contains the required information or a time interval must be monitored. The class of time-to-digital converters is based on principles resembling the counting analog-to-digital converter. A high frequency clock is
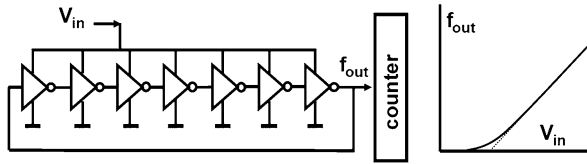
Fig. 8.80   VCO as voltage-to-frequency converter
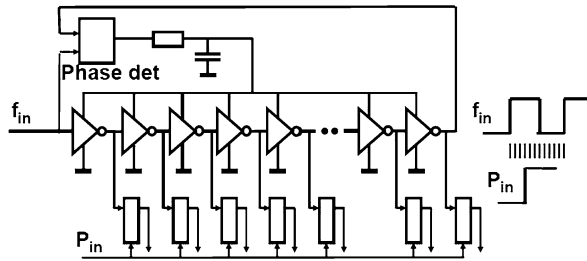
Fig. 8.81   Time-to-digital converter [205]
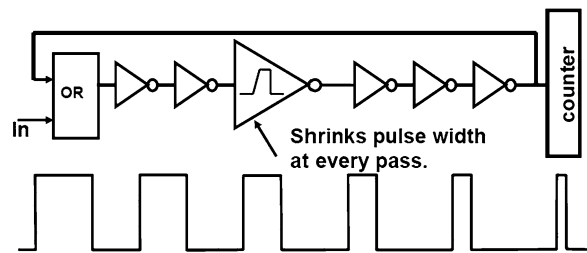
Fig. 8.82   Starving pulse time-to-digital converter [206]

counted during the interval of interest. In case this interval becomes too small, such a simple technique is not practical as the counting frequency is too high. Various structures as in Fig. 8.81 allow to subdivide the fast clock pulse and the resulting set of time-shifted pulses is used to clock the pulse P_{in} at a resolution of e.g. 1/32 of the clock period. The accuracy is limited to jitter in this system, but can achieve a 10 ps range [205].

An elegant implementation is the starving pulse converter [206]. The pulse that has to be measured is entered into a ring of inverters. If the inverters show perfectly symmetrical rising and falling edges without delay, the pulse would travel indefinitely in the inverter ring. One inverter is deliberately modified. Its rising edge is slow, so at each pass the pulse will become a time fraction shorter. The original pulse width is measured by counting the number the pulse before the signal is extinguished, see Fig. 8.82.

Fig. 8.83 The Vernier or
Nonius principle
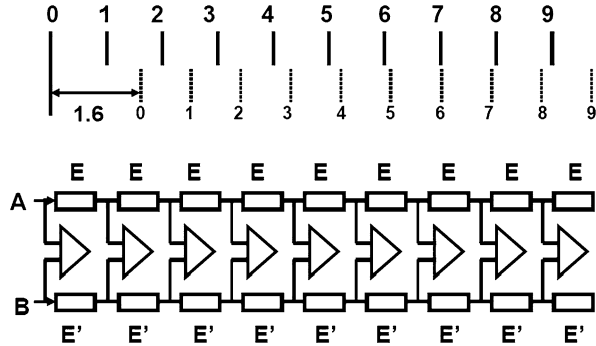
8.10.4 The Vernier/Nonius Principle

The need for higher resolution is not unique to the field of analog-to-digital conversion. In mechanics the Nonius or Vernier scale (after 16th and 17th century Spanish and French mathematicians) are widely used to determine accurately the subdivision in a primary scale. Figure 8.83 (upper) shows a primary scale and a secondary scale (dashed). The secondary scale has a subdivision of 0 to 9 and spans 90% of the primary scale. The unknown distance offsets the zero point of the primary scale and the zero point of the secondary scale. The distance equals the entire number of primary scale units (1 in this example) plus that fraction which is denoted by the number in the secondary scale where the marks of the primary and secondary scale are in line (in this example 6). This principle can be used in electronic designs as is shown in Fig. 8.83 (lower). The chains with elements E and E′ form the primary and secondary scales while the comparators determine the position where both scales are "in-line". This principle can be used to convert a time period [207]. The elements are implemented as timing cells with slightly different delays. On terminals A and B the start and stop of the interval under measurement are applied. The same technique can also be used with elements E implemented as resistors forming a flash-like converter [208].

8.10.5 The Floating-point Converter

Most analog-to-digital converters operate on the assumption that over the entire range the same resolution is required. In some systems such as sensors, the required resolution varies with the amplitude of the signal. In computing, this sort of problems is addressed by means of floating-point arithmetic. Also in analog-to-digital conversion floating-point conversion is possible, although most system engineers prefer a fixed-point converter. Figure 8.84 shows an example of a floating point analog-to-digital converter. The actual analog-to-digital converter can use any architecture. The floating point mechanism is around the converter and consists in its basic form of an analog pre-scaler like a variable gain amplifier and a digital
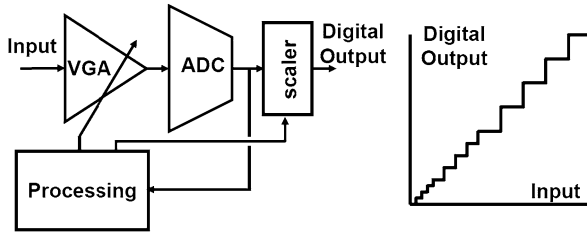
Fig. 8.84 The floating point principle and its transfer curve

post-scaler. A processing unit detects whether the signal is sufficiently large to use the analog-to-digital converter optimally. If the signal is too large or too small, the processing unit will adapt the input and output scaling. If both these units run with inverse amplifications, the transfer curve will show a resolution-amplitude ratio that remains within certain limits. The difficult point in this design is the accuracy of the pre-scaler. The output level with a gain of "2A" should be the exactly double of the gain "A". Offsets and gain errors are limiting this concept. A more detailed analysis is found in e.g. [209]. The application of this sort of converters is in sensor interfaces.

Chapter 9
Sigma-delta Modulation

Analog-to-digital converters for broad-band signals are realized using the various converter principles of the previous chapter. In many applications the bandwidth is limited, e.g. in various forms of communication. In advanced CMOS processes sampling and processing frequencies of hundreds of MHz are at the disposal of the designer. This allows a different set of principles for analog-to-digital conversion. De Jager [210], Cutler [211] and Widrow [212] were the first researchers to propose the idea that a reduced accuracy per sample can be compensated by using a high sample rate to convert a limited bandwidth. Inose [213] proposed to place a filter operation in the loop, thereby creating a sigma-delta modulator topology.

In the next sections the various steps towards efficient narrow band analog-to-digital conversion are discussed: oversampling, noise-shaping and sigma-delta conversion.

9.1 Oversampling

In the process of sampling and quantization an error signal is generated which is fundamentally a distortion component. In order to handle this quantization error, the multitude of distortion components folded back by various multiples of the sampling frequency is approximated as white noise as long as the resolution N is sufficiently large and no correlation appears between the input signal and the sample rate (e.g. one is not an integer multiple of the other). This quantization energy has a fixed amount of power $V_{\text{LSB}}^2/12$ for a given resolution N and is distributed evenly over the band from 0 to $f_s/2$ and mirrored into the alias bands.

If the sample rate of an analog-to-digital converter is increased from the sample rate needed to fulfill the Nyquist criterion $f_{s,ny}$ to a new frequency f_s the noise power density (power per Hertz) is reduced with the ratio of the sample rates. Figure 9.1 shows the increase of the sampling rate by a factor of four. As the noise power density is reduced by a factor of four, the amplitude noise density is reduced by a factor of two. In a fixed bandwidth the signal-to-noise ratio will increase by this
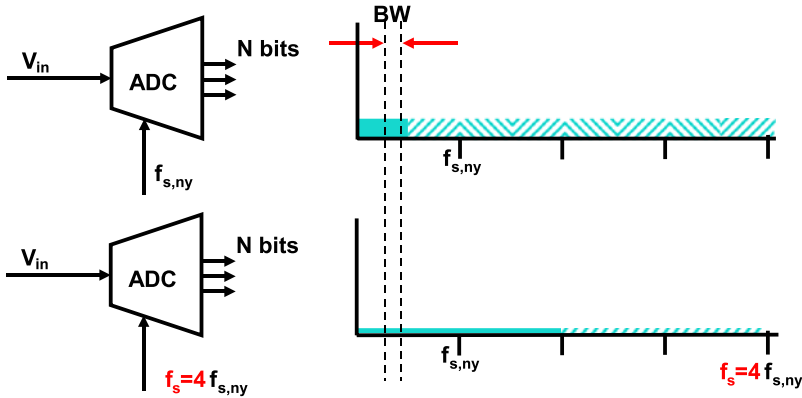
Fig. 9.1 Oversampling in analog-to-digital conversion. In the derivations the lower boundary of the bandwidth is set to 0 Hz

factor of four in power density and the effective resolution will artificially increase by one bit. The ratio

$$\text{OSR} = \frac{f_s}{f_{s,ny}} = \frac{f_s}{2f_b} \tag{9.1}$$

is called the oversampling ratio (OSR). The bandwidth BW of the signal is for the ease of the calculation assumed to range from 0 to f_b. Similar results are obtained for bands at other frequency locations.

In a Nyquist converter the quantization energy Q^2 is supposed to be uniformly distributed in the frequency band till half of the sample rate. The corresponding noise power density is equal to the total noise divided by half of the sampling rate $f_s/2$:

$$\frac{Q^2}{f_s/2} = \frac{V_{\text{LSB}}^2/12}{f_s/2} = \frac{V_{\text{LSB}}^2}{6f_s} \tag{9.2}$$

In a band from 0 to f_b the total noise power Q_b^2 is found by integrating the noise density over the band from 0 to f_b resulting in a Nyquist noise power:

$$Q_b^2 = \frac{V_{\text{LSB}}^2 f_b}{6f_s} = \frac{V_{\text{LSB}}^2}{12} \frac{1}{\text{OSR}} \tag{9.3}$$

The above relation will be used to compare the upcoming results for noise shaping. The noise power in a fixed bandwidth reduces proportionally to the oversampling rate. The gain in signal-to-noise ratio and in effective number of bits in a fixed bandwidth is:

$$\Delta \text{SNR} = {}^{10}\log\left(\frac{f_s}{f_{s,ny}}\right) = {}^{10}\log(\text{OSR}) \tag{9.4}$$

$$\Delta \text{ENOB} = \frac{1}{2}{}^2\log\left(\frac{f_s}{f_{s,ny}}\right) = \frac{1}{2}{}^2\log(\text{OSR}) \tag{9.5}$$
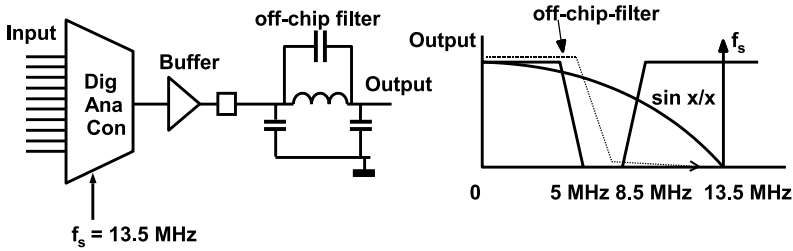
Fig. 9.2 A digital-to-analog converter running at Nyquist rate with the output filtered by an external filter

Oversampling is not effective for signals where the assumption that the quantization process can be approximated by white noise is not valid. The signal-to-noise ratio of the conversion of DC-signals cannot be improved by oversampling alone. A helper signal must be present to create the necessary conditions, e.g. in the form of dither, see Sect. 5.3.3.

The main advantage of oversampling in analog-to-digital conversion is the extra frequency range that is obtained between the alias band and the band of interest, see Fig. 3.4. The specifications of the alias filter can thereby be relaxed, although the higher sample rate will lead to higher power consumption on the digital side.

In a digital-to-analog conversion it is even more beneficial to increase the sample rate and avoid an operation mode close to the Nyquist limit. Figure 9.2 shows an example of a Nyquist digital-to-analog converter succeeded by an off-chip filter. At signal frequencies close to half of the sample rate, large transient steps will occur at the input of the on-chip buffer and the output pin. These steps cause slewing and distortion in the buffer. The buffer has to be designed with additional bandwidth and the input stages will need large bias currents. This set-up has a relatively poor performance close to $f_s/2$ due to $\sin(x)/x$ signal loss. The passive filter requires some 3–7 poles and is expensive to produce, especially if $\sin(x)/x$ compensation is also needed.

In order to create an oversampling version of the digital signal, a number of processing steps must be taken, see Fig. 9.3. Between the original samples at f_s new sample points are inserted with a value "0", in the example at $t = T_s/4, 2T_s/4, 3T_s/4, \ldots$. Although the frequency spectrum is still the same, this means that the hardware runs at $4f_s$, which is formally the new sample rate. Now a digital filter (see Sect. 3.2.1) removes the alias bands. In the time domain the filter operation will change the value of the inserted sample points to new values. Some authors refer to this method as "up-sampling" in order to emphasize the contrast with down-sampling of sigma-delta output described in Sect. 3.2.3.

Figure 9.4 shows an integrated circuit solution: the sample rate is locally doubled and the alias bands in the frequency spectrum are removed by digital filtering. Now the large transients in the output are more than halved in amplitude, and relatively simple non-critical post-filtering (first order) is sufficient to restore the analog signal. The inherent $\sin(x)/x$ is reduced by an order of magnitude, so no compensation
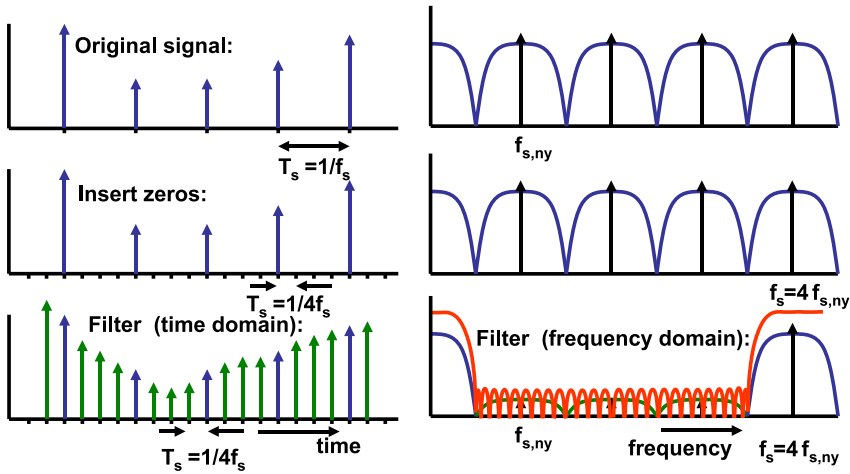
Fig. 9.3 Four-fold digital oversampling in time and frequency domain
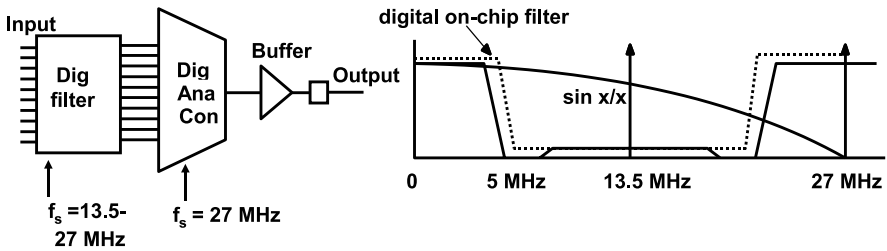
Fig. 9.4 A digital-to-analog converter driven by digital pre-filter that performs a two-time over-sampling

Table 9.1 Comparison of the two digital-to-analog conversion solutions in Figs. 9.2 and 9.4

Standard solution	Oversampling solution
External filter needed	Internal digital CMOS filter
10 mA current in driver	5 mA current in driver power for dig. filter
$sin(x)/x$ loss of 4 dB	$sin(x)/x$ loss $= 0.5$ dB $2\times$ clock needed

is required. From a power perspective a trade-off must be made between the additional power and area for the filter and the quality loss and additional power in the buffer (see Table 9.1). This trade-off is more in favor of an oversampling solution for more advanced CMOS processes. First the area and power of the digital filter shrinks, secondly the switching speed of the short-channel transistors allows high oversampling frequencies.

Fig. 9.5 Signal-to-noise improvement with noise shaping
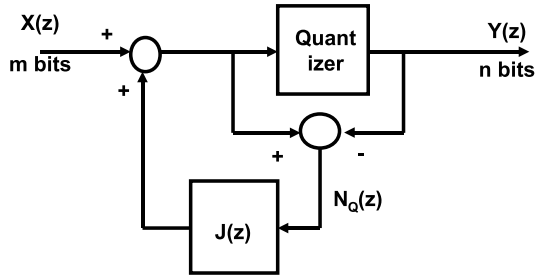
9.2 Noise Shaping

Oversampling is a useful concept for relaxing the alias filter and buffer specifications in the total conversion chain. A second reason for applying oversampling is that it creates an additional frequency span. This additional frequency range is used in noise shaping for shifting the quantization energy out of the wanted signal band into a part of the frequency span where it no longer affects the signal.

Figure 9.5 shows the basic structure of a noise shaper circuit. The quantization error is formed by subtracting the input and output signals from the quantizer. This quantization error is passed through a filter and fed-back to the input. This structure is not a classical feedback loop. There is no signal that is fed back to its own origin. The inverse of the quantization error is delayed and added to the signal. Low-frequency components in the error signal are effectively suppressed. At high frequencies the delay will cause a phase shift leading to an enhanced error energy level. In the z-domain the transfer of input and noise into the output is:

$$Y(z) = X(z) + [1 - J(z)]N_Q(z) \tag{9.6}$$

If the filter function $J(z)$ is chosen as a unit delay:

$$Y(z) = X(z) + [1 - z^{-1}]N_Q(z) \tag{9.7}$$

this transfer function can be visualized in the frequency domain via the transformation $z = e^{j\omega T_s}$:

$$Y(\omega) = X(\omega) + (1 - e^{-j\omega T_s})N_Q(\omega) \tag{9.8}$$

In order to determine the noise transfer to the output (NTF), the absolute value of the function in brackets is evaluated:

$$|\text{NTF}(\omega)|^2 = \left| \frac{Y(\omega)}{N_Q(\omega)} \right|^2 = |1 - e^{-j\omega T_s}|^2 = 2 - 2\cos(\omega T_s) \tag{9.9}$$

In the noise shaper the quantization energy density is shaped with the function $(2 - 2\cos(\omega T_s))$. Figure 9.6 shows the resulting noise power density for the filter function $J(z) = z^{-1}$. Integration of the noise power density function over the band from 0 to $f_s/2$ gives:

$$\frac{V_{\text{LSB}}^2}{6 f_s} \int_{f=0}^{f=f_s/2} 2 - 2\cos(\omega T_s)\, df = \frac{V_{\text{LSB}}^2}{6} \tag{9.10}$$

Fig. 9.6 The noise shaper pushes the quantization energy to higher frequencies
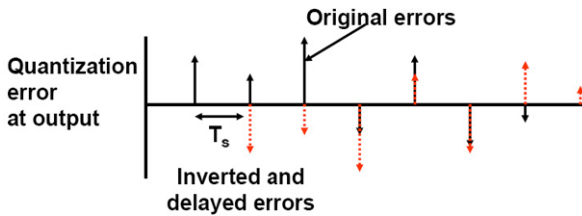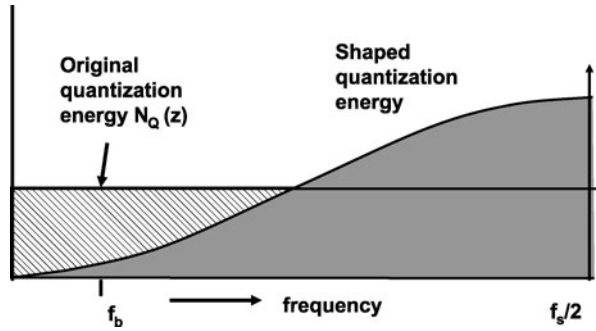
Fig. 9.7 The output of the noise shaper contains both the original error sequence and its inverted and delayed version. Both are not visible at the output as in this plot, because they are part of the quantized signal. This plot helps in realizing that the DC-content of the error combination is low

which equals twice the quantization noise power. This can be understood by considering that the amplitudes of succeeding quantizations are uncorrelated. The extracted quantization error is delayed and combined with the present quantization error to yield the output value, see Fig. 9.7. With two (in amplitude) uncorrelated quantization powers at the output, the total quantization error energy doubles, although it is shaped in the frequency domain.

For small values of the cosine argument the approximation $\cos(x) \approx 1 - x^2/2$ results in a noise power density:

$$\frac{V_{\text{LSB}}^2}{6 f_s}(2 - 2\cos(\omega T_s)) \approx \frac{V_{\text{LSB}}^2}{6 f_s}\left(\frac{2\pi f}{f_s}\right)^2 \tag{9.11}$$

The noise power density near DC is strongly reduced. Close to $f_s/2$ the power density is four times the Nyquist power density, so the noise shaper does not eliminate the noise but shifts it to a different frequency region.

The total noise power in a band from 0 to f_b in the noise shaper is found by integrating the noise power density over that frequency range:

$$\int_{f=0}^{f_b} \frac{V_{\text{LSB}}^2}{6 f_s}|1 - e^{-2j\pi f T_s}|^2 \, df = \frac{V_{\text{LSB}}^2}{3}\left[\frac{f_b}{f_s} - \frac{\sin(2\pi f_b/f_s)}{2\pi}\right]$$

$$\approx \frac{V_{\text{LSB}}^2}{12}\frac{\pi^2}{3}\left(\frac{2 f_b}{f_s}\right)^3 = \frac{V_{\text{LSB}}^2}{12}\frac{\pi^2}{3\text{OSR}^3} \tag{9.12}$$

Fig. 9.8 Noise shaping in
digital rounding
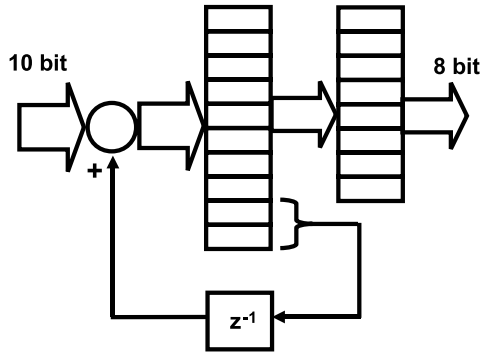
The total in-band noise of a noise-shaper with respect to only oversampling conversion is found by dividing this result by the total in-band noise of an oversampling conversion, (9.3):

$$\text{Noise power reduction} = \left[2 - 2\frac{\sin(2\pi f_b/f_s)}{2\pi f_b/f_s} \right] \tag{9.13}$$

Approximating the sine with $\sin(x) = x - x^3/3!$:

$$\text{Noise power reduction} = \frac{4}{3}(\pi f_b/f_s)^2$$
$$\text{Noise amplitude reduction} = \frac{2}{\sqrt{3}}(\pi f_b/f_s) \tag{9.14}$$

The total in-band noise power is related to the oversampling ratio OSR by a cube power in (9.12). One OSR term comes from the oversampling mechanism, the remaining OSR^2 term comes from the noise shaping loop. Doubling the oversampling ratio reduces the noise power by a factor 8, or 9 dB or 1.5 effective number of bits. Thereby noise-shaping is a powerful mechanism to improve the effective resolution of a converter.

A simple application of a first-order noise shaper is shown in Fig. 9.8. The problem here is to truncate an 10-bit digital data word to 8 bits. The truncation is a form of quantization. The two residue bits form the quantization error. These two bits are fed back via a delay and are added into the original signal. The DC-content of the signal is preserved, as all bit information is added together. However the consequence is that the carry-bit due to the successive addition of the 9th and 10th bit occurs at irregular intervals in the new output word.

Next to first order noise shaping, also higher order shaping is possible by extending the filter $J(z)$. If this function is chosen as an n-th order polynomial:

$$1 - J(z) = (1 - z^{-1})^n \tag{9.15}$$

the frequency transfer function becomes:

$$Y(\omega) = X(\omega) + (1 - e^{-j\omega T_s})^n Q(\omega) \tag{9.16}$$
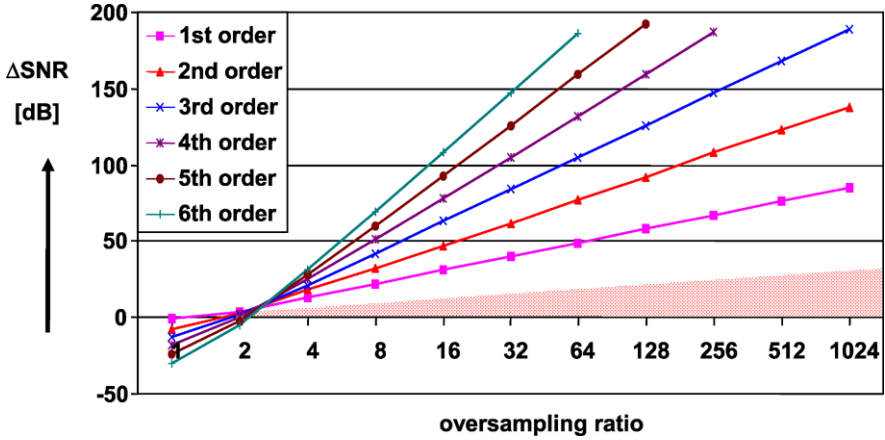
Fig. 9.9 The gain in signal to noise ratio for an *n*-th order noise shaper as calculated in (9.18). The *shaded triangle* is the portion that originates from oversampling alone. This plot is equally valid for sigma-delta conversion

This is certainly not the only possible higher-order filter for a noise shaper. However this choice results in a manageable mathematical description, showing the importance of the additional order of the filter. The total noise power in a band from 0 to f_b in an *n*-th order noise shaper is found in a similar way as the first order [8, formula 300]:

$$\int_{f=0}^{f_b} \frac{V_{\text{LSB}}^2}{6 f_s} |1 - e^{-2j\pi f T_s}|^{2n} \, df = \int_{f=0}^{f_b} \frac{V_{\text{LSB}}^2}{6 f_s} (2\sin(\pi f/f_s))^{2n} \, df$$

$$= \frac{V_{\text{LSB}}^2}{6\pi} \frac{(2n)!}{(n!)^2} \left[\pi \frac{f_b}{f_s} - \cos\left(\pi \frac{f_b}{f_s}\right) \sum_{k=0}^{k=n-1} \frac{2^{2k}(k!)^2 (2n)! \sin^{2k+1}(\pi f_b/f_s)}{(2k+1)!} \right] \tag{9.17}$$

The signal-to-noise ratio gain is found by dividing this result by the noise in a Nyquist band $V_{\text{LSB}}^2/12$:

$$\text{SNR gain} = \frac{2}{\pi} \frac{(2n)!}{(n!)^2} \left[\pi \frac{f_b}{f_s} - \cos\left(\pi \frac{f_b}{f_s}\right) \sum_{k=0}^{k=n-1} \frac{2^{2k}(k!)^2 (2n)! \sin^{2k+1}(\pi f_b/f_s)}{(2k+1)!} \right] \tag{9.18}$$

Figure 9.9 shows the signal-to-noise gain for 1-st till 6-th order noise shaping.

Assuming that f_b/f_s is sufficiently small, a simpler expression can be obtained for the noise power:

$$\int_{f=0}^{f_b} \frac{V_{\text{LSB}}^2}{6 f_s} |1 - e^{-2j\pi f T_s}|^{2n} \, df \approx \int_{f=0}^{f_b} \frac{V_{\text{LSB}}^2}{6 f_s} (2\pi f/f_s)^{2n} \, df$$

$$= \frac{V_{\text{LSB}}^2}{12\pi} \frac{(2\pi f/f_s)^{2n+1}}{2n+1} = \left(\frac{V_{\text{LSB}}^2}{12}\right) \times \left(\frac{1}{\text{OSR}}\right) \times \left(\frac{\pi^{2n}}{(2n+1)\text{OSR}^{2n}}\right) \tag{9.19}$$

In the last formulation the different processes have been made explicit: the quantization energy, the signal-to-noise gain by mere oversampling and the signal-to-noise gain by noise shaping. The noise power reduces by $(2n + 1) \times 6$ dB for a doubling of the oversampling rate. The SNR gain of n-th order noise shaping over Nyquist conversion is given as:

$$\Delta\text{SNR} = 10^{10} \log\left((2n + 1)\frac{\text{OSR}^{2n+1}}{\pi^{2n}} \right)$$
$$= 20^{10} \log\left(\frac{\text{OSR}^{n+0.5}\sqrt{2n + 1}}{\pi^n} \right) \tag{9.20}$$

This expression has an error of less than 1 dB with respect to (9.18) and Fig. 9.9 for oversampling ratio's larger than 4.

This expression is based on the assumption in (9.15) that $J(z^{-1})$ is a Butterworth filter. Other filters lead to different curves. Yet, this analysis gives a first impression on what improvement can be reached.

Noise shaping has been demonstrated here on abstract input and output signals. Noise shaping is used to quantize analog signals to the digital domain. Due to the change of domain around the noise shaper, the digital output signal must be converted back to the analog domain before the subtraction from the analog input signal can take place. In case of a gain error the output signal Y is (slightly) attenuated by $(1 - c)$ and the transfer becomes:

$$Y(z) = \frac{X(z)}{1 - cz^{-1}} + \frac{1 - J(z)}{1 - cz^{-1}} Q(z) \tag{9.21}$$

which implies a linear signal distortion for the input signal and an less accurate filter function for the quantization error. This application of the noise shaping idea requires some tuning of the analog signals, in order to obtain the correct transfer function.

9.3 Sigma-delta Modulation

An important solution for converting an analog signal into a low bit-width pulse stream is the sigma-delta[1] converter[2] is as shown in Fig. 9.10 [213–217]. Sigma-delta conversion as a form of signal quantization that can be used in analog-to-digital, digital-to-analog and digital-to-digital conversions of signals. In contrast to the noise shaper, both the signal component and the quantization error are circulated

---

[1]Is the term "sigma-delta" or "delta-sigma" more correct? Inose [213] uses in 1962 delta-sigma. The other argument is that the basic form was originally a delta-modulator which was extended with an summing function: a sigma-delta modulator.

[2]More language issues: "modulator" or "converter". In this book the circuitry around the quantizer is referred to as the modulator. The same circuit is called a converter if its system function is the dominant feature.
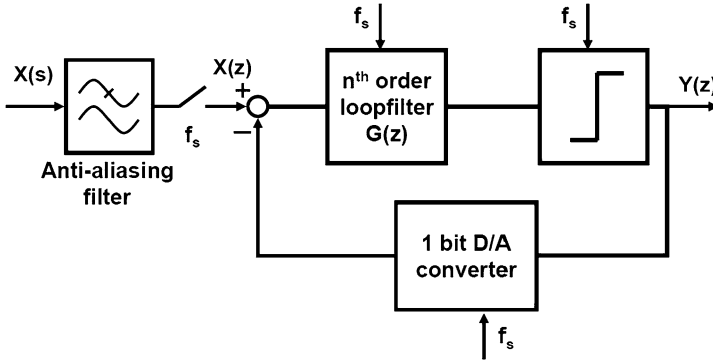
Fig. 9.10 Basic scheme of a sigma-delta modulator
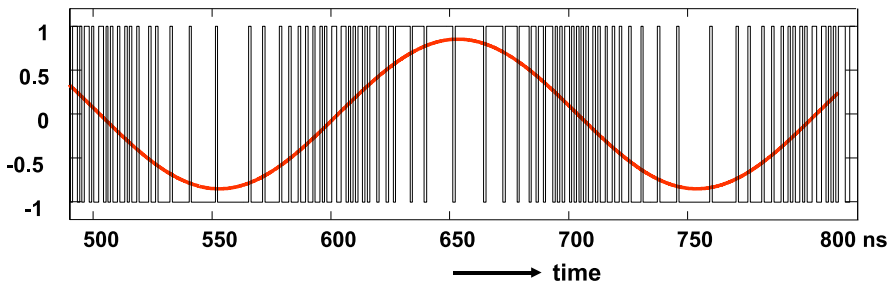
Fig. 9.11 A sine wave with an amplitude of 85% of the full scale is represented by a pulse density signal of the digital-to-analog converter

in the feedback path of the sigma-delta converter. The stability of the loop is the critical design challenge. The quantizer can vary from a simple comparator to a flash converter. In this scheme the quantizer operates on the filtered residue signal and reduces this residue signal to a multi-level discrete amplitude signal. The sigma-delta modulator produces a stream of bits even if the input signal is a DC-level. The observation that a sigma-delta modulator has some similarity with a voltage-controlled oscillator is helpful in understanding some aspects of the modulators behavior.

In the early realizations of sigma-delta conversion the quantizer is a comparator. The output signal consists of a one-bit signal with two levels: either $0, 1$ or more symmetrical $-1, +1$. In Fig. 9.11 a sine wave is shown with such a one-bit representation. The sine wave is here at 85% of the level of the feedback digital-to-analog converter. It is clear that a sine wave that reaches the maximum levels of the digital-to-analog converter cannot be properly represented, as the modulator has no modulation freedom anymore at the maximum excursion of the signal. The maximum peak-peak level of input signal with respect to the range of the feedback digital-to-analog converter that is correctly converted is called the maximum overload level $\alpha_{OL} \leq 1$. The overload level is mostly expressed in dB below the maximum and ranges between -2 and -4 dB. On top of that there is another factor
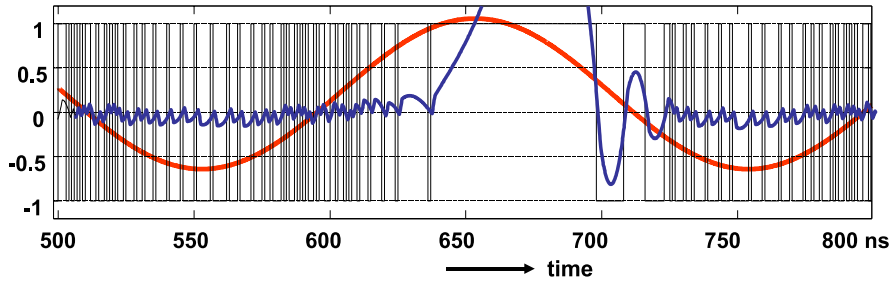
Fig. 9.12 A sine wave with an amplitude of 85% of the full scale and 20% bias creates an overload situation with instability. In blue is the signal after the filter

of 0.7 (-3 dB) between the power of a block wave and the power of an equal amplitude sine wave. The actual achievable overload level depends on the implementation of the converter. Input levels that (temporarily) exceed the overload level will cause the internal signals to clip. The output signal cannot be switched in a sequence that corresponds to the input. This situation is called instability and often manifests itself as a fixed output value, or a slow oscillation. During overload the integrators that form the filter will be driven into a saturated state and will recover only slowly. A short period of overload may result in a much longer instability, see Fig. 9.12.

The stable situation in a sigma-delta converter is a controlled oscillation with a frequency close to half of the sample rate. Stability in a sigma-delta is defined as:

$$V_i(nT_s) \text{ is bounded by } \pm V_{\text{DAC,max}}, \forall n, \forall i \qquad (9.22)$$

Every internal node V_i must be bounded at every time moment nT_s to the maximum voltages the feedback loop can provide.[3]

Similar to the situation in time-division digital-to-analog conversion in Sect. 7.2.6 the one-bit representation of a sine wave generated by sigma-delta modulation shows a lot of unwanted products. Comparing the power consumed in a resistor R of a digital sequence switching between $+V_a$ and $-V_a$ to the power of an equal amplitude sine wave:

$$\text{Power of one-bit digital} \quad \leftrightarrow \quad \text{Power of maximum sine wave}$$

$$P = \frac{V^2}{R} = \frac{(\pm V_a)^2}{R} = \frac{V_a^2}{R} \quad \leftrightarrow \quad P = \int \frac{(\alpha_{OL} V_a \sin(\omega t))^2}{R} dt \qquad (9.23)$$
$$= \frac{\alpha_{OL}^2}{2} \frac{V_a^2}{R}$$

shows that more than half of the power in a one-bit signal consists of quantization or unwanted distortion. Next to this unwanted power, the overload factor causes another -2 to -4 dB loss in the practical implementation.

---

[3]Normally a designer will try to use maximum signals throughout the entire circuit. In case scaled signals are used in some part of the circuit the value of V_i must be scaled as well.
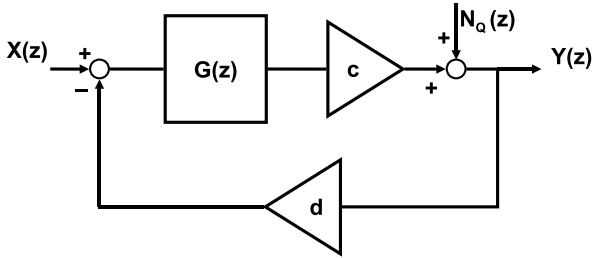
Fig. 9.13 The non-linear quantizer has been replaced by a gain factor and a noise injection point

The one-bit representation of a sigma-delta output signal is therefore largely composed of high-order distortion products. Any non-linearity in the feedback of the converter can easily create intermodulation between these higher order sampled distortions. Spurs will appear in the signal path. Also the digital processing of such a signal in the succeeding signal path requires attention for all non-linear operations such as rounding.

Unfortunately the analog summation point in the sigma-delta analog-to-digital converter must process these high-frequency components at high linearity. The input summation point is often the most critical circuit block in the design. Issues are:

- At the summation node the digital-to-analog converted signal is subtracted from the analog input. The combination of both signals can result in spikes that are more than double the input amplitude, due to the nature of the sampled and de-layed signals.
- Bandwidths and signal swings of opamps or OTAs must match these linearity and noise requirements.
- The dynamic aspects have to be taken into account such as slewing of the currents.
- The impedance levels have to be designed to a thermal noise level that does not limit the performance.

Designing the input with a sufficiently low-distortion requires to spend a lot of power.

As the quantizer is a non-linear element, it cannot be fully modeled in a linear system. In every model description the quantizer has to be approximated by linear components. Figure 9.13 shows that the quantizer is replaced by a gain factor c and the quantization errors have been made explicit by adding them with a summation node N_Q. When the quantizer's effective gain is modeled as c and the digital-to-analog converter gain as d, the input-output relation becomes:

$$Y(z) = \frac{cG(z)}{1 + cdG(z)} X(z) + \frac{1}{1 + cdG(z)} N_Q(z) \tag{9.24}$$

The gain factor c in one-bit quantizers is mostly established from the simulation of the input power versus the output power, see also Sect. 9.5. For multi-bit quantizers the ratio between the input step size V_{LSB} and the corresponding digital-to-analog converters output is a good measure for c. The feedback loop shapes the noise power

Fig. 9.14 Example of a frequency spectrum at the output of an audio sigma-delta modulator. The inverse third-order filter characteristic can be observed in the noise pattern
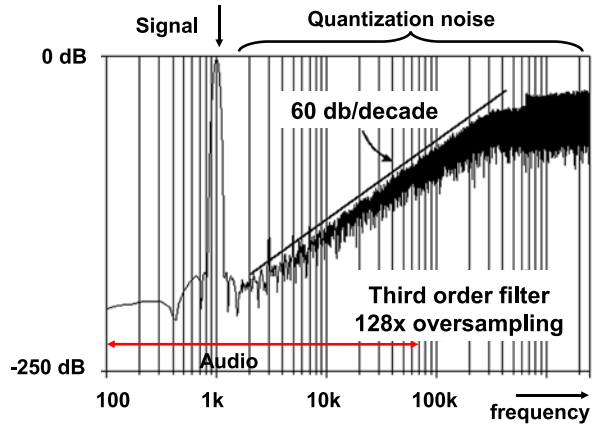
Fig. 9.15 The ideal signal transfer and noise transfer function
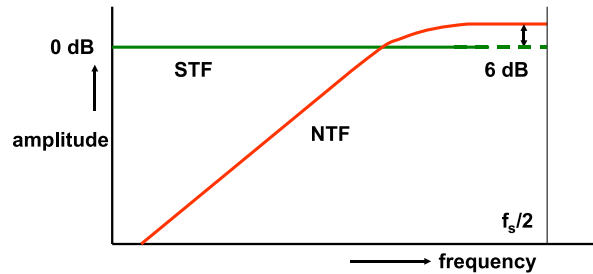
from a flat spectrum into a shape which is the inverse of the loop filter $G(z)$. The amplification of the digital-to-analog conversion is set to 1. The substitution of $z \leftrightarrow e^{j\omega T_s}$ results in a frequency domain representation

$$Y(\omega) = \frac{cG(\omega)}{1 + cG(\omega)} X(\omega) + \frac{1}{1 + cG(\omega)} N_Q(\omega) \tag{9.25}$$

At high values of $cG(\omega)$ the quantization errors are suppressed and the output signal equals the input signal. In the frequency band where the filter does not produce much gain, the quantization power increases, see Fig. 9.14.

The previous description of an output signal as a sum of two input sources, is often split in separate transfer functions: a Signal Transfer Function (STF) and a Noise Transfer Function (STF):

$$Y(\omega) = STF(\omega)X(\omega) + NTF(\omega)N_Q(\omega)$$

$$STF(\omega) = \frac{cG(\omega)}{1 + cG(\omega)} \tag{9.26}$$

$$NTF(\omega) = \frac{1}{1 + cG(\omega)}$$

Figure 9.15 shows both transfer functions. Often designers want to see a flat STF: the signal passes through the sigma-delta modulator without experiencing filtering.
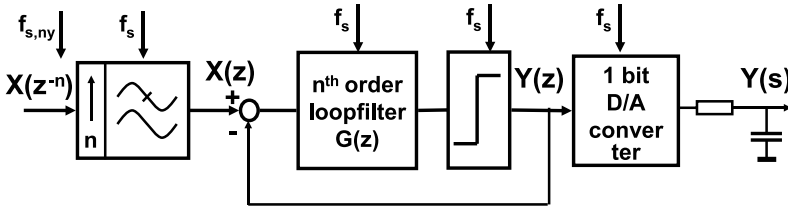
Fig. 9.16 A digital-to-analog sigma-delta modulator

However, at high frequencies where the active elements lose gain, bumps and other irregularities may appear. The NTF should suppress the noise at the band of interest. An increase of noise is visible close to the half of the sample rate.

9.3.1 Sigma-delta Digital-to-Analog Conversion

In a digital-to-analog converter based on sigma-delta conversion a lot of elements return that were present in the analog-to-digital sigma-delta converter. The loop filter is now implemented in the digital domain, just as the quantizer. The main function of such a sigma-delta converter is the conversion from one digital word width to a smaller word width, see Fig. 9.16. As no time-continuous input is present, first the low sample rate input signal must be up-converted to a higher sample rate, see Sect. 9.1. At the higher sample rate either noise-shaping or sigma-delta modulation can be used to reduce the word width (quantizing), while keeping the spectral signal-to-noise ratio in a limited bandwidth in tact. The resulting bit stream, e.g. a one-bit stream contains the required signal. In some applications with a high over-sampling ratio, a passive filter is sufficient to obtain an analog signal of high quality. As mentioned before, the one-bit output signal contains a lot of unwanted energy and care has to be taken that this energy is removed. Various examples show the potential to reach the 16–18 bit level for audio bandwidths [218, 219] or even 19 bit [220] with various optimization tricks. A full 20-bit signal in a 20 kHz bandwidth corresponds to a jitter of 8 ps_{rms}, which is realizable for a consumer equipment price level. Another interesting option to create an analog signal while filtering off the high-frequency components is the semi-digital converter in Sect. 7.2.4.

9.4 Time-discrete Sigma-delta Modulation

9.4.1 A First Order Modulator

Comparison of the transfer function of the noise shaper and the sigma-delta modulator makes clear that principally the filter function $G(z)$ can be obtained by rewriting $J(z)$, however this is not optimum. Figure 9.17 shows a first order time-discrete sigma-delta modulator. The loop filter is implemented as a first-order integrator.

Fig. 9.17 A first-order
sigma-delta modulator
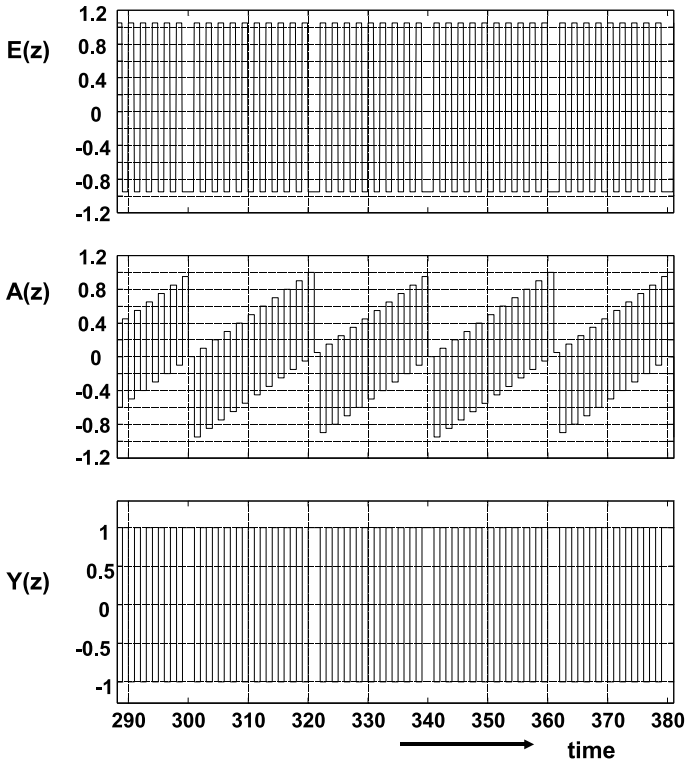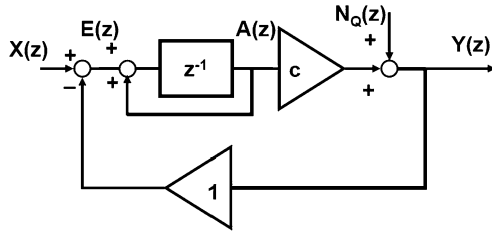
Fig. 9.18 The signals in a first order time-discrete sigma-delta modulator with a small DC in-
put signal. *Top*: the signal after the summation node. *Middle*: the integrator output. *Bottom*: the
comparator output

A small DC-input signal is applied in the simulation shown in Fig. 9.18. The dom-
inant signal in the loop is a block wave with frequency components of $f_s/2$. This
block wave at the output is inverted when it passes the input summation point and
is then integrated. The resulting signal provides the input for the comparator. The
combination of the delay in the filter and the inversion in the loop creates sufficient
delay to implement a stable oscillation. The DC input signal is superimposed on this
oscillation and a varying signal modulates the oscillation. The integrator slightly in-
creases its output value after repeated passes of the oscillation signal until it reaches
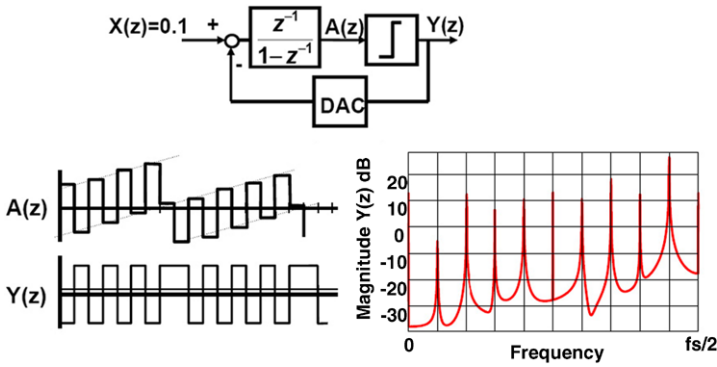
Fig. 9.19 A small DC-offset in a first-order sigma-delta modulator creates a pattern with frequency components in the signal band. These idle tones occur in low-order sigma-delta modulators

a level where the incoming block wave is not sufficient to reach the threshold for the comparator. The oscillation wave skips one transition. In the long-term output pattern these skipped transitions result after averaging for the representation of the DC-input signal.

The above transfer function for this first order sigma-delta configuration is split in a signal transfer function (STF) and a noise-transfer function. (NTF) With $c = d = 1$ and the integrator filter $G(z) = z^{-1}/(1 - z^{-1})$ these functions reduce to:

$$
\begin{aligned}
\text{STF}(z) &= \frac{Y(z)}{X(z)} = \frac{cG(z)}{1 + cdG(z)} \approx z^{-1} \\
\text{NTF}(z) &= \frac{Y(z)}{N_Q(z)} = \frac{1}{1 + cdG(z)} \approx 1 - z^{-1}
\end{aligned}
\tag{9.27}
$$

The signal passes unaltered through the modulator, it is delayed by one sample pulse. The noise is filtered, as can be seen in the frequency domain:

$$
|\text{NTF}(\omega)|^2 = (2\sin(\omega T_s/2))^2 = 2 - 2\cos(\omega T_s/2) \tag{9.28}
$$

and the total noise power in a band from 0 to f_b is found in the same way as in the noise shaper:

$$
\int_{f=0}^{f_b} \frac{V_{\text{LSB}}^2}{6f_s} |1 - e^{-2j\pi f T_s}|^2 \, df \approx \frac{V_{\text{LSB}}^2}{12} \frac{\pi^2}{3} \left(\frac{2f_b}{f_s}\right)^3 = \frac{V_{\text{LSB}}^2}{12} \frac{\pi^2}{3\text{OSR}^3} \tag{9.29}
$$

The Noise Transfer Function is in this example a first-order high-pass filter: suppression of noise close to DC while the amplitude of the noise goes above 0 dB in close to half of the sampling rate.

Figure 9.18 reveals another phenomena in sigma-delta modulation. The lower trace shows an output pattern with a dominant $f_s/2$ component. However the regular interruption creates low frequency signal components. These frequency components are called: "idle tones".
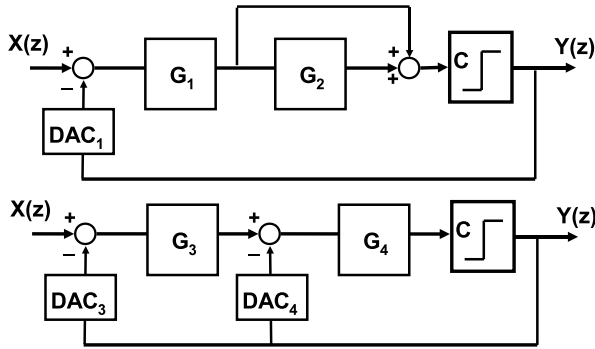
Fig. 9.20 Two methods to extend a first order sigma-delta modulator: the feedforward and feedback topologies

Idle tones are the major draw back of first-order converters. These correlated components or idle tones, see Fig. 9.19, can present serious problems in applications.[4]

9.4.2 A Second Order Modulator

A first order sigma-delta modulator is stable as long as the signal is within the range of the feedback digital-to-analog converter. The addition of a second integrator stage creates a loop where the negative feedback (180°) plus the contribution of two integrator stages (maximum $2 \times 90°$) equals 360°. It is clear that a small extra delay in the loop will push the modulator over 360° and create unwanted behavior. A second order modulator is conditionally stable. Proper conversion depends on input signal level, feedback and delay. The filter topologies combine first-order and second-order signals to achieve proper operation.

The first-order filter topology is extended to second order or higher by means of a feed-forward or feedback topology. The feed-forward topology cascades two first order integrator sections, see Fig. 9.20. The filter function is shaped by weighted addition of the outputs of the successive integrators. This summation point is loading the previous stages and requires careful design. In addition to the internal signals also the input signal itself can be fed into the summation point. This topology is discussed in Sect. 9.7.3.

The same filter function can also be realized by a feedback topology. Now a digital-to-analog converter is needed for every pole in the filter. The two-fold subtraction allows more freedom to choose higher signal levels. The feedback factor is not exactly unity, so some peaking in the STF can occur.

---

[4]In the audio range these components are audible and are called "whistles". Trained listeners can hear idle tones down to 100 dB below full-signal.
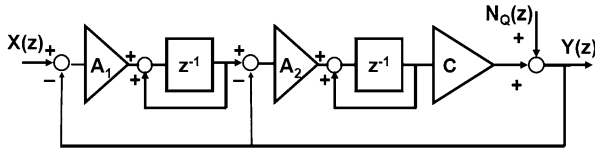
Fig. 9.21 A second order sigma-delta modulator with two forward coefficients. Weighting coefficients can also be placed in the feedback path
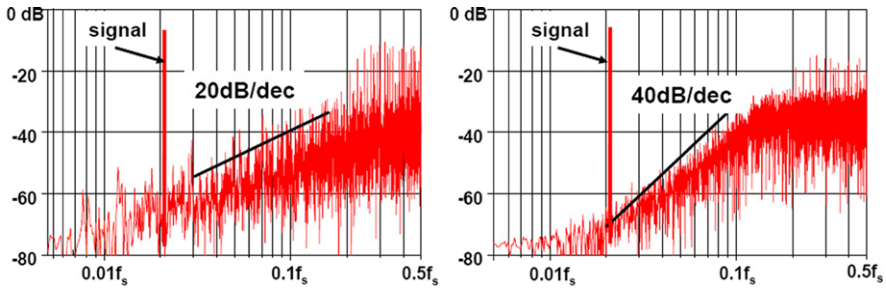
Fig. 9.22 The spectra of a first and second order sigma-delta modulator

Figure 9.21 shows the feedback extension for a time-discrete second-order topology. The coefficients A_1 and A_2 are often chosen to scale the signal after the summation point back into the range. The noise transfer function is of the form

$$NTF(z) = \frac{Y(z)}{N_Q(z)} \propto (1 - z^{-1})^2 \tag{9.30}$$

resulting in a behavior equivalent to a second order noise shaper, with its noise power approximated by:

$$\int_{f=0}^{f_b} \frac{V_{\text{LSB}}^2}{6 f_s} |1 - e^{-2j\pi f T_s}|^4 \, df \approx \frac{V_{\text{LSB}}^2}{12} \frac{\pi^2}{5} \left(\frac{2 f_b}{f_s}\right)^5 \tag{9.31}$$

which is a similar description as was found for a second-order noise shaper, see Fig. 9.9.

The output spectrum of a first-order and second-order sigma-delta modulator are compared in Fig. 9.22. The first order spectrum still contains a lot of distinct frequency components pointing to a large correlation in the unwanted components. The second order spectrum shows a 40 dB/dec slope and much less idle tones. The dual integration of the signal and the fed-back quantizer result create a much more complex pattern, which appears to contain less correlated products. Higher order converters scramble the pattern even more due to the extra time constants in the filters. Third and fourth order modulators show therefore hardly any idle tones, see Fig. 9.14. Yet it is possible with the right sort of input signal to see short periods of idle tones in second and third order modulators.

Figure 9.23 shows the realization of a second order switched-capacitor sigma-delta modulator. The input network samples the input voltage and transfers this
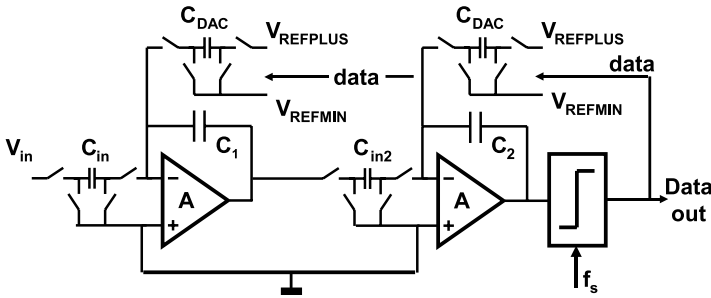
Fig. 9.23 A switched capacitor realization of a second-order sigma-delta modulator

charge in integrator capacitor C_1. A following stage implements the second integrator. The output of the comparator determines whether a positive or negative charge packet is fed back to both integrators. In switched capacitor technique the capacitor ratios determine the coefficients.

Extending the time-discrete sigma-delta converter to a third or higher order is possible [221], however the loop-filter has to be designed in such a way that a stable oscillation is guaranteed [222, 223]. This can be done by inserting zeros in the transfer function. Yet, stability problems may occur as the gain model of the quantizer in this analysis is assumed to be linear. The quantizer gain is a non-linear function and varies with the input amplitude. Lee's rule is based on simulations: $|\text{NTF}(f)| < 1.6$ for all frequencies.

With cascaded sigma-delta converters higher order performance is possible without higher-order loop filters.

9.4.3 Cascaded Sigma-delta Modulator

If more noise power must be suppressed in a fixed bandwidth either the oversampling rate must be increased or a higher order loop filter must be designed without the stability problems. The cascaded sigma-delta approach [224] solves this issue, see Fig. 9.24. A first sigma-delta converter uses a first or second order filter $G_1(z)$ in its quantization loop. The overall transfer of this sigma-delta modulator is given by (9.24). Although this sigma-delta loop suppresses the quantization errors $N_{Q1}(z)$, this noise can be further suppressed by considering that this noise is a deterministic signal that can be measured and subtracted from the output, see Fig. 9.24. The noise of the first sigma-delta converter is isolated by creating an analog copy of the output signal (via "d") and then subtracting the input signal of the quantizer to yield: $X_2(z) = N_{Q1}(z)$. The quantization errors of the first modulator are now isolated in the analog domain and are converted into the digital domain via a second sigma-delta modulator. $Y_2(z)$ is the digital version of $N_{Q1}(z)$. Before this noise can be subtracted from the output signal $Y_1(z)$ the noise has to be shaped in frequency in a same way as $N_{Q1}(z)$ is shaped in the loop. $Y_2(z)$ has to be divided by the loop
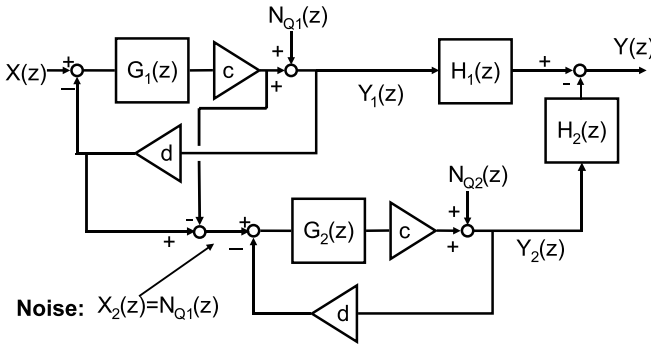
Fig. 9.24 The cascaded sigma-delta modulator

filter characteristic $G_1(z)$ and $Y_1(z)$ has to be delayed to match the delay in the second modulator. The overall noise function is found as ($d = 1$):

$$Y_1(z) = \frac{G_1(z)}{1 + G_1(z)} X(z) + \frac{1}{1 + G_1(z)} N_{Q1}(z)$$

$$Y_2(z) = \frac{G_2(z)}{1 + G_2(z)} N_{Q1}(z) + \frac{1}{1 + G_2(z)} N_{Q2}(z)$$

$$Y(z) = \frac{H_1(z)G_1(z)}{1 + G_1(z)} X(z) + \left(\frac{H_1(z)}{1 + G_1(z)} - \frac{H_2(z)G_2(z)}{1 + G_2(z)} \right) N_{Q1}(z)$$

$$- \frac{H_2(z)}{1 + G_2(z)} N_{Q2}(z)$$

(9.32)

With G_1, G_2 sufficiently large and $H_2(z)G_1(z) = H_1(z) = z^{-k}$:

$$Y(z) = X(z) + \frac{1}{G_1(z)G_2(z)} N_{Q2}(z) \qquad (9.33)$$

The term z^{-k} represents the delay of k sample pulses. The amount of noise suppression is determined by the successive filter functions and depends on the cancellation of the two paths that the noise of the first quantizer follows (the direct path and the path via the second quantizer and filter $H_2(z)$). If this cancellation is perfect, the remaining quantization errors originate from the second modulator. This noise can be reduced in the same way in a third loop. Yet, if a 1% mismatch occurs in the cancellation, the first-order noise will contribute on a level of 40 dB below the original first modulator performance (which still can be an acceptable performance).

A cascaded converter with a second order filter in the main loop and a first order filter in the second is denoted as a "2-1 cascaded sigma delta converter". The total noise transfer is equivalent to the noise shaper as in Fig. 9.9. The problem in the realization of analog-to-digital converters with cascaded sigma-delta modulation is therefore in the matching of the analog filter $G_1(z)$ with the digital counterpart $H_2(z)$. In switched-capacitor technique the matching of the analog and digital filter functions depends on the quality of the capacitor ratios. A time-continuous cascaded
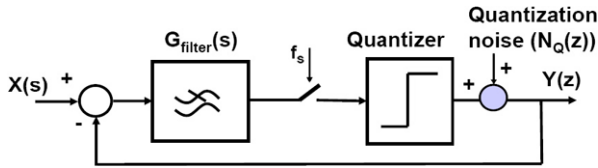
Fig. 9.25 A time-continuous sigma-delta modulator samples after the filter
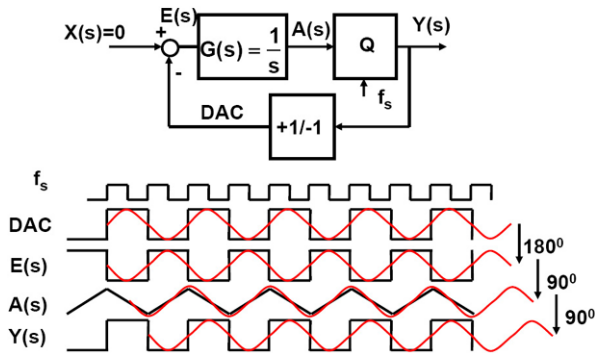
Fig. 9.26 The signal flow in a first order time-continuous sigma-delta modulator showing the phase relations

sigma-delta converter needs a calibration mechanism to trim the digital filter to the analog filter [227] or careful design [229].

9.5 Time-continuous Sigma-delta Modulation

9.5.1 A First-order Modulator

Time-discrete sigma-delta modulation assumes a sampled data stream at its input. All internal processes are synchronized to the sample rate. The alias filtering necessary for sampled data systems must be performed before the data enters the sigma-delta modulator. In contrast, time-continuous sigma-delta modulation uses loop filters based on time-continuous components. The sampling takes place after the filter, see Fig. 9.25. Mostly the sampling is implemented by the decision making of the quantizer. The conversion system is now partly time-continuous and partly time-discrete. Figure 9.26 shows the wave forms for a situation with zero input signal. In contrast to the synchronous mode of working in a time-discrete sigma-delta converter, here the timing in the loop is only synchronized by the quantizer. The quantizer sends out a digital signal that is converted to analog levels by the digital-to-analog converter without loss of time. This signal gets inverted at the summing node and is integrated to a triangular shape $A(s)$. The fundamental frequency of the
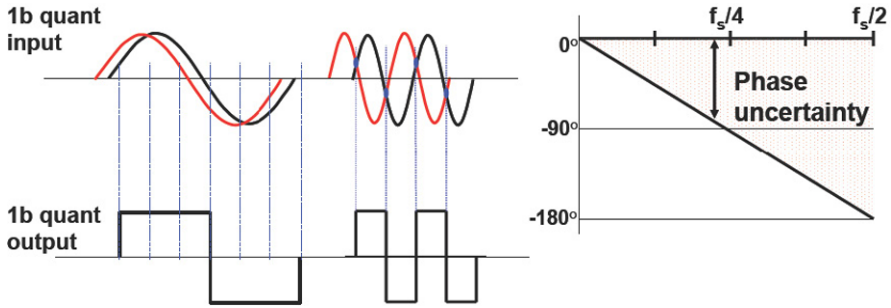
Fig. 9.27 Two examples of two signals that give result in the same comparator decisions. The time uncertainty of one sample cycle translates in a linear phase deviation with signal frequency

signal $A(s)$ before and after the comparator $Y(s)$ is 90° shifted due to the sampling of a triangular shape to a block wave. Part of the phase shift to reach 360° for letting the loop oscillate, is accomplished through the waiting time before the clock pulse activates the quantizer. There is some phase margin or waiting time after the filter. The comparator decides at the edge of the sample pulse. With respect to this sample pulse signals while be early, see Fig. 9.27. This margin can be anything between zero and the period of one sample pulse. The margin is often referred to as "phase-uncertainty", which is a misleading term as there is no uncertainty involved. The waiting time in a first-order modulator ideally corresponds to a quarter of half of the sample rate as the integrator delays for 90°. This implies that the delay of a second integrator in the loop would just fit. Any additional delay would corrupt the stability.

The Signal Transfer Function (STF) and a Noise Transfer Function (STF) are formulated in the s domain:

$$Y(s) = \text{STF}(s)X(s) + \text{NTF}(s)N_Q(s)$$

$$\text{STF}(s) = \frac{cG(s)}{1 + cG(s)}$$
$$\text{NTF}(s) = \frac{1}{1 + cG(s)}$$

(9.34)

where c is the gain of the comparator, see Fig. 9.13. The complication is the sampling process after the time-continuous filter. Up to half of the sampling rate this description is mostly applicable.

In the presence of some DC-offset at the input, Fig. 9.28, the wave forms in the time-continuous sigma-delta converter resemble the wave forms in the time-discrete case in Fig. 9.18. The integration of the high-frequency component and the input signal is now well visible.

A second-order time-continuous sigma-delta modulator is created by adding one more integration stage. The higher order filter will suppress the quantization errors more. This becomes visible in less correlated error signals as can be seen from Fig. 9.29. This configuration has as the immediate advantage that (a part of) the alias filtering is performed by the loop-filter. A related advantage for communication
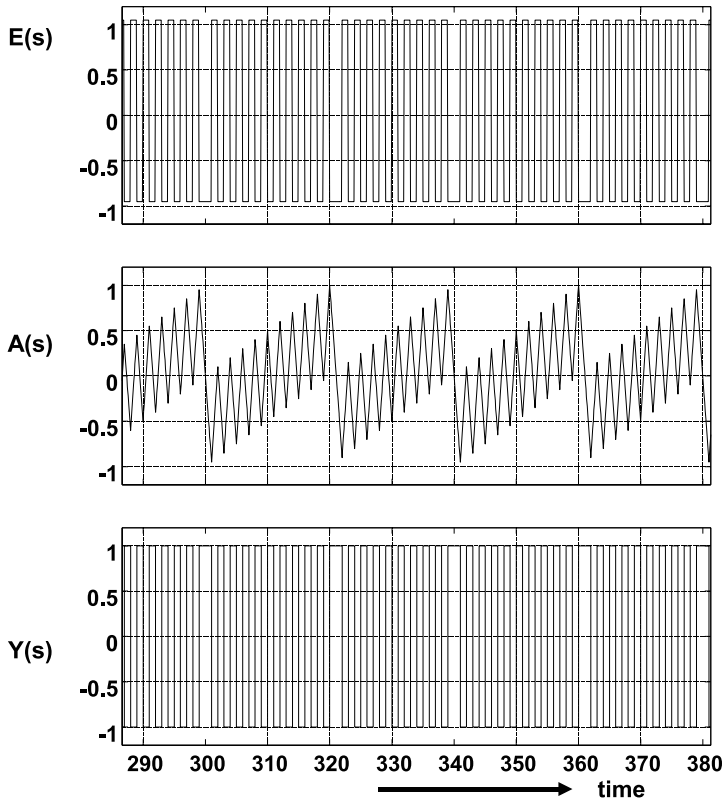
Fig. 9.28 The signals in a first order time-continuous sigma-delta modulator with a small DC input signal

systems is that high-frequency interferers or strong adjacent channels are suppressed by the system in a similar fashion as alias products before they are sampled. This may lead to a more relaxed system design.

The time-continuous loop filter has to meet the specifications within the desired bandwidth, which is by definition much lower than the sample rate. In the time-discrete domain, where the filter is built in switched capacitor technique, the switches run at the sample rate and the opamps need to settle at that speed. This normally leads to a higher power consumption in the opamps. The disadvantage of the time-continuous filter is that its shape or the relative position of poles and zero's can be well controlled, but not the absolute position in the frequency band. A tolerance of 10% in modern processes has to be taken into account and translates in some guard-banding in the specification of the bandwidth of interest.

A second disadvantage of time-continuous conversion relates to the digital-to-analog converter. In a time-discrete switched-capacitor implementation the feedback charge is defined by the capacitor and the reference voltage. The timing is of
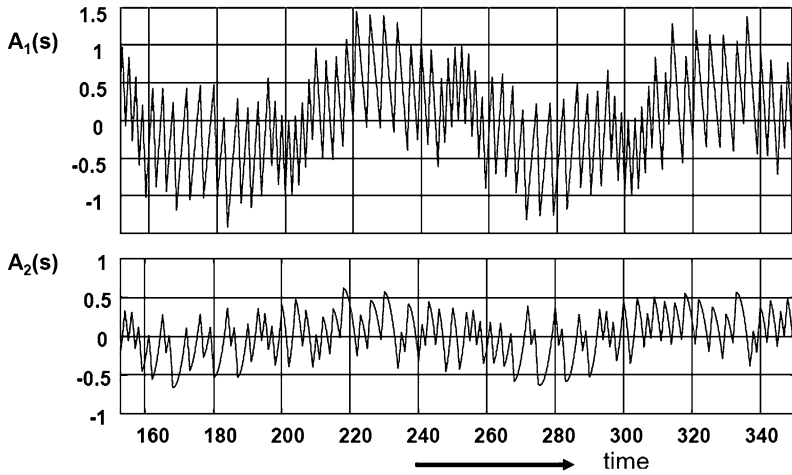
Fig. 9.29 A comparison of the signal after the filter in a first order and a second-order sigma delta modulator
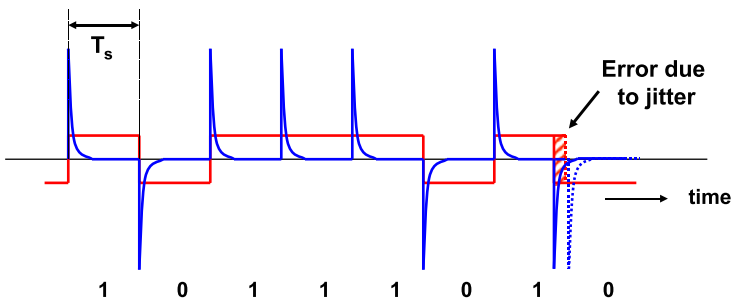
Fig. 9.30 A time-continuous DAC feedback signal is a continuous pulse. A switched-capacitor DAC generates charge spikes. Jitter just shifts the position of a switched-capacitor pulse, while it modulates the volume of a time-continuous pulse

secondary importance as the charge will anyhow be transferred into the summation point. In a time-continuous converter the feedback from the digital-to-analog converter is in the form of a current multiplied by a time period. In this implementation, jitter in the timing pulses directly affects the performance, see Fig. 9.30. Asymmetry in the feedback path, see Fig. 9.31, can be caused by differences in rising and falling edges. In this example the data pulse of two successive symbols "1" contains less charge than two single symbols. This problem of a non-return-to-zero (NRZ) pulse sequence can be circumvented by a "return-to-zero" implementation technique. Now every pulse, with or without successor, will return to a zero-value. In a differential design a similar technique is possible. More advanced schemes employing dynamic element matching have been published in [225].
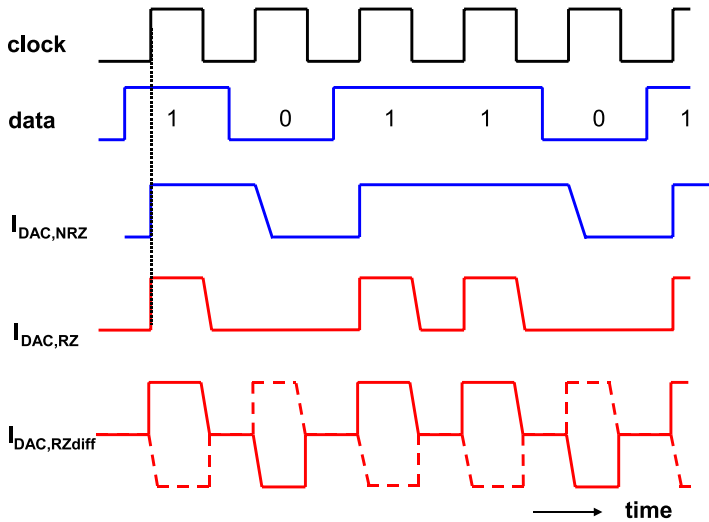
Fig. 9.31 Asymmetry in the digital-to-analog conversion pulses $I_{DAC,NRZ}$ can be mitigated with a return-to-zero technique $I_{DAC,RZ}$ or a differential return to zero technique $I_{DAC,RZdiff}$ [226]

9.5.2 Higher Order Sigma-delta Converters

A prerequisite for the application of sigma-delta conversion is a sufficiently high oversampling factor of f_s over the signal bandwidth. Sigma-delta conversion allows to use small quantizers (e.g. a comparator and a few switches for a digital-to-analog conversion) and is power-efficient. The sigma delta modulator does not reduce the quantization power as a whole, it only moves undesired quantization power from one frequency band into another band. So the necessary condition for the application of a sigma delta converter is a sufficiently large difference between the relevant bandwidth and the sampling rate. If the oversampling factor is large, there is more frequency band available to deploy a large gain of the filter, and there is a frequency band where the quantization power is moved to. With sample rates of 5 to 10 Ms/s, resolutions of 16 bits (98 dB) can be achieved within the audio band from 0 to 20 kHz (e.g. [226]).

For achieving low noise levels in wider bandwidths without increasing the sample rate into the GHz range, the filter order must be increased.

Figure 9.32 shows the implementation of a fourth order feed-forward sigma-delta converter. The loop filter consists of a cascade of four integrators whose outputs are summed via coefficients. The comparator decides on the polarity of the output of the filter. A simple switch reversing the polarity of a reference current source, is used for the digital to analog feedback. In case of instability due to an excess voltage at the input, the output of the last integrator may saturate. The clipper effectively switches off the contribution of this integrator to the overall filter function and the modulator will behave as a third order circuit. If the third integrator clips, the modulator is only second order. Although the performance is degraded, some signal conversion
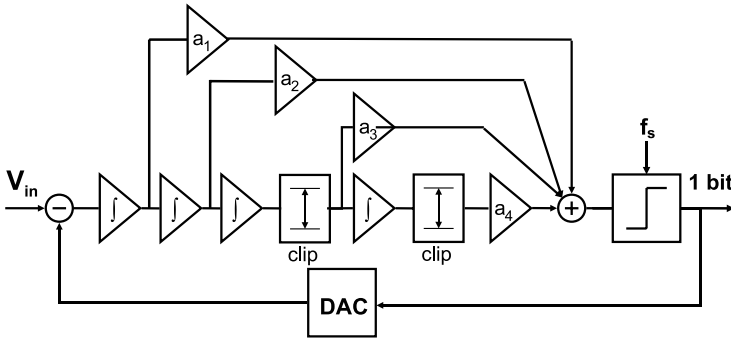
Fig. 9.32 A fourth-order feed-forward sigma delta modulator [226]

is performed. This behavior is called "graceful degradation". In this scheme the clippers will prevent too much saturation of the analog circuits. After removing the excess signal the integrator will recover and help again to shape the filter function.

The combination of a time-continuous filter and a time-discrete quantization creates a few additional problems in the Laplace analysis. Not only the definition of the gain factors c and d becomes more complicated, but also the time relation between the comparator's input and output is less trivial. The gain in the loop is relevant for the loop behavior. As the comparator interfaces between the analog and digital domain and is highly non-linear, the description of the gain is not a simple mathematical expression. It is mostly allowed to describe the total gain in the path from the input of the comparator to the output of the digital-to-analog converter. As these signals have no linear or linearizable relation, the gain must be determined by comparing the power functions. If cd incorporates the comparator and the gain of the digital-to-analog converter, where $T \gg T_s$:

$$(cd)^2 = \frac{\int v_{\text{quantizer,out}}^2 \, dt}{\int v_{\text{quantizer,in}}^2 \, dt} = \frac{(\pm 1)^2}{(1/T) \int_T v_{\text{quantizer,in}}^2 \, dt} \tag{9.35}$$

It can be necessary to do a numerical comparison of the input power with the rather simple output power of the comparator. In simulation [226] the gain does not vary much over the input amplitude range, see Figs. 9.33 and 9.34. Moreover the feedback loop concept tolerates some inaccuracy in this parameter.

The rather small deviation in gain can be dealt with by requiring a gain guard band in the stability analysis.

The design of the filter function has to take into account that the requirement for stable oscillation in the loop requires a feedback path where the filter of order one or two and the delay before the quantizer equal the remaining 180°. Therefore, the design of higher order filters is possible if the filter is designed in such a way that it shows a high order for low frequencies, while turning back to a first order response near the passing of the 0-dB loop gain. In Fig. 9.35 the low-frequency noise shaping part of the filter is fourth order. The filter is designed to turn back to a first order and 90° phase at the frequencies where the 0 dB gain is reached. At this gain a
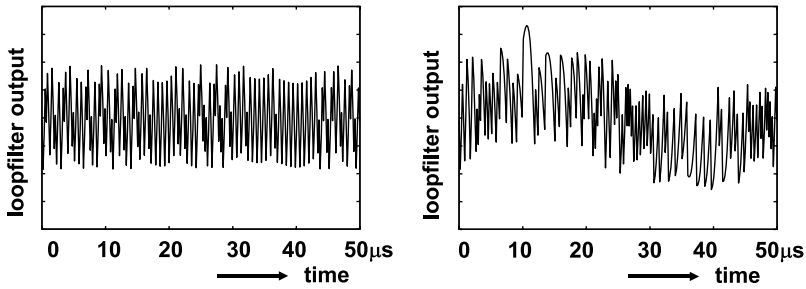
Fig. 9.33   The simulated input for the quantizer, *left* with 0.1% signal input to the modulator, *right*: with maximum input signal, from [226]

Fig. 9.34   The gain as function of the input signal to the modulator, from [226]
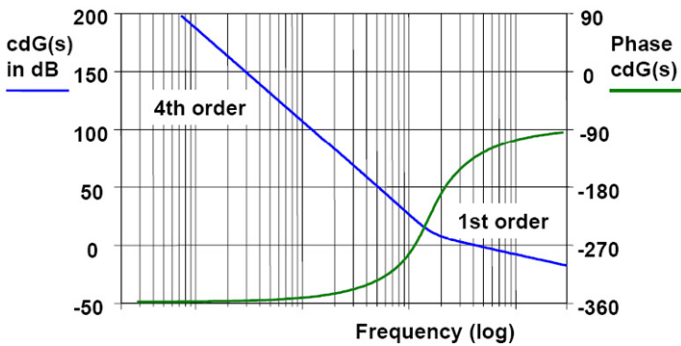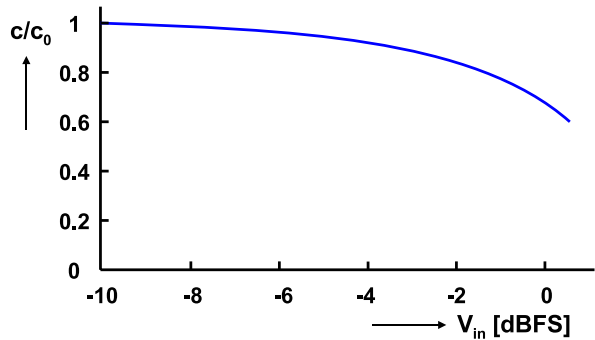
Fig. 9.35   The fourth-order filter transfer function turns back to a first order transfer function close to the frequency region where the modulator acts as an oscillator

360° overall phase shift must be reached: 180° from the inversion at the summation node, >90° from the filter and the rest is consumed in the delay before the sampling moment, the phase uncertainty.

Figure 9.36 shows the resulting design of a fourth-order loop filter, including some guard banding for the gain inaccuracy. Also the phase uncertainty is added to the phase of the filter. The region between the two phase lines contains a frequency
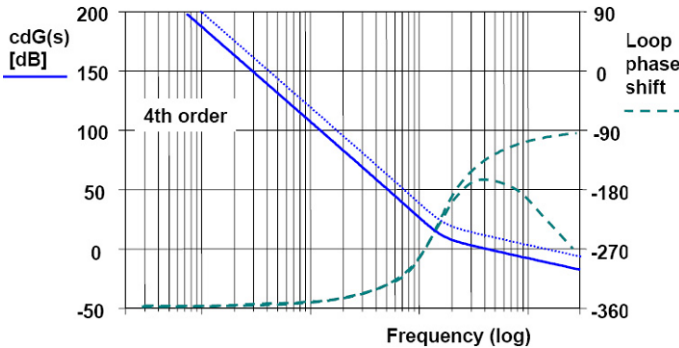
Fig. 9.36 The fourth-order filter transfer function with gain tolerance and the phase plot including phase uncertainty

where the overall gain is 1 and a phase of 180° is available. This is a necessary but not sufficient condition for the loop to oscillate. Applying common practices such as Lee's rule, does not eliminate the need to rigorously simulate any non-linear circuit to verify proper functionality.

The shape of the filters in a time-continuous sigma-delta converter depends on the rations of the frequency determining elements. The values of these components determine the pole and zero frequencies. The tolerances on the filter components require unlike the time-discrete topologies, some additional margin in the oversampling factor.

The latency of a sigma-delta converter is often seen as one of its main disadvantages. Yet, this argument has to be judged carefully. The total delay between the applied input signal and the resulting Nyquist-format digital output is composed of the delay from input to output of the modulator, plus the delay in the succeeding down-sample filter. The first component is mainly determined by the delay in the analog loop filter, which is dominated by the first pole. The down-sample filter, see Sect. 3.2.3, will convert the oversampled bit stream from the sigma-delta modulator into digital PCM code at a Nyquist-rate sample rate. The delay in the down-sample filter can run into many clock periods, especially if there is a requirement on linear phase and FIR filters.

9.5.3 Time-discrete and Time-continuous Sigma Delta Conversion

Figure 9.37 summarizes a number of architectural choices for the designer of a sigma-delta converter. The overall system specifications result in a desired conversion bandwidth. Based on this bandwidth Fig. 9.9 allows to determine the required oversampling factor and the resulting sample rate in relation to the filter order. In first instance it seems that the difference between time-discrete and time-continuous sigma-delta conversion is limited to the implementation of the filter. Yet this choice has some major consequences.
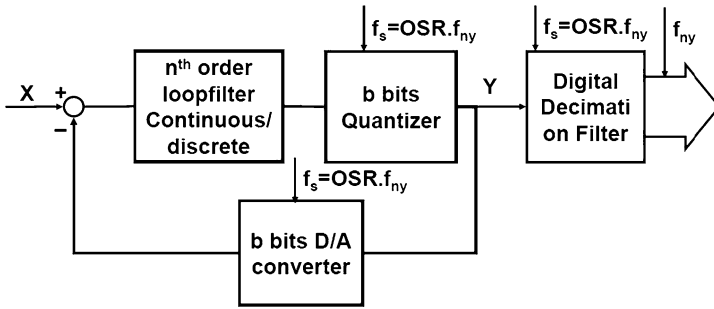
Fig. 9.37 Several design choices play a role during the implementation of a sigma-delta converter: the oversampling factor OSR, the levels in the quantizer, the type and order of filter

From a technological point of view the optimum choice for realizing time-discrete filters in standard CMOS is the switched capacitor technique. The poles and zero's are in the frequency domain determined by the capacitor ratio and the sampling frequency and an accuracy of 0.1% is here achievable. Switched-capacitor sigma-delta converters are therefore a natural choice for realizing cascaded structures as the matching of filter characteristics between the loop filter and the digital filter in the noise branch is essential. In time-continuous filters, the filter sections can be designed with mutually matched $g_m C$ sections which will guarantee a well-defined shape of the curve over process variations. The absolute position, however, is determined by the actual process parameter values and can vary $\pm 10\%$. Breems [227] showed that a calibration mechanism can make a time-continuous filter match over a certain frequency range to the characteristics of the digital noise filter, although a time-discrete digital filter can only match a time-continuous filter over a limited part of the frequency band.

The choice for a time-continuous filter eliminates the sampling mechanism at the input of the sigma-delta modulator and shifts the sampling to the quantizer circuit. This configuration allows to use the time-continuous filter as an alias filter, see Fig. 3.4. This feature is efficient from a system point of view especially with higher-order filters. In various communication systems one of the major system issues is the interferer suppression. In time-discrete converters the sampling of an out-of-band interferer takes place before filtering. The frequency band in which the alias products return, depends on the position of the interferer frequency with respect to the sample rate. The time-continuous filter first reduces the amplitude of an interferer signal before the sampling process takes place [230, 231]. A similar argument holds for signals that are injected through on-chip substrate coupling or EMC.

A switched-capacitor time-discrete sigma-delta modulator shifts charges around in its analog circuits. These charges are transferred from one opamp-capacitor section into another and require sufficient settling for the switched-capacitor circuit. The processing speed of these sections is equal to the sampling rate. Multiplexing over a few filter branches is possible, although the matching between the branches becomes an issue. A significant amount of power is needed in the opamps to allow a fast settling.

Circuit designers want to increase the oversampling ratio for optimum sigma-delta conversion. However, a large bandwidth-to-sample rate ratio requires a filter capacitor ratio of a similar ratio. Due to matching requirements and parasitic components there is a minimum-size for the capacitor value. This implies that the counter-part capacitor in the filter is potentially large and requires a lot of power in its driving opamps. This dilemma is less pronounced in time-continuous converters. Many authors therefore argue that a power advantage of time-continuous sigma-delta converters exists for the same specification. Also the on-resistance of the switches become important to consider if large capacitors are needed in the switched-capacitor filter. In time-continuous filters it is mainly the bandwidth and dynamic range of the wanted signal that sets the requirements for the filter design. An exception is the input summing circuit, where the analog input signal meets the time-discrete feedback signal. If both are in the voltage domain, this node will experience large voltage excursions and is consequently sensitive to distortion. The summing circuit and successive stages amplify the remaining signal, so the thermal and $1/f$ noise performance of the input circuit is most critical. The noise and distortion arguments make that the input summing circuit is the most challenging part of the time-continuous design.

An advantage of shifting charges in a switched-capacitor circuit is that the result of this process, the total amount of transferred charge, is independent of the timing. Shorter or longer sample periods due to jitter will modulate on the settling tail only, but not affect the amplitude to a serious degree. In a time-continuous sigma-delta converter often the feedback path is implemented in the current or voltage domain. The feedback signal equals the integral of the voltage or current over the sample period. This means that any jitter or asymmetry in the timing edges of the feedback pulse is heavily impacting the performance. Figure 7.26 shows the impact of jitter on pulse-width modulated signals. This impact of jitter in a time-continuous converter is a major limitation. As the jitter is multiplied by the amplitude of the feedback digital-to-analog converter, the single bit implementations (with the largest feedback amplitude) are most sensitive.

Both approaches have advantages and disadvantages. Some interesting approaches combine elements of both techniques in the feedback path [227, 232].

9.6 Multi-bit Sigma-delta Conversion

In sigma-delta conversion the signal-to-noise improvement by oversampling and subsequent noise shaping is huge. A multi-bit quantizer from that perspective has little advantage over a single-bit quantizer. In addition a single-bit quantizer uses only two levels for the feedback path, which is inherently linear. Yet multi-bit quantizers with output wave forms as in Fig. 9.38, enjoy an increasing popularity [227, 228]. The major motivation for multi-bit quantizers comes from the implementation issues that arise with a single-bit quantizer. The discussion can be summarized as follows:

- A single-bit quantizer generates a lot of noise in the loop, see Sect. 7.2.6. This noise has to be treated in a linear way, because non-linearities will cause mixing
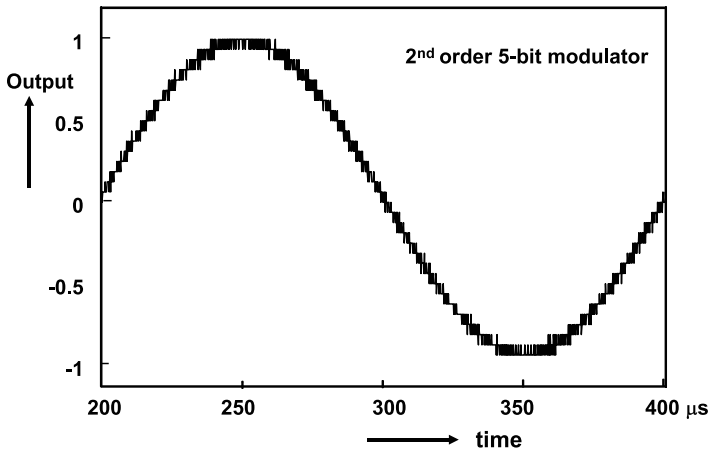
Fig. 9.38 The reconstructed output wave form of a 5-bit quantizer in a sigma-delta modulator

Table 9.2 Simulated noise level reduction for multi-level quantization

Resolution	ΔSNR
$1 \rightarrow 2$	7.0 dB
$2 \rightarrow 3$	6.2 dB
$3 \rightarrow 4$	6.1 dB
$4 \rightarrow 5$, etc.	6.0 dB

products in the desired band. The lower noise level of a multi-level converter alleviates this requirement.

- The starting signal-to-noise level of a single-bit quantizer before the noise reduction is low. Every bit of extra quantization lowers the level of the overall noise spectrum and increases the signal-to-noise ratio *SNR*. Table 9.2 is based on Table 5.2:
- The summation node has to deal with a multitude of requirements. Single-bit quantizers cause large and steep feedback pulses. A multi-bit approach relieves this issue for the summing node and to a lesser extent for the succeeding filter stages.
- The sensitivity to jitter and timing asymmetries is lower for the smaller steps of a multi-level digital-to-analog converter in the feedback loop.
- The swing of the residue signal in the first integrator of a multi-level quantizer is reduced, so a larger integrator gain is possible reducing distortion.
- A multi-level quantizer requires more comparators, more digital hardware and a more extensive digital-to-analog converter, which translates in more power.
- The gain factor in a single-bit quantizer has to be determined by comparing the input and output power. In a multi-level quantizer the gain factor is much better defined. The entire definition of filter and stability is better controlled and less gain margin has to be built into a multi-level quantizer.

Fig. 9.39 Signals close to
overload in a 2-level and
3-level quantizer
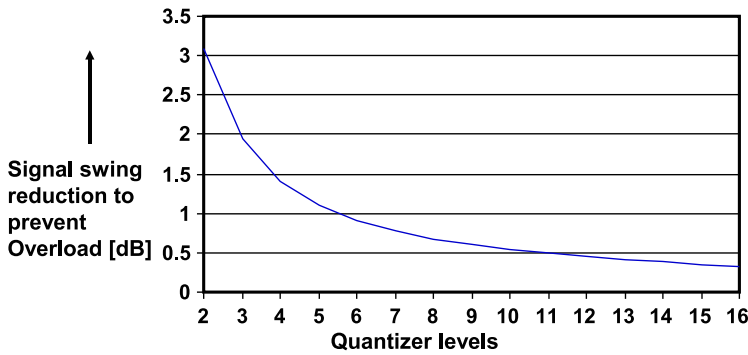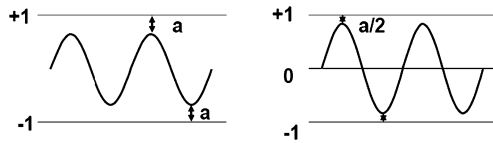
Fig. 9.40 The required signal reduction to prevent overload reduces with increasing number of
quantization levels

- The overload level of a multi-level modulator improves [233] and thereby in-
  creases the signal-to-noise ratio. In Fig. 9.39 the 2-level quantizer reconstructs a
  sine wave with pulses switching between $+1$ and -1. For proper operation the
  sine wave is limited to a fraction $\alpha_{o,2}$ of the range defined by the levels $+1$ to -1.
  [233] assumes that in a multi-level quantizer the same relative distance between
  the signal and a single step of the feedback path is required for proper operation.
  If this distance is a on both sides of a single transition between two levels, it can
  be $a/2$ for a three levels. In general the signal swing as a fraction of the total
  range for a quantizer with n_q quantization levels is:

$$\alpha_{o,nq} = \frac{n - 2 + 2\alpha_{o,2}}{n} \tag{9.36}$$

Figure 9.40 shows a graph of this formula.

As a sigma-delta modulator is a feedback system, the quality of the entire conversion
is determined by the properties of the feedback path. The linearity of the digital-
to-analog converter in this path limits the overall achievable linearity. Even with
a two-bit quantizer and a two-bit digital-to-analog converter, this digital-to-analog
converter must exhibit a 16-bit (timing) accuracy to be able to achieve a 16-bit
resolution sigma-delta conversion. The main challenge of multi-level sigma delta
conversion is therefore the design of a digital-to-analog converter in the loop that
meets the overall specification. With the introduction of accurate digital-to-analog
conversion techniques, such as data weighted averaging [136], see Sect. 7.4.3, the
application space for multi-bit quantizers has considerably increased. Yet, the error-
shaping in the digital-to-analog converter must be carefully designed in order not
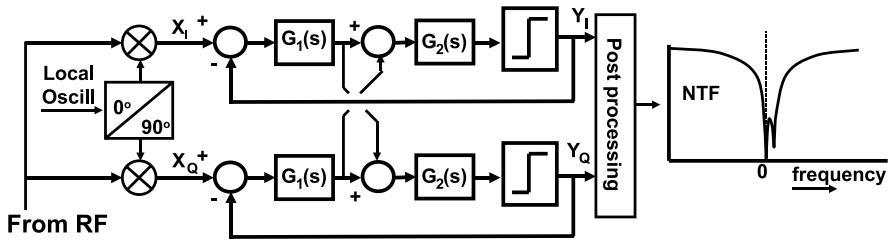
Fig. 9.41 The complex sigma delta converter consists of two coupled sigma-delta modulators, e.g. [227, 234]

to dominate the overall noise reduction. Moreover the delay in the processing must be kept low. Also the tones that can occur in simple DWA algorithms must be suppressed by means of additional randomizing steps.

9.7 Various Forms of Sigma-delta Modulation

9.7.1 Complex Sigma-delta Modulation

In digital communication systems signals are modulated on the in-phase carrier and the 90° shifted quadrature component. These I/Q signals are converted to the digital domain with a complex sigma-delta converter Fig. 9.41. Such a converter is built from two identical sigma-delta converters. The input signal from the RF circuits is down-modulated with two phases of the local oscillator. The X_I and X_Q signals are fed in the two sigma-delta modulators. The phase difference between the two signals means that the quadrature signal is a quarter of a carrier period delayed with respect to the in-phase signal. This property allows to form a complex frequency domain where the positive frequency domain is not a mirrored version of the negative side. Moreover, the mutual coupling between the two signal paths allows to define single-sided filter characteristics for the Noise-Transfer Function. These properties are beneficial for designing an optimum converter for various communication systems.

9.7.2 Asynchronous Sigma-delta Modulation

The asynchronous sigma-delta converter [240] in Fig. 9.42 uses the same topology as the time-continuous converter with the exception of the quantizer. The clocked quantizer is replaced by an non-clocked non-linear element, such as a continuous comparator, or a comparator with some hysteresis. As there is initially no quantization involved, there is no quantization energy created and neither to be shaped. The loop operates as a kind of multi-vibrator on a self-oscillation frequency. This
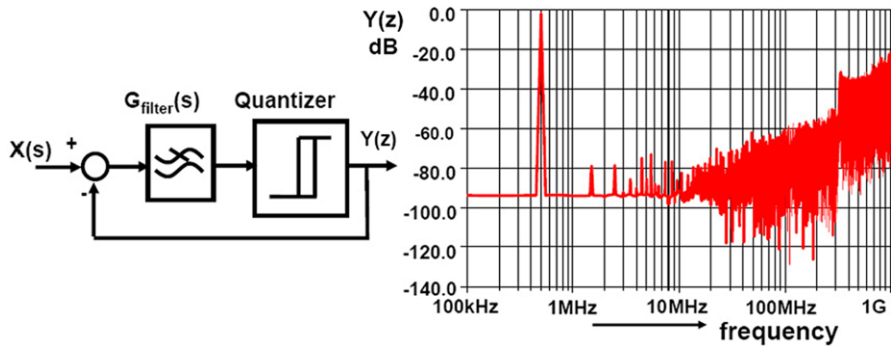
Fig. 9.42 The asynchronous sigma-delta converter and output spectrum

Fig. 9.43 A feed-forward sigma-delta modulator
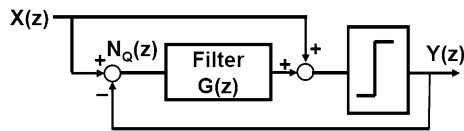
frequency is determined by the loop filter and the hysteresis amplitude. Nevertheless the comparator generates high-order distortion products, that also modulate the oscillation frequency as in an FM-modulator. Some similarity with pulse-density modulation (PDM) is certainly present.

Higher up in the spectrum, see Fig. 9.42 (right), these products are visible. Another way of looking at an asynchronous sigma-delta modulator is as a synchronous modulator running at a sufficiently high sample rate. The main problem with an asynchronous sigma-delta modulator is equivalent to more asynchronous systems: application is possible in a small system without clocked interfaces, or the data must at some point be processed by a synchronous component. Then quantization will occur in a similar way as in a level-crossing analog-to-digital converter, Sect. 8.10.1.

9.7.3 Input Feed-forward Modulator

In the standard topology of a sigma-delta converter the summation node at the input causes many problems. The combination of the full-signal and the feedback signal creates a large residue amplitude. This signal requires special attention to avoid distortion and saturation effects such as slewing. A solution to this problem is the input feed-forward architecture as shown in Fig. 9.43. The input signal is fed into the chain just before the quantizer. This leads to a transfer function ($c = d = 1$):

$$Y(z) = X(z) + \frac{N_Q(z)}{G(z) + 1} \tag{9.37}$$
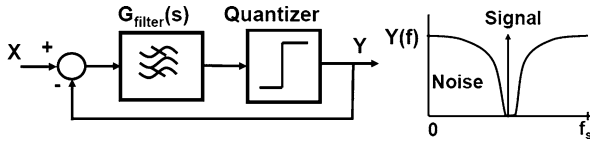
Fig. 9.44   A band-pass sigma-delta modulator

The summation node at the input creates the difference between input and output, which ideally contains only the quantization error and no signal. Although this swing can be large as well, the effect on signal distortion is less. Consequently the gain in the first stages can be somewhat larger. This comes at the cost of the introduction of the second summation point. A recent example is [236] with BW $= 25$ kHz, SNDR $= 95$ dB, P $= 0.87$ mW.

9.7.4 Band-pass Sigma-delta Converter

The dominant direction in the implementation of communications systems is the shift of more functions from the analog domain into the digital domain. As a result designers are looking for methods to implement the analog-to-digital interface at the intermediate frequency (IF) level of a superheterodyne receiver architecture. The RF signal is first converted to an intermediate frequency to allow sufficient rejection of the image frequency and to filter the wanted signal at a fixed frequency. Various forms of demodulation will result in the base-band signal. Depending on the system, the intermediate frequency is 450–500 kHz for AM radio, 10.7 MHz for FM radio, 33–38 MHz for European television and tens for MHz for various communication systems. Sigma-delta conversion allows to convert the narrow-band signal at the IF frequency by means of band-pass sigma delta conversion, Fig. 9.44. In stead of a low-pass filter a band-pass filter will clear the IF frequency band from quantization noise. Engelen et al. [237] showed a sixth-order sigma delta analog-to-digital converter with an IF of 10.7 MHz, BW $= 200$ kHz, ENOB $= 10.2$, P $= 60$ mW and Su and Wooley [129] implemented a digital-to-analog converter at one quarter of the sample rate. On system level there is ambiguity in this concept. In the analog domain dedicated filtering has to remove interferers and image frequencies in order to reduce the requirements for the modulator. During the band-pass sigma-delta conversion these clean frequency ranges are filled again with quantization energy. After the band-pass converter there is another filter needed to remove that unwanted spectrum. Low-IF or zero-IF topologies with a down-mixer combined in the first stage of the sigma-delta modulator [225, 234, 235, 238, 239] have proven to be preferred system solutions. van der Zwan et al. [238] showed a sigma delta analog-to-digital converter with an IF of 10.7 MHz, BW $= 200$ kHz, carrier-noise $= 79$ dB, P $= 11$ mW.
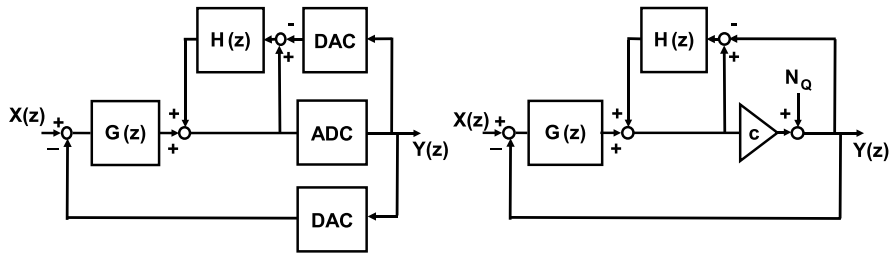
Fig. 9.45 A sigma-delta converter with embedded noise shaper. *Left* is the design and *right* the equivalent scheme [216]

9.7.5 Sigma Delta Loop with Noise-shaping

Higher order noise suppression can be achieved with high-order loop filters or with cascading of sigma-delta modulation loops. Higher-order filters require a higher oversampling ratio to allow the transfer top turn back to a first order behavior around unity gain. The cascaded sigma-delta needs more hardware to implement higher order noise shaping and depends on the matching between the analog and digital filters.

A third technique combines the advantages of a relative low-order sigma delta loop and an embedded noise shaper, see Fig. 9.45 (left). The noise shaper is constructed around the low-resolution analog-to-digital converter that serves as a quantizer. The difference between the input of the quantizer and the analog version of the output is fed into a filter H and re-inserted into the loop. This quantizer with noise-shaper serves as a modified quantizer in the overall sigma delta loop. With the help of the equivalent scheme in Fig. 9.45 (right) the transfer is calculated as:

$$Y(z) = \frac{cG}{(1+cG) + H(c-1)} X(z) + \frac{1-H}{(1+cG) + H(c-1)} N(z)$$

$$\approx X(z) + \frac{1-H}{G} N(z) \tag{9.38}$$

Depending on the time-continuous or time discrete implementation the signals and operations are in the s or z domain. The approximation is valid for $c = 1$ and $G \gg 1$. The term $(1 - H)$ reflects the shaping of the noise in the inner part of the circuit, while the term $1/G$ is the overall sigma delta function. This set-up allows to gain some additional noise suppression without the disadvantages of the other solutions.

9.7.6 Incremental Sigma-delta Converter

The idea behind sigma-delta conversion is to approximate a time-continuous input signal by many samples of a relatively low resolution digital-to-analog converter. If
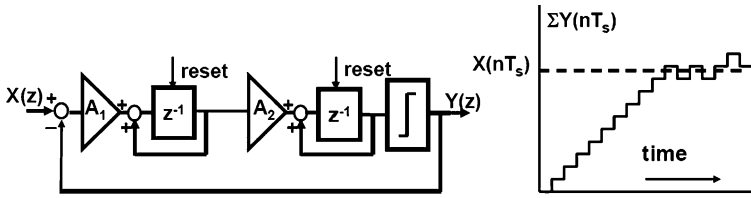
Fig. 9.46 A second-order incremental sigma-delta modulator

the input signal is a wave form where the useful information is only available during a fraction of the sample period, such as the output of a sensor, sigma-delta conversion is not directly applicable. A technique that allows some sigma-delta treatment of periodically available signals is the incremental sigma-delta converter [241], see Fig. 9.46. The main difference with a normal sigma-delta converter is the reset that sets the integrator values back to a starting value. If a signal $x(i) = x_{in}$, $i = 0 \ldots n$ is applied the integrators will sum the difference between the input signal and the generated output pulses. After processing n sample periods after a reset, the outputs of the first and second integrator equal:

$$\sum_{i=0}^{n-1}(x_{in} - y(i))$$

$$\sum_{m=0}^{n-1}\sum_{i=0}^{m-1}(x_{in} - y(i))$$

(9.39)

where $y(i)$ is the output for clock cycle i. With sufficient samples processed, the converter will reach an equilibrium. The comparator will be toggling around the output level of the last integrator, generating that pulse density whose average value matches the applied input signal. If the converter reaches a stable pattern, the toggle level of the comparator is assumed to be zero. The resulting transfer function of a converter with one integrator becomes:

$$\frac{1}{n}\sum_{i=0}^{n-1} y(i) \approx x_{in}$$

(9.40)

The first-order incremental sigma-delta converter can be easily read out by summing the positive and negative output pulses. This relatively easy converter circuit demands no stringent requirements on component accuracy or timing and is somewhat faster than linear Nyquist rate converters such as dual-slope. The accuracy improves with a larger number of samples n. The accuracy measured in effective bits improves only slowly with $^2\log(n)$. The main source of concern are sources of time varying behavior, such as drift in the components or interferer signals (unless the total period of operation can be chosen in synchronicity with the interferer).

A second-order incremental sigma-delta converter has a transfer function:

$$\frac{2}{n(n-1)} \sum_{m=0}^{n-1} \sum_{i=0}^{m-1} y(i) \approx x_{in} \tag{9.41}$$

In second-order incremental sigma-delta converters the two integration steps result in a quadratic relation with n and the accuracy of the total system is proportional to $^2\log(n^2)$, which means that for $n = 300\ldots600$ a 16-bit resolution is possible.

Chapter 10
Characterization and Specification

Every system designer of analog-to-digital conversion techniques will use the specified parameters to determine the usefulness of a converter for the application. The definition of parameters is mostly well described, however the actual measurement of the parameter can still leave margins for variation [93, 242, 243]. A supplier may restrict himself to the typical value of a parameter, however also the maximum and minimum values can be part of a specification. Moreover, it may be useful to examine the temperature and supply voltage ranges in which the specified value for a parameter is guaranteed.

The application will determine which parameters of a converter are most relevant. For high quality audio equipment the distortion is relevant. In communication equipment intermodulation is important as well as the spurious free dynamic range in a certain frequency span. Next to the standard specification points of analog-to-digital converters that are listed in Table 10.1 also a range of secondary qualifications can be taken into account. Examples are: phase behavior, package (height, volume, pin), tolerance to light (e.g. in an optical sensor), available references (internal or external), input impedance, etc.

The correct application of a converter depends crucially on a good characterization, test set-up and parameter extraction.

10.1 The Test Hardware

A correct evaluation of a converter starts at the beginning of the chip design. Interfaces to and from the test equipment must be defined. These analog or digital drivers and buffers should not interfere with the test or jeopardize the signal quality. High sampling frequencies require low-jitter buffers and high frequency analog output signals require wide bandwidth buffers.

Sub-sampling can be used to test a sampling device such as a track-and-hold circuit. Two devices are cascaded as in Fig. 10.1. The first device is operated at a high sample and signal rate. The second track-and-hold samples the output of the first device at an integer fraction of the sample rate. The original output signal of

Table 10.1 Main characterization parameters of an analog-to-digital converter

Specification	Symbol	Unit
Nominal amplitude resolution	N	1
Sample frequency	f_s	Hz, s^{-1}
Bandwidth	BW	Hz
Integral linearity	INL	LSB
Differential linearity	DNL	LSB
Monotonicity		
Missing codes		
Harmonic distortion	THD	dB
Intermodulation distortion	IM2,3	dB
Spurious free dynamic range	SFDR	dB
Signal-noise and distortion ratio	SINAD	dB
Signal-noise ratio	SNR	dB
Effective number of bits	ENOB	1
Dynamic range	DR	dB
Jitter	σ_t	ps
Power consumption	P	W
Temperature range	T	°C
Power supply	V_{DD}	V
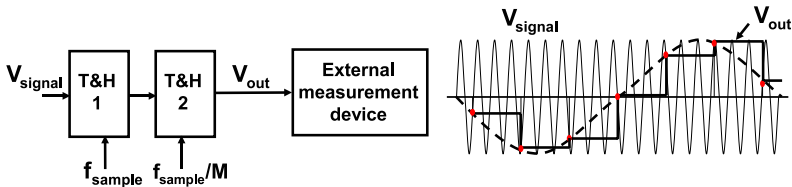
Fig. 10.1 The first track-and-hold circuit is tested at high input and high sample rate. The second device runs a sample rate that is an integer factor slower. The resulting output signal contains signals that correspond to the first track-and-hold output signal and all harmonics

the first track-and-hold circuit and its harmonics are sub-sampled to a low output frequency that can be easily measured. The "sub-sample" method requires quite some skills to interpret correctly the resulting components.

Every experiment starts with a PCB on which the device-under-test (DUT) is mounted. Some more points must be considered when designing a test board:

- Analog and digital power supplies and signal sources should be kept separate and only connected together on one single node. Be aware of coupling of earth loops via the mains plug.
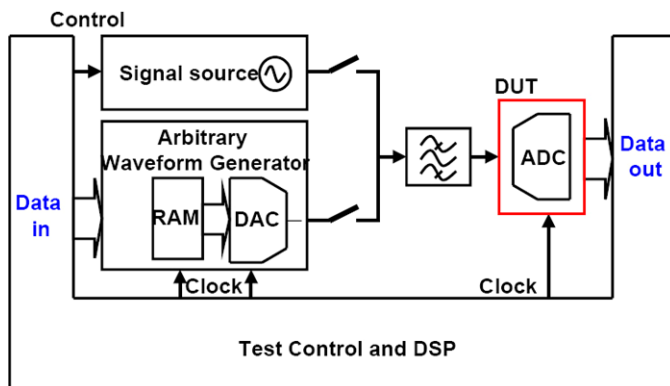
Fig. 10.2   Test set-up for digital-to-analog conversion

- Provide sufficient decoupling: microfarad electrolytic capacitors for the low-frequencies and metal or ceramic capacitors for the high frequencies. Mount them as close to the package as possible.
- In the evaluation phase it may be tempting to use a tool that allows to exchange the samples easily. However these tools add a lot of distance between the die and the PCB and therefore add many nanohenries of inductance. In Fig. 10.3 a technique is shown where the surface mounted device is pushed onto the connection electrodes of the printed-circuit card. This set-up keeps the distances short.
- High frequency connections must be laid-out keeping in mind that every wire is a transmission line. PCB lay-out packages have options to design wires and surrounding grounding in such a way that a defined impedance is achieved.
- Every signal must be properly terminated close to the test device. This certainly holds for digital signals. Non-terminated digital signals will ring and inject spurious charges into the substrate.

A professional measurement set-up for characterizing an analog-digital converter uses a computer to control the set-up and to analyze the measurement results, see also the relevant IEEE standardization documents [92–94]. Many professional evaluation set-ups are constructed with racks of measurement equipment connected by some interface bus. A computer equipped with test software will control the equipment, set-up voltages and currents, step through the signal range and capture the data in a data-logger of several Gigabytes storage. The signal source and the generator for the sample rate have to comply to more stringent specifications than the device under test. Modern signal sources are equipped with an extensive user interface, which goes sometimes at the cost of signal distortion and purity. Old "analog" generators are often to be preferred over the modern equipment of the same price level. A well-known method to obtain a high quality measurement signal uses passive filters, such as the anti-alias filter in Fig. 10.2. This set-up avoids that remaining distortion components, noise of various origins as well as cross-talk of the generators internal processing, disturb the measurement.
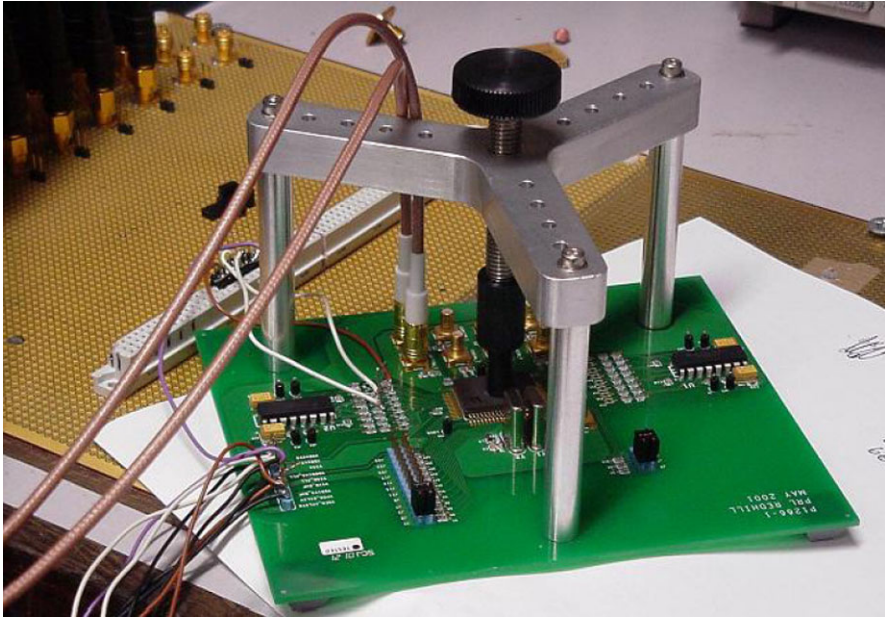
Fig. 10.3 Test board for measuring an analog-to-digital converter, *courtesy*: R.v. Veldhoven

The analog-to-digital converter is normally mounted on a load-board: a printed-circuit board adapted for connection to the main test equipment, see Fig. 10.3. As a direct coupling of the converter to the test equipment may result in long wires, loading and ground loops, the (digital) side of the converter is buffered near the device. This buffer will act as a decoupling of signal ringing over the long connection lines. The buffer shields the converter from the high energies that are associated with driving the tester. In extreme cases the connection between the tester and the device is made via an optical fiber, so that a perfect electrical separation between the converter and the tester is achieved. In the tester a data storage device (data-logger or data-grabber) will store the high-speed data that comes from the converter. The computer can then in a second phase analyze the data at a convenient speed. The postprocessing results in an output as shown in Fig. 10.6.

An important part of the requirements for analog-to-digital testing hold similarly for digital-to-analog converters. Figure 10.4 show a potential set-up for the test. In accordance with the principle of coherent testing of the next section, the computer generates and stores a number of data samples in the storage. A cyclic process reads the data at the desired sample rate and feeds the digital-to-analog converter. The required measurement equipment must exceed the specifications of the to-be-tested device. By applying a passive filter, the main component of the output signal can be suppressed so that the measurement equipment only needs to have sufficient resolution for the remaining components. The signal analysis will involve a spectrum analyzer or another form of analog-to-digital conversion.
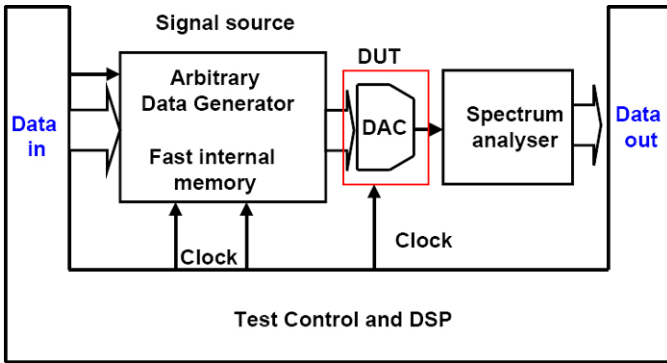
Fig. 10.4 Measurement set-up for a digital-to-analog converter

10.2 Measurement Methods

10.2.1 INL and DNL

A simple way to evaluate the behavior of a converter is to apply a sawtooth signal to the input. For converters with specifications on absolute accuracy a programmable voltage source is a good choice. Less demanding applications can start with a generator or a self-built circuit. If the sawtooth is sufficiently slow, there will be enough sample moments to determine the DC-parameters as INL, DNL and monotonicity. If these parameters need to be established at a 0.1 LSB accuracy level, a data storage of 10×2^N samples is necessary.

A sawtooth signal is not the most critical signal for fast converters and is not easy to generate at high precision. Nyquist analog-to-digital converters with maximum input frequencies ranging from tens of Megahertz into the Gigahertz range require the use of statistical methods. The input signal for measuring dynamic specifications is then preferably a sine wave. This test signal can be obtained with a relatively high quality through the use of passive filters. These fast and relatively accurate methods for determining INL and DNL use the statistical properties of sine waves.

When a full-amplitude sine wave with period T_{sw} is applied to a converter, there is a probability for every code to be hit a number of times. A sine wave will hit more levels in the upper and lower range than in the middle. If the conversion range is defined mathematically between 0 and 1, a full-amplitude sine wave takes the form:

$$y(t) = 0.5 - 0.5\cos(2\pi t / T_{sw}) \tag{10.1}$$

The signal will go from the lowest level to the highest level in half of a cosine period. Δy is a fraction of the range (e.g. 1 LSB) at conversion level y and is called a "data bin". Δt_y is the corresponding fraction of time of the half cosine wave. Δt_y corresponds to the hits in bin Δy while half of the cosine period ($T_{sw}/2$) corresponds to

Fig. 10.5 Example of the ideal distribution of hits when a 5-bit analog-to-digital converter is fed with a sine wave
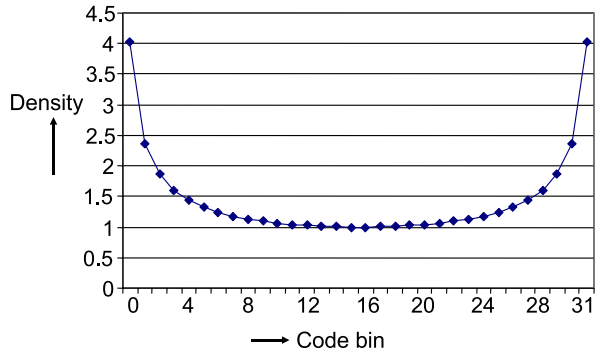
the total amount of samples. The ratio $\Delta t_y/(T_{sw}/2)$ is now the fraction of hits that end up in bin Δy.

$$t = \frac{1}{\omega_{sw}} \arccos(1 - 2y)$$

$$\frac{dt}{dy} = \frac{1}{\omega_{sw}\sqrt{y - y^2}} \tag{10.2}$$

$$\frac{\Delta t_y}{T_{sw}/2} = \frac{2}{T_{sw}} \frac{dt}{dy} \Delta y = \frac{\Delta y}{\pi\sqrt{y - y^2}}$$

Δy is chosen as 1 LSB. Figure 10.5 shows a characteristic distribution of the number of hits per level, or binning of levels. A test run generates the actual measured distribution of the converted values of a sine wave. This measured distribution is compared to this theoretical curve and the deviations (scaled to the same level) result in an INL and DNL plot, as is shown in Fig. 10.6.

This "histogram" method can be used at any frequency. It provides also information on the linearity problems at higher signal frequencies. The DNL measurement is not optimum as non-monotonicity in this measurement method is not found. Non-monotonicity just changes the DNL value of the corresponding code, the associated step-back is missed. An additional saw-tooth test is required.

The above calculation allows to determine the minimum number of samples that must be generated to get one hit in the middle bin. There the level corresponds to $y = 0.5$ and

$$\frac{\text{hits in bin at } y = 0.5}{\text{total samples}} = \frac{\Delta t_{y=0.5}}{T_{sw}/2} = \frac{\Delta y}{\pi\sqrt{y - y^2}} = \frac{\Delta y}{\pi/2} = \frac{1}{\pi 2^N/2} \tag{10.3}$$

Applying a saw-tooth signal means that ideally 2^N sample periods are needed to hit each level once. With a sine wave the number of samples has to be $\pi/2$ times larger. To obtain an accuracy of 0.1 LSB in INL and DNL a minimum of $10 \times \pi 2^N/2$ samples is required.

The calculation above, suggests that the input amplitude of the sine wave must accurately match the analog-to-digital converter range. In advanced test packages,
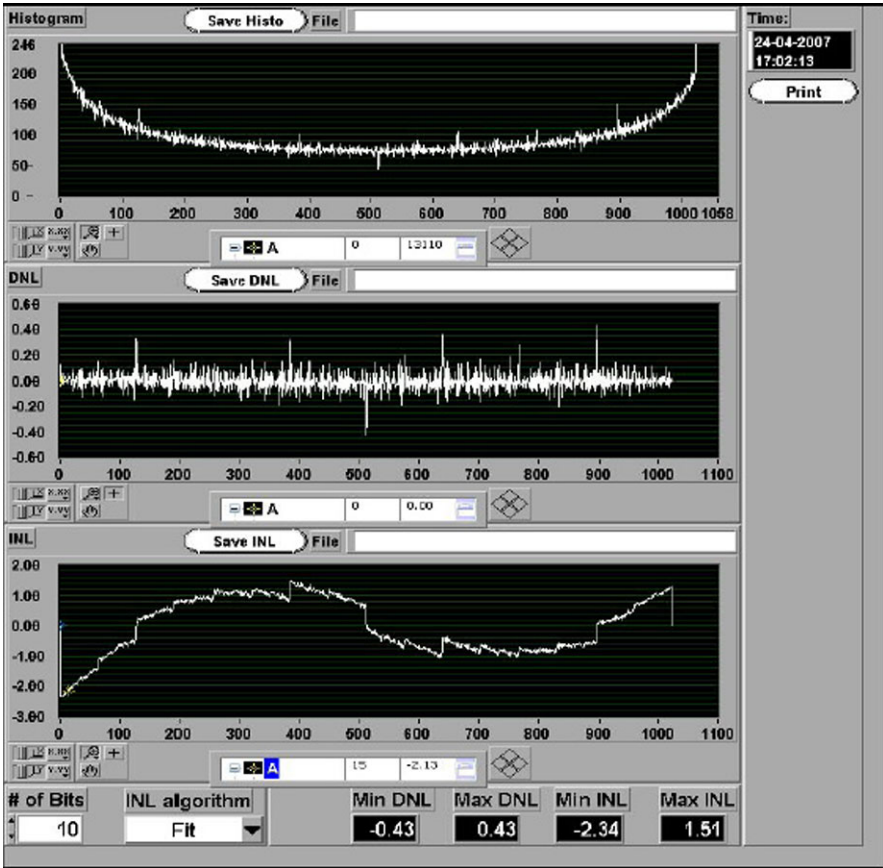
Fig. 10.6 Output of an automated test set-up. *Top*: histogram output, *middle*: DNL, *bottom*: INL

routines exist that will allow also amplitudes that extend over the input range. A re-
construction of the input is also possible, see Fig. 10.7.

10.2.2 Harmonic Behavior

The same sine waves allow to measure harmonic distortion and related qualities
(intermodulation, spurious-free dynamic range, etc.) as well as the signal-to-noise
ratio. Fourier transformation of the output sample series allows to generate a fre-
quency diagram, Fig. 10.8. Many Fourier algorithms require that the period in which
the data is measured contains both an integer number of signal periods as well as an
integer number of sample periods. If this condition is not met, the resulting signal
will show side lobes, making the interpretation of the Fourier result tedious. This
phenomenon is called frequency leakage and is illustrated in Fig. 10.9.
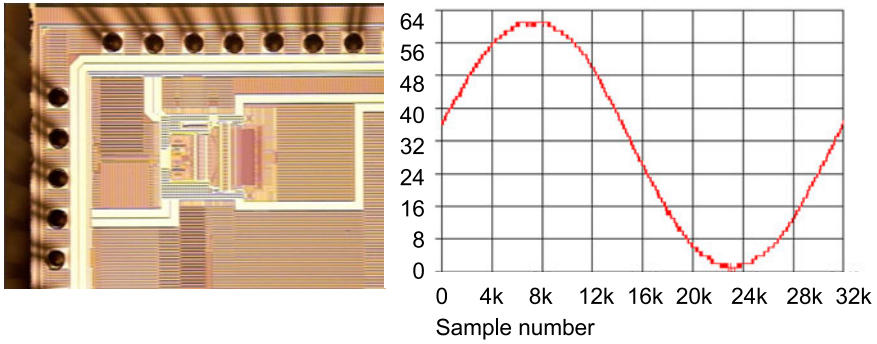
Fig. 10.7 Reconstructed wave form of a 311 MHz signal sampled at 1.44 Gs/s [158]

Fig. 10.8 Dynamic measurement of an analog-to-digital converter on intermodulation at $f_s = 100$ MHs/s
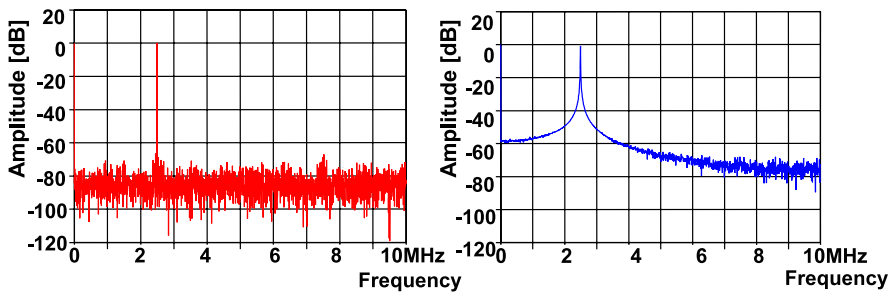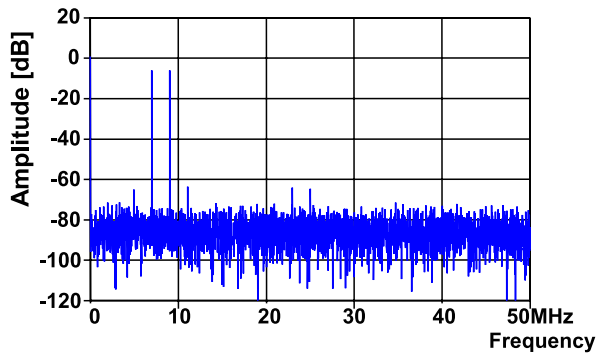
Fig. 10.9 Frequency leakage because of 1 missing sample, *left*: 4000 samples, *right plot*: 3999 samples

A second pitfall can occur if the sample rate is a simple multiple of the signal frequency. Under stable signal conditions only a limited number of the levels in the conversion process will be used. The evaluation of the converter is based on the repetition of the same limited sequence of measurements, and adds no information on the levels that are missed, see Fig. 10.10.
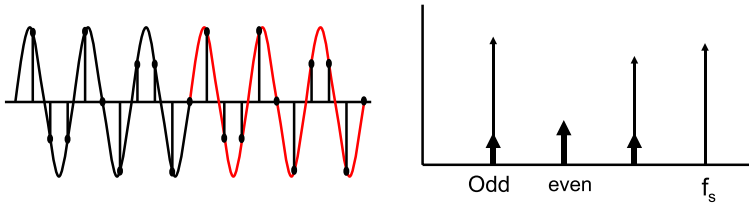
Fig. 10.10 If the sample rate is an integer multiple of the sample rate, a part of the measured samples are simple duplicates of the earlier sequence. In the frequency domain this may lead to the masking of harmonics behind other harmonics or behind the fundamental frequency

The basic requirement for a good test that avoids both problems, is called the "coherent testing" condition:

$$T_{\text{meas}} = \frac{M_s}{f_s} = \frac{M_{\text{signal}}}{f_{\text{signal}}} \tag{10.4}$$

Where M_{signal} equals the number of input signal periods and M_s the number of sample periods. If both integers are mutually prime no repetition of test sequences will occur. Mutual prime or co-prime means that the largest common divisor of M_{signal} and M_s is 1. The total measurement period is given by T_{meas}. The measurement period is inversely proportional to the frequency resolution or the frequency "binning" of the Fourier transform. It is therefore necessary to choose M_{signal}, M_s and T_{meas} sufficiently large.

The discrete Fourier transform creates $M_s/2 + 1$ frequency bins of a size $1/T_{\text{meas}} = f_s/M_s$. Both bins at 0 and at $f_s/2$ are counted. A spectrum analyzer often provides the option to define the bin size by means of the "resolution bandwidth" parameter. If this value is set, automatically the measurement period will be adjusted.

The energy in the time discrete signal is distributed over these frequency bins. If energies from different phenomena (e.g. a harmonic component and a folded component) end up in the same bin, the signal strength of these components will add up or extinguish. A finer frequency grid can be obtained by increasing the number of samples by increasing the measurement period T_{meas}.

As an example: The 8-bit analog-to-digital converter in Fig. 10.11 has a theoretical maximum signal-to-noise ratio of $1.76 + 8 \times 6$ dB $= 49.8$ dB. A measurement and Fourier transform with 200 samples will result in 101 bins. The quantization energy is distributed over these bins, so the "noise floor" in the spectrum is expected at $49.8 + ^{10}\log(100)$ dB ≈ 70 dB below the fundamental frequency. A tenfold increase will lead to a 10 times lower amount of energy per bin. In a spectral plot the noise floor will drop by 10 dB. Figure 10.11 compares spectra taken with 200 and 2000 samples.
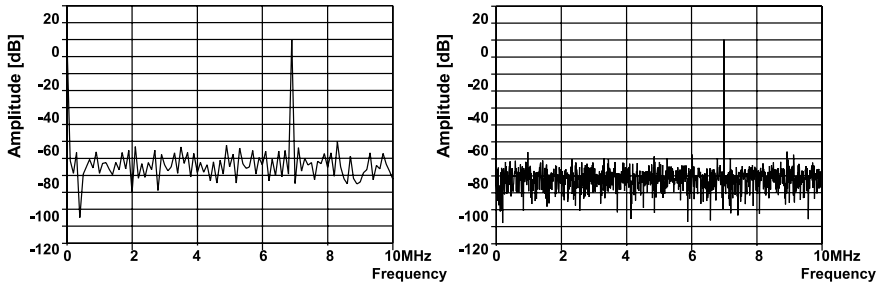
Fig. 10.11 Increasing the measurement period and the number of samples by a factor of 10, reduces the bin size with that factor and lowers the noise floor by 10 dB

10.3 Self Testing

In complex systems sometimes forms of self testing are necessary. Think of sensor systems that need calibration in places that are difficult to reach. In another example there is a liability aspect to the measurement equipment and the usability of the converter must be established in-situ (e.g. in a drilling head at 2 kilometer below the earth's surface).

Considerations for the implementation of self test are:

- Complexity versus functionality: it may be sufficient to establish correct connectivity of the converter. A simple block wave may be sufficient to test.
- Independence: no test may lead to a positive result because one error has the same effect on the test circuit as on the converter. Using the same reference for the converter as for the test circuit, will disable proper detection of reference deviations.
- The cost of error detection: are repair facilities present, or can redundancy lead to a solution (e.g. take a two-out-of-three vote).
- A parametric test can only be performed if somehow accuracy of the test signal is provided. So self-tests creates the need for having somewhere a more accurate reference.

Self-testing can be implemented in systems where both a receive and a send chain are present. In a 2.4 GHz transceiver, such a "loop-back" facility feeds a fraction of the transmit power into the receiver. Proper test sequences applied to the digital-to-analog converter input in the send chain allow a functional self-test and also a few parameters can be evaluated.

Another example of self-testing comes from systems where it is impossible to approach the converter. For seismic purposes ships drag large seismic arrays of cables with sensor interfaces. These arrays span several hundreds of meters. Before a measurement is taken the quality of the total interface chain is tested by means of build-in self test circuits.

It is expected that these professional developments of self-testing in some years will result in a considerable improvement of the performance of self-test methods.

Chapter 11
Technology

The technology in which a converter is designed, is crucial to the obtainable performance. A full discussion on all aspects of technology fills a series of books. In this chapter just a few relevant aspects are discussed: the consequences of the CMOS process evolution, variability with systematic and random variations, packaging and substrate interference.

11.1 Technology Roadmap

Once the system aspects of an IC have become clear, some important technological choices are still open that influence the performance of the analog-to-digital or digital-to-analog converter.

The technological parameters that are most important for an analog designer are: power supply (limiting the signal swing), feature size (speed), process options and tolerances. Many of these specifications in CMOS are derived from the ITRS road map [19]. The International Technology Road map for Semiconductors is a consortium of semiconductor companies all over the world that updates every year the most important parameters of CMOS technology and projects them into the future. The projection of the ITRS serves as a focus point for equipment supplier, IC manufacturers and electronics appliances producers as an indication of what is seen as a realistic progress in this field. Table 11.1 presents some data.

The ITRS road map displays the view of industry on various aspects of IC technology and design. However, the quoted numbers should be read with care. The node identifier ("90-nm technology") was in older technologies identical to half the pitch of the critical layers (poly-silicon) and a direct indication for the lithographical needs. For technologies above 130 nm, this number indeed reflects the half-pitch of the lithography. At 90-nm and below the relation with lithography is less strict. The 65-nm node allows for irregular design, such as analog-to-digital converters, a half pitch of 80–85 nm. The effective gate length that directly determines the digital speed, on the other hand is reduced to less than 40 nm by applying various channel-length reduction techniques. Similar remarks hold for other numbers in the road

Table 11.1 Some characteristic data from the ITRS road map 2005, supplemented with older data

Year (ITRS analog/RF)	1995	1998	2000	2002	2004	2007	2010	2013
ITRS node (nm)	350	250	180	130	90	65	45	32
V_{DD} (V)	3.3	2.5	1.8	1.2	1.2	1.1	1.1	1
Min. gate length (nm)	350	250	100	75	53	37	28	22
d_{ox} (nm)	7.0	5.7	3.6	3.2	2.2	1.9	1.4	1.3
g_m/g_{ds} (5× min. gate-length)	200	150	120	47	32	30	30	30
$1/f$ noise ($\mu V^2\,\mu m^2$/Hz)	500	300	180	180	160	100	80	70
A_{VT} matching (mV μm)	7	6	6	6	6	5	5	5
Peak f_T (GHz)	20	40	70	100	170	240	320	400
Nfmin (dB)					0.33	0.25	0.2	<0.2
DRAM size	64M	128M	256M	512M	1G	4G	8G	32G
Tr density logic (tr/cm^2)				39M	77M	154M	309M	617M
Logic density (gates/mm^2)	20k	40k	100k	200k	400k	800k	1.6M	3M
SRAM density (bits/mm^2)				350k	700k	1.4M	3M	6M
Digital clock speed (GHz)				1.68	3.99	6.74	11.5	19

map. The logic gate density is measured on an eight-pitch height 2-input NAND. Such a low-height pitch would in practical designs lead to a massive wiring congestion. Normally cell heights between 9 and 14 are used. With a fill factor between 70 and 90%, the effective logic gate density in a real design is about half of the quoted number. As a reference: a powerful ARM9 embedded processor requires some 300k gates, excluding the memories and special functions.

11.1.1 Power Supply and Signal Swing

Many analog circuits are limited by signal swing. Figure 11.1 shows the expected power supply voltage according to the ITRS. The available single-sided signal swing (shaded area) is derived by subtracting from the minimum power supply ($0.9V_{DD}$) the maximum threshold voltage and a gate bias of 0.2 V. The graph clearly shows that this form of signal handling has reached its limits. Besides a low threshold-voltage option or a local power supply boost, the most in suitable form of representation for signals on mixed signal chips is the differential mode. This form of signal representation is less sensitive to substrate noise and provides inherent cancellation of even harmonic distortion, however it requires differential inputs and outputs: two-pin analog I/Os. Despite these measures, signal swing is expected to decrease from 2 V to some 0.3–0.5 V peak-peak in minimum power supply applications. Further repercussions due to the reduced signal swing will be discussed in the next sections.

Fig. 11.1 The power supply evolution as a function of process generation. The signal swing (*shaded*) is derived by using the minimum power supply minus a threshold and a bias voltage
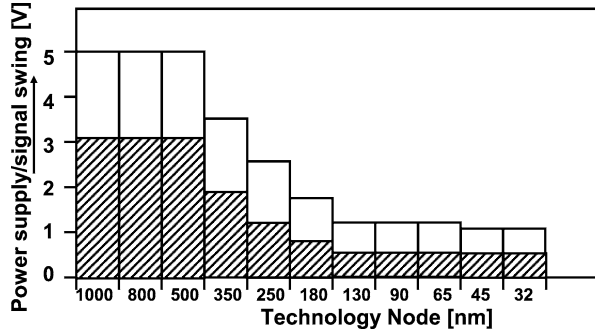
11.1.2 Feature Size

In digital CMOS technology the drive for smaller feature sizes is of course a dominant factor as a short gate length directly improves the speed performance. For low-operating power the development in capacitances is of particular importance, because low capacitance values lower the required power. Digital power consumption in CMOS logic for a single inverter is [244]:

$$\text{Power} = N_i f (C_{\text{load}} + C_{\text{gate}}) V_{DD}^2 + P_{sc} + P_{\text{leak}} \tag{11.1}$$

The power is composed of a capacitor charging component, a short circuit and leakage component. N_i is the number of transitions (activity) that in worst case, may exceed 1 due to e.g. carry or ripple-through mechanisms. C_{load} represents the node capacitance which is strongly affected by the technology and the feature size. In active circuits the dynamic power consumption ($f C V^2$) is dominant, while in stand-by mode the leakage current is the main power surge. Leakage originates mostly from the sub-threshold current in switch-off mode. Short-circuit dissipation is caused when during the transition of a gate, the NMOS and PMOS transistors are both switched on. Only in rare situation this causes a relevant contribution to the power dissipation.

Next to the technological need for reducing the electrical fields in the MOS transistor, the quadratic relation for the supply voltage V_{DD} explains the drive towards lower supply voltages (see Fig. 11.1).

In analog circuits, feature size reduction promises a higher cut-off frequency f_T for the same current. Figure 11.2 shows the cut-off frequency as a function of technology node. The cut-off frequency is the ratio of transconductance and load capacitance ($f_T = g_m/2\pi C$). These cut-off frequencies are determined by RF-designers at a high gate voltage of $V_G \approx V_{DD}$. Analog-to-digital converter designers usually cannot use gate overdrive voltages V_{GT} over 0.2–0.4 V. Consequently the speed values from Fig. 11.2 are only a relative indication for process improvement.

Fig. 11.2 Evolution of the
cut-off frequency f_T for
minimum gate length NMOS
transistor, from [245]. LOP
means low operating power,
and LSBP means low
stand-by power
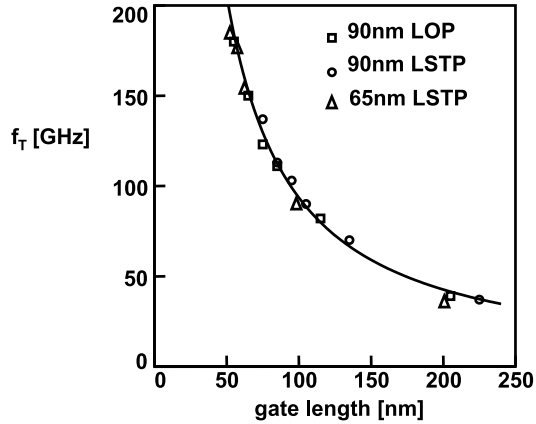
11.1.3 Process Options

Specific application demands, availability, experience and cost dominate the choice
of the baseline process technology. One level below the baseline technology there is
more freedom: most foundries offer two or more generations of a process: a mature
process with large feature sizes and a advanced shrink version. Baseline processes
also allow adding a variety of process options: wiring layers, additional components
(capacitors, resistors, fuses) and device parameters. Process options are often used
to circumvent the inherent drawbacks of baseline digital CMOS process. Moreover,
options allow designers to map the function better on the process, reducing area
and the amount of power consumed. Process options can be subdivided into four
categories [246]:

- No extra masks, no process adaptations.
  The extent of analog characterization and monitoring of specific process pa-
  rameters. Examples are: MOST noise, matching, temperature and voltage depen-
  dencies of passive and active components. Proper characterization (of e.g. match-
  ing) reduces margins and allows to design for minimum power consumption. The
  reduction in process tolerances also falls in this category, see Table 11.2.
- No extra masks, minor process adaptations.
  The definition of MOST thresholds is crucial for low-voltage operation. Ana-
  log designers want high-threshold transistors for input stages and low-threshold
  transistors for switches. Structures that are normally considered parasitics in the
  base-line process are useful in an analog design. Examples: vertical PNP transis-
  tors, large MOSTs as capacitors, resistors and multi-layer interconnect capacitors.
  Stacked-layer capacitors show a minimum parasitic component. This option does
  not introduce new mask steps, but may require another combination of masks.
- Extra masks, minor process adaptations.
  Examples are: multiple thresholds in order to optimize for analog switching,
  Lightly-Doped Drain (LDD) suppression on the source side of a MOS transistor,

Table 11.2 Increase in the width of a minimum length transistor due to parameter tolerances

Parameter variation	Nominal	Small	Medium	Large
C_{load} wiring/diffusion	0%	+10%	+20%	+30%
L minimum length	0%	+10%	+20%	+30%
T_{clock} clock skew	0%	−10%	−20%	−30%
Temperature T	25 °C	85 °C	120 °C	140 °C
μ mobility $\propto T^{-2}$	0%	−30%	−42%	−48%
$V_{DD} - V_T$ gate drive	0%	−5%	−10%	−15%
Effect on gate width	1×	2.2×	4.1×	7.1×

definition of resistors by removing the silicide layer, a special di-electric layer to create so-called metal-interconnect-metal (MIM) capacitors.
- Extra masks, major process adaptations.

  The addition of new components may necessitate new process steps and masks. This is certainly the most critical category of process options as it increases the costs. Examples: second gate oxide for high analog voltages, second poly-silicon layer. A double poly-silicon capacitor has a decade less parasitics than a stacked wiring layer capacitor, realizing the same function at a lower power consumption.

The lists of examples of the above categories can be expanded. The first category is needed in all industrial analog design: in that sense analog design always requires options. The second category is also needed for converter designs. The third and particularly the fourth categories are disputable: here the analog designer's wishes will inflict a cost penalty upon the digital part of the design. This will rarely be acceptable for a commercially viable application.

11.2 Variability: an Overview

Variability[1] is generally interpreted as a collection of phenomena characterized by uncontrolled parameter variation between individual transistors or components in a circuit. This collection is populated with a large number of effects ranging from offset mechanisms to reliability aspects.

Variability effects can be subdivided along three main axes:

- Time independent versus time variant effects.
- Global variations versus local variations.
- Deterministic versus stochastic (statistical) effects.

---

[1]The following four sections also appear in a modified form in: "Compact Modeling: Principles, Techniques and Applications" by G. Gildenblat (ed.). The original publication is available at http://www.springerlink.com and was co-authored with Hans Tuinhout and Maarten Vertregt.

Table 11.3 Time dependencies of some variability effects. The *lower row* lists the methodologies the designer has at his disposal to mitigate these effects

Static	Seconds	Micro-seconds	Pico-seconds
Process corners	Supply voltage	Wiring IR drop	Dyn. IR drop
	Temperature	Temp gradient	Jitter
Lithography	Hot carrier		Substrate noise
Line edge roughness	NBTI, drift		
Dopant fluctuation	$1/f$ noise	$1/f$ noise	kT noise
WPE, STI Stress	Soft breakdown		
Design margins, calibration	Compensation circuitry	Stretched design margins	Stretched design margins

Table 11.4 From global to local variability

	IC	Sub circuit	Transistor	Atomic
Electrical	Supply voltage	Substrate noise	Wiring IR drop, jitter, cross-talk	
Thermal	Temperature	Temperature gradient	Local heating	Thermal noise
Technology	Process corner	CMP density	Well-proximity	Mechanical stress
Lithography	Line width	Layer density	Lithographic proximity effects	Line-edge roughness
Physics		Wafer gradient	NBTI, soft-breakdown, hot-carrier, stress	Random dopant, mobility variations

Every axis has specific properties that must be considered before defining a successful model and choosing the appropriate design solutions. The first axis that is considered in the control of variability effects is the behavior in the time-domain. Table 11.3 subdivides the variability effects in static, slow, medium and fast changing events. This subdivision is linked to the methodologies a designer has available to mitigate performance loss or yield loss.

Static effects allow a one-time correction of the associated devices in the circuit and can be performed during final test of the product. These corrections can be implemented via non-volatile memories, laser trim or polysilicon fuses. Slow time-dependent mechanisms must be corrected during operation and can be addressed via circuit compensation tricks, such as auto-zeroing and on-chip calibration. However, variations with a time span comparable to the maximum speed of the process cannot be handled in this way. Stretched margins or more power are required.

In the scope of this chapter only phenomena that can be considered static within the observation period are discussed.

Table 11.4 lists variability effects along the physical dimension axis from global to local variations. This list is not complete and new process generations add new

Table 11.5 Deterministic and stochastic effects

Deterministic	Stochastic
Proximity effects	Line edge roughness
Electrical offsets	Dopant fluctuation
WPE, Stress, STI	Mobility fluctuation

phenomena. The global-to-local axis is subdivided into four levels of granularity. Variations on the level of an integrated circuit are normally considered as standard design space variables. Designers are used to incorporate these variations in their Process-Voltage-Temperature "PVT" analysis. Extensive models of components and well-characterized parameter sets cover these aspects.

The variations that affect sub-circuits are less commonly incorporated in models and design software. Often the only mechanism available to a designer is a set of design rule checks (DRC) or electrical rule checks (ERC).

Most effects at transistor level are well-modeled as the transistor is the focus point of circuit modeling. Various forms of reliability and hot-carrier effects can be predicted based on the physics involved. Electrical effects on this level of granularity are well understood, but the problem is to judge their relevance within the complexity of the entire circuit. The current variation in a transistor due to substrate noise is trivial if the magnitude of the substrate noise is known. However, establishing that magnitude for a multi-million devices circuit is practically impossible.

An increasing number of transistor-related and atomic-scale effects become relevant in nanometer devices. Most of these effects are well-described and modeled in physics on a micrometer level, but the associated statistical effects are not an integral part of the design flow.

From a statistics point of view these (time-independent) effects can be subdivided in two classes: deterministic and stochastic. In designer's terms: offsets and random matching. As simple as this division seems, there is a complication: a number of phenomena is from a physics point of view deterministic, but due to circuit complexity, a statistical approach is used to serve as a (temporary) fix. An example is wiring stress, where the complexity of concise modeling is too cumbersome.

From a philosophical perspective, the listed random effects are not truly stochastic effects. Here a practical point of view will be used: all effects that are reproducible from die-to-die will be categorized as deterministic. If the variation source changes the behavior of every individual device with respect to the average, the effect will be described with stochastic means, see Table 11.5.

11.3 Deterministic Offsets

Systematic offsets between pairs of resistors, CMOS transistors or capacitors are due to electrical biasing differences, mechanical stress or lithographic and technological effects in the fabrication process [247].
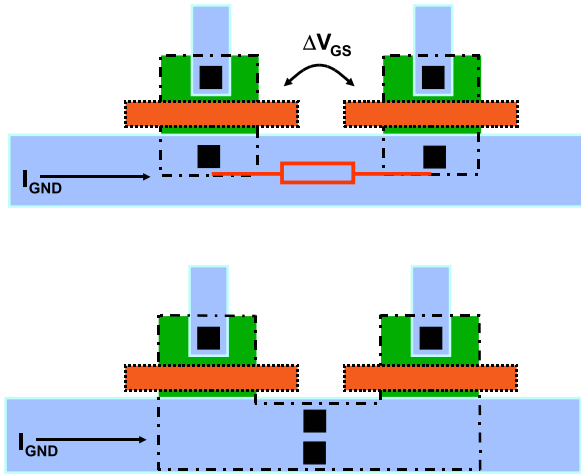
Fig. 11.3 A small resistance in the source connection wire will result in offset in the gate voltage. A common connection point is better from an offset point of view

11.3.1 Offset Caused by Electrical Differences

It may seem trivial, but good matching requires in the first place that matched structures are built of devices made from the same material and are of equal size. This implies that a 2 : 1 ratio is constructed with 3 identical elements, connected in a manner that will result in the desired ratio. The required 2 : 1 ratio applies for every aspect of the combination: area, perimeter, etc. A pitfall can occur when an existing lay-out is scaled and the resulting dimensions rounded off to fit the new lay-out grid. The rounding operation can result in unexpected size deviations in originally perfectly matched devices.

For electrical matching, the voltages on all the elements must be identical. Node voltages are affected by voltage drops in power lines, leakage currents in diodes, substrate coupling, parasitic components, etc. Electrically derived effects must also be considered, e.g. heat gradients due to power dissipation and temporary charging effects due to phenomena as Negative Bias Temperature Instability [248].

Figure 11.3 shows a surprisingly often encountered example of electrical mismatch: a pair of transistors with a wire connecting the sources. The current I_{GND} in this wire will cause a voltage drop between the two source terminals of the transistors. This voltage drop shows up as an offset ΔV_{GS} in the gate-drive voltage. Such problems increase when additional currents are routed via this wire. A star-like connection is well-suited to avoid this problem. A rigorous inspection of the lay-out in which the main current paths have been identified, is always needed when offset problems are suspected.
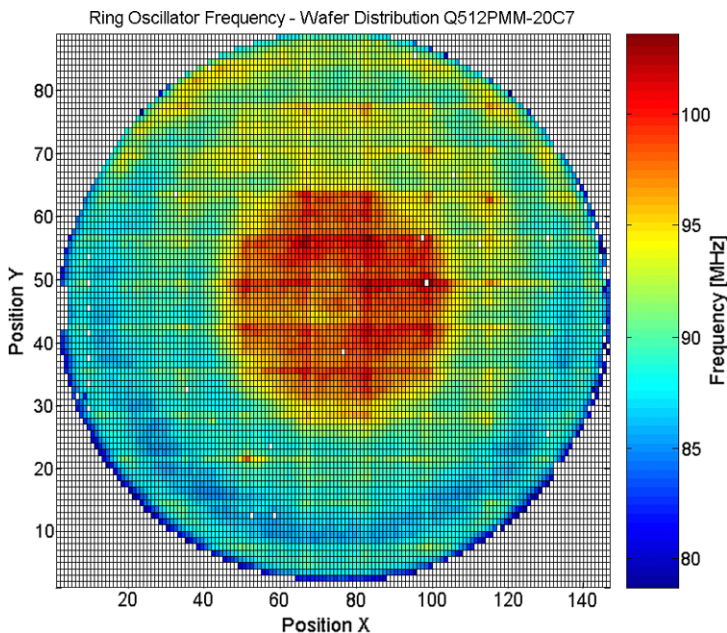
Fig. 11.4 A wafer from a 90-nm CMOS production lot on which a free-running oscillator frequency is measured. The frequency deviation (80–100 MHz) is largely an indication for the variation in gate length due to mask-plate dimensional errors, lithography and the processing steps that determine the electrical gate length. *Courtesy*: B. Ljevar, NXP

11.3.2 Offset Caused by Lithography

During the lithography process the structures drawn in the lay-out are transferred to a mask and in the resist. In the next step physical structures are etched into a wafer. During this process there are many crucial details that will affect the quality of the patterning. Figure 11.4 shows the frequencies of free-running ring oscillators of circuits that are measured on a 300-mm wafer. Ring oscillators are most sensitive to gate-length variation, so the frequency map indirectly represents a gate-length map. Most prominent is the middle area, where a 20% higher frequency is measured with respect to the rest. A potential cause is the heat non-uniformity during the high-temperature dopant activation step. A second phenomena is visible as a pattern of rectangular shapes. The wafer stepper uses a reticle containing 16 by 7 devices. This $\approx$400 mm^2 reticle is repeated to pattern the entire wafer. Random effects are averaged over the nearly one hundred stages of the oscillator and are hardly visible.

These large distance effects do not have a direct impact on equality of transistor pairs. Most of this effect is a global variation and becomes part of the tolerance budget in the definition of the parameter corners of a process. In some digital design environment a specific guard band is used to make sure that the potential loss of current drive capability becomes visible to the designer. Yet the example makes clear what the significance is of lithography variations on the overall performance.
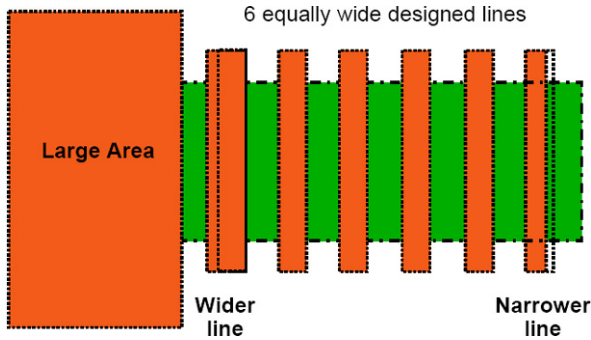
Fig. 11.5 The proximity effect: lines with large neighboring structures grow in size, while lines next to open space shrink

Lithography critically depends on the flatness of the wafer surface. Damascene wiring technology has been developed to avoid height differences. In a design with non-regular wiring patterns, the design rules will prescribe filling patterns: "tiling". If the design tool is allowed to automatically generate tiling patterns, undesired side effects may occur. It is clear that the proximity of a tiling structure will affect capacitor ratios. Also stress patterns and thickness variations can occur. A safe approach is to define and position the tiling patterns during the lay-out phase by hand.

11.3.3 Proximity Effects

Figure 11.5 visualizes proximity effects on a group of lines. The proximity effect is caused by the diffused light from neighboring fields. The line width in the open field will become narrower. Large neighboring structures cause lines to expand. In a precision lay-out dummy structures are placed at distances up to 20–40 μm. Proximity effects can be caused by lithography, but also by depletion of etch liquids or gases.

In 65-nm technologies it is practically impossible to define minimum width lines with acceptable tolerance at random positions. In order to create minimum width patterns pre-distortion is applied to the mask in the form of optical proximity correction (OPC). An example of dimensional deformations of an advanced lithographic tool is visible in the lithography simulation of Fig. 11.6.

Patterns in one layer may sometimes affect the patterning in other layers. During the spinning of the resist fluid, resist may accumulate against altitude differences of previous layers on a partly processed wafer. This results in circular gradients and is therefore often not easily recognized as a systematic offset.

An ion beam from an implanter that is perfectly aligned with the crystallographic orientation of the substrate will result in ions to penetrate deeply into the lattice. This effect is called "channeling". During the ion-implantation steps in

Fig. 11.6 The tips of the structures are enhanced by the OPC tool in this lithography simulation
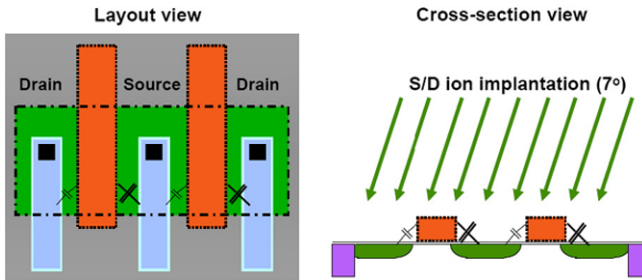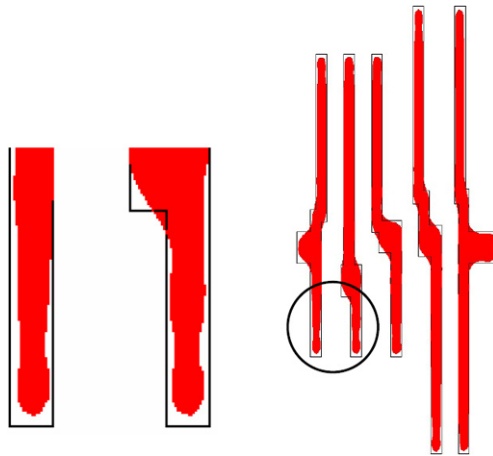
Fig. 11.7 The source and drain diffusions are implanted under an angle. This causes asymmetry for drain and sources

older processes the implantation beam is tilted by some 5–8°. As a result of this non-perpendicular implantation, channeling is avoided but source and drain diffusions will be asymmetrical. The diffusion on one side may extend further underneath the gate than on the other side, see Fig. 11.7. In order to prevent inequalities in currents or overlap capacitors, the directions in which the MOS currents flow must be chosen to run parallel, and not rotated or anti-parallel. In integrated circuit manufacturing there are more processing steps that can cause similar asymmetries.

The well-proximity effect in Fig. 11.8 has no direct relation with lithography. This effect is believed to be caused during the implantation of the well in the substrate. The implanted ions interact with the photo-resist boundary and cause a horizontal gradient in the well implantation dose. Variations ranging from 1 µm [249] to 2 µm [250] have been reported. It is advisable to use an overlap well beyond the minimum lay-out design rule for implantation masks, where possible.
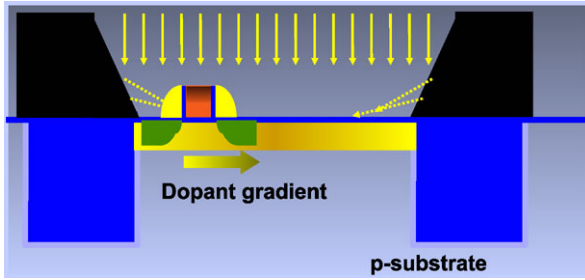
Fig. 11.8 The well-proximity effect occurs during the well implant. The drawn transistor is fabricated afterwards, but resides in a well with a horizontal doping gradient [249, 250]

Fig. 11.9 Temperature difference can be visualized by means of a liquid crystal technique. Here the track-and-hold circuit generates the heat, which spreads out via the wiring into the ladders of this analog-to-digital converter
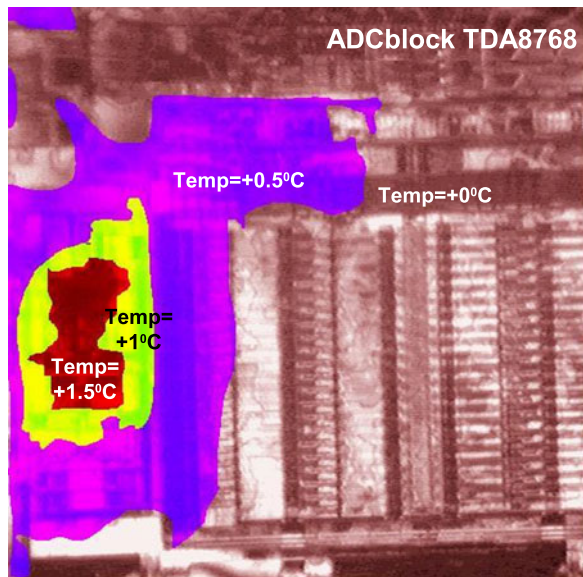
11.3.4 Temperature Gradients

Gradients can exist in doping, resistivity, and layer thickness. Although structures tend to decrease in dimensions, situations may occur in which equality is required in a distance in the order of 1 mm. In older processes CMOS thresholds were observed to deviate up to 5 mV over this distance. Resistivity gradients can reach a relative error of several percent over this distance. In advanced processes ($<0.18\ \mu$m) process control is much better and technology gradients are hardly present.

The temperature distribution across a circuit in operation can be a reason for parameter gradients, see Fig. 11.9. In a System-on-Chip the different blocks show
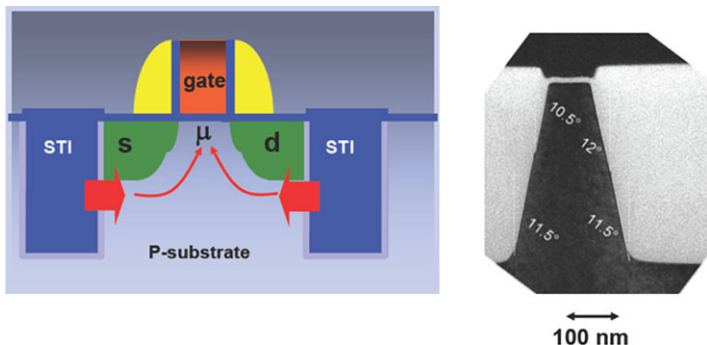
Fig. 11.10 Stress in integrated circuits is caused by thermal expansion of various materials. The blocks of shallow trench isolation (STI) create a considerable stress in the active area of the transistor, causing current variations up to 10%. The photo shows a cross-section, where the slopes of the STI are indicated. *Courtesy*: C. Detcheverry, NXP

a great variety of power dissipations. On-chip memories shows a relatively low power density. Output drivers, transmitters, high-speed processors, power regulators (LDO) and input stages (LNA) may consume much more. In larger chips (50–100 mm^2) with 2–5 Watt power dissipation temperature differences up to 20 °C can occur. Local temperature gradients of 2–5 °C/mm are possible. With threshold-voltage and diode temperature coefficients of −2 mV/°C, an offset in the order of several millivolts is well possible. It is therefore important to consider the power distribution when temperature sensitive circuits and heat sources are placed on the same die. The circuit designer can position the critical devices on equal-temperature lines or use cross-coupling.

11.3.5 Offset Caused by Stress

During the fabrication of a circuit, layers are deposited on the substrate. These layers are built of different materials with different thermal expansion coefficients. After the devices have been cooled to room temperature the differences in thermal expansion coefficients lead to mechanical stress. This stress can result in a positive or negative change of the local parameter value. A secondary effect is that the global stress pattern is locally affected by neighboring lay-out features, causing stress modulation in the surrounding components. In high-precision analog design this will lead to undesired systematic offsets.

In resistors and transistors stress predominantly impacts the mobility of carriers. Tensile stress increases the electron mobility and reduces the hole mobility. Compressive stress works opposite and is used to enhance mobility in PMOS transistors. An effect on threshold voltage does occur as well, see Fig. 11.12. Some major causes for stress are:

Fig. 11.11 A transistor is placed asymmetrically with respect to the reference device. The edges between active and STI inflict stress in the channel region, however due to the asymmetry the effect of the stress is different for both devices
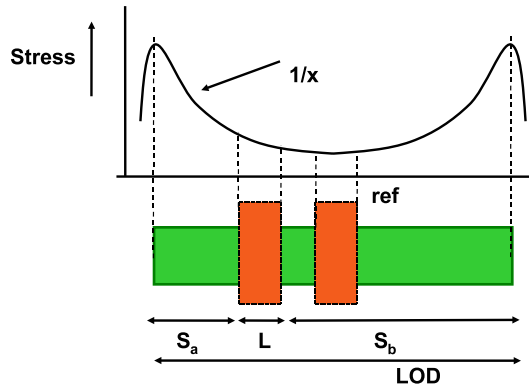
- The presence of the die boundary close to sensitive devices. Typically a distance of several hundreds of microns is safe to avoid die-edge related stress effects.
- Plastic packages are molded around the die. After cooling these packages create rather severe mechanical stress. Special gels or polyimide coatings on top of the die relieve this problem. A simple way to detect package related stress is to heat the package with a hot air flow.
- In modern isolation techniques a trench is etched in the substrate and silicon dioxide is deposited and planarized: Shallow Trench Isolation (STI), Fig. 11.10. The different thermal coefficients of silicon dioxide and the substrate cause mechanical stress. A transistor in the substrate with its diffusion areas is surrounded by STI and experiences this stress. This effect is called "STI-stress" or "LOD-stress" (Length Of Diffusion).
- In a densely packed circuit, there will be many active-to-STI edges. Unrelated edges that are close to a device will influence the stress pattern, this effect is known as "OD-to-OD" stress or "OD-spacing" effect. The Oxide-Definition (OD) mask defines the inverse of the active area.
- In advanced processes an etch-stop layer can be used to create stress that increases the current in a high-performance MOS transistor. The proximity of other structures will influence this effect, called "PS-to-PS" stress or "poly-space" effect.
- Aluminum has a different thermal expansion coefficient from the dielectric that surrounds it. Asymmetries in wiring will result in stress related offset.

Stress related to Shallow Trench Isolation has been subject of various studies [250, 252–254]. At temperatures of over 1000 °C areas of silicon dioxide are formed in the substrate. At that temperature the structure is free of stress or relaxed. After cooling the difference in thermal expansion coefficient causes a compressive mechanical stress that peaks at the border of the active area and the STI formation [252]. The mechanical stress deforms the lattice and causes the mobility in the transistors to vary.
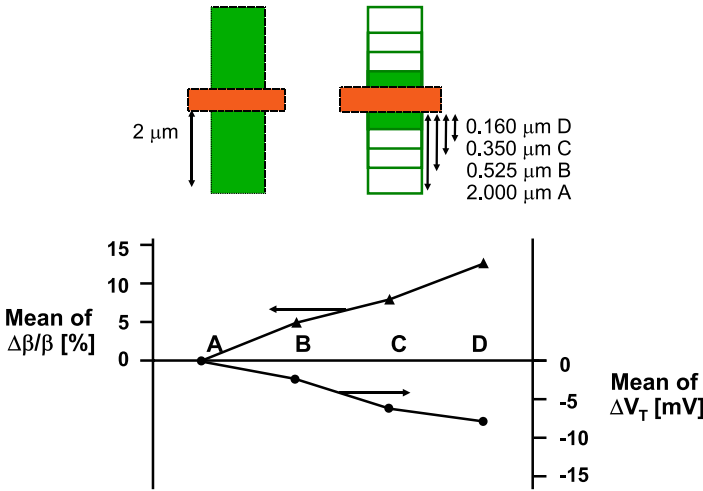
Fig. 11.12 An experiment shows the influence of the STI edge on the drain current and threshold voltage, [254]. *Top*: a 65-nm 2/0.5 μm NMOS reference transistor is designed with the STI edge of the source and drain at 2.0 μm. A second device has a similar STI distance (A) or at 0.525 (B), 0.35 (C), and 0.16 μm (D)

Figure 11.11 shows two transistors that are placed asymmetrically with respect to the edges of the active area and the STI isolation. The mobility of the device μ_{eff} is usually modeled with an inverse distance model [252]:

$$S = \frac{1}{S_a + L/2} + \frac{1}{S_b + L/2}$$

$$\mu_{\text{eff}} = \frac{1 + K_{s,\mu} S}{1 + K_{s,\mu} S_{\text{ref}}} \mu_{\text{ref}} \tag{11.2}$$

$$V_T = V_{T,\text{ref}} + K_{s,VT}(S - S_{\text{ref}})$$

where the suffix "ref" indicates the reference device and K_s is a process parameter. The parameter S reflects the distance relation. The stress from the STI edges affects also the doping profile under the transistor. This phenomenon is not fully understood, however the idea that a lattice deformation changes the diffusion of the doping atoms, seems plausible.

In [254] an experiment is reported where the STI-to-active edge of the source and drain is varied. Figure 11.12 shows current factor deviations up to 12% and threshold voltage variations of 10 mV. These observations are technology specific but similar effects are reported in [250, 252].

Next to the effect of the STI-to-active edge of the transistor itself, also neighboring edges will modulate the mechanical stress pattern. This may be less relevant in
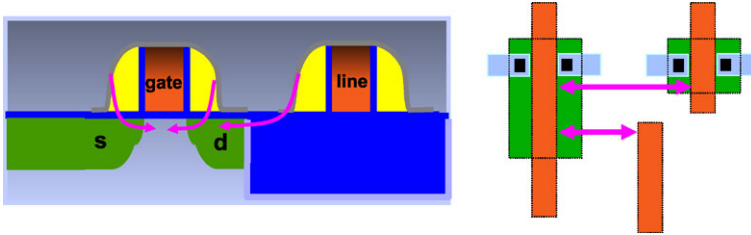
Fig. 11.13 The poly-space effect is caused by the etch-stop layer

Fig. 11.14 An aluminum wire placed at a certain distance of the emitter causes current deviations, that can be measured up to 40 μm
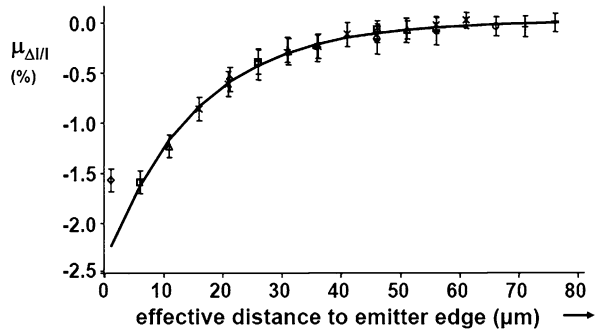
a digital circuit, however, in precision analog design these effects must be taken into account. The generalized model uses:

$$S = \sum_{i=1}^{n} \frac{\pm W_i}{2W} \frac{1}{S_i + L/2} \tag{11.3}$$

where the first term relates the width of the edge to the gate-width and the $\pm$ sign to an STI-active edge or an active-STI edge.

The poly-space effect [255] in Fig. 11.13 is caused by the stress related to the use of the etch-stop layer. This layer applies compressive or tensile stress to the transistor in order to increase the current drive capabilities. Also stress coming from neighboring devices will influence the stress pattern under the critical transistors. Again an inverse distance model applies.

The wiring pattern causes transistors to show offset. The coverage of transistors with metal layers can lead to mobility reduction due to incomplete annealing of interface states [257] and to variations in the stress pattern. When a wiring pattern is placed at different spacings on one side of a bipolar pair [258], the resulting current variation is in the order of 1%. Even wiring on top level (e.g. tiling patterns) can cause stress [259].

Figure 11.14 suggests that the impact of the wiring pattern halves for every 10 μm distance up to 40 μm. This example again shows that a regular, symmetrical and consistent lay-out is required for analog circuits that should yield offsets below 1%.

Table 11.6 Guide lines for the design of equal components

1	Equal components are of the same material, have the same form and dimensions
2	The potentials, temperatures, pressures and other environmental factors are identical
3	Currents in components run in parallel, not anti-parallel or perpendicular
4	Only use cross-coupled structures if there is a clear reason for that (e.g. temperature gradient)
5	Keep wiring away from the components
6	Use star-connected wiring for power, clock and signal
7	Apply symmetrical dummy structures up to 20 μm away from sensitive structures
8	Keep supply and ground wiring together and take care that no other circuits dump their return current in a ground line
9	Check on voltage drops in power lines
10	Stay 200 μm away from the die edges to reduce stress from packaging
11	Tiling patterns are automatically inserted and can lead to unpredictable coupling, isolation thickness variations, stress. Do not switch of the tiling pattern, but define one that causes no problems

11.3.6 Offset Mitigation

Systematic deviations are mostly identified during the initial design trials of a process. Sometimes optimizations in an established process flow still lead to unexpected deviations. Measures for overcoming them are found in an extensive study of the fabrication process. As dimensions shrink, some offset effects are incorporated in the device model descriptions that quantify the impact [36]. Despite the complex nature of some variations, a number of guidelines can be formulated to minimize the effect of these offset causes, see Table 11.6 [260].

Common centroid structures are used to reduce the gradient effects [261]. Applying a common centroid geometry is not trivial, an asymmetry in the wiring scheme can easily cause more problems than are solved, see Fig. 11.15. On the left side is a standard cross-coupled differential pair with common source. The right side shows in-line common centroid structures. The lower structure is exactly common centroid with the disadvantage that the outer devices need dummy structures to compensate for their lack of neighbors. The upper-right structure needs no dummy structures, at the cost of a small spacing between the common centroid points.

By following these design guidelines the effects of systematic errors can be significantly reduced. The obtainable limits in a production environment differ per component and can be summarized as:

- Resistors: The absolute value suffers from process variations and temperature. Yet, the relative accuracy of matched resistors is in the order of 10^{-3}–10^{-4} depending on type (diffused is better than polysilicon), size, and environment. Sensitive to substrate coupling.
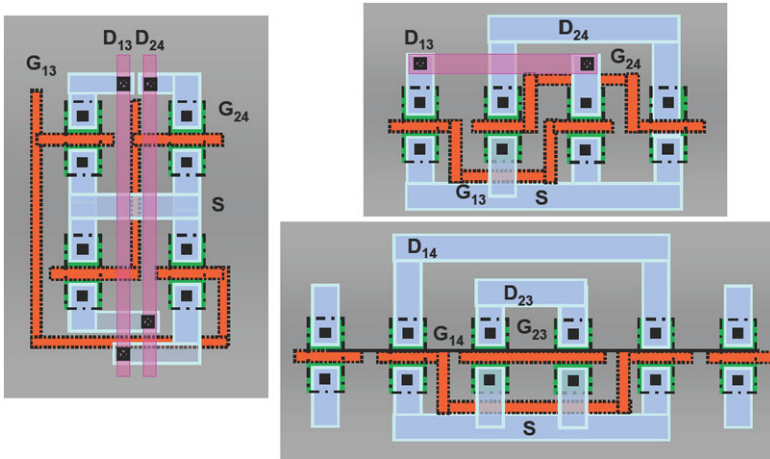
Fig. 11.15 Some common centroid arrangements

- Capacitors: The absolute value is usually well-defined in a double poly-silicon or MIM process. Also horizontally arranged capacitors such as fringe capacitors reach excellent performance. The relative accuracy of capacitors is in the order of 10^{-4} for >1 pF sizes. Minimum usable sizes in design are limited by parasitic elements, relative accuracy or the kT/C noise floor. In the application the net effect of the capacitor is sensitive to different parasitic couplings, which can be mitigated with stray-capacitor insensitive circuit topologies. Often capacitors are seen as a low-power solution, but handling charges requires large peak currents during transfers, so the power of the surrounding circuits limits the low-power ambition of capacitor based circuit solutions.
- Transistors: The current is sensitive to temperature, process spread and variability effects. The relative accuracy in current is in the order of 10^{-3}. Back-gate modulation by substrate noise and $1/f$ noise must be considered.
- Time: with a more or less fixed timing variation or jitter (1–5 ps$_{rms}$), the best accuracy is achieved for low signal bandwidths.

11.4 Random Matching

11.4.1 Random Fluctuations in Devices

Parameters that reflect the behavior of devices, are the result of the combination of a large number of microscopic processes. The conductivity of resistors and transistors and the capacitance of capacitors is built up of a large number of single events: e.g. the presence of ions in the conduction path, the local distances between the polysilicon grains that form capacitor plates, etc. Already in 1961 W. Shockley recognized that these atomic processes can lead to random fluctuations of device parameters.

Fig. 11.16 Definition of the
area function $h(x, y)$
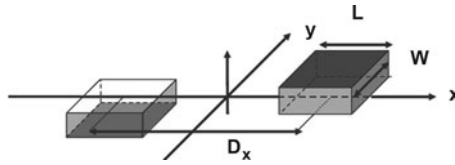
Various authors have investigated random effects in specific structures: capacitors
[23–25, 262, 263], resistors [251, 264], MOS transistors [265–268] and bipolar de-
vices [269].

In a general parameter fluctuation model [28] a parameter P describes some
physical property of a device. P is composed of a deterministic and random varying
function resulting in varying values of P at different coordinate pairs (x, y) on the
wafer. The average value of the parameter over any area is given by the weighted
integral of $P(x, y)$ over this area. The actual difference between two parameters P
of two identically sized areas at coordinates (x_1, y_1) and (x_2, y_2) is:

$$\Delta P(x_{12}, y_{12}) = P(x_1, y_1) - P(x_2, y_2)$$
$$= \frac{1}{\text{area}} \left[\int\int_{\text{area}(x1,y1)} P(x', y')\, dx'\, dy' \right.$$
$$\left. - \int\int_{\text{area}(x2,y2)} P(x', y')\, dx'\, dy' \right] \qquad (11.4)$$

This integral can be interpreted as the convolution of double box functions formed
by the integral boundaries (or the device dimensions) with the "mismatch source"
function $P(x, y)$. In the Fourier domain the convolution transforms in a multiplica-
tion and allows separating the geometry-dependent part from the mismatch source:

$$\Delta\mathcal{P}(\omega_x, \omega_y) = \mathcal{G}(\omega_x, \omega_y)\mathcal{P}(\omega_x, \omega_y) \qquad (11.5)$$

Now the mismatch generating process $\mathcal{P}(\omega_x, \omega_y)$ can be regarded as a source that
generates spatial frequencies that are spatially filtered by the device geometry de-
pendence function $\mathcal{G}(\omega_x, \omega_y)$. These two components are analyzed separately.

The geometry function as shown in Fig. 11.16 for a pair of rectangular devices
with area WL is defined as:

$$h(x, y) = \begin{cases} \dfrac{1}{WL}, & (D_x/2 - L/2) < x < (D_x/2 + L/2),\ -W/2 < y < W/2 \\[2mm] \dfrac{-1}{WL}, & (-D_x/2 - L/2) < x < (-D_x/2 + L/2),\ -W/2 < y < W/2 \\[2mm] 0, & \text{elsewhere} \end{cases}$$

$$(11.6)$$

For convenience it has been assumed that both areas are at a distance D_x along the
x-axis. Some mathematical manipulation (see the appendix) results in a geometry
function for the difference in paired transistor parameters Fig. 11.16:

$$\mathcal{G}(\omega_x, \omega_y) = \frac{\sin(\omega_x L/2)}{\omega_x L/2} \frac{\sin(\omega_y W/2)}{\omega_y W/2} [2\sin(\omega_x D_x/2)] \qquad (11.7)$$

This geometry function has a zero value for $\omega_x = 0$, thereby eliminating the global value of the parameter from the calculations. The geometry functions for other geometries are found in the same way, e.g. a cross-coupled group of four transistors as in Fig. 11.14 (left), has a geometry function where the last term in brackets in (11.7) is replaced by $[\cos(\omega_x D_x/2) - \cos(\omega_y D_y/2)]$.

After this analysis of the geometry dependence the specification of the random contribution to $P(x, y)$ or $\mathcal{P}(\omega_x, \omega_y)$ has to be formulated.

Two classes of distinct physical mismatch causes are considered here as examples of local and global variations. Every mismatch-generating physical process that fulfills the mathematical properties of these classes, results in a similar behavior at the level of mismatching transistor parameters.

The first class is a random process on a parameter P characterized by:

- The total mismatch of parameter P is composed of mutually independent events of the mismatch generating process.
- The effects on the parameter are so small that the contributions to the parameter are linear.
- The correlation distance between the events is small compared to the size of the device (basically saying that boundary effects can be ignored).

Statistically the values of parameter ΔP are described by a Poisson process that converges for a large number of events to a Gaussian distribution with zero mean. In the frequency domain this type of spatial random processes is modeled as spatial "white noise". A process with these properties is described in the Fourier domain as a constant value for all spatial frequencies.

The combination of this mismatch generating process and the paired-transistor geometry function results in a description of the power or the variance of the difference ΔP in parameter P between the two instances [28]:

$$\sigma_{\Delta P}^2 = \frac{A_P^2}{WL} \tag{11.8}$$

A_P is the area proportionality constant for parameter ΔP. The proportionality constant can be measured and used to predict the mismatch variance of a circuit.

Many known processes that cause mismatching parameters fulfill in first order the above-mentioned mathematical constraints: distribution of ion-implanted, diffused or substrate ions (Random Dopant Fluctuations), local mobility fluctuations, polysilicon and oxide granularity, oxide charges, etc.

Equation (11.8) describes the statistical properties of area *averaged* or *relative* values of parameter P. The *absolute* number of events (like the charge in a MOS channel) is proportional to the area of the device WL. Therefore differences in the sums of atomic effects obey a Gaussian distribution with zero mean and

$$\sigma_{\Delta P} = A_P \sqrt{WL} \tag{11.9}$$

In analyzing statistical effects it is important to consider whether the parameter is an absolute quantity (e.g. total amount of ions) or is relative (averaged) to the device area (e.g. threshold voltage).

Apart from theoretical derivations and measurements, 3-D device simulations are applied to analyze the impact of random dopants, line-edge roughness and polysilicon granularity in advanced processes [270].

The assumption of a short correlation distance in the above process implies that no relation exists between matching and the distance D_x between two transistors. Wafer maps show, however, all sorts of parameter-value distributions that originate from wafer fabrication. This effect is observed in Fig. 11.4. The extremes of a parameter are often located in the middle and at the edges of a wafer. A rough approximation is obtained by dividing the maximum process spread e.g. 100 mV for a threshold spread or 10% for the current factor by half of the wafer diameter. Gradients in the order of 1 mV/mm result.

Another approach includes distance effects in the stochastic model. This second class of mismatch effects is a deterministic process but, as the original placement of dies on a wafer is unknown after packaging, the effect of this parameter value distribution is modeled as an *additional* stochastic process with a long correlation distance. In the Fourier domain this effect is described as a fixed low-frequency contribution with a spatial frequency inversely proportional to the wafer diameter. The representation of parameter fluctuations in the Fourier domain allows easy determination of the power contents, which in turn can be interpreted as the variance (σ^2) of the stochastic parameter [28].

$$\sigma^2_{\Delta P} = \frac{A_P^2}{WL} + S_P^2 D_x^2 \tag{11.10}$$

S_P describes the variation of parameter P with the spacing. Note that the probability density function of e.g. a circular parameter pattern is not Gaussian and only randomized because the link to the position on the wafer will be lost after packaging.[2]

For a group of four cross-coupled transistors is found in a similar way:

$$\sigma^2_{\Delta P} \approx \frac{A_P^2}{2WL} + S_P^2 D_x^2 \frac{D_y^2}{\text{wafer diameter}^2} \tag{11.11}$$

The effect of the doubled gate area and the reduction of the linear components in the gradient is obvious.

11.4.2 MOS Threshold Mismatch

The threshold voltage is given by:

$$V_T - V_{FB} = \frac{Q_B}{C_{ox}} = \frac{q N_x z_d}{C_{ox}} = \frac{\sqrt{2q\epsilon N_x \phi_b}}{C_{ox}} \tag{11.12}$$

---

[2]The importance of this model extension is more in signaling a potential problem in the process than in accurately modeling a phenomenon.

Fig. 11.17 A cross-section through a MOS transistor indicating the depletion region
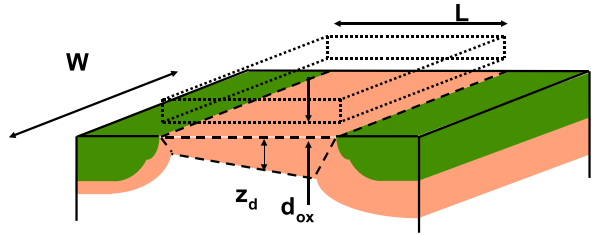
where ϵ is the permittivity, N_x the dope concentration and ϕ_b the Fermi potential. If the depletion region of a transistor (see Fig. 11.17) is defined by its width W, length L and a depletion region depth $z_d = \sqrt{2\epsilon\phi_b/qN_x}$, then the volume of the depletion region is (in first order): WLz_d. Different impurities are active in this region, with concentrations around 10^{16}–10^{18} cm^{-3}. N_x contains acceptor and donor ions from the intrinsic substrate dope, the well, threshold adjust, punch-through implantation.[3] In the variance analysis it is important to note that the total number of charged ions and other charge contributions must be considered, not the net resulting charge.

The variance in the number of atoms is now approximated by Poisson statistics:

$$\sigma_c^2 = \mu_c \quad \mu_c = WLz_dN_x \quad \Rightarrow \quad \sigma_c = \sqrt{WLz_dN_x} \qquad (11.13)$$

The threshold variance can now be derived from (11.12) by considering that the variance of a threshold voltage equals the variance of the charge in the depletion region multiplied by the partial derivative of the threshold versus the charge

$$\sigma_{\text{single VT}} = \sigma_c \frac{\partial(V_T)}{\partial(WLz_dN_x)} \qquad (11.14)$$

As matching usually occurs between pairs of transistors, the variance of the *difference between two transistors* is [28, 256]:

$$\sigma_{\Delta VT} = \sqrt{2}\sigma_{\text{single VT}} = \frac{qd_{ox}\sqrt{2N_xz_d}}{\epsilon_{ox}\sqrt{WL}} = \frac{A_{VT}}{\sqrt{WL}} \propto \frac{d_{ox}\sqrt[4]{N_x}}{\sqrt{WL}} \qquad (11.15)$$

This function is commonly depicted as a linear relation between $\sigma_{\Delta VT}$ and $1/\sqrt{\text{area}}$. Figure 11.18 shows an example of the measured dependence for $\sigma_{\Delta VT}$ versus $1/\sqrt{\text{area}}$. The slope of the line equals the parameter A_{VT}. In this plot the lay-out dimensions were used. For the smallest sizes the effective gate area is smaller due to underdiffusion and channel encroachment. In order to test the hypothesis that depletion charge is the dominant factor in threshold matching, Table 11.7 compares the A_{VT} coefficients as measured and as calculated using the above formula. The quantity N_xz_d was derived from process simulation which was tuned with accurate C/V measurements [246]. In these relatively straight-forward 0.6 μm and 0.8 μm CMOS processes the fit is good for three out of four A_{VT} coefficients. The deviation

---

[3]For ease of understanding only a uniformly distributed dopant is assumed, more complicated distributions must be numerically evaluated.
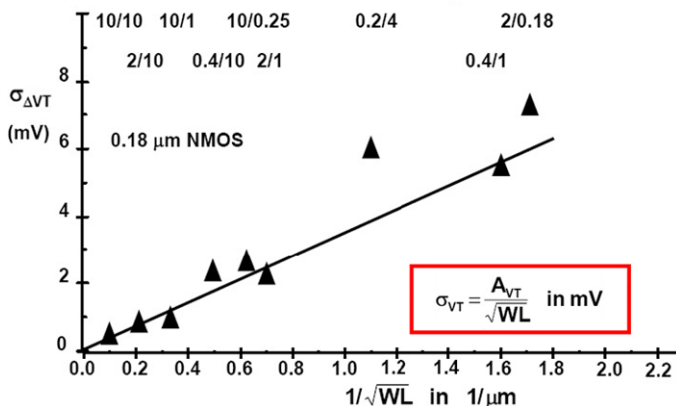
Fig. 11.18 The standard deviation of the NMOS threshold and the relative current factor versus the inverse square root of the area, for a 0.18 μm CMOS process

Table 11.7 Comparison of measured and calculated threshold mismatch coefficients

	A_{VT} measured	A_{VT} calculated
0.8 μm data:		
NMOST	10.7 mV μm	10.6 mV μm
PMOST	18.0 mV μm	18.6 mV μm
0.6 μm data:		
NMOST	11.0 mV μm	7.4 mV μm
PMOST	8.5 mV μm	8.6 mV μm

of the 0.6 μm NMOST shows that other factors except random dopant fluctuation play a role. Such increases can be attributed to insufficient annealing of interface states [257].

The basic threshold-voltage mismatch model has been extended by various authors. More geometry dependence factors can be included to address deep-submicron effects [271]. A fundamental limit for dopant fluctuation related mismatch was derived in [272]. Andricciola and Tuinhout [273] show that the threshold-voltage mismatch coefficients are not affected by temperature. Work reported in [254] indicates that there is no relation between deterministic variations and random dopant fluctuations.

In deep sub-micron processes the short channel effects in the channel are controlled by means of "halo" or "pocket" implants. These implants are self-aligned with the gate stack and can introduce some significant variations in the local doping profiles. Next to their own variation, the self-aligned feature prints any line-edge roughness in the doping profile. The pocket implants defy the uniform dopant hy-

pothesis for the calculation of the threshold mismatch of (11.8). An additional term can be included for the variation due to the pocket implant:

$$\sigma_{\Delta VT}^2 = \frac{A_{VT}^2}{WL} + \frac{B_{VT}^2}{f(W, L)} \tag{11.16}$$

where the function $f(W, L)$ of the width and length still needs to be established.

11.4.3 Current Mismatch in Strong and Weak Inversion

The matching properties of the current factor are derived by examining the mutually independent components W, L, μ and C_{ox}:

$$\frac{\sigma_{\Delta \beta}^2}{\beta^2} = \frac{\sigma^2(W)}{W^2} + \frac{\sigma^2(L)}{L^2} + \frac{\sigma^2(C_{ox})}{C_{ox}^2} + \frac{\sigma^2(\mu_n)}{\mu_n^2} \tag{11.17}$$

The mismatch-generating processes for the gate oxide and the mobility are treated in accordance with (11.8). The variations in W and L originate from line-edge roughness. The analysis of edge-roughness is a one-dimensional variant of the analysis in the previous section and leads to $\sigma^2(L) \propto 1/W$ and $\sigma^2(W) \propto 1/L$.

$$\frac{\sigma_{\Delta \beta}^2}{\beta^2} = \frac{A_W^2}{W^2 L} + \frac{A_L^2}{WL^2} + \frac{A_\mu^2}{WL} + \frac{A_{C_{ox}}^2}{WL} \approx \frac{A_\beta^2}{WL} \tag{11.18}$$

where A_W, A_L, A_μ and $A_{C_{ox}}$ are process-related constants.

A significant contribution of gate-oxide thickness variation would lead to a negative correlation between the threshold voltage mismatch and the current factor mismatch. Such a correlation is generally not observed. If W and L are large enough the respective contributions will also disappear. At gate-lengths below 65-nm, simulations [270] indicate some role for edge roughness. This role is in measurements hard to identify in the presence of large threshold mismatch. In [28] it was assumed that the matching of the current factor is determined by local variations of the mobility. Many experiments show that mobility affecting measures (e.g. placing the devices under an angle) indeed lead to a strong increase in current factor mismatch. The relative mismatch in the current factor can be approximated by the inverse-area description as seen in the last part of (11.18).

In contrast to the threshold random fluctuation the absolute current factor variation is a function of temperature. However, the relative current factor mismatch as formulated in (11.18) is according to [273] much less sensitive to temperature.

Considering only the threshold and current factor variations,[4] the variance of the difference in drain currents ΔI between two equally sized MOS devices can be calculated. Using the generalized statistical method described in [10]:

$$\left(\frac{\sigma_{\Delta I}}{I} \right)^2 = \left(\frac{dI}{dV_T} \right)^2 \sigma_{\Delta VT}^2 + \left(\frac{dI}{d\beta} \right)^2 \sigma_{\Delta \beta}^2 \tag{11.19}$$

---

[4]The contribution of mobility reduction factor θ is next in line.
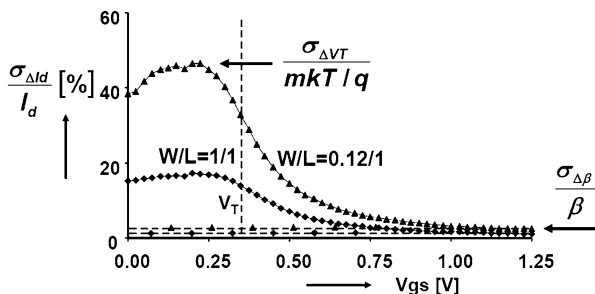
Fig. 11.19 The relative current mismatch for two 65-nm technology transistor geometries swept over the full voltage range. Measurements by N. Wils

For strong inversion this equation is combined with the simple square-law current model:

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{2\sigma_{\Delta VT}}{V_{GS} - V_T}\right)^2 + \left(\frac{\sigma_{\Delta\beta}}{\beta}\right)^2 \qquad (11.20)$$

Equation (11.20) suggests an infinite current mismatch if the gate voltage equals the threshold voltage, however, in that mode of operation the weak inversion model is applicable and levels off the maximum current mismatch. In weak inversion the current is modeled as an exponential function. Due to the low current level, the current factor mismatch is of less importance:

$$\left(\frac{\sigma_{\Delta I}}{I}\right)^2 = \left(\frac{q\sigma_{\Delta VT}}{mkT}\right)^2 \qquad (11.21)$$

Although the threshold mismatch itself shows no temperature dependence, the current mismatch certainly will vary due to the denominator term. Figure 11.19 shows an example of the current mismatch relative to the drain current. At high gate-source voltages the current factor mismatch in the strong inversion equation (11.20) dominates. At lower gate-source voltages the threshold-related term in this equation gains importance.

In a 65-nm process $A_{VT} = 3.5$ mV µm. So $\sigma_{VT} = 10$ mV for a 0.12/1 device. Equation (11.21) predicts a relative current mismatch of 40%, which is confirmed by the measurement in Fig. 11.19. Except for extremely low current densities where the depletion layer width is shrinking, the observation can be made that the same value for $\sigma_{\Delta VT}$ applies for both the strong and the weak inversion regime. This example shows that the operating regime where the mismatch parameters of the transistor are extracted, has a marginal effect on the accuracy[5] of the prediction in other regimes, as confirmed in e.g. [267, 268, 274].

The standard deviation of the current difference between the 0.12/1 µm transistors reaches ≈40% of the current in the sub-threshold regime. Obviously this would

---

[5]Accuracy means that the standard deviation of a circuit parameter is within 10% of the prediction, see Sect. 11.4.6.

Fig. 11.20 Development of the threshold mismatch factor A_{VT} for NMOS and PMOS transistors as a function of the nominal oxide thickness of the process. The processes span 65-nm up to 1.6 μm CMOS
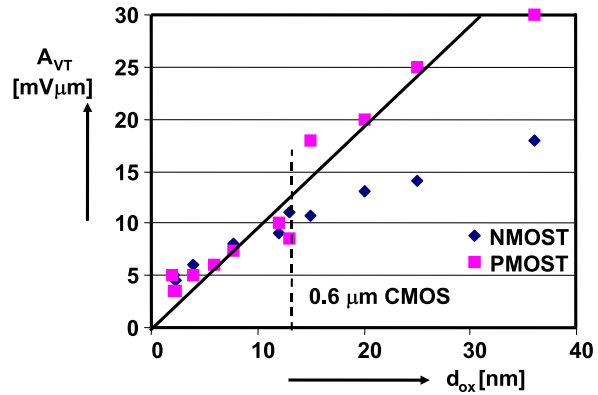
imply a reverse drain current below -2.5σ of a Gaussian distribution. At these levels of mismatch the assumption that a small Gaussian distributed threshold mismatch voltage will turn into a Gaussian approximation of the mismatch current is not valid anymore. The probability density function of the current mismatch needs here a log-normal distribution.

11.4.4 Mismatch for Various Processes

In Fig. 11.20 the threshold mismatch coefficient A_{VT} is plotted as a function of the nominal oxide thickness. As predicted by (11.15) the mismatch coefficient becomes lower for thinner gate-oxide thickness. Of course many more changes in the device architecture took place, still the oxide-thickness seems to be the dominant parameter. The large PMOS transistor coefficients for >0.6 μm CMOS generations is caused by the compensating implants: the N- and PMOS transistor threshold adjust and n-well implants. The quantity $N_x = (N_a + N_d)$ is relevant for matching, while the net value $(N_a - N_d)$ determines the threshold in the PMOS transistor. Beyond the 0.6 μm node a twin well construction with a dedicated well implant for the PMOS transistor is used that avoids compensating charges.

In Fig. 11.20 the diagonal line indicates an A_{VT} factor increase of 1 mV μm for every nm of gate insulator thickness. This line is a first order estimate of what a well-engineered process should bring.

Over the same process range the current mismatch factor A_β varies between 1.2% μm and 2% μm.

Figure 11.21 shows a simulation of the threshold voltage of 200 NMOS and PMOS 90-nm devices in their process corners. Although the corner structure is still visible, it is clear that for small transistors in advanced processes the mismatch is of the same magnitude as the process corner variation. The plot is somewhat misleading as it combines global process corner variation and local mismatch. A simple "root-mean-square" addition of these two variation sources ignores the fundamental difference between the two.
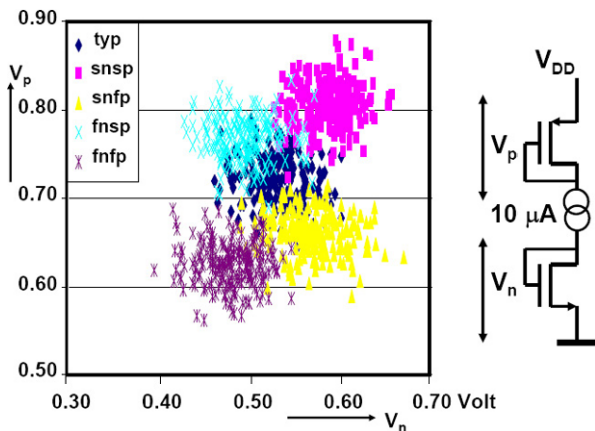
Fig. 11.21 Simulation of 200 0.2/0.1 P- and NMOS transistors in their 90-nm process corners. The notation "snfp" refers to slow NMOS and fast PMOS transistors. The variation due to mismatch is of equal importance as the process variation

Fig. 11.22 Development of power supply voltage and the measured NMOS threshold matching of a transistor 1.5 times the minimum size through various process generations
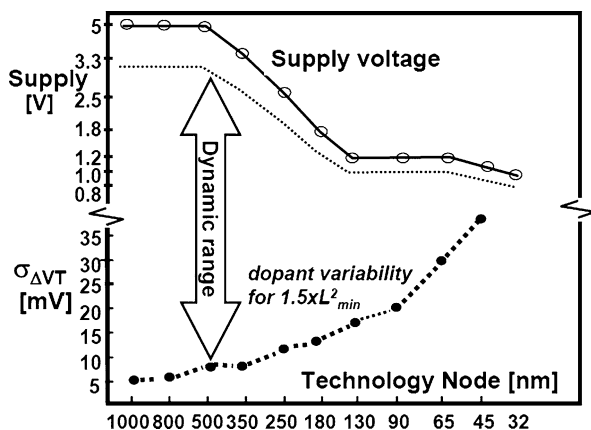
The analysis of matching in various processes allows to compare in Fig. 11.22 the MOS mismatch to the development of power supply voltage. A transistor with a size of $1.5L^2_{min}$ was chosen. During the process development from the 2.5 μm process to the 0.35 μm process both the mismatch and minimum gate-length have reduced. The power supply remained fixed at 5 V for the micron range process generations with a signal swing of 2.5–3 V. Circuits that relied on analog CMOS performance such as analog-to-digital converters could improve their performance in these process generations by not fully following the line-width reduction.

At the 0.35 μm CMOS node the maximum electrical fields in intrinsic transistors were reached for both the vertical gate-oxide field and the lateral field controlling the charge transport. For this reason and in order to reduce power consumption, the power supply voltage was lowered. On the other hand, the need to tailor the internal

Table 11.8 An overview of matching models and value ranges

MOS transistors	$\sigma_{\Delta VT} = \dfrac{A_{VT}}{\sqrt{WL}}$	$A_{VT} = 1 \text{ mV } \mu\text{m/nm}$
MOS transistors	$\dfrac{\sigma_{\Delta\beta}}{\beta} = \dfrac{A_\beta}{\sqrt{WL}}$	$A_\beta = 1\text{--}2\% \ \mu\text{m}$
Bipolar transistors (BJT)	$\sigma_{\Delta Vbe} = \dfrac{A_{Vbe}}{\sqrt{WL}}$	$A_{be} = 0.3 \text{ mV } \mu\text{m}$
Diffused/poly resistors	$\dfrac{\sigma_{\Delta R}}{R} = \dfrac{A_R}{\sqrt{WL}}$	$A_R = 0.5/5\% \ \mu\text{m}$
Plate, fringe capacitors	$\dfrac{\sigma_{\Delta C}}{C} = \dfrac{A_R}{\sqrt{C}}$	$A_C = 0.3\%\sqrt{\text{fF}}$

fields in the transistor, has led to less uniform and higher implantation channel dope. As can be expected from the theoretical background, the slower scaling of the gate-oxide thickness made that the threshold matching factor A_{VT} stopped decreasing. This became especially pronounced in 65–32 nm technologies, where pocket implants create an additional mismatch source. Shrinking the area of analog blocks in sub-micron processes is clearly an important economical issue, but in combination with a rising mismatch coefficient this will lead to lower performance. The reduction in the signal-to-matching coefficient ratio in sub-micron CMOS will necessitate changes in the system, design or technology. In order to maintain high quality signal processing, some enhancements to the standard processes are needed, such as the use of high-voltage devices or precision resistors and capacitors.

11.4.5 Application to Other Components

In Table 11.8 matching parameters of various components are listed. In the previous paragraphs the mismatch parameters for the MOS transistor have been extensively discussed.

The behavior of the bipolar transistor is dominated by the number of dopants in the base that are not depleted. The fluctuation of this number, comparable to the fluctuation of the charge in the depletion layer of the MOS transistor, causes the base-emitter voltages between two bipolar devices to mismatch. Therefore a variance can be defined for ΔV_{be}. In [269] various experiments have confirmed the validity of this mismatch model.

Resistors for high precision analog design are formed by polysilicon or diffused n- or p-doped areas. In advanced processes these layers are covered with a silicide layer to lower their impedance to the 2–5 $\Omega/\square$ level. A special mask is applied to prevent the deposition of silicide in order to obtain sheet resistances of 50–500 $\Omega/\square$. Polysilicon resistors are enclosed by silicon dioxide acting as a thermal isolator, see Sect. 2.2.1. Dissipated heat may destroy such a resistor, a small amount of dissipated heat will affect the grain boundary structure and lead to a resistance value shift after cooling.

Resistors suffer from area related mismatch and from edge roughness. The general description for the relative mismatch is therefore:

$$\frac{\sigma_{\Delta R}^2}{R^2} = \frac{A_W^2}{W^2 L} + \frac{A_L^2}{W L^2} + \frac{A_\mu^2}{W L} \approx \frac{A_R^2}{W L} \tag{11.22}$$

The mobility variation mechanism includes impurity/doping scatter and in the case of polysilicon resistors also includes grain boundary disturbance. The last mechanism is important as the mismatch coefficient A_R increases to 5% µm, while the diffused resistors allow to use 0.5% µm. An additional factor in resistor design are the head-end connections [251, 264]. These connections introduce edge roughness, but more importantly also impose stress on the entire structure. A careful design or an additional margin is needed.

In [24] mismatch of capacitors was attributed to edge effects and to area non-uniformities. The first error would result in a line variation $\sigma_{\Delta C} \propto \sqrt{L},\ \sqrt{W}$. The area non-uniformities comply to the mathematical conditions of the general model resulting in $\sigma_{\Delta C} \propto \sqrt{WL}$. At small dimensions the edge effects will dominate, but at reasonable sized capacitors > 0.1 pF, the area effects become dominant. The description of capacitor mismatch in Table 11.8 is different from the resistor model. For resistors the W/L ratio determines the value allowing to choose the device area independent of the resistor value. The capacitor value is always proportional to the area WL.

11.4.6 Modeling Remarks

The present model describes random variations of devices by means of two statistical variables on the device-parameter level. This intuitive description allows easy communication between foundries, designers and architects to balance power, area and performance of mismatch-sensitive circuits. The standard deviations of Monte-Carlo simulated circuit parameters and from measured circuits agree within approximately 10%, see e.g. Fig. 6.7. This number is based on quantities of around hundred measured samples of a circuit fabricated in a stable process, without foundry liability margins. A significantly better accuracy requires quadratically more samples and e.g. corrections for the effective channel width and length.

In [28] the mismatch contributions of the back-bias factor and the mobility reduction coefficient have been added. Also other authors [266, 271] have proposed methods to get a closer prediction of mismatch in the various operating modes. One problem is that extra mismatch factors are not easy separable from others, requiring high-precision measurements [275]. E.g. the mismatch in the mobility reduction coefficient θ and the source series resistance R_s are to a certain extent interchangeable: $\theta \Leftrightarrow \beta R_s$.

From a mathematical perspective, the threshold mismatch coefficient in an $I_D - V_{GS}$ plot corresponds to the variation in the back-extrapolated zero crossing of the drain current, while the current factor mismatch coefficient describes the slope

of the curve. The zero-order and first-order mismatch components are absorbed in threshold voltage and current factor mismatch coefficients. Identifying other mismatch effects requires a large statistical effort to realize sufficient significance.

Alternative model approaches concentrate on the physical parameters e.g. mobility in stead of current factor. Michael and Ismail [266] analyzes the individually extracted (BSIM) model parameters of test devices by means of a Principal Component Analysis (PCA). This technique assigns the overall mismatch to the most relevant model parameter and than re-iterates to assign the remaining variation to other model parameters. Depending on the data points, this approach can lead to unexpected cross-correlations with little physical significance. In [276] the measured I-V plots are directly used as the input for a PCA analysis. A large numerical effort seems imperative. The work indicates that just a few principal components sufficiently describe the statistical behavior [276, Fig. 5]. This observation is in line with e.g. [277].

11.5 Consequences for Design

A designer has several options to deal with global variations in his circuit. Next to rigorous simulation, the circuit itself can be designed in a way that effects of global variations are minimized. These techniques include differential design and replica-biasing. Unfortunately these methods are ineffective for battling local variations. Here the designer has to rely on statistical simulations.

The threshold and current factor based model that was defined in the previous sections describes the statistical properties of random mismatch in devices. In the design phase of a circuit these abstract notions of statistical variations on parameters must be translated into design decisions. Application to electronic design includes techniques as Monte-Carlo simulation, hand calculation and various other techniques.

11.5.1 Analog design

In Sect. 6.2.2 an example of a bandgap circuit has been discussed. The caveat in the circuit is in the random offset of the operational amplifier input. This offset is also amplified and will create significant differences in the output voltage of the circuit. In Fig. 6.7 the measured output voltages of a few hundred devices is compared to the Monte-Carlo simulation of a similar number of samples. The MOS model in the simulation is equipped with a mismatch model as in (11.15).

A second example is shown in Fig. 11.23. The differential non-linearity curve (DNL) is a measure for the error that an analog-to-digital converter makes at every code transition. If the DNL reaches the ± 1 limits non-monotonicity in the analog-to-digital transfer curve can occur. In high speed converters MOS threshold mismatch
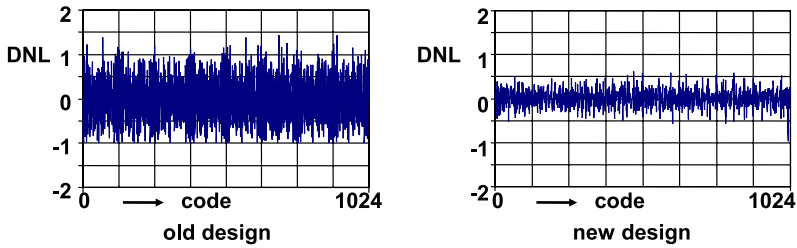
Fig. 11.23 The differential non-linearity is a quality parameter of this 10-bit analog-to-digital converter

Table 11.9 The simulated standard deviation of the difference in arrival time of the two pulses in the inverter chain of Fig. 11.24. The *bottom line* shows the random skew for scaled feature size

Process node	0.25 μm	0.18 μm	0.13 μm	90 nm	65 nm
Clock period	10 ns	5 ns	2 ns	1 ns	0.5 ns
$\sigma_{\Delta T2}$ $C_{\text{load}} = 50$ fF	16 ps	21 ps	38 ps	68 ps	88 ps
$\sigma_{\Delta T2}$ $C_{\text{load}} = 50 \dots 15$ fF	16 ps	16 ps	22 ps	33 ps	32 ps

of the input pair of the comparators is the dominant contributor to differential non-linearity. It is imperative that this error must be made small. The measurement of the first prototype (left) shows significant deviations with excursions down to -1. After analyzing the design with a Monte-Carlo simulation, the major contributors were located and correctly dimensioned, resulting in the measured curve on the right. Today the Monte-Carlo analysis is a standard tool for high-performance circuits in analog design.

The consequences of mismatch on the yield of analog-to-digital converters is discussed in Fig. 8.22 [199]. The shape of the curve is characteristic for yield plots. The transition from excellent yield numbers to commercially uninteresting levels is sharp. It shows importance of using the correct models and parameter values in precision designs.

Power, speed and accuracy span the design space e.g. [278]. The idea that accuracy must be balanced against power can be easily understood by considering that the voltage uncertainty on the gate capacitance can be described as an energy term [279, 280] and Sect. 11.5.4.

11.5.2 Digital Design

Also digital designers experience that for small devices the random component can exceed the process corner variation. An example is shown in Fig. 11.24 and Table 11.9 [281]. A pulse is applied to two sets of inverters and it is expected that the outputs will change state simultaneously. Due to mismatch between the transistors

Fig. 11.24 An input pulse is applied to two chains of inverters. Due to mismatches in the transistors there is a time difference at the two outputs. This example mimics a critical timing path

$W_p = 2W_n = 8L_{min}$
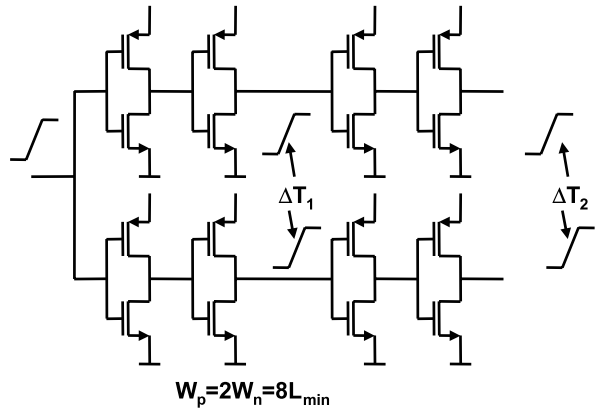
of both rows of inverters, a random skew will exist in the difference of arrival time between various samples of this circuit. In Table 11.9 the standard deviation of this random skew is compared to the clock period in five technologies. From an effect in the 0.1% range the random skew can take up to 10% of the clock period in a 65 nm process.

In digital timing chains, an increasing chain length will linearly increase the deterministic skew, while the random component will go up via a square-root relation. The relative impact of random variations reduces. It should also be noted that the "root-mean-square" addition mechanism favors large contributors. The random timing skew in a long chain can be dominated by just one small-sized inverter.

Also in memory structures statistical effects play a major role. On the level of a SRAM cell, these effects can be evaluated as in any other analog circuit. The additional problem in memory structures is the large amount of cells. This requires simulating statistical distributions up to 7σ. This is not practical in Monte-Carlo simulations. Special statistical acceleration mechanisms (importance sampling, Latin Hypercube sampling) allow to sample tails of statistical distributions [283]. Memory designs are affected by mismatch in several ways. Threshold variations influence the margins for the read and write operations. Moreover low-threshold devices create (exponentially) large leakage currents. The choice for the size of the transistors in an SRAM cell and in the sense amplifier critically depends on an accurate prediction of the mismatch coefficients.

11.5.3 Drift

In literature the term "drift" refers to various phenomena in electronics circuits. Drift is a long-term parameter change that can be caused by aging effects, mechanical influences and temperature. The most frequent appearance of drift in analog-to-digital conversion is a temperature-dependent shift of input offset. The magnitude of this drift is in the order of several μV/°C measured at differential inputs of circuits.

In a completely symmetrical and differential circuit with identical components there is no reason for drift. However, mismatch can cause unbalance that translates in drift. In a simple differential input pair inequalities will exist due to random threshold mismatch and current factor mismatch. The threshold mismatch is a charge based phenomenon and will create a temperature independent offset. The current factor mismatch does cause drift. Assume that a current factor difference $\Delta\beta$ exists in a differential pair due to mismatch. This current factor difference must be compensated by an input referred voltage. The temperature dependence of the voltage between the inputs of this differential pair is the drift of this differential pair.

The offset current is proportional to the current factor difference and can be translated into an input referred offset ΔV_{in} via the transconductance:

$$\Delta V_{\text{in}} = \frac{I_d}{g_m}\frac{\Delta\beta}{\beta} \tag{11.23}$$

Andricciola and Tuinhout [273] show that the ratio $\Delta\beta/\beta$ is only marginally dependent on temperature.

If a constant bias current is provided the only temperature dependence is due to the transconductance. With the help of (2.111) an expression for the temperature dependent drift at the input of a differential pair due to current factor mismatch is found:

$$\text{Drift} = \frac{d\Delta V_{\text{in}}}{dT} = \frac{\alpha}{T}\frac{I_d}{g_m}\frac{\Delta\beta}{\beta} = \frac{\alpha}{T}\frac{V_{GS}-V_T}{2}\frac{\Delta\beta}{\beta} = \frac{\alpha\Delta V_{\text{in}}}{T} \tag{11.24}$$

11.5.4 Limits of Power and Accuracy

One of the main questions in low-power conversion design is the ultimate limit of power consumption. Mathematicians would claim that the mapping of analog values on a discrete amplitude scale should not change the signal and therefore be zero-power.

In physics, however, a lower limit can be derived from quantum-mechanical effects, concerning the minimum number of electrons required to represent one bit. Another limit is given by Dijkstra [282] for $\Sigma\Delta$ converters. His analysis assumes that thermal noise on the input impedance or transconductance is dominant. This approach results in a energy limit based on the product of SNR and thermal kT noise. These limits are however four to five decades away from practical realizations. This is partly due to the fact that much "overhead" power has to be incorporated in real designs: transconductances in MOS need large amounts of current, parasitics have to be overcome, safety margins for all kinds of unexpected variations have to be accounted for.

In this section an approximation for power consumption in parallel-circuit structures as in high-speed converters will be derived based on random variations in the circuit components. This may be the case in multiplexed circuits or in circuits in

which the signal path is level-dependent, as in full-flash converters. The component variations between the various paths will result in unwanted signals: fractions of sample rates, fixed pattern noise, etc. These component variations will decrease when the area of a MOS transistor is increased. On the other hand, the loading of the signal will also increase when gate areas increases:

$$\text{Capacitive load} \qquad \text{Threshold variance}$$

$$C_{\text{gate}} = WLC_{ox} \qquad \sigma_{\Delta VT} = \frac{A_{VT}}{\sqrt{WL}} \tag{11.25}$$

The voltage uncertainty on the gate capacitance can be described as an energy term [279, 280]:

$$E_{\sigma_{\Delta VT}} = C_{\text{gate}}\sigma^2(V_T) = C_{ox}A_{VT}^2 = 4.5 \times 10^{-19} \text{ J} \tag{11.26}$$

which is independent of the transistor size and corresponds to about 100 kT at room temperature. This energy can be seen as the energy required to toggle a latch pair of transistors in meta-stable condition into a desired position with a one-σ certainty. In circuits with parallel data paths the energy required to overcome component mismatch may hence dominate over kT noise by two orders of magnitude.

In analog-to-digital converters mismatch may manifest itself as noise, but is also observable in other specification points. The most critical point is differential non-linearity, because this effect relates to the maximum error in the parallel structure, while noise is only a power average value over all the error terms.

The charge standard deviation, (11.13), can be used to derive an estimate for a minimum LSB size. In order to be sure that an LSB change in the input signal is always detected, the minimum charge for an LSB is α times the charge standard deviation. In a normal distribution a sufficiently low probability is obtained with $\alpha = 7$–10. So the resulting minimum LSB charge is:

$$Q_{\text{LSB}} = \alpha \times \sigma(Q_d) = \alpha q \sqrt{WL z_d N_a} \tag{11.27}$$

The minimum LSB charge combined with the signal speed results in a current. It is assumed that no slewing of the comparator stage is allowed and the current is delivered in class-A operation. In converter terminology: the DNL performance must be reached at a signal frequency f_{signal}. The minimum amount of time required for an input signal to change V_{LSB} is t_{LSB}. A distinction can be made between time-continuous mode and sample-and-hold mode:

$$\text{Time-continuous mode} \qquad \text{Sample-and-hold mode}$$

$$t_{\text{LSB}} = \frac{V_{\text{LSB}}}{\partial V_{\text{signal}}/\partial t} \qquad t_{\text{LSB}} = \frac{V_{\text{LSB}}}{\Delta V_{SH}/\tau_{SH}} \tag{11.28}$$

$$t_{\text{LSB}} = \frac{1}{2^N \pi f_{\text{signal}}} \qquad t_{\text{LSB}} = \frac{1}{2^N N \ln(2) f_s} \tag{11.29}$$

In the context of this analysis the results of both modes are comparable, so the analysis will proceed with the time-continuous result. The current needed for supplying Q_{LSB} at the steepest point of an input signal at a frequency f_{signal} is:

$$i_{\max} = \frac{Q_{\text{LSB}}}{t_{\text{LSB}}} = 2^N \pi f_{\text{signal}} \alpha q \sqrt{W L z_d N_a} \tag{11.30}$$

Example $N = 8$, $f_{\text{signal}} = 10^7$ Hz, $\alpha = 10$, $N_a = 10^{16}$ cm^{-3}, $W = L = z_d = 1$ μm Results in $i_{\max} = 1.3$ μA for a 1/1 μm transistor.

Multiplication with V_{DD} and rearrangement of terms this formula yields the power per resolution and bandwidth. The proportionality constant is then:

$$\frac{\text{Power}}{2^N 2BW} = \pi \alpha q V_{DD} \sqrt{W L z_d N_a} \tag{11.31}$$

which amounts to about 10^{-4} pJ per signal-carrying 1/1 transistor. On the assumption of some 100 matching critical transistors per high-speed analog-to-digital converter, this (rough) approximation results in 0.01 pJ, which is about two orders removed from the realized converters in 1 μm technology.

Closer examination of this discrepancy shows that the following factors contributed to this gap:

- It was assumed that only the transistor was charged. In the design the parasitic capacitors require roughly the same amounts of charge.
- More effects than the depletion charge alone contribute to the uncertainty of an LSB: W, L dependencies, mobility variations, etc.
- Analog-to-digital converters are designed to operate under worst-case circumstances: temperature, power supply and process variations require an overdesign of about 2–3 times.

In the above analysis the basic problem of reaching accuracy at high speed was concentrated in the comparator. This implies that N comparators for an N-bit digital word is the lowest-power configuration for analog-to-digital converters. In oversampling a single comparator is used, however that comparison requires to incorporate the power of the digital filtering.

11.6 Packaging

An integrated circuit becomes available to a user in the form of a packaged device which allows to connect the die to the other electronics circuitry. In the seventies and eighties the dominant form of packaging was the "Dual-in-Line" package (DIL or DIP), see Fig. 11.25 top. The die is attached to a lead-frame. Thin gold wires (bonds) connect the bondpads on the die to two rows of pins that are spaced at a pitch of 2.54 mm. These pins connect the die to the printed-circuit board wires. Typical packages are made of ceramic or plastic material and count 14, 16, 24, 40, 64 pins. The pins or leads are inserted in the printed-circuit board via through-holes. Ceramic packages are expensive and mostly used in a prototype phase, while plastic
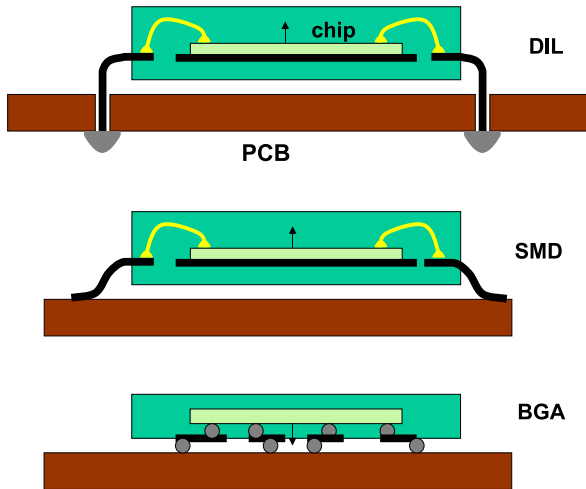
Fig. 11.25 Three types of packages: Dual-in-Line, Surface Mounted Device and Bump-bonded device

molded packaged are used for mass production. Unfortunately a circuit mounted in a ceramic package can behave differently from the same circuit molded in plastic. Some reasons are:

- When a die is molded the surrounding plastic is warm. After cooling the plastic shrinks and creates mechanical stress on the die. This stress influences the mobility of sensitive structures.
- In ceramic packages the die can be soldered to the frame. This allows a good drainage of heat and a low-ohmic connection to the substrate. In a plastic package the die is normally glued. The thermal and electrical conductivity of the glue is not optimum.
- In volume production the wafer is thinned from 500 to 100 μm, resulting in different a substrate conductivity.

With the increasing integration of functions and the growing size and number of interfaces, packages are required with more pins and smaller size. The first step is to use all four sides of the package, indicated with the term "Quad", see Table 11.10. Pin-grid arrays use the entire bottom of the package for connecting hundreds of signals to the printed-circuit board.

In "surface mounted devices" or SMD technology the package is mounted on the surface of the printed-circuit board in stead via through-hole connections, see Fig. 11.25 middle. The pitch of the leads is reduced to 1.27 mm and for some packages even down to 0.8 mm.

In a next step to miniaturization the connecting leads are replaced by electrodes under the package, Fig. 11.25 bottom. Special balls connect the "Ball-Grid Array" to the printed-circuit board. This Chip-Scale Packaging technique allows an effective package that is area-wise only 20% larger than the encapsulated die. In Fig. 11.25

Table 11.10 Commonly used packaging abbreviations. Often these terms are followed by the number of pins

BGA:	Ball Grid Array
CERDIP:	Ceramic DIP
CFP:	Ceramic Flat Pack
CPGA:	Ceramic Pin Grid Array
CQFP:	Ceramic Quad Flat Pack
CSP:	Chip Scale Package
LCC:	Leaded/Leadless Chip Carrier
LQFP:	Low Profile Quad Flat Package
PGA:	Pin Grid Array
PLCC:	Plastic Leaded Chip Carrier
PQFP:	Plastic Quad Flat Pack
SIP:	Single In-line Package
SOP:	Small-Outline Package

(bottom) the die is placed upside-down and a special redistribution pattern is used to move the bondpads to the correct position in the package. Many more combinations of wiring techniques exist.

The package plays an important role in the thermal management. The overall thermal behavior of a circuit on a printed-circuit board is defined as:

$$\theta_{JA} = \frac{\text{Junction temperature} - \text{Ambient temperature}}{\text{Power}} \qquad (11.32)$$

See also Sect. 2.2.1. The value of θ_{JA} varies with many factors: the sizes of the die and package, the package material, the number of leads, the number layers of the printed-circuit board and lay-out, the air-flow, etc. A die of 2 mm^2 in a small QFP44 package on a cheap printed-circuit board can reach $\theta_{JA} > 100\,°C/W$. A 100 mm^2 die in a 500 pin BGA allows $\theta_{JA} < 10\,°C/W$.

The package connects the signals to the circuit via a bondpad structure, see Fig. 11.26. The bondpad consists of a top-level metal flap of dimensions between $70 \times 70\ \mu m^2$ for advanced bonding and $130 \times 130\ \mu m^2$ for somewhat older processes. Next to this connection electrode, protection structures are used to avoid electro-static discharges (ESD) that could damage the connected circuitry. These protections mostly consist of diode-like structures towards both power supplies and act as voltage clamping devices. Moreover a resistor in series with the signal source prevents excessive currents to enter the circuit.

The protection devices are connected to shared supply rails that run along the bond pads. These supply lines are also connected to an external ground via (inductive) impedances. The consequence is that a coupling mechanism exists between all pads connected to one set of protection supplies. To avoid coupling between the analog and digital side of the chip the analog and digital protection supply lines have been separated.

Two ESD test methods are common practice: the human body model (HBM) consists of a 100 pF capacitor charged to 2, 4 or 8 kV, which is then discharged

Fig. 11.26 A schematic
lay-out of the bondpads for
analog and digital and their
protection devices
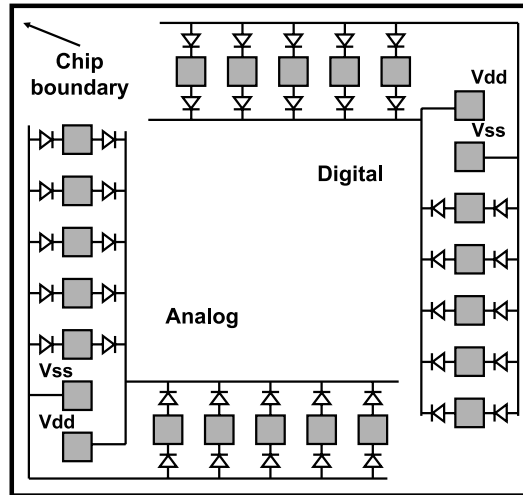
Table 11.11 Some values for the components in the equivalent circuit diagram of Fig. 11.27

Bond wire	$L \approx 1$ nH/mm, $L = 1$ nH(CSP)–10 nH(DIL)
Bond wire	$R \approx 0.3$–1 Ω
Bond pad capacitance to substrate	0.8–2 pF for 0.1×0.1 mm^2
Diode capacitance to supply	0.3–0.7 pF
Protection resistor	100–400 Ω

in the bondpad via a 1500 Ω resistor. The second model mimics contact to some
piece of equipment. Testing with the Machine-Model requires to discharge a 200 pF
capacitance into a circuit connection with no more series impedance than a little
bit of inductance (0.5 μH). It will be clear that the protection diodes and resistor
cannot be of minimum size and some considerable loading due to the parasitics
associated to the protection devices on both inputs and outputs must be considered
during the design. Next to that, the non-linear nature of the diodes in combination
with a high source impedance can create distortion. E.g. a diffusion area of 250 μm^2
in a 0.18 μm CMOS process will vary in capacitance between 0.15 and 0.3 pF over
1.8 V. Figure 11.27 and Table 11.11 show a simple equivalent circuit diagram and
some data of the components.

11.7 Substrate Noise

The trend to integrate more functionality in one package has led to ICs with a
large number of different functions in one die or one package. These functions
pose completely different requirements on the environment they operate in. In a
receiver chip, the combination of microvolt sensitive Low-Noise Amplifiers, analog-
to-digital converters and Digital Signal Processors running at hundreds of MHz is
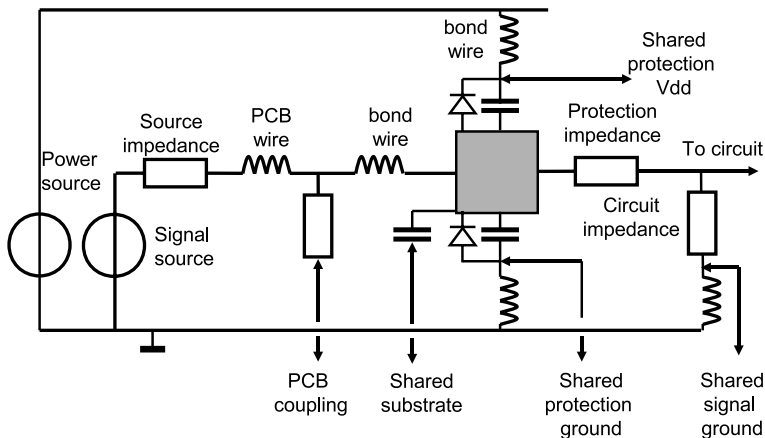
Fig. 11.27 Equivalent diagram showing the most important parasitic elements in connecting a signal to the circuit on the IC

Fig. 11.28 A general model for describing substrate noise generation, transport and effect
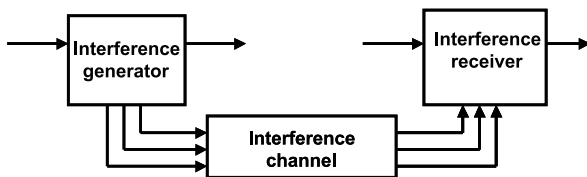
not unusual. Control over the mutual interference between these blocks is essential to avoid that a design after production does not meet the specifications. Interference is important in all analog ICs, where it can affect crucial specifications like SNR, jitter, etc.

An analog-to-digital converter operates by definition on the boundary of two domains. An essential aspect of an analog-to-digital converter is the interference between these two domains. Another relevant aspect is the sampling that occurs in an analog-to-digital converter. A converter is not only sensitive to noise in its direct baseband, but also can alias substrate noise contributions from the higher frequency bands into the baseband.

More aspects play a role. Coupling via the package and substrate are most relevant for the MHz range. For RF design also the coupling via air must be considered.

Figure 11.28 shows the generally accepted model for analyzing and describing interference.[6] The interference generator and the interference receiver are functional blocks or IP cores on a die. The generator can be a digital block: a group of I/O cells, a processor. Also analog cells can act as generator, e.g. an oscillator.

---

[6]This research was carried out together with Sergei Kapora and Jacob Bakker. The model was suggested by Joost Briaire.
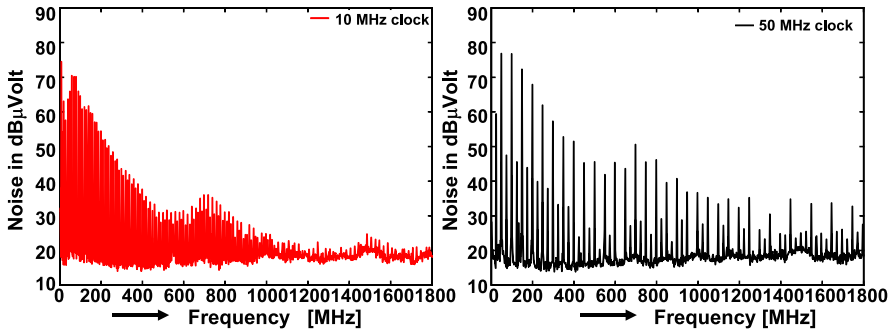
Fig. 11.29 A spectrum of the substrate potential measured close to a 0.5 mm^2 65-nm CMOS digital shift-register block. The clock runs at 10 and 50 MHz and the input is a half-clock signal. The *vertical axis* is in dB µV: 60 dB µV equals 1 mV. *Courtesy*: S. Kapora and J. Bakker, NXP

The receiver will normally be an analog block; nevertheless also digital blocks can become victim of interference. Sensitive memories or the performance of low-power logic can be affected by strong signals in the substrate. Although they are placed here in an order appropriate for the model of substrate noise propagation, the actual placement of these blocks in the signal processing chain can be completely different.

The interference generator will mostly consist of many individual signal sources. In most practical situations these signals will not be known to the outside world, and are considered to be random. This is a dangerous assumption, as unexpected correlation between the signals and concentration in limited frequency bands may easily result in misjudged behavior. In the generator there are many aspects to consider:

- Frequency, amplitude and shape of the internal signals.
- Lay-out: symmetrical shapes, differential signaling, etc.
- Parasitic components: junctions, coupling capacitors.
- Substrate connectivity: location of substrate contacts and density.
- Power wiring: resistivity.
- The amount of simultaneous switching: area, synchronous versus asynchronous, the bus width, etc.

Figure 11.29 shows the spectrum of the substrate potential measured close to the digital shift register at clock frequencies of 10 and 50 MHz. The input signal toggles every clock cycle and runs at half of the clock frequency. The plot shows the that main source of interference is related to the clock and its harmonics. It is also obvious that the shape of the spectrum remains at the same position. The generated noise is perhaps potentially more wide band, but the filtering by the capacitors and inductors and the shape of the current pulse is of greater importance.

The interference channels are the substrate, coupling of signals via the air, parasitic wiring, coupling via the package and parasitic leakage paths, etc. The effect of the impedance of package wiring is essential to consider as this effect contributes significantly to the effects in the substrate.
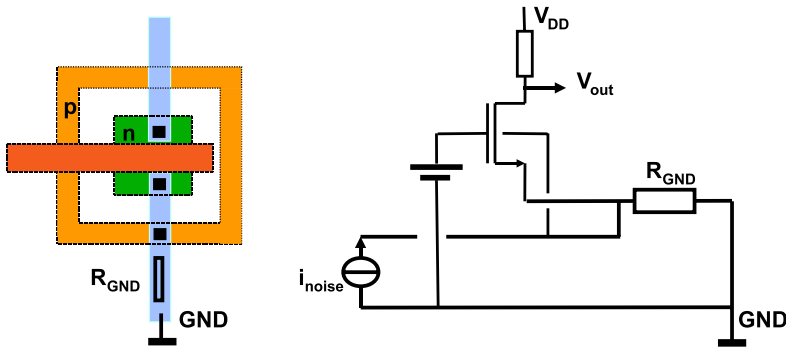
Fig. 11.30 *Left*: a lay-out of a substrate-noise sensitive NMOS transistor surrounded by a p-doped substrate connection. *Right*: the equivalent circuit diagram

When analyzing the coupling of interference via the substrate, three types of substrates must be distinguished:

- An epitaxial layer on top of a low-ohmic substrate with resistivity in the order of 0.01 Ω cm. In this case the substrate must be considered as a copper plate, the substrate acts, even for large circuits, as one single node.
- An epitaxial layer on high-ohmic substrate, or a bulk high-ohmic substrate. Special modeling of the substrate is required.
- Pseudo isolating substrates as used in silicon-on-insulator processes. Here the substrate is only capacitively coupled.

The substrate-noise receiver is mostly a circuit where transistors operate in an amplification mode. The back-bias voltage of one or both types of transistors is modulated by substrate noise and enters the signal path. Nodes are also capacitively coupled to the substrate. The receiver sees many paths to the substrate, and therefore substrate noise analysis is a multi-input problem. Again a similar list of issues can be set-up as for the noise transmitter: frequency, amplitude and shape of the internal signals, lay-out: symmetrical shapes, differential signaling, etc., parasitic components: junctions, coupling capacitors, substrate connectivity: location of substrate contacts and density, guard rings, deep well isolation and other technological options, power wiring: resistivity.

An error is easily made in substrate noise prevention. Figure 11.30 shows a transistor were the designer has attempted to connect the substrate to the ground via a p-ring. From the equivalent circuit diagram it is clear that if the resistor in the ground line is negligible, the noise current in the substrate will be diverted to ground. However, if this resistor is a few Ohms, there will not only be a voltage on the back-bias terminal of the NMOS transistor but also a modulation on the source. With increasing impedance in the source line (e.g. due to the inductive bondwire) the magnitude of the noise will increase. This example shows that it is generally not advisable to combine the wires for protective rings with wires that feed the circuit.

Next to a qualitative description of the noise generator, channel and receiver also a quantitative description is necessary. Moreover the accuracy with which each part
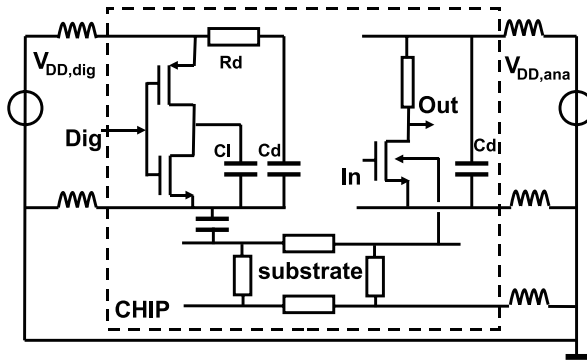
Fig. 11.31 The most simple form of coupling: *on the left* is an inverter whose switching noise propagates via the substrate to the analog amplifier

of the chain is modeled must fit to the overall accuracy: it is useless to aim for 0.1% channel modeling accuracy while the generator is only with 10% accuracy predictable.

Figure 11.31 shows a simple example of substrate noise generation. During the switching of the inverter, current flows into or out of the load capacitor C_L. The charging current spike will run from the power supply via the package and bondwire impedance, the power network and the PMOS switch to the capacitor and back via the ground network. Capacitors connected to the power supply will generate a similar current if the NMOS switch is activated. In both cases the switching current produces a voltage spike over the power and ground network. This voltage spike is in the case of CMOS logic the dominant noise generating mechanism. The voltage spike couples via the wire capacitance to the substrate network. The effect increases if the local source of the NMOS transistor is directly connected to the substrate for latch-up reasons. Parasitic capacitive coupling of the bottom plate of the load capacitor or the source and drain diffusions of the CMOS transistors form another coupling mechanism.

In the substrate the noise propagates to the analog circuit, where it can modulate the back-bias voltage of the NMOS transistor or enter via the capacitive coupling of drains.

This simple scheme allows already some conclusions:

- The package impedances and the power network are crucial for the generation of digital and analog noise. Low-ohmic implementation is mandatory. The use of bump-bonding lowers the package impedance considerably.
- It will be difficult to achieve a good noise suppression if the power supplies of analog and digital are combined. The mutual path will certainly contain impedances over which the circuits can couple.
- It is recommended to avoid digital cells with direct connection of the source to the substrate. Especially the I/O drivers generate large current spikes and can often be considered the main noise source.
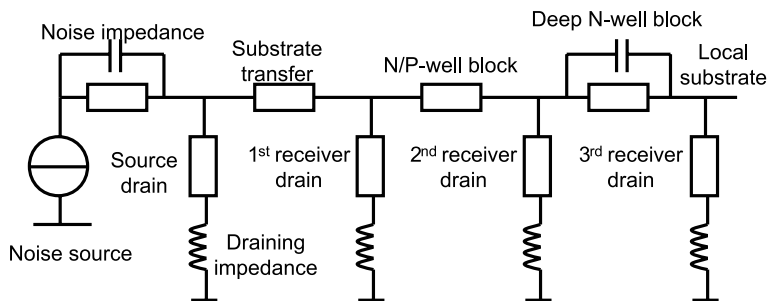
Fig. 11.32 A equivalent circuit diagram of the noise propagation route

- The decoupling of the digital cells is important. This decoupling capacitor supplies part of the load current, thereby lowering the current spike via the power network. Switching of CMOS inverters can be fast: 100 ps range. The time constant of the decoupling network $R_D C_D$ should be lower than that.
- Also decoupling on the analog side helps to minimize coupling. On one hand the power network impedance and the decoupling capacitor form a low-pass filter foe external noise. And on the other hand the internal power supply is stabilized.
- Use all free space for decoupling capacitors.

The path from noise generator to noise receiver is modeled in Fig. 11.32. The impedance of the generator is either resistive or capacitive. From the generator to the receiver several measures can be taken to minimize the transfer. In essence, substrate noise reduction is a matter of increasing the series impedances between generator and receiver and reducing the draining impedances.

- Distance between generator and receiver increases the impedance only marginally. At short distance the width dimension of the generator will play a role. Nevertheless it can help to place silent blocks (like memories) between noisy processors and the analog design.
- Already close to the generator a p-diffusion can connect the substrate to a clean supply. This requires a significantly lower impedance than the impedance of the noise source. This measure poses a danger as a low impedance noise generator can spoil a clean supply.
- In CMOS processes on high-ohmic epitaxial material or substrates often a channel stop implantation is used to prevent parasitic inversion under wires. This low-ohmic surface conductor will be the main connector at short distances. Most processes provide a blocker layer to break this path.
- An alternative is to use a ring-shaped n-well.
- At distances comparable to the thickness of the wafer or longer, the main conduction path is via the substrate. This conductivity path can be broken by means of a deep n-well implant, that will completely surround the sensitive NMOS transistors. Unfortunately this implant will connect together all n-wells of the PMOS transistors.

- Last but not least the use of a differential design style in the generator (current-mode logic) and receiver will suppress common contributions of the substrate noise.
- The package and bondpads influence the substrate noise in two ways: first the connections have series inductance and resistance, secondly the bondpad forms a capacitive coupling to the substrate.

The values of the components in Fig. 11.32 depend on the technology and the lay-out. Although every lay-out is specific some orders of magnitude can be estimated. A typical path through the substrate is of the order of 200–1000 Ω. A p-ring connecting the substrate locally to an external ground can have an impedance of 10 Ω and a few nH inductance. However the path from the p-ring to the area right under the sensitive transistor will add another 10–50 Ω. So many substrate noise transfer plots show a 60–80 dB suppression at DC and a first order increase starting somewhere between 100 MHz and 1 GHz.

Chapter 12
System Aspects of Conversion

Figure 12.1 shows the block diagram of consumer electronic equipment. Dedicated analog interface circuits (sensors, tuners, etc.) provide signals to the analog-to-digital converter and the processing core. The results of the processing are fed to digital-to-analog converters and from there to analog output interfaces (power amplifiers, display drivers, etc.). The CMOS signal processing chain formed by analog-to-digital converters, DSPs and ASICs, is controlled by CPUs or micro-controller cores such as ARM and Cortex processors. Many system chips are designed in standard digital CMOS. The addition of analog circuitry results in a strong focus on CMOS mixed-signal implementation where to use of specialized technology steps is prohibited for cost reasons. Placing the analog-to-digital converters on the digital die avoids some digital signal interfacing between chips. This helps to minimize the I/O pin count and reduces electro-magnetic compatibility problems caused by the high-speed digital busses.

Advanced CMOS technology allows the full integration of a lot of system functionality. This includes most of the analog interfacing and sometimes even the RF part in many application fields: video, audio, communication, etc. As a result of this trend, more and more functionality has to fit within the maximum power budget of an IC package. Low-cost packages for large-volume markets (consumer, telecom) allow around 1–5 W dissipation. This requirement and the need for longer battery life time in portable equipment has led to a general trend to lower the power needed for analog and digital circuits.

CMOS scaling was over the last two decades more effective for digital area and power reduction than in analog designs, see Fig. 12.2. The reduction of technological dimensions has led to digital power supply reductions from 5 to 1.2 V and will culminate in around 1.0 V in the years beyond 2010 [19]. This has led to a significant power reduction in the digital domain. The technological advances have not resulted in an equally dramatic improvement of the performance/power ratio of analog circuits, often culminating in the situation that the analog blocks are now the largest consumers of on-chip power. The analog-to-digital converters and the digital-to-analog converters, their support circuits such as variable gain amplifiers, references, timing circuits and the postprocessing filters present a contribution typ-
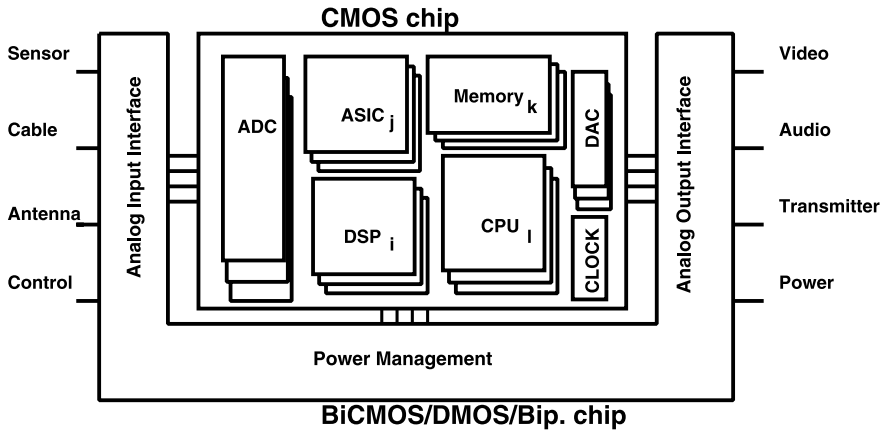
Fig. 12.1 General set-up of an electronic system-on-chip
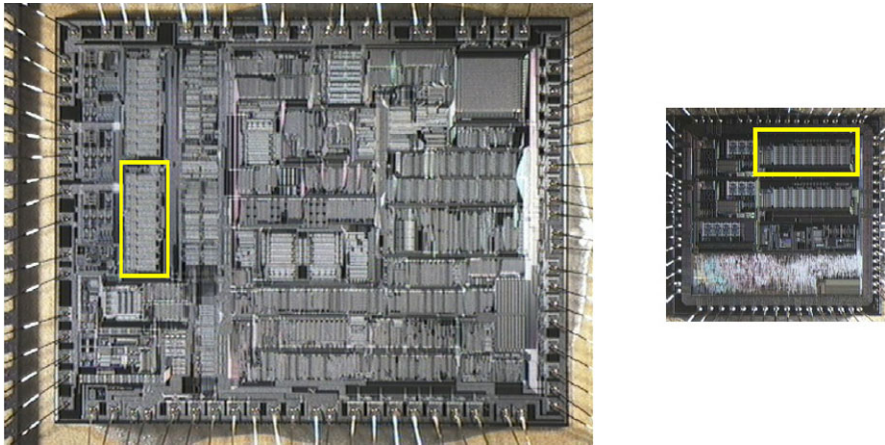
Fig. 12.2 Evolution of digital television processing, *left*: a die in 1 μm CMOS (1995), and *right*: a die with equivalent processing power in 0.35 μm CMOS (2002). The digital circuitry has shrunk by a factor of 10, while the analog-to-digital converter (*light box*) is approximately of the same size. *Courtesy*: Philips Semiconductors Hamburg

ically in the order of 30–50% of the power for the total system-on-chip. This has created a lot of push towards low-power implementations.

A second force towards analog low power comes from portability requirements e.g. for image sensors for cameras, Fig. 12.3. The entire system power budget must be lowered for longer battery lifetimes. This form of low power requires additional considerations, e.g. stand-by power, caused by leakage currents, must be lowered to a few microWatts per function.

Fig. 12.3 An analog-to-digital converter (on lay-out) is an important part of this CMOS image sensor (on wafer). Photo is courtesy of the Philips Semiconductor Imaging Sensor group
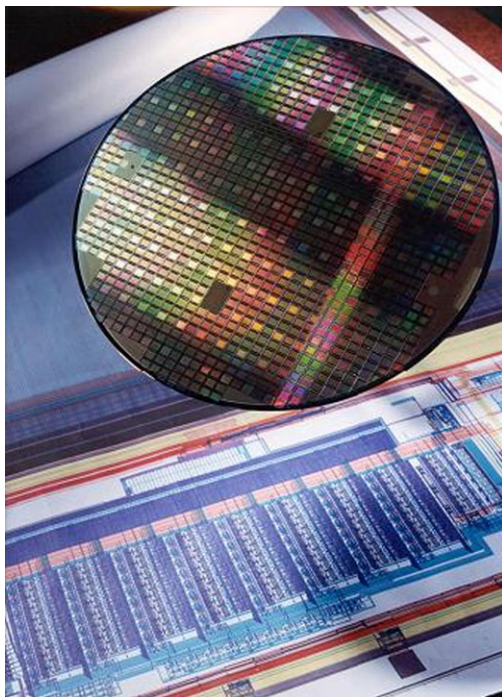
The drive towards lower power consumption in combination with the increased functionality requirements has led to an increased attention for power efficiency as a metric for performance.

This chapter will touch on a few aspects of the relation between system design and converter solutions from a low-power perspective.

12.1 System Aspects

The analog-to-digital conversion function is an integral part of the system: it is therefore always necessary to compare a number of alternative solutions on system level before choosing a conversion solution for a specific system. Next to the required bandwidth, resolution and power, a number of side considerations can play a role:

- Which conversion topology fits logically into the system? Some input devices (sensors) generate sampled-data information. Here an analog-to-digital conversion topology that normally requires a sample-and-hold circuit, can save some area and power.
- Which characteristics of the system fit well to the converter properties. A band-limited signal is well suited for oversampling. A control loop requires a short loop, where low-latency converters are a natural choice.
- What are the critical boundary conditions: is power more important than performance?

Table 12.1 Some system requirements on analog-to-digital conversion. Here some indicative data is given. Implementation choices can lead to shifts in requirements

Application	SNR or DR (dB)	Sample rate f_s (Ms/s)	Bandwidth (MHz)	Remarks
Consumer				
CD audio embedded	90–110	5.6	0.02	
Analog TV A/D	50–55	13.5	5	50 mW power
Analog TV D/A	60	27	5	30 mW power
Cable TV	40	32	15	50 mW power
Camcorder	55	3.5–20	1–2	10–30 mW
PC monitor D/A	60	100–400	40	Output drive
Digital radio	60–70	40–100	10.7–38.9	IM3 = 80 dB
Communication				
GSM, basestation	100	80	25	70 MHz carrier
GSM handset, 2G	80–85		0.2	Low power
GSM, EDGE	80–85		0.27	
CDMA2000	75		1.23	
UMTS	60–70		3.84	
Bluetooth 1.2	66		0.55	2.4–2.5 GHz
WiFi 802.11b	50–55		5.5	
Industrial				
Ultrasound imaging	60–70	40	20	300 mW
Oscilloscope front-end	50	100–500	250	
Seismic sensors	140		0.001	Reliability
Monitoring oil drilling	50		0.1	Temp = 300 °C

- How well do the impedances and signal levels match? Small signal levels require to be amplified before a flash converter is used. Signals in the form of charges fit well with switched-capacitor implementations.
- Are strong or weak out-of-band signals to be expected? Oversampling creates free bandwidth to accommodate such interferers. Time-continuous sigma-delta modulation provides alias filtering.
- Which element determining the resolution is most critical: noise, distortion, spurious?

Table 12.1 compares a few systems and their requirements on converters.

12.1.1 Specification of Functionality

Many trade-offs have to be made during the specification process of a complex system (e.g. a television set, mobile phone, medical imaging equipment). Today's

Fig. 12.4 Analog-to-digital
converters for CCD
interfaces, baseband video
and data transmission
applications
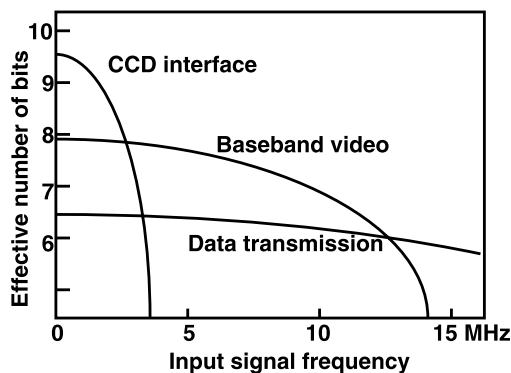
equipment is so complicated that design decisions have to be taken by large teams
of experts. Specification points for every individual function have to be negotiated
and fixed. Over-specification particularly leads to unnecessary power consumption
and must be avoided. Analog-to-digital conversion is particularly vulnerable to over-
specification due to its "natural" position between the analog and digital disciplines.
A frequently encountered approach is to specify a converter by asking the team
working on the analog side of the system to determine signal bandwidth and ampli-
tude, SNR and THD specs and combine this data with the specs of the team working
on the digital side: clock speed and duty cycle, word width, etc. This "collision" of
the analog and digital worlds leads to over-specification and must be avoided in the
design.

The requirement within one class of systems may look alike, yet strong differ-
ences can exist. Figure 12.4 shows three types of analog-to-digital converters used in
video signal processing. The plot shows the relation between the required Effective
Number Of Bits and the signal frequency. The CCD image-sensor interface analog-
to-digital converters require a relatively high degree of accuracy at low frequencies
in order to match the CCD dynamic range quality. The application in portable digital
cameras will set the focus on low-power and low-voltage [168]. Data transmission
for a digital satellite television signal needs some 6–7 effective bits, but requires
proper conversion at the high frequencies.

Between these two extremes is the baseband video converter, used in the signal
paths of high-end television sets. Its requirements range from good DC linearity
(DNL < 0.5 LSB) in order to prevent contouring, some THD < −55 dB at the color
carrier frequency (3.57–4.43 MHz) for color decoding, and a minimum 7 bit perfor-
mance up to 10–12 MHz signal frequencies for (QAM) HDTV-like applications.

Although all three converters are used for video signal processing, the specifica-
tions are hardly interchangeable. Several sub-optimizations exist, for each of these
three converters, which have resulted in a wide range of power-performance combi-
nations.

The trade-off on system level demands the optimization of the entire signal chain,
and is not only a quest for the lowest power analog-to-digital converter. An im-
portant aspect of this optimization is flexibility. Unlike specification tolerances,

Table 12.2 Balancing between the digital backs-draw and the analog benefits of a high sampling rate

Digital drawbacks	Analog advantages
More dynamic power consumption	SNR gain due to oversampling
Longer digital filter structures	Alias filtering is simpler
Less ripple-through time for the logic	Less steep voltage steps (slewing)
Larger storage units for fixed time delays	Less $\sin(x)/x$ loss

flexibility demands the adaptation of the system to pre-determined (large) system-parameter shifts. Some of these shifts are necessitated by the wish to serve several product lines of a manufacturer, e.g.: compatibility with several power supply voltages, availability of various output formats. In other cases the flexibility is implemented in the system itself, as the system has a multi-standard nature, comprising of several transmission standards, interfaces with different sources, etc. Flexibility mostly implies more hardware and more power than a single point optimized solution. Sometimes the increase in power can be reduced by designing the device in a clever way (e.g. by dividing the power of the circuit in a way that only the required parts of the circuit are connected). A proper balance has to be found in incorporating flexibility in a system with minimum power or component overheads.

12.1.2 Signal Processing Strategy

The main parameters of analog signal processing on a system level are dynamic range, signal-to-noise ratio and bandwidth. These quantities translate into resolution and sampling rate requirements in the digital domain.

On the system level the sampling rate of an analog-to-digital converter is locked to or derived from the system clock. The choice of clock (and sampling) frequency used on the system level is important with respect to specification and power. The main criterion for the sampling frequency is given by the Nyquist theorem ($f_s > 2\text{BW}$). For high bandwidth circuits the sampling frequency is in practice 20 to 50% higher than the minimum rate required by the Nyquist theorem, see Fig. 3.4. The main reason is the trade-off between complexity and cost of the analog alias filter versus the consequences of a higher clock rate in the digital signal processing, see Table 12.2. This balance between signal quality and filter complexity in the analog domain and power and area in the digital domain shifts rapidly with the advancement in technology.

In the signal processing strategy it is important to choose the appropriate conversion position in order to minimize the power required. There will usually be some freedom in choosing the position of the data converter within the signal chain. As
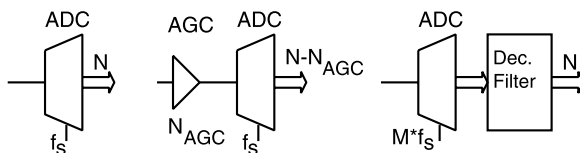
Fig. 12.5 (**a**) Standard analog-to-digital converter, (**b**) analog-to-digital converter with gain control and (**c**) reduced resolution analog-to-digital converter with oversampling and decimation filter
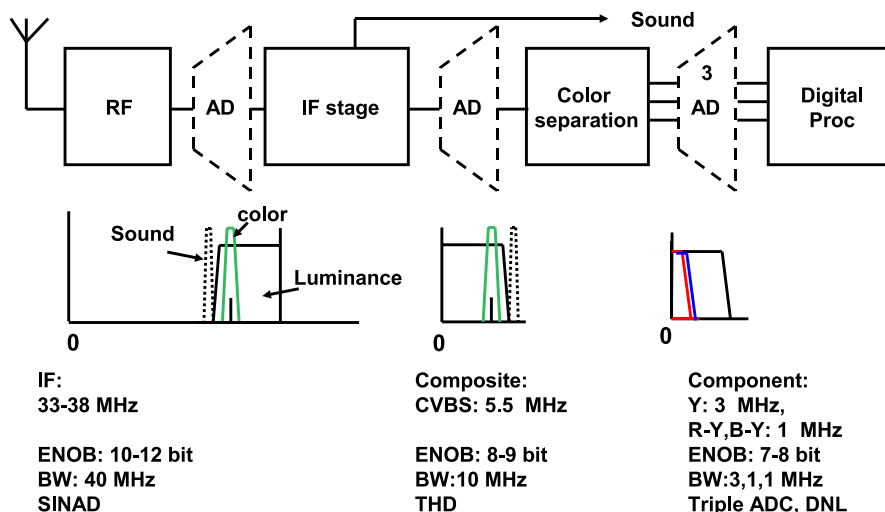
Fig. 12.6 In a television signal processing chain there are three potential positions for the analog-to-digital converter: in the IF domain, on composite signal level or on component level. Spectra for all three positions are indicated and some global specifications of the converters

analog-to-digital conversion[1] is quantization in the amplitude and in time domains trade-offs can be made between these two domains.

In the amplitude domain the optimum signal-to-noise ratio and the dynamic range have to be determined. In those cases in which a lower SNR can be used with respect to the dynamic range, the preferred system solution is a gain controlled amplifier followed by a minimum SNR analog-to-digital converter, see Fig. 12.5. Gain control does require some form of signal analysis, which is mostly a cheap function in the digital domain. Another option in this case is a companding analog-to-digital converter that uses non-uniform quantization levels.

In specific system architectures the advantages and disadvantages of these solutions have to be investigated. Generally a reduction in analog-to-digital resolution outweighs the area and power costs of gain control or of a digital filter.

---

[1]The arguments here presented apply to analog-to-digital conversion, but hold mostly also for digital-to-analog conversion.

Fig. 12.7 Four examples of input structures for analog-to-digital converters
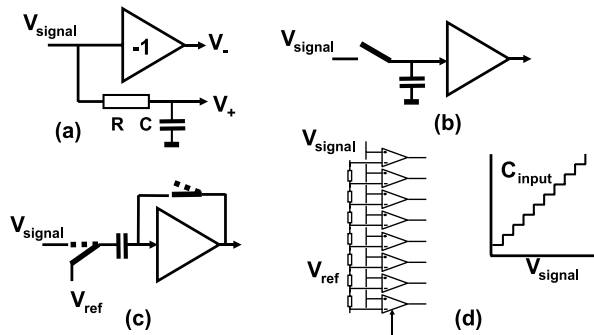
Figure 12.6 shows the three positions where an analog-to-digital converter can be placed in the analog-television signal processing chain. The most aggressive scenario is a converter in the Intermediate Frequency (IF) domain (Europe, U.S. and Japan use IF frequencies up to 60 MHz). The converter has to deal with a bandwidth up to 60 MHz and a dynamic range equivalent to 12 bit.

If the IF demodulation is left to analog circuits, the next position is on the composite level. The video signal is now demodulated but still uses the frequency interleaved color modulation scheme (with a carrier at 4.43 MHz (Europe) or 3.57 MHz (U.S.)). The converter needs 9–10 bits resolution, but is sensitive to interference from distortion products, see Sect. 12.1.4. In the last position the video signal has been split into its luminance (Y) and color components. Now three converters are required, or with some multiplexing two converters. The requirements here are pretty straight-forward.

12.1.3 Input Circuits

An analog-to-digital converter in a system is often regarded as a black box, and it is supposed that implementation choices remain without consequences for surrounding electronics.

However, due to economical and power constraints, many converters do not obey to this rule. The choice of the conversion principle can result in severe consequences for the circuitry driving the input. Power can be reduced in certain circuits, but this power reduction may lead to larger power consumption in other parts of the system. Such power shifts occur via the interfaces: input and output terminals, clock requirements, references, output representation, etc.

Figure 12.7 shows four implementations of input circuits of commercially available analog-to-digital converters. Example (a) shows a way to convert a single-ended input signal into a differential signal. The RC-network has a cut-off frequency identical to the inverting buffer, but does load the input driver. Moreover that driver needs a well defined output impedance. Example (b) shows a simple approach to a sample-and-hold circuit. However the input driver needs to be low-ohmic and

should not distort when supplying the switching currents. Example (c) is often used in pipeline converters. Next to the previous remarks it must be noted that the switching operation sends a current from the input driver into the reference generator. This reference source must be able to accommodate these charge pulses. Example (d) is typical in full-flash converters. The differential input pair is part of the comparator. Due to overdrive or underdrive this pair and many of its neighboring pairs will be either in a saturated "on" or "off" state. Thereby these gate capacitances will show as a capacitive load to the input or not. Consequently the input capacitance of a full-flash converter depends on the signal level and is therefore input voltage dependent. In case of a large number of input pairs the variation in input loading can amount to several picofarads. Driving such a variable impedance with high-frequency signals creates distortion.

   All four circuits effectively require a driver with a 20–50 Ω impedance over a wide voltage range. A circuit application diagram will show high-power external drivers that often require the same amount of power as the analog-to-digital converter! From a system point of view, it may be more efficient to use a input-buffered converter than a low-power analog-to-digital converter with high input driver demands.

   Next to the unwanted loading of the input also other disturbances at this input and clock terminal can influence the overall performance:

- Insufficient suppression of the alias filters results in unexpected side bands or spurious signals and spoils the accuracy of the converter.
- Unwanted signals like electro-magnetic interference can enter the system via ground or reference connections.
- The jitter of the sampling clock should be in line with the specifications of the converter: treat the sample clock as if it where a signal line. Do not use digital buffers that are fed from disturbed digital power supplies.

In the case of low signal bandwidths the ratio of the sample rate and the bandwidth is chosen so that it optimizes the entire conversion chain. Figure 3.4 shows the trade-off between the filter order and the sample rate/bandwidth ratio.

   Proper signal analysis allows to use the conversion properties in a functional way. Examples are:

- The inherent sampling performs de-modulation (see Sect. 12.1.4), which can be implemented in several analog-to-digital converter architectures. Particularly suitable are structures with high-performance sample and hold circuits.
- The inverse process is equally applicable: use upper bands generated by the digital-to-analog function for e.g. Direct Digital Synthesis of radio frequency signals.
- Multiple input signals can be multiplexed on the (S/H) input of an analog-to-digital converter.
- Alias filtering can be combined with system-required filtering.
- A local sample rate increase can be used to relax the requirements in other parts of the system (see Sect. 9.1).

Fig. 12.8 Converter
arrangement for IF
conversion
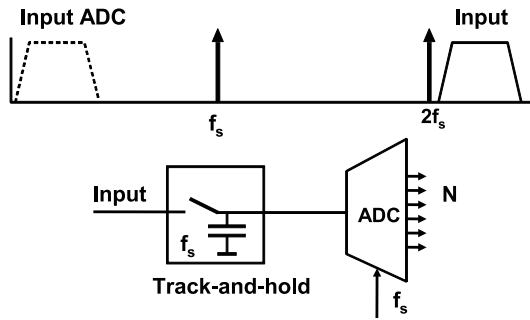
Fig. 12.9 Spectrum of a
sampled modulation signal
with distortion: (**a**) Sample
rate equals 3.1 times the
modulation frequency,
(**b**) sample rate equals 3.5
times the modulation
frequency
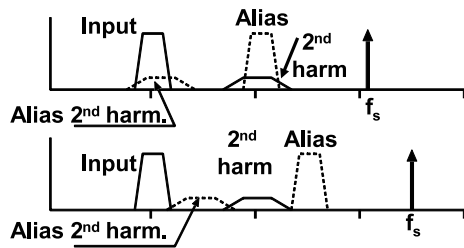
12.1.4 Conversion of Modulated Signals

Figure 12.8 shows an example of the use of the inherent S/H function for realizing
a system need: down-mixing of modulated signals. The information content of the
input signal is rather band-limited, but it is modulated on a relatively high carrier.
In this case it is power efficient to implement the conversion starting with a sample-
and-hold. The sampling function is used to modulate the signal band to a much lower
frequency. This process is called "sub-sampling". In this example down-modulation
is performed around twice the clock frequency. Now the conversion task for the
analog-to-digital converter core is much simpler.

Input signals containing carrier-modulated components imply additional diffi-
culties: aliases of the harmonic distortion of the signals will fold back during the
sampling process in the converter. If the sampling rate is close to an integer multiple
of that carrier frequency, the harmonics will interfere with the original signal (see
Fig. 12.9). This occurs in many transmission systems, e.g. in PAL video in which
the color modulation frequency is 4.433 MHz while the preferred sampling rate is
13.5 Ms/s. A sample rate close to an integer of the modulation frequency may ne-
cessitate additional suppression of the amplitude of the distortion component. This
will be achieved at the expense of more power in the converter.

In this section several aspects have been shown of system choices that influence
the power budget of analog-to-digital conversion. Low power requires thorough sig-
nal and system analysis.
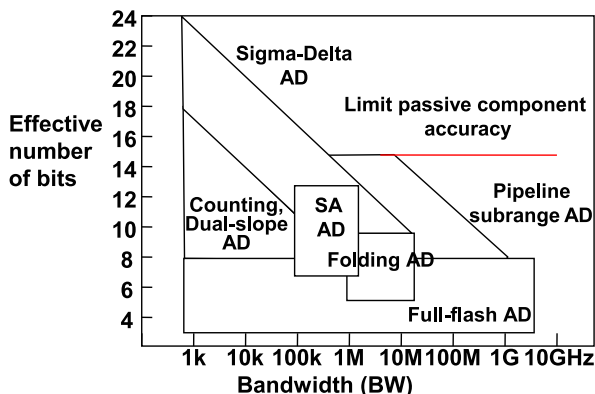
Fig. 12.10 Typical bandwidth and accuracies for various analog-to-digital converters

Table 12.3 Comparing analog-to-digital converter architectures

Type of analog-to-digital converter	Clock cycles for N bit conversion	Specification
Full-flash converter	1	very fast BW = 1 GHz, $N < 6$–8, power hungry
Folding converter	1	$N < 8, 9$
Pipeline	N	$N < 12$–14, fast BW = 10–200 MHz, efficient, latency of $> N$ clock cycles
Successive approx	N	Compact, BW = 2–5 MHz, $N < 12$, low power
Sigma-delta	20–50	N up to 24, BW = 100 Hz–5 MHz
Dual-slope	2^N	$N = 14$–20, BW = 10 kHz

12.2 Comparing Converters

Table 12.3 briefly summarizes some analog-to-digital converter characteristics and Fig. 12.10 shows various converter architectures in a resolution-bandwidth space. With the previous boundary conditions in mind a first estimate for an architecture can be made. A more quantitative approach can further assist in the discussion on system level and allows comparing converter implementations with a similar architecture.

In complex systems and in portable applications the main architectural decisions are often determined based on bandwidth, resolution and the available power. Despite the fact that there is no universal law for analog-to-digital power consumption, a practical approach is certainly possible. This approach is based on the observation that more accuracy or more bandwidth both require more power and that the combination of bandwidth, power and accuracy in a Figure-of-Merit is useful [285, 286].
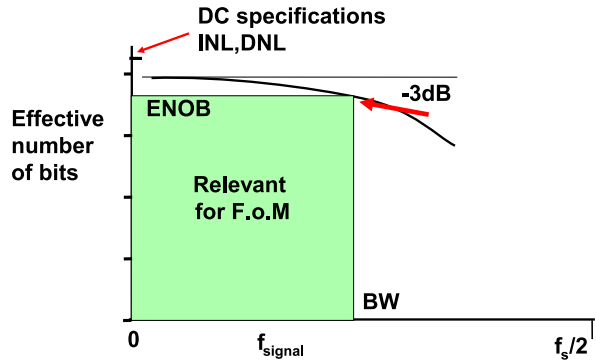
The starting point for a performance versus power efficiency comparison is the amount of signal power needed to overcome thermal noise at a certain signal-to-noise ratio (SNR):

$$P_{\text{sig}} = 4kTBW \times \text{SNR} \tag{12.1}$$

$$\text{Power efficiency} = \frac{\text{Power consumed}}{P_{\text{sig}}} = \frac{\text{Power consumed}}{4kTBW\text{SNR}} \tag{12.2}$$

This relation is often used to evaluate the efficiency of filters, opamps, etc. Some authors use this relation as a starting point for analyzing the lower limits of conversion efficiency, e.g. [226, 282, 284].

Figure 12.11 shows the behavior of the Effective Number of Bits (comprising Signal-to-Noise and Distortion) as a function of the applied input frequency. At higher frequencies the conversion becomes less accurate due to the increased distortion and other unwanted effects. As a measure for the performance of the converter a combination of bandwidth BW and resolution ENOB is chosen that is roughly 3 dB or 0.5 effective bits below the low-frequency value. As the shape of this curve in this region often is dictated by a first order decay, the product of resolution and bandwidth is not sensitive to the actual choice. The maximum bandwidth is limited to half of the sample rate due to Nyquist's theorem. This combination of power, bandwidth and the associated measured resolution are the right ingredients for measuring the efficiency of the conversion. The ENOB is a substitute for the signal-to-noise ratio, (5.13), SNR $= \frac{3}{2}2^{2\text{ENOB}}$, leading to:

$$\text{Power efficiency} = \frac{\text{Power consumed}}{4kTBW\frac{3}{2}2^{2\text{ENOB}}} \tag{12.3}$$

However, using this substitution in the theoretical relation of (12.2) does not fit well to the available data on analog-to-digital converters. Comparing many converters as in Table 12.4, Figs. 12.12 and 12.13, indicates that a better fit is obtained with:

$$\text{F.o.M.} = \frac{\text{Power}}{2^{\text{ENOB}} \times \text{Minimum}(2BW, f_s)} \text{ [energy per level]} \tag{12.4}$$

Table 12.4 Power efficiency of some ISSCC published analog-to-digital converters. Some care must be taken in comparing the numbers as the measurements conditions are not identical

Year	Author	Architecture	N bit	f_s Ms/s	Power W	BW MHz	ENOB bit	F.o.M. pJ/conv
2008	Schvan	Mux flash	6	24 000	1.2	12 000	3.5	4.4
2002	Scholten	Flash	6	1600	0.328	660	5	7.7
2002	Sushihara	Flash	7	450	0.05	75	5.8	5.9
2003	Poulton	Mux pipe	8	20 000	10	500	6.5	110
2007	Craninckx	SAR	9	50	0.00029	10	7.4	0.065
2009	Naraghi	Linear	9	1	0.000014	0.5	7.9	0.1
2004	Hernes	Pipe	10	220	0.135	100	8.1	2.4
2005	Yoshioka	Pipe	10	125	0.04	62	8.15	1.1
2008	Elzakker	Charge r.	10	1	0.0000019	0.5	8.5	0.004
2002	Miyazaki	Pipe	10	30	0.016	15	8.7	1.3
2004	Stroeble	Pipe	10	80	0.033	10.7	9	3.0
2006	Geelen	Pipe	10	100	0.035	50	9.3	0.55
2007	Hsu	Interleaved	11	800	0.35	400	8.7	1.1
2005	Doerrer	TC ΣΔ	12	10	0.0033	2	8.4	2.5
2001	Ploeg	3-stage	12	54	0.295	10	9.5	20
2009	Brooks	Pipe	12	50	0.0045	25	10.0	0.09
2007	Christen	ΣΔ	12	240	0.021	10	10.2	0.9
2006	Shimizu	Sub-range	12	40	0.03	20	10.5	0.52
2003	Murrman	Pipe	12	75	0.314	37.5	10.8	2.3
2000	Singer	Pipe	12	65	0.47	32	11	3.6
2001	Singer	Pipe	14	75	0.34	37	11.7	1.4
2004	Chiu	Pipe	14	10	0.112	5	11.8	3.1
2007	Hesener	SAR	14	40	0.66	0.96	13.8	24.2
2006	Schreier	BP ΣΔ	15	264	0.375	8	12.3	4.5
2006	Silva	TC ΣΔ	19	42	0.105	0.2	14.5	11.0
2005	Morrow	TC-SC ΣΔ	17	6.1	0.033	0.02	15.5	18

The differences are attributable to architecture (number of comparators), technology, noise or limited matching, etc.

In popular literature such a number is called "Figure of Merit". In this F.o.M. a factor of $2 \times BW$ is used to allow comparison with older F.o.M. numbers that use the nominal resolution N and the sample rate f_s. A low F.o.M. indicates that a converter uses less power for a certain measured specification, or delivers a better specification for the same power.
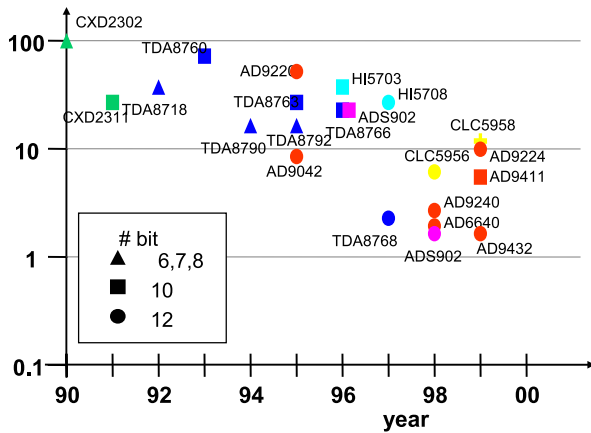
Fig. 12.12 Figure of merit for various industrial analog-to-digital converters in the time frame 1990–2000
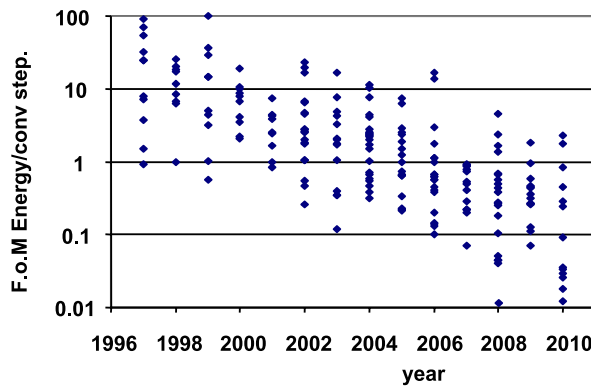
Fig. 12.13 Figure of merit for various analog-to-digital converters as published on International Solid-State Circuits conferences. (From: B. Murmann, "ADC Performance Survey 1997–2010," [Online]. Available: http://www.stanford.edu/~murmann/adcsurvey.html)

The dominant difference between the theoretical F.o.M. and the practically chosen F.o.M. is in the exponent of the signal-to-noise ratio or the effective number of bits. Although a solid analysis is lacking, part of the explanation can be found in the observation that the theoretical formula allows any choice of currents and voltages needed to form the power. In practice the voltage choice is restricted to the available power supply, which may explain the multiplier of the *ENOB* term.

This F.o.M. allows to compare various converter principles and resolution/bandwidth combinations. If the F.o.M is plotted over time as in Figs. 12.12 and 12.13,

Fig. 12.14 In this bandwidth/resolution field the global specifications of some consumer and communication analog-to-digital interfaces are shown. Also a power estimation based on a F.o.M. of 1 pJ/conv step is indicated
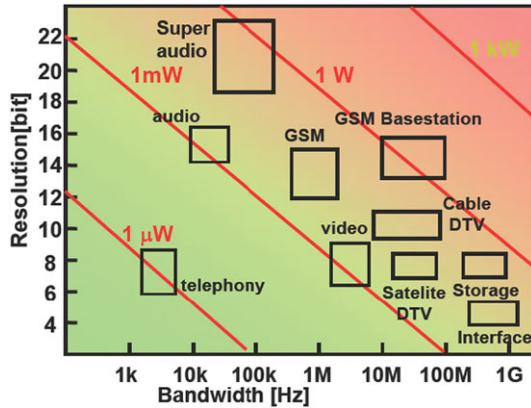
a clear reduction of power for a certain specification is observed. This plot implies a rate of 1 bit per 3 years improvement.[2]

Next to a comparative function the F.o.M. also can be used to predict the conversion power for a specific architecture choice. In the year 2010 an efficient converter uses according to Fig. 12.13 less than 1 pJ per conversion step. This result is obtained by comparing various analog-to-digital converter architectures in various stages of industrialization. It may be useful to limit to just one architecture and compare equivalent stages of development.

This value of F.o.M. = 1 pJ/conv can now be used to calculate the allowable power for a design target. Of course this estimate is based on some crude assumptions and is merely an indication for the order of magnitude that one can expect.

$$\text{Estimated power} = \text{F.o.M.} \times 2BW \times 2^{\text{ENOB}} \qquad (12.5)$$

Figure 12.14 shows the projected power dissipation in a field spanned by a resolution and bandwidth axis. Moving the analog-to-digital conversion from e.g. direct telephony speech level to the GSM baseband digital level means a shift from (8b/3 kHz) to (12b/200 kHz), but also costs three orders of magnitude in power consumption.

12.3 Limits of Conversion

The previous metrics also allow some thoughts on the potential limits of analog-to-digital conversion. Figure 12.15 combines a number of limits that have been discussed. Thermal noise is in any system the fundamental lower limit. In Fig. 12.15 the thermal noise is plotted as formulated in the denominator of (12.3). The graphs assumes a hypothetical analog-to-digital converter based on just one resistive component that carries the signal and that generates thermal noise.

---

[2]Compared to Moore's law for digital circuit where speed doubles and area and power halves for every generation (2 years) this is a maigre result.
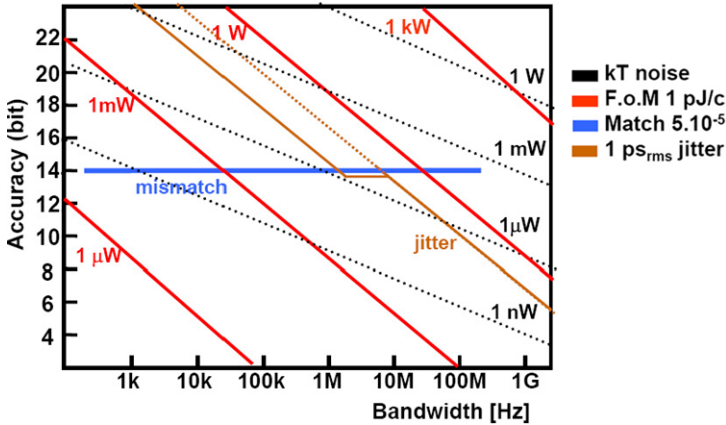
Fig. 12.15 A few limits to analog-to-digital conversion: noise, matching and jitter are compared to the F.o.M. values

At lower frequencies and high resolutions the (fitted) curves of the Figure-of Merit come closer to this thermal noise lower boundary. Yet at 20-bit there are three orders of magnitude difference on power level.

Two more lines are indicating limits. Components such as well-designed resistors, can be matched up to a 14-bit accuracy level or $2-5 \times 10^{-5}$. This limit poses a restriction on the best achievable accuracy relying on components. In order to improve this accuracy data-weighted averaging, dynamic element or similar techniques are requires. Or, component matching is simply removed from the comparison by means of (one-bit) sigma-delta modulation. These techniques rely on time accuracy and are limited by jitter. The noise introduced by jitter in the sampling process, creates a diagonal barrier in the plot. Above 14 bit accuracy, single bit feedback is assumed with the pronounced jitter influence (Fig. 7.26).

As system integration demands more aggressive combinations of accuracy and bandwidth, it remains necessary to improve on component matching techniques and on jitter performance. Next to that the enhancement of power efficiency will remain a continuing quest in analog-to-digital conversion.

Exercises

Chapter 2.

1. A resistor of 100 Ω is connected in parallel to a resistor of 500 Ω. What is the equivalent resistance? What is the noise spectral density in V/$\sqrt{\text{Hz}}$? The 100 Ω resistor heated to 127 °C, the 500 Ω resistor is cooled to -173 °C. What is the noise spectral density in V/$\sqrt{\text{Hz}}$?

2. A capacitor with value C is discharged (0 V). A second capacitor of the same size is charged to a voltage V. What is the total energy stored? Both capacitors are connected in parallel. What is the voltage over the two capacitors? What is the total energy stored? Where is the lost energy?

Chapter 3.

1. A sampling system runs at 10 Ms/s, the input is uncorrelated white noise with a total effective amplitude of 1 mVrms in the band limited to 120 MHz. What is the noise density? What is the rms-amplitude after sampling?

2. The input stage of a sampling system is generating distortion; plot the first five harmonics in the band between 0 and the sampling frequency of 100 Ms/s if an input signal of 3.3 MHz is applied. Once more, but now with 10 Ms/s sampling.

Chapter 4.

1. If the third order signal component of a sample-and-hold is given by AC_{hold}^3 and the noise signal level as $BC_{\text{hold}}^{-0.5}$ what is the optimum C_{hold} for best SINAD performance?

2. Design a track and hold circuit. The process is 0.18 μm CMOS with a nominal power supply of 1.8 V. The NMOST has a threshold of 0.4 volt and a $\beta_\square = 350$ μA/V^2. The PMOST has a threshold of 0.45 volt and a $\beta_\square = 80$ μA/V^2. A generator with an impedance of 50 Ω generates a 0.2 Volt$_{\text{peak}}$ or 0.4 Volt$_{\text{peak-peak}}$ signal and any DC input level can be set. The output of the track and hold should be able to drive a 1 pF load. The sample rate is 10 Ms/s, and a 10-bit accuracy is needed at the end of the sampling phase. Determine the appropriate architecture, the size of the hold-capacitor, the switch(es), and design the opamp. A bias current of 100 μA is available. If this track-and-hold is used to deglitch a DAC and drive an analog circuit, what should change?

Chapter 5.

1. An ideal 8-bit quantizer samples at 60 Ms/s, in the following digital circuit a filter section limits the band of interest to 1 MHz, what is the expected signal to noise ratio in this band due to the quantization white noise. What combinations of resolution and sampling speeds are possible for digital representation of this signal next to an 8-bit sample at 60 Ms/s.

2. A 10-bit ADC is not perfect: at the desired signal and sampling frequency, the DNL is 0.7 bit, while a 2nd order distortion component folds back at a relative amplitude of -56 dB, moreover a fixed clock component at 1/3 of the sampling rate appears at -60 dB. Calculate the effective bits of this ADC. Advise whether the LSB can be removed, so that the succeeding processing runs with 9-bit samples.

3. Suppose an INL spec of 1 bit is given, is there a limit to the DNL spec? And the other way around?

Chapter 7.

1. In an R-2R ladder each resistor has an uncorrelated standard deviation of 1%. If a DA converter is constructed, with this ladder, how many bits of resolution can be processed without running into monotonicity problems.

2. 65 current sources of 0.1 mA each are arranged in a line and connected with their negative terminal to an aluminum return wire. This wire shows a 0.1 Ω impedance between two adjacent current sources. So the total line impedance is 6.4 Ω. Calculate the maximum voltage drop over this return line, if:

– the return line is connected to ground on one extreme
– the return line is connected to ground on both extremes
– the return line is connected to ground in the middle.

3. A standard bandgap configuration has a diode ratio of 8 in a 0.18 μm technology. What is the minimum size of the input transistors of the opamp to keep the 1-sigma spread of the output voltage smaller than 1%.

4. A segmented current DAC is build with current sources of 1 sigma = 1% random mismatch, independent of the current value. How many bits can be designed as a binary section when a DNL of 0.5 LSB must be reach for 99.7% (or -3 to $+3$ sigma) of the produced devices.

5. A unary current matrix with 1024 current sources produces a maximum current of 20 mA in a 50 Ω load. The current sources are build in 0.18 μm CMOS with a gate length of 4 μm with an output impedance of 100 kΩ. A single-transistor cascode stage is used. What is the total output impedance of one current source if the gate cascode transistor measures 1 μ/0.18 μ? What will be the distortion (THD) if a maximum amplitude sine wave is applied. What must be done to reduce the distortion to -60 dB. If a parasitic capacitance of 100 fF is present parallel to each current source, what will be the frequency where the distortion is raised by 3 dB?

Chapter 8.

1. In a sub-range ADC the buffers between the coarse and the fine ladder have 3 mV and -2 mV offsets. Sketch the resulting INL for a 10 bit 2 V ADC in case

the buffers are switched directly or in a "monkey" way. What is in both cases the largest DNL?

2. Compare at flash converters, successive approximation converters and dual-slope converters with respect to DNL, INL and absolute accuracy in case of comparator threshold mismatch.

3. A differential NMOST input pair ($W/L = 50/2$) is loaded with a PMOST current mirror ($W/L = 36/1$ μm) in a process with $A_{VT,N} = A_{VT,P} = 6$ mV μm, and the ratio between the beta square (for $W/L = 1$) is $N/P = 3$. Calculate the input referred mismatch.

4. In an AD converter with an input range of 1 V the comparator can have an input referred error of maximum/minimum $+2/-2$ mV. What is the best resolution a full-flash converter can reach if a DNL of maximum 0.5 LSB is required? Which converter type could reach a better resolution?

5. In an IC process input pairs (as used in comparators, gain stages, etc.) suffer from sigma-$V_{in} = 50$ mV maximum uncorrelated errors, while resistors can be made with 0.1% accuracy. Which ADC architectures can be made advantageously in this process (give resolution indication).

6. What happens to an external input signal near to the chopping frequency? And what happens to an input signal close to the DEM frequency?

Chapter 9.

1. Determine a suitable order of the noise shape filter and oversampling ratio to obtain 100 dB SNR gain.

2. In a noise shaper the feed back path function $J(z) = z^{-2}$. Sketch the noise transfer function of this noise shaper.

3. What efficiency can be expected from a PWM class-D output stage for 4 Ω load impedance with 0.2 Ω resistance per switch, 2 nF switch gate capacitance, and 0.5 MHz clock frequency.

4. What signal indicates "instability" in a sigma-delta converter? What measures can be taken to recover from instability?

5. A second-order sigma delta modulator runs at a sampling rate of 5 Ms/s. A signal band of 20 kHz is required. The comparator and DAC can be assumed ideal, with only the gain uncertainty and the sampling uncertainty of the comparator needs to be taken into account. Set-up a 2nd order filter using capacitors, resistors and ideal opamps, that gives a maximum SNR ratio in this band.

Chapter 10.

1. Testing an ADC during 10 μs at a sampling speed of 20 MHz, the performance at 3 MHz signal frequencies is measured. Calculate an appropriate set of test conditions.

2. A 10 bit ADC needs to be tested dynamically at 40 Ms/s in a DSP based environment. (a) Determine the minimum test time needed to have accessed all codes. (b) Why is it important that all codes have been accessed? There is 1 ms test time available for the FFT. (c) Determine the input frequency at the Nyquist edge for a

good test. (d) What is the number of bins in the FFT? (e) Determine the approximate noise level of the FFT. (f) What can be the technical disadvantage of a long test time?

Chapter 11.

1. In a simple flash converter the size of the input pair of transistors is 20/5 in a process with $A_{VT} = 15$ mV μm. The expected input signal is 2 Volt$_{\text{peak-peak}}$. What resolution limit (monotonicity) do you expect?

2. What improvement in DNL can you expect for a matching-sensitive ADC if the oxide thickness of a process reduces by a factor 2? And what is the improvement if also the power supply (and signal amplitude) decreases by a factor 2?

Chapter 12.

1. Find on the web, or in a data sheet two AD converters: one high speed and one high resolution. Determine the figure of merit.

Bibliography

1 Introduction

1. D. Seitzer, G. Pretzl, N.A. Hamdy, *Electronic Analog-to-Digital Conversion* (Wiley-Interscience, New York, 1983), ISBN: 0-471-90198-9
2. R. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters* (Kluwer Academic Publishers, Dordrecht, 1994), ISBN: 0-7923-9436-4, 2nd edn. (2003), ISBN: 1-4020-7500-6
3. B. Razavi, *Principles of Data Conversion System Design* (Wiley–IEEE Press, New York, 1994), ISBN: 978-0-7803-1093-3
4. P. Jespers, *Integrated Converters D-to-A and A-to-D Architectures, Analysis and Simulation* (Oxford Press, London, 2001), ISBN: 0-19-856446-5
5. F. Maloberti, *Data Converters* (Springer, Berlin, 2007), ISBN: 0-38-732485-2

2 Components and Definitions

6. R.D. Carmichael, E.R. Smith, *Mathematical Tables and Formulas* (Dover Publications, New York, 1931), ISBN: 486-60111-0
7. M. Abramovic, I.A. Stegun (eds.), *Handbook of Mathematical Functions* (Dover Publications, New York, 1965), ISBN: 0-486-61272-4
8. W.H. Beyer, *CRC Standard Mathematical Tables*, 28th edn. (CRC Press, Boca Raton, 1987), ISBN: 0-8493-0628-0
9. W.J. Rey, *Introduction to Robust and Quasi-robust Statistical Methods* (Springer-Verlag, Berlin, 1983), ISBN: 0-387-12866-2
10. A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* student edn. (McGrawHill, New York, 1965). 4th edn. (McGrawHill, New York, 2001), ISBN: 0-073-66011-6
11. R.C. Weast (ed.), *CRC Handbook of Chemistry and Physics*, 64th edn. (CRC Press, Boca Raton, 1984), ISBN: 0-8493-0464-4
12. L.J. van der Pauw, A method of measuring specific resistivity and Hall effect of discs of arbitrary shape. Philips Research Reports **13**, 1–9 (1958)
13. S.M. Sze, *Physics of Semiconductor Devices*, 2nd edn. (John Wiley & Sons, New York, 1981), 3rd edn. (2006), ISBN: 978-0-471-14323-9
14. J.R. Black, Electromigration: A brief survey and some recent results. IEEE Transactions on Electron Devices **ED-16**, 338–347 (1969)

15. E.T. Ogawa, K.-D. Lee, V.A. Blaschke, P.S. Ho, Electromigration reliability issues in dual-damascene Cu interconnections. IEEE Transactions on Reliability **51**, 403–419 (2002)

16. A. van der Ziel, *Noise in Solid-State Devices and Circuits* (Wiley-Interscience, New York, 1986), ISBN: 0-471-832340

17. R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman Lectures on Physics, vols. 1, 2, and 3* (Addison-Wesley, Reading, 1977), ISBN: 0-201-02010-6-H, sixth printing

18. E.B. Rosa, The self and mutual inductances of linear conductors. Bulletin of the Bureau of Standards **4**, 301–344 (1908)

19. ITRS, *The national technology roadmap for semiconductors, technology needs*, 1994–2009. Updates: http://www.itrs.net

20. J.Y.W. Seto, The electrical properties of polycrystalline silicon films. Journal of Applied Physics **46**, 5247–5254 (1975)

21. M. Bely et al., Capacitive integrated circuit structure, US patent 7-170-178, 2007

22. C. Wei, R.F. Barrington, J.R. Mautz, T.K. Sarkar, Multiconductor transmission lines in multi-layered dielectric media. IEEE Transactions on Microwave Theory and Techniques **32**, 439–450 (1984)

23. J.L. McCreary, Matching properties, and voltage and temperature dependence of MOS capacitors. IEEE Journal of Solid-State Circuits **16**, 608–616 (1981)

24. J.-B. Shyu, G.C. Temes, K. Yao, Random errors in MOS capacitors. IEEE Journal of Solid-State Circuits **17**, 1070–1076 (1982)

25. R. Aparicio, Capacity limits and matching properties of integrated capacitors. IEEE Journal of Solid-State Circuits **27**, 384–393 (2002)

26. H.C. de Graaff, F.M. Klaassen, *Compact Transistor Modeling for Circuit Design* (Springer-Verlag, Vienna, 1990), ISBN: 3-211-82136-8

27. M. Vertregt, The analog challenge of nanometer CMOS, in *Technical Digest International Electron Devices Meeting* (2006), pp. 1–8

28. M.J.M. Pelgrom, A.C.J. Duinmaijer, A.P.G. Welbers, Matching properties of MOS transistors. IEEE Journal of Solid-State Circuits **24**, 1433–1440 (1989)

29. P.H. Woerlee, M.J. Knitel, R. van Langevelde, D.B.M. Klaassen, L.F. Tiemeijer, A.J. Scholten, A.T.A. Zegers-van Duijnhoven, RF-CMOS performance trends. IEEE Transactions on Electron Devices **48**, 1776–1782 (2001)

30. A.J. Scholten, L.F. Tiemeijer, P.W.H. De Vreede, D.B.M. Klaassen, A large signal non-quasi-static MOS model for RF circuit simulation, in *Technical Digest International Electron Devices Meeting* (1999), pp. 163–166

31. J.D. Meindl, Low power microelectronics: retrospect and prospect. Proceedings of the IEEE **83**, 619–635 (1995)

32. F.M. Klaassen, W. Hes, On the temperature coefficient of the MOSFET threshold voltage. Solid-State Electronics **29**, 787–789 (1986)

33. A.J. Scholten, L.F. Tiemeijer, R. van Langevelde, R.J. Havens, A.T.A. Zegers-van Duijnhoven, V.C. Venezia, Noise modeling for RF CMOS circuit simulation. IEEE Transactions on Electron Devices **50**, 618–632 (2003)

34. J.R. Brews, MOSFET hand analysis using BSIM. IEEE Circuits and Devices Magazine **21**, 28–36 (2006)

35. C.C. Enz, F. Krummenacher, E.A. Vittoz, An analytical MOS transistor model valid in all regions of operations and dedicated to low-voltage and low-current applications. Analog Integrated Circuits and Signal Processing Journal **8**, 83–114 (1995)

36. G. Gildenblat, X. Li, W. Wu, H. Wang, A. Jha, R. van Langevelde, G.D.J. Smit, A.J. Scholten, D.B.M. Klaassen, PSP: An advanced surface-potential-based MOSFET model for circuit simulation. IEEE Transactions on Electron Devices **53**, 1979–1993 (2006)

37. T. Sakurai, A.R. Newton, Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. IEEE Journal of Solid-State Circuits **25**, 584–594 (1990)

38. M.S.L. Lee, B.M. Tenbroek, W. Redman-White, J. Benson, M.J. Uren, A physically based compact model of partially depleted MOSFETs for analog circuit stimulation. IEEE Journal of Solid-State Circuits **36**, 110–121 (2001)

39. K. Martin, A. Sedra, Effects of the op amp finite gain and bandwidth on the performance of switched-capacitor filters. IEEE Transactions on Circuits and Systems **CAS-28**, 822–829 (1981)
40. Q. Huang, W. Sansen, Design techniques for improved capacitor area efficiency in switched-capacitor biquads. IEEE Transactions Circuits Systems **CAS-34**, 1590–1599 (1987)
41. P.E. Allen, E. Sanchez-Sinencio, *Switched Capacitor Circuits* (Van Nostrand-Reinhold, New York, 1984), ISBN: 0-4422-0873-1
42. D. Allstot, W. Black, Technological design consideration for monolithic MOS switched-capacitor filtering systems. Proceedings of the IEEE **71**, 967–986 (1983)
43. R. Gregorian, G.C. Temes, *Analog MOS Integrated Circuits for Signal Processing* (John Wiley & Sons, New York, 1986), ISBN: 0-471-09797-7
44. D. Johns, K.W. Martin, *Analog Integrated Circuit Design* (John Wiley & Sons, New York, 1997), ISBN: 0-471-14448-7
45. R. Schreier, J. Silva, J. Steensgaard, G.C. Temes, Design-oriented estimation of thermal noise in switched-capacitor circuits. IEEE Transactions on Circuits and Systems I **15**, 2358–2368 (2005)
46. R.P. Sallen, E.L. Key, A practical method of designing RC active filters. IRE Transactions on Circuit Theory 2 **CT-2**, 74–85 (1955)
47. B. Nauta, *Analog CMOS Filters for Very High Frequencies* (Kluwer Academic Publishers, Dordrecht, 1992), ISBN: 0792392728
48. P.R. Gray, R.G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3rd edn. (John Wiley & Sons, New York, 1993). 4th edn. (John Wiley & Sons, New York, 2001). ISBN: 0-471-32168-0
49. P. Allen, D. Holberg, *CMOS Analog Circuit Design* (Holt, Rinehart & Winston, New York, 1987)
50. J.J.F. Rijns, CMOS low-distortion high-frequency variable-gain amplifier. IEEE Journal of Solid-State Circuits **31**, 1029–1034 (1996)
51. S. Ghandhi, Darlington's compound connection for transistors. IRE Transactions on Circuit Theory **4**, 291–292 (1957), see also U.S. patent 2-663-806
52. K. Bult, G.J.G.M. Geelen, A fast-settling CMOS op amp for SC circuits with 90-dB DC gain. IEEE Journal of Solid-State Circuits **25**, 1379–1384 (1990)
53. B.Y.T. Kamath, R.G. Meyer, P.R. Gray, Relationship between frequency response and set-tling time of operational amplifiers. IEEE Journal of Solid-State Circuits **9**, 347–352 (1974)
54. J. Solomon, The monolithic op amp: a tutorial study. IEEE Journal of Solid-State Circuits **9**, 314–332 (1974)
55. Y. Tsividis, P. Gray, An integrated NMOS operational amplifier with internal compensation. IEEE Journal of Solid-State Circuits **11**, 748–754 (1976)
56. Y. Tsividis, Design consideration in single-channel MOS analog integrated circuits—a tuto-rial. IEEE Journal of Solid-State Circuits **13**, 383–391 (1978)
57. P.R. Gray, R. Meyer, MOS operational amplifier design a tutorial overview. IEEE Journal of Solid-State Circuits **17**, 969–982 (1982)
58. W. Redman-White, A high bandwidth constant gm and slew-rate rail-to-rail CMOS input circuit and its application to analog cells for low voltage VLSI systems. IEEE Journal of Solid-State Circuits **32**, 701–712 (1997)
59. E.M. Cherry, D.E. Hooper, The design of wide-band transistor feedback amplifiers. Proceed-ings of the IEE **110**(2), 375–389 (1963)
60. C. Hermans, M.S.J. Steyaert, A high-speed 850-nm optical receiver front-end in 0.18 μm CMOS. IEEE Journal of Solid-State Circuits **41**, 1606–1614 (2006)
61. D.B. Leeson, A simple model of feedback oscillator noise spectrum. Proceedings of the IEEE **54**, 329–330 (1966)
62. A. Demir, Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and $1/f$ noise. IEEE Transactions on Circuits and Systems I **53**, 1869–1884 (2006)
63. A. Hajimiri, T. Lee, A general theory of phase noise in electrical oscillators. IEEE Journal of Solid-State Circuits **33**, 179–194 (1998)

64. B. Razavi, A study of phase noise in CMOS oscillators. IEEE Journal of Solid-State Circuits **31**, 331–343 (1996)

65. J.D. van der Tang, *High frequency oscillator design for integrated transceivers*, Ph.D. thesis, Technical University Eindhoven, 2002

66. Q. Huang, Phase noise to carrier ratio in LC oscillators. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications **47**, 965–980 (2000)

67. E. Vittoz, M. Degrauwe, S. Bitz, High-performance crystal oscillator circuits: Theory and application. IEEE Journal of Solid-State Circuits **23**, 774–783 (1988)

68. W. Thommen, An improved low-power crystal oscillator, in *25th European Solid-State Circuits Conference* (1999), pp. 146–149

69. P. Geraedts, E. van Tuijl, E. Klumperink, G. Wienk, B. Nauta, A 90 μW 12 MHz relaxation oscillator with a −162 dB FOM, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2008), pp. 348–349

70. F. Sebastiano, L. Breems, K. Makinwa, S. Drago, D. Leenaerts, B. Nauta, A low-voltage mobility-based frequency reference for crystal-less ULP radios. IEEE Journal of Solid-State Circuits **44**, 2002–2009 (2009)

3 Sampling

71. L.R. Rabiner, B. Gold, *Theory and Application of Digital Signal Processing* (Prentice Hall, Inc., Englewood Cliffs, 1975), ISBN: 0-139-141014

72. A.W.M. van den Enden, N.A.M. Verhoeckx, *Discrete Time Signal Processing, an Introduction* (Prentice Hall, New York, 1989), ISBN: 0-132-167557

73. H. Nyquist, Certain topics in telegraph transmission theory. Transactions of the AIEE **47**, 617–644 (1928), reprinted in Proceedings of the IEEE **90**, 280–305 (2002)

74. C.E. Shannon, A mathematical theory of communication. The Bell System Technical Journal **27**, 379–423 and 623–656 (1948)

75. C.E. Shannon, Communication in the presence of noise, *Proceedings of the IRE*, pp. 10–21 (1949), reprinted in Proceedings of the IEEE **86**, 447–457 (1998)

76. M. Unser, Sampling-50 years after Shannon. Proceedings of the IEEE **88**, 569–587 (2000)

77. E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on Information Theory **52**, 489–509 (2006)

78. M. Shinagawa, Y. Akazawa, T. Wakimoto, Jitter analysis of high-speed sampling systems. IEEE Journal of Solid-State Circuits **25**, 220–224 (1990)

79. J.H. McClellan, T.W. Parks, L.R. Rabiner, A computer program for designing optimum FIR linear phase digital filters. IEEE Transactions on Audio Electroacoustics **21**, 506–526 (1973)

80. C.H. Séquin, M.F. Tompsett, *Charge Transfer Devices* (Academic Press, New York, 1975), Supplement 8 to *Advances in Electronics and Electron Physics*

4 Sample-and-Hold

81. J. Crols, M. Steyaert, Switched-opamp: An approach to realize full CMOS switched-capacitor circuits at very low power supply voltages. IEEE Journal of Solid-State Circuits **29**, 936–942 (1994)

82. A. Baschirotto, R. Castello, A 1-V 1.8-MHz CMOS switched-opamp SC filter with rail-to-rail output swing. IEEE Journal of Solid-State Circuits **32**, 1979–1986 (1997)

83. A. Keramat, Z. Tao, A capacitor mismatch and gain insensitive 1.5-bit/stage pipelined A/D converter, in *Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems* (2000), pp. 48–51

84. J. Dickson, On-chip high-voltage generation MNOS integrated circuits using an improved voltage multiplier technique. IEEE Journal of Solid-State Circuits **11**, 374–378 (1976)

85. R.W. Knepper, Dynamic depletion mode: An E/D mosfet circuit method for improved performance. IEEE Journal of Solid-State Circuits **13**, 542–548 (1978)

86. A.M. Abo, P.R. Gray, A 1.5-V, 10-bit, 14.3-MS/s CMOS pipe-line analog-to-digital converter. IEEE Journal of Solid-State Circuits **34**, 599–606 (1999)

87. S. Limotyrakis, S.D. Kulchycki, D.K. Su, B.A. Wooley, A 150-MS/s 8-b 71-mW CMOS time-interleaved ADC. IEEE Journal of Solid-State Circuits **40**, 1057–1067 (2005)

88. B.A. Song, M.F. Tompsett, K.R. Lakshmikumar, A 12-bit 1-MS/s capacitor error-averaging pipelined A/D converter. IEEE Journal of Solid-State Circuits **23**, 1324–1333 (1988)

89. W. Yang, D. Kelly, I. Mehr, M.T. Sayuk, L. Singer, A 3-V 340-mW 14-b 75-MS/s CMOS ADC with 85-dB SFDR at Nyquist input. IEEE Journal of Solid-State Circuits **36**, 1931–1936 (2001)

90. P. Vorenkamp, J.P.M. Verdaasdonk, Fully bipolar, 120-MS/s 10-b track-and-hold circuit. IEEE Journal of Solid-State Circuits **27**, 988–992 (1992)

5 Quantization

91. A. Harley Reeves, Electric signaling system, U.S. Patent 2-272-070, issued February 3, 1942. Also French Patent 852-183 issued 1938, and British Patent 538-860 issued 1939

92. IEEE Std 1057–1994, *IEEE Standard for Digitizing Waveform Recorders* (1994)

93. IEEE 1241–2000, *Standard for Terminology and Test Methods for Analog-to-Digital Converters*, IEEE Std1241 (IEEE, New York, 2000), ISBN: 0-7381-2724-8, revision 2007

94. S.J. Tilden, T.E. Linnenbrink, P.J. Green, Overview of IEEE-STD-1241 standard for terminology and test methods for analog-to-digital converters, in *Instrumentation and Measurement Technology Conference* (1999), pp. 1498–1503

95. W.R. Bennett, Spectra of quantized signals. Bell System Technical Journal **27**, 446–472 (1948)

96. N. Blachman, The intermodulation and distortion due to quantization of sinusoids. IEEE Transactions on Acoustics, Speech and Signal Processing **ASSP 33**, 1417–1426 (1985)

97. M.S. Oude Alink, A.B.J. Kokkeler, E.A.M. Klumperink, K.C. Rovers, G. Smit, B. Nauta, Spurious-free dynamic range of a uniform quantizer. IEEE Transactions on Circuits and Systems II: Express Briefs **56**, 434–438 (2009)

98. S. Lloyd, Least squares quantization in PCM. IEEE Transactions on Information Theory **28**, 129–137 (1982) (transcript from 1957 paper)

99. J. Max, Quantizing for minimum distortion. IRE Transactions on Information Theory **6**, 7–12 (1960)

100. R.A. Wannamaker, S.P. Lipshitz, J. Vanderkooy, J.N. Wright, A theory of nonsubtractive dither. IEEE Transactions on Signal Processing **48**, 499–516 (2000)

6 Reference Circuits

101. D. Hilbiber, A new semiconductor voltage standard, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (1964), pp. 32–33

102. R.J. Widlar, New developments in IC voltage regulators. IEEE Journal of Solid-State Circuits **6**, 2–7 (1971)

103. K.E. Kuijk, A precision reference voltage source. IEEE Journal of Solid-State Circuits **8**, 222–226 (1973)

104. M.A.P. Pertijs, J.H. Huijsing, *Precision Temperature Sensors in CMOS Technology* (Springer, Berlin, 2006), ISBN: 140205257X

105. B.S. Song, P.R. Gray, A precision curvature-compensated CMOS bandgap reference. IEEE Journal of Solid-State Circuits **18**, 634–643 (1983)

106. H. Banba, H. Shiga, A. Umezawa, T. Miyaba, T. Tanzawa, S. Atsumi, K. Sakui, A CMOS band-gap reference circuit with sub 1-V operation. IEEE Journal of Solid-State Circuits **34**, 670–674 (1999)

107. V. Petrescu, M.J.M. Pelgrom, H.J.M. Veendrick, P. Pavithran, J. Wieling, Monitors for a signal integrity measurement system, in *32nd European Solid-State Circuits Conference* (2006), pp. 122–125

108. V. Petrescu, M.J.M. Pelgrom, H.J.M. Veendrick, P. Pavithran, J. Wieling, A signal-integrity self-test concept for debugging nanometer CMOS ICs, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2006), pp. 544–545

109. C.J.B. Fayomi, G.I. Wirth, H.F. Achigui, A. Matsuzawa, Sub 1 V CMOS bandgap reference design techniques: a survey. Analog Integrated Circuits and Signal Processing **62**, 141–157 (2010)

110. A.-J. Annema, Low-power bandgap references featuring DTMOSTs. IEEE Journal of Solid-State Circuits **34**, 949–955 (1999)

111. R.A. Blauschild, P.A. Tucci, R.S. Muller, R.G. Meyer, A new NMOS temperature-stable voltage reference. IEEE Journal of Solid-State Circuits **13**, 767–774 (1978)

112. H.-J. Song, C.-K. Kim, A temperature-stabilized SOI voltage reference based on threshold voltage difference between enhancement and depletion NMOSFET's. IEEE Journal of Solid-State Circuits **28**, 671–677 (1993)

7 Digital-to-Analog Conversion

113. R.J. van de Plassche, Dynamic element matching for high-accuracy monolithic D/A converters. IEEE Journal of Solid-State Circuits **21**, 795–800 (1976)

114. J.A. Schoeff, An inherently monotonic 12 bit DAC. IEEE Journal of Solid-State Circuits **24**, 904–911 (1979)

115. J.R. Naylor, A complete high-speed voltage output 16-bit monolithic DAC. IEEE Journal of Solid-State Circuits **28**, 729–735 (1983)

116. H.J. Schouwenaars, E.C. Dijkmans, B.M.J. Kup, E.J.M. van Tuijl, A monolithic dual 16-bit D/A converter. IEEE Journal of Solid-State Circuits **21**, 424–429 (1986)

117. D.W.J. Groeneveld, H.J. Schouwenaars, H.A.H. Termeer, C.A.A. Bastiaansen, A self-calibration technique for monolithic high-resolution D/A converters. IEEE Journal of Solid-State Circuits **24**, 1517–1522 (1989)

118. H.J. Schouwenaars, D.W.J. Groeneveld, H.A.H. Termeer, A low-power stereo 16-bit CMOS D/A converter for digital audio. IEEE Journal of Solid-State Circuits **23**, 1290–1297 (1988)

119. C.-H. Lin, K. Bult, A 10-b 250-M sample/s CMOS DAC in 1 mm^2, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (1998), pp. 214–215

120. A. Van Den Bosch, M. Borremans, M. Steyaert, W. Sansen, A 12 b 500 MS/s current-steering CMOS D/A converter, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2001), pp. 366–367

121. A. Van Den Bosch, M. Borremans, M. Steyaert, W. Sansen, A 10-bit 1-G sample/s Nyquist current-steering CMOS D/A converter. IEEE Journal of Solid-State Circuits **36**, 315–324 (2001)

122. C.A.A. Bastiaansen, D.W.J. Groeneveld, H.J. Schouwenaars, H.A.H. Termeer, A 10-b 40-MHz 0.8-μm CMOS Current-Output D/A Converter. IEEE Journal Solid-State Circuits **26**, 917–921 (1991)

123. K. Doris, J. Briaire, D. Leenaerts, M. Vertregt, A. van Roermund, A 12 b 500 MS/s DAC with >70 dB SFDR up to 120 MHz in 0.18 μm CMOS, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2005), pp. 116–588

124. G.A.M. Van der Plas, J. Vandenbussche, W. Sansen, M.S.J. Steyaert, G.G.E. Gielen, A 14-bit intrinsic accuracy Q^2 random walk CMOS DAC. IEEE Journal of Solid-State Circuits **34**, 1708–1718 (1999)

125. B. Jewett, J. Liu, K. Poulton, A 1.2 GS/s 15 b DAC for precision signal generation, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2005), pp. 110–111

126. B. Schafferer, R. Adams, A 3V CMOS 400 mW 14 b 1.4 GS/s DAC for multi-carrier applications, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2004), pp. 360–361

127. W. Schofield, D. Mercer, L.S. Onge, A 16 b 400 MS/s DAC with < -80 dBc IMD to 300 MHz and < -160 dBm/Hz noise power spectral density, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2003), pp. 126–127

128. D.K. Su, B.A. Wooley, A CMOS oversampling D/A converter with a current-mode semidigital reconstruction filter. IEEE Journal of Solid-State Circuits **28**, 1224–1233 (1993)

129. D.B. Barkin, A.C.Y. Lin, D.K. Su, B.A. Wooley, A CMOS oversampling bandpass cascaded D/A converter with digital FIR and current-mode semi-digital filtering. IEEE Journal of Solid-State Circuits **39**, 585–593 (2004)

130. R.E. Suarez, P.R. Gray, D.A. Hodges, All-MOS charge-redistribution analog-to-digital conversion techniques. II. IEEE Journal of Solid-State Circuits **10**, 379–385 (1975)

131. B.-S. Song, S.-H. Lee, M.F.A. Tompsett, 10-b 15-MHz CMOS recycling two-step A/D converter. IEEE Journal of Solid-State Circuits **25**, 1328–1338 (1990)

132. K. Philips, J. van den Homberg, C. Dijkmans, PowerDAC: A single-chip audio DAC with a 70%-efficient power stage in 0.5 μm CMOS, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (1999), pp. 154–155

133. P.G. Blanken, S.E.J. Menten, A 10 μV-offset 8 kHz bandwidth 4th-order chopped ΣΔ A/D converter for battery management, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2002), pp. 388–389

134. O.J.A.P. Nys, R.K. Henderson, An analysis of dynamic element matching techniques in sigma-delta modulation, in *IEEE International Symposium on Circuits and Systems* (1996), pp. 231–234

135. R.K. Henderson, O.J.A.P. Nys, Dynamic element matching techniques with arbitrary noise shaping function, in *IEEE International Symposium on Circuits and Systems* (1996), pp. 293–296

136. L. Carley, A noise-shaping coder topology for 15+ bit converters. IEEE Journal of Solid-State Circuits **24**, 267–273 (1989)

137. M.J. Story, Digital to analog converter adapted to select input sources based on a preselected algorithm once per cycle of a sampling signal, US patent 5-138-317, 1992

138. A. Maloberti, *Convertitore digitale analogico sigma-delta multilivello con matching dinamico degli elementi*, Tesi di Laurea, Universita degli Studi di Pavia, 1990–1991 (this thesis was not available)

139. M.R. Miller, C.S. Petrie, A multibit sigma-delta ADC for multimode receivers. IEEE Journal of Solid-State Circuits **38**, 475–482 (2003)

140. A.G.F. Dingwall, V. Zazzu, An 8-MHz CMOS subranging 8-bit A/D converter. IEEE Journal of Solid-State Circuits **20**, 1138–1143 (1985)

141. A. Abrial, J. Bouvier, J. Fournier, P. Senn, M. Viellard, A 27-MHz digital-to-analog video processor. IEEE Journal of Solid-State Circuits **23**, 1358–1369 (1988)

142. M.J.M. Pelgrom, A 10 b 50 MHz CMOS D/A converter with 75 Ω buffer. IEEE Journal of Solid-State Circuits **25**, 1347–1352 (1990)

143. T. Miki, Y. Nakamura, M. Nakaya, S. Asai, Y. Akasaka, Y. Horiba, An 80-MHz 8-bit CMOS D/A converter. IEEE Journal of Solid-State Circuits **21**, 983–988 (1986)

144. M.J.M. Pelgrom, M. Roorda, An algorithmic 15 bit CMOS digital-to-analog converter. IEEE Journal of Solid-State Circuits **23**, 1402–1405 (1988)

145. H. Matsumoto, K. Watanabe, Switched-capacitor algorithmic digital-to-analog converters. IEEE Transactions on Circuits and Systems **33**, 721–724 (1986)

8 Analog-to-Digital Conversion

146. H.L. Fiedler, B. Hoefflinger, W. Demmer, P. Draheim, A 5-bit building block for 20 MHz A/D converters. IEEE Journal of Solid-State Circuits **26**, 151–155 (1981)
147. J.-T. Wu, B.A. Wooley, A 100-MHz pipelined CMOS comparator. IEEE Journal of Solid-State Circuits **23**, 1379–1385 (1988)
148. B. Nauta, A.G.W. Venes, A 70 Ms/s 110 mW 8-b CMOS folding and interpolating A/D converter. IEEE Journal of Solid-State Circuits **30**, 1302–1308 (1995)
149. A.G.W. Venes, R.J. van de Plassche, An 80-MHz, 80-mW, 8-b CMOS folding A/D converter with distributed track-and-hold preprocessing. IEEE Journal of Solid-State Circuits **31**, 1846–1853 (1996)
150. W. Ellersick, K.Y. Chih-Kong, M. Horowitz, W. Dally GAD, A 12-GS/s CMOS 4-bit A/D converter for an equalized multi-level link, in *Symposium on VLSI Circuits, Digest of Technical Papers* (1999), pp. 49–52
151. J. Montanaro et al., A 160 MHz, 32 b, 0.5 W CMOS RISC microprocessor. IEEE Journal of Solid-State Circuits **31**, 1703–1714 (1996)
152. B. Verbruggen, J. Craninckx, M. Kuijk, P. Wambacq, G. Van der Plas, A 2.2 mW 5 b 1.75 GS/s folding flash ADC in 90 nm digital CMOS, in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers* (2008), pp. 252–611
153. N. Fukushima, T. Yamada, N. Kumazawa, Y. Hasegawa, M. Soneda, A CMOS 40 MHz 8 b 105 mW two-step ADC, in *International Solid-State Circuits Conference, Digest of Technical Papers* (1989), pp. 14–15
154. J.H. Atherton, H.T. Simmonds, An offset reduction technique for use with CMOS integrated comparators and amplifiers. IEEE Journal of Solid-State Circuits **27**, 1168–1175 (1992)
155. H. Reyhani, P. Quinlan, A 5 V, 6-b, 80 Ms/s BiCMOS flash ADC. IEEE Journal of Solid-State Circuits **29**, 873–878 (1994)
156. P. Vorenkamp, J.P.M. Verdaasdonk, A 10 b 50 MHz pipelined ADC, in *International Solid-State Circuits Conference, Digest of Technical Papers* (1992), pp. 32–33
157. K. Kattmann, J. Barrow, A technique for reducing differential nonlinearity errors in flash A/D converters, in *International Solid-State Circuits Conference, Digest of Technical Papers* (1991), pp. 170–171
158. P.C.S. Scholtens, M. Vertregt, A 6-b 1.6-G sample/s flash ADC in 0.18 μm CMOS using averaging termination. IEEE Journal of Solid-State Circuits **37**, 1599–1609 (2002)
159. K. Bult, A. Buchwald, An embedded 240-mW 10-b 50-MS/s CMOS ADC in 1-mm^2. IEEE Journal of Solid-State Circuits **32**, 1887–1895 (1997)
160. K. Uyttenhove, J. Vandenbussche, E. Lauwers, G.G.E. Gielen, M.S.J. Steyaert, Design techniques and implementation of an 8-bit 200-MS/s interpolating/averaging CMOS A/D converter. IEEE Journal of Solid-State Circuits **38**, 483–494 (2003)
161. R.E.J. Van De Grift, I.W.J.M. Rutten, M. van der Veen, An 8-bit video ADC incorporating folding and interpolation techniques. IEEE Journal of Solid-State Circuits **22**, 944–953 (1987)
162. P. Vorenkamp, R. Roovers, A 12-b, 60-MS/s cascaded folding and interpolating ADC. IEEE Journal of Solid-State Circuits **32**, 1876–1886 (1997)
163. R.J. Van De Plassche, R.E.J. van der Grift, A high-speed 7 bit A/D converter. IEEE Journal of Solid-State Circuits **14**, 938–943 (1979)
164. G. Hoogzaad, R. Roovers, A 65-mW, 10-bit, 40-MS/s BiCMOS Nyquist ADC in 0.8 mm^2. IEEE Journal of Solid-State Circuits **34**, 1796–1802 (1999)

165. M.J. Choe, B.-S. Song, K. Bacrania, A 13 b 40 MS/s CMOS pipelined folding ADC with background offset trimming, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2000), pp. 36–37

166. S.H. Lewis, P.R. Gray, A pipelined 5-MS/s 9-bit analog-to-digital converter. IEEE Journal of Solid-State Circuits **22**, 954–961 (1987)

167. H. van der Ploeg, M. Vertregt, M. Lammers, A 15-bit 30-MS/s 145-mW three-step ADC for imaging applications. IEEE Journal of Solid-State Circuits **41**, 1572–1577 (2006)

168. K. Kusumoto, A. Matsuzawa, K. Murata, A 10-b 20-MHz 30-mW pipelined interpolating CMOS ADC. IEEE Journal of Solid-State Circuits **28**, 1200–1206 (1993)

169. M. Haas, D. Draxelmayr, F. Kuttner, B. Zojer, A monolithic triple 8-bit CMOS video coder. IEEE Transactions on Consumer Electronics **36**, 722–729 (1990)

170. C. Moreland, F. Murden, M. Elliott, J. Young, M. Hensley, R. Stop, A 14 b 100 MS/s sub-ranging ADC. IEEE Journal of Solid-State Circuits **35**, 1791–1798 (2000)

171. R. McCharles, D. Hodges, Charge circuits for analog LSI. IEEE Transactions on Circuits and Systems **25**, 490–497 (1978)

172. P.W. Li, M.J. Chin, P.R. Gray, R. Castello, A ratio-independent algorithmic analog-to-digital conversion technique. IEEE Journal of Solid-State Circuits **19**, 828–836 (1984)

173. A.N. Karanicolas, H.-S. Lee, K.L. Barcrania, A 15-b 1-MS/s digitally self-calibrated pipeline ADC. IEEE Journal of Solid-State Circuits **28**, 1207–1215 (1993)

174. K. Nagaraj, H.S. Fetterman, J. Anidjar, S.H. Lewis, R.G. Renninger, A 250-mW, 8-b, 52-MSs/s parallel-pipelined A/D converter with reduced number of amplifiers. IEEE Journal of Solid-State Circuits **32**, 312–320 (1997)

175. Y. Chiu, P.R. Gray, B. Nikolic, A 14-b 12-MS/s CMOS pipeline ADC with over 100-dB SFDR. IEEE Journal of Solid-State Circuits **39**, 2139–2151 (2004)

176. X. Wang, P.J. Hurst, S.H. Lewis, A 12-bit 20-MS/s pipelined analog-to-digital converter with nested digital background calibration. IEEE Journal of Solid-State Circuits **39**, 1799–1808 (2004)

177. B. Murmann, B.E. Boser, A 12-bit 75-MS/s pipelined ADC using open-loop residue amplification. IEEE Journal of Solid-State Circuits **38**, 2040–2050 (2003)

178. E. Iroaga, B. Murmann, A 12-Bit 75-MS/s pipelined ADC using incomplete settling. IEEE Journal of Solid-State Circuits **42**, 748–756 (2007)

179. G. Geelen, E. Paulus, D. Simanjuntak, H. Pastoor, R. Verlinden, A 90 nm CMOS 1.2 V 10 b power and speed programmable pipelined ADC with 0.5 pJ/conversion-step, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2006), pp. 214–215

180. B.M. Min, P. Kim, F.W. Bowman, D.M. Boisvert, A.J. Aude, A 69-mW 10-bit 80-MS/s pipelined CMOS ADC. IEEE Journal of Solid-State Circuits **38**, 2031–2039 (2003)

181. S. Bardsley, C. Dillon, R. Kummaraguntla, C. Lane, A.M.A. Ali, B. Rigsbee, D. Combs, A 100-dB SFDR 80-MSPS 14-bit 0.35-μm BiCMOS pipeline ADC. IEEE Journal of Solid-State Circuits **41**, 2144–2153 (2006)

182. B.G. Lee, B.M. Min, G. Manganaro, J.W. Valvano, A 14 b 100 MS/s pipelined ADC with a merged active S/H and first MDAC, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2008), pp. 248–249

183. H. van de Vel, B. Buter, H. v. d. Ploeg, M. Vertregt, G. Geelen, E. Paulus, A 1.2 V 250-mW 14-b 100 MS/s digitally calibrated pipeline ADC in 90-nm CMOS. IEEE Journal of Solid-State Circuits **44**, 1047–1056 (2009)

184. I. Mehr, L. Singer, A 55-mW 10-bit 40-MS/s Nyquist-rate CMOS ADC. IEEE Journal of Solid-State Circuits **35**, 318–323 (2000)

185. T. Sepke, J.K. Fiorenza, C.G. Sodini, P. Holloway, H.-S. Lee, Comparator-based switched-capacitor circuits for scaled CMOS technologies, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2006), pp. 812–821

186. H. Wang, X. Wang, P.J. Hurst, S.H. Lewis, Nested digital background calibration of a 12-bit pipelined ADC without an input SHA. IEEE Journal of Solid-State Circuits **44**, 2780–2789 (2009)

187. J.L. McCreary, P.R. Gray, All-MOS charge redistribution analog-to-digital conversion techniques I. IEEE Journal of Solid-State Circuits **10**, 371–379 (1975)

188. A. Agnes, E. Bonizzoni, P. Malcovati, F. Maloberti, A 9.4-ENOB 1 V 3.8 μW 100 kS/s SAR ADC with time-domain comparator, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2008), pp. 246–247

189. M. van Elzakker, E. van Tuijl, P. Geraedts, D. Schinkel, E. Klumperink, B. Nauta, A 1.9 μW 4.4 fJ/conversion-step 10 b 1 MS/s charge-redistribution ADC, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2008), pp. 244–245

190. F. Kuttner, A 1.2 V 10 b 20 MS/s non-binary successive approximation ADC in 0.13 μm CMOS, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2002), pp. 176–177

191. B. Ginetti, P. Jespers, A. Vandemeulebroecke, A CMOS 13-b cyclic. A/D converter. IEEE Journal of Solid-State Circuits **27**, 957–964 (1992)

192. M. Mase, S. Kawahito, M. Sasaki, Y. Wakamori, M. Furuta, A wide dynamic range CMOS image sensor with multiple exposure-time signal outputs and 12-bit column-parallel cyclic A/D converters. IEEE Journal of Solid-State Circuits **40**, 2787–2795 (2005)

193. S. Naraghi, M. Courcy, M.P. Flynn, A 9 b 14 μW 0.06 mm^2 PPM ADC in 90 nm digital CMOS, in *IEEE International Solid-State Circuits Conference Digest of Technical Papers* (2009), pp. 168–169

194. B.K. Howard, Binary quantizer, United States patent 2-715-678, 1955

195. H. van der Ploeg, G. Hoogzaad, H.A.H. Termeer, M. Vertregt, R.L.J. Roovers, A 2.5 V, 12 b, 54 MS/s, 0.25 μm CMOS ADC, in *International Solid-State Circuits Conference, Digest of Technical Papers* (2001), pp. 132–133

196. M.J.M. Pelgrom, A. Jochijms, H. Heijns, A CCD delay line for video applications. IEEE Transactions on Consumer Electronics **33**, 603–609 (1987)

197. N. Kurosawa, H. Kobayashi, K. Maruyama, H. Sugawara, K. Kobayashi, Explicit analysis of channel mismatch effects in time-interleaved ADC systems. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications **48**, 261–271 (2001)

198. M. Vertregt, M.B. Dijkstra, A.C. v. Rens, M.J.M. Pelgrom, A versatile digital CMOS video delay line with embedded ADC, DAC and RAM, in *19th European Solid-State Circuits Conference* (1993), pp. 226–229

199. M.J.M. Pelgrom, A.C. v. Rens, M. Vertregt, M.B. Dijkstra, A 25-Ms/s 8-bit CMOS A/D converter for embedded application. IEEE Journal of Solid-State Circuits **29**, 879–886 (1994)

200. B. Murray, H. Menting, A highly integrated D2MAC decoder, in *IEEE International Conference on Consumer Electronics, Digest of Technical Papers* (1992), pp. 56–57

201. J.W. Mark, T.D. Todd, A nonuniform sampling approach to data compression. IEEE Transactions on Communications **29**, 24–32 (1981)

202. E. Allier, J. Goulier, G. Sicard, A. Dezzani, E. Andre, M. Renaudin, A 120 nm low power asynchronous ADC, in *International Symposium on Low-Power Electronics and Design* (2005), pp. 60–65

203. C.-S. Lin, B.-D. Liu, A new successive approximation architecture for low-power low-cost CMOS A/D converter. IEEE Journal of Solid-State Circuits **38**, 54–62 (2003)

204. S.-W.M. Chen, R.W. Brodersen, A 6-bit 600-MS/s 5.3-mW asynchronous ADC in 0.13-μm CMOS. IEEE Journal of Solid-State Circuits **41**, 2669–2680 (2006)

205. J.-P. Jansson, A. Mantyniemi, J. Kostamovaara, A CMOS time-to-digital converter with better than 10 ps single-shot precision. IEEE Journal of Solid-State Circuits **41**, 1286–1296 (2006)

206. P. Chen, S.-L. Liu, J. Wu, A CMOS pulse-shrinking delay element for time interval measurement. IEEE Transactions on Circuits and Systems **47**, 954–958 (2000)

207. T.E. Rahkonen, J.T. Kostamovaara, The use of stabilized CMOS delay lines for the digitization of short time intervals. IEEE Journal of Solid-State Circuits **28**, 887–894 (1993)

208. H. van der Ploeg, *The nonius analog-to-digital converter*, Internal Philips Research Report/University Twente B.Sc. report, supervisor M. Pelgrom, 1997

209. V.Z. Groza, High-resolution floating-point ADC. IEEE Transactions on Instrumentation and Measurement **50**, 1812–1829 (2001)

9 Sigma-delta Conversion

210. F. de Jager, Delta modulation, a method of PCM transmission using the 1-unit code. Philips Research Reports **7**, 442–466 (1952)
211. C. Cutler, Transmission systems employing quantization, U.S. Patent 2-927-962, 1960
212. B. Widrow, A study of rough amplitude quantization by means of Nyquist sampling theory. IRE Transactions on Circuit Theory **CT-3**, 266–276 (1959)
213. H. Inose, Y. Yasuda, J. Murakami, A telemetering system by code modulation-$\Delta\Sigma$ modulation. IRE Trans. Space Electronics and Telemetry **SET-8**, 204–209 (1962), or: Proceedings of the IEEE **51**, 1524–1535 (1963)
214. J.C. Candy, G.C. Temes (eds.), *Oversampling Delta-Sigma Data Converters: Theory, Design and Simulation* (IEEE, New York, 1992)
215. S.R. Norsworthy, R. Schreier, G.C. Temes (eds.), *Delta-Sigma Data Converters: Theory, Design, and Simulation* (IEEE Press, Piscataway, 1997), ISBN: 0-7803-1045-4
216. R. Schreier, G.C. Temes, *Understanding Delta-Sigma Data Converters* (John Wiley & Sons Inc., New York, 2004), ISBN: 0-471-46585-2
217. P.J.A. Naus, E.C. Dijkmans, Multi-bit oversampled $\Sigma\Delta$ A/D converters as front-end for CD players. IEEE Journal of Solid-State Circuits **26**, 905–909 (1991)
218. B.M.J. Kup, E.C. Dijkmans, P.J.A. Naus, J. Sneep, A bit-stream digital-to-analog converter with 18-b resolution. IEEE Journal of Solid-State Circuits **26**, 1757–1763 (1991)
219. P.J.A. Naus, E.C. Dijkmans, E.F. Stikvoort, D.J. Holland, W. Bradinal, A CMOS stereo 16-bit D/A converter for digital audio. IEEE Journal of Solid-State Circuits **22**, 390–395 (1987)
220. R. Adams, K.Q. Nguyen, A 113-dB SNR oversampling DAC with segmented noise-shaped scrambling. IEEE Journal of Solid-State Circuits **33**, 1871–1878 (1998)
221. S. Gambini, J. Rabaey, A 100-kS/s 65-dB DR sigma-delta ADC with 0.65V supply voltage, in *33th European Solid State Circuits Conference* (2007), pp. 202–205
222. W.L. Lee, C.G. Sodini, A topology for higher order interpolative coders, in *Proceedings of the IEEE International Symposium on Circuits and Systems* (1987), pp. 459–462
223. K.C.-H. Chao, S. Nadeem, W.L. Lee, C.G. Sodini, A higher order topology for interpolative modulators for oversampling A/D converters. IEEE Transactions on Circuits and Systems **37**, 309–318 (1990)
224. L.A. Williams III, B.A. Wooley, A third-order sigma-delta modulator with extended dynamic range. IEEE Journal of Solid-State Circuits **29**, 193–202 (1994)
225. L.J. Breems, E.C. Dijkmans, J.H. Huijsing, A quadrature data-dependent DEM algorithm to improve image rejection of a complex $\Sigma\Delta$ modulator. IEEE Journal of Solid-State Circuits **36**, 1879–1886 (2001)
226. E.J. van der Zwan, E.C. Dijkmans, A 0.2 mW CMOS $\Sigma\Delta$ modulator for speech coding with 80 dB dynamic range. IEEE Journal of Solid-State Circuits **31**, 1873–1880 (1996)
227. L.J. Breems, R. Rutten, R.H.M. van Veldhoven, G. van der Weide, A 56 mW continuous-time quadrature cascaded $\Sigma\Delta$ modulator with 77 dB DR in a near zero-IF 20 MHz band. IEEE Journal of Solid-State Circuits **42**, 2696–2705 (2007)
228. R.W. Adams, Design and implementation of an audio 18-bit analog-to-digital converter using oversampling techniques. Journal Audio Engineering Society **34**, 153–166 (1986)
229. S.D. Kulchycki, R. Trofin, K. Vleugels, B.A. Wooley, A 77-dB dynamic range, 7.5-MHz hybrid continuous-time/discrete-time cascaded sigma delta modulator. IEEE Journal of Solid-State Circuits **43**, 796–804 (2008)
230. K.J.P. Philips, $\Sigma\Delta$ *AD conversion for signal conditioning*, Ph.D. thesis, Technical University Eindhoven, 2005
231. K. Philips, P.A.C.M. Nuijten, R.L.J. Roovers, A.H.M. van Roermund, F.M. Chavero, M.T. Pallares, A. Torralba, A continuous-time SD ADC with increased immunity to interferers. IEEE Journal of Solid-State Circuits **39**, 2170–2178 (2004)
232. K. Nguyen, R. Adams, K. Sweetland, H. Chen, A 106-dB SNR hybrid oversampling analog-to-digital converter for digital audio. IEEE Journal of Solid-State Circuits **30**, 2408–2415 (2005)

233. R.H.M. van Veldhoven, B.J. Minnis, H.A. Hegt, A.H.M. van Roermund, A 3.3-mW ΣΔ modulator for UMTS in 0.18-μm CMOS with 70-dB dynamic range in 2-MHz bandwidth. IEEE Journal of Solid-State Circuits **37**, 1645–1652 (2002)

234. P.G.R. Silva, L.J. Breems, K. Makinwa, R. Roovers, J.H. Huijsing, An IF-to-baseband ΣΔ modulator for AM/FM/IBOC radio receivers with a 118 dB dynamic range. IEEE Journal of Solid-State Circuits **42**, 1076–1089 (2007)

235. R.H.M. van Veldhoven, A triple-mode continuous-time ΣΔ modulator with switched-capacitor feedback DAC for a GSM-EDGE/ CDMA2000/UMTS receiver. IEEE Journal of Solid-State Circuits **38**, 2069–2076 (2003)

236. H. Park, K.Y. Nam, D.K. Su, K. Vleugels, B.A. Wooley, A 0.7-V 870-μW digital-audio CMOS sigma-delta modulator. IEEE Journal of Solid-State Circuits **44**, 1078–1088 (2009)

237. J. Engelen, R. van de Plassche, E. Stikoort, A. Venes, A sixth-order continuous-time bandpass sigma-delta modulator for digital radio IF. IEEE Journal of Solid-State Circuits **34**, 1753–1764 (1999)

238. E.J. van der Zwan, K. Philips, C.A.A. Bastiaansen, A 10.7-MHz IF-to-baseband ΣΔ A/D conversion system for AM/FM radio receivers. IEEE Journal of Solid-State Circuits **35**, 1810–1819 (2000)

239. H.J. Bergveld, K.M.M. van Kaam, D.M.W. Leenaerts, K.J.P. Philips, A.W.P. Vaassen, G. Wetzker, A low-power highly-digitized receiver for 2.4-GHz-band GFSK applications, in *IEEE Radio Frequency Integrated Circuits (RFIC) Symposium* (2004), pp. 347–350

240. S.F. Ouzounov, E. Roza, J.A. Hegt, G. van de Weide, A.H.M. van Roermund, Analysis and design of high-performance asynchronous sigma delta modulators with binary quantizer. IEEE Journal of Solid-State Circuits **41**, 588–596 (2006)

241. S. Kavusi, H. Kakavand, A. El Gamal, On incremental sigma-delta modulation with optimal filtering. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications **53**, 1004–1015 (2006)

10 Characterization and Specification

242. F.H. Irons, The noise power ratio—Theory and ADC testing. IEEE Transactions on Instrumentation and Measurement **49**, 659–665 (2000)

243. L.S. Milor, A tutorial introduction to research on analog and mixed-signal circuit testing. IEEE Transactions on Circuits and Systems-II **45**, 1389–1407 (1998)

11 Physical Restrictions

244. H. Veendrick, *Deep Submicron CMOS ICs* (Kluwer, Deventer, 1998), ISBN: 90-557-612-81. Revised edition: H. Veendrick, *Nanometer CMOS ICs* (Springer, Deventer, 2008), ISBN: 978-1-4020-8332-7

245. R. v. Langevelde, RF performance and modeling of CMOS devices, in *Educational Sessions Workbook of Custom Integrated Circuits Conference* (2003)

246. M.J.M. Pelgrom, M. Vertregt, CMOS technology for mixed signal ICs. Solid-State Electronics **41**, 967–974 (1997)

247. R.W. Gregor, On the relationship between topography and transistor matching in an analog CMOS technology. IEEE Transactions on Electron Devices **39**, 275–282 (1992)

248. J.H. Stathis, S. Zafar, The negative bias temperature instability in MOS devices: A review. Microelectronics Reliability **46**, 270–286 (2006)

249. T.B. Hook, J. Brown, P. Cottrell, E. Adler, D. Hoyniak, J. Johnson, R. Mann, Lateral ion implant straggle and mask proximity effect. IEEE Transactions on Electron Devices **50**, 1946–1951 (2003)

250. P. Drennan, M. Kniffin, D. Locascio, Implications of proximity effects for analog design, in *IEEE Custom Integrated Circuits Conference* (2006), pp. 169–176

251. H.P. Tuinhout, G. Hoogzaad, M. Vertregt, R. Roovers, C. Erdmann, Design and characterisation of a high precision resistor ladder test structure, in *IEEE International Conference on Microelectronic Test Structures* (2002), pp. 223–228

252. R.A. Bianchi, G. Bouche, O. Roux-dit-Buisson, Accurate modeling of trench isolation induced mechanical stress effects on MOSFET electrical performance, in *Technical Digest International Electron Devices Meeting* (2002), pp. 117–120

253. K.-W.i Su et al., A scaleable model for STI mechanical stress effect on layout dependence of MOS electrical characteristics, in *Proceedings of the IEEE Custom Integrated Circuits Conference* (2003), pp. 245–248

254. N. Wils, H.P. Tuinhout, M. Meijer, Characterization of STI edge effects on CMOS variability. IEEE Transactions on Semiconductor Manufacturing **22**, 59–65 (2009)

255. L. Ge, V. Adams, K. Loiko, D. Tekleab, X.-Z. Bo, M. Foisy, V. Kolagunta, S. Veeraraghavan, Modeling and simulation of poly-space effects in uniaxially-strained etch stop layer stressors, in *IEEE International SOI Conference* (2007), pp. 25–26

256. T. Mizuno, J. Okamura, A. Toriumi, Experimental study of threshold voltage fluctuation due to statistical variation of channel dopant number in MOSFETs. IEEE Transactions on Electron Devices **41**, 2216–2221 (1994)

257. H.P. Tuinhout, M.J.M. Pelgrom, R. Penning de Vries, M. Vertregt, Effects of metal coverage on MOSFET matching, in *Technical Digest International Electron Devices Meeting* (1996), pp. 735–739

258. H.P. Tuinhout, A. Bretveld, W.C.M. Peters, Measuring the span of stress asymmetries on high-precision matched devices, in *International Conference on Microelectronic Test Structures* (2004), pp. 117–122

259. H.P. Tuinhout, M. Vertregt, Characterization of systematic MOSFET current factor mismatch caused by metal CMP dummy structures. IEEE Transactions on Semiconductor Manufacturing **14**, 302–310 (2001)

260. M.J.M. Pelgrom, M. Vertregt, H.P. Tuinhout, Matching of MOS transistors, in *MEAD Course Material* (1998–2009)

261. L. Mao-Feng, A. Tammineedi, R. Geiger, Current mirror layout strategies for enhancing matching performance. Analog Integrated Circuits and Signal Processing **28**, 9–26 (2001)

262. H.P. Tuinhout, H. Elzinga, J.T. Brugman, F. Postma, Accurate capacitor matching measurements using floating-gate test structures, in *IEEE International Conference on Microelectronic Test Structures* (1994), pp. 133–137

263. H.P. Tuinhout, F. van Rossem, N. Wils, High-precision on-wafer backend capacitor mismatch measurements using a benchtop semiconductor characterization system, in *IEEE International Conference on Microelectronic Test Structures* (2009), pp. 3–8

264. P.G. Drennan, Diffused resistor mismatch modeling and characterization, in *Bipolar/BiCMOS Circuits and Technology Meeting* (1999), pp. 27–30

265. K.R. Lakshmikumar, R.A. Hadaway, M.A. Copeland, Characterization and modeling of mismatch in MOS transistors for precision analog design. IEEE Journal of Solid-State Circuits **21**, 1057–1066 (1986)

266. C. Michael, M. Ismail, Statistical modeling of device mismatch for analog MOS integrated circuits. IEEE Journal of Solid-State Circuits **27**, 154–166 (1992)

267. F. Forti, M.E. Wright, Measurement of MOS current mismatch in the weak inversion region. IEEE Journal of Solid-State Circuits **29**, 138–142 (1994)

268. J.A. Croon, W. Sansen, H.E. Maes, *Matching Properties of Deep Sub-micron MOS Transistors* (Springer, Dordrecht, 2005), ISBN: 0-387-24314-3

269. H.P. Tuinhout, Improving BiCMOS technologies using BJT parametric mismatch characterisation, in *Bipolar/BiCMOS Circuits and Technology Meeting* (2003), pp. 163–170

270. A.R. Brown, G. Roy, A. Asenov, Poly-si-gate-related variability in decananometer MOS-FETs with conventional architecture. IEEE Transactions on Electron Devices **54**, 3056–3063 (2007)
271. J. Bastos, M. Steyaert, R. Roovers, P. Kinget, W. Sansen, B. Graindourze, A. Pergoot, E. Janssens, Mismatch characterization of small size MOS transistors, in *IEEE International Conference on Microelectronic Test Structures* (1995), pp. 271–276
272. P.A. Stolk, F.P. Widdershoven, D.B.M. Klaassen, Modeling statistical dopant fluctuations in MOS transistors. IEEE Transactions on Electron Devices **45**, 1960–1971 (1998)
273. P. Andricciola, H.P. Tuinhout, The temperature dependence of mismatch in deep-submicrometer bulk MOSFETs. IEEE Electron Device Letters **30**, 690–692 (2009)
274. J.A. Croon, H.P. Tuinhout, R. Difrenza, J. Knol, A.J. Moonen, S. Decoutere, H.E. Maes, W. Sansen, A comparison of extraction techniques for threshold voltage mismatch, *IEEE International Conference on Microelectronic Test Structures* (2002), pp. 235–240
275. H.P. Tuinhout, *Electrical characterisation of matched pairs for evaluation of integrated circuit technologies*, Ph.D. thesis, Delft University of Technology, 2005, http://repository.tudelft.nl/file/82893/025295
276. K. Takeuchi, M. Hane, Statistical compact model parameter extraction by direct fitting to variations. IEEE Transactions on Electron Devices **55**, 1487–1493 (2008)
277. B. Cheng, S. Roy, A. Asenov, Statistical compact model parameter extraction strategy for intrinsic parameter fluctuation, in *Simulation on Semiconductor Processes and Devices*, ed. by T. Grasser, S. Selberherr (Springer, Vienna, 2007), pp. 301–304
278. M. Vertregt, P.C.S. Scholtens, Assessment of the merits of CMOS technology scaling for analog circuit design, in *30th European Solid-State Circuits Conference* (2004), pp. 57–64
279. M.J.M. Pelgrom, Low-power high-speed A/D conversion, in *20th European Solid-State Circuits Conference, Low-power Workshop* (1994)
280. P. Kinget, M. Steyaert, Impact of transistor mismatch on the speed accuracy power trade-off, in *Custom Integrated Circuits Conference* (1996)
281. M.J.M. Pelgrom, H.P. Tuinhout, M. Vertregt, Transistor matching in analog CMOS applications, in *International Electron Devices Meeting* (1998), pp. 915–918
282. E. Dijkstra, O. Nys, E. Blumenkrantz, Low power oversampled A/D converters, in *Advances in Analog Circuit Design*, ed. by R.J. van de Plassche (Kluwer Academic Publishers, Norwell, 1995), p. 89
283. T.S. Doorn, E.J.W. ter Maten, J.A. Croon, A. Di Bucchianico, O. Wittich, Importance sampling Monte Carlo simulations for accurate estimation of SRAM yield, in *34th European Solid-State Circuits Conference* (2008), pp. 230–233

12 System Aspects

284. E. Vittoz, Low power low-voltage limitations and prospects in analog design. In *Advances in Analog Circuit Design*, ed. by R.J. van de Plassche (Kluwer Academic Publishers, Norwell, 1995), p. 3
285. M.J.M. Pelgrom, Low-power CMOS data conversion, in *Low-voltage Low-power Integrated Circuits and Systems*, ed. by E. Sanchez-Sinencio, A. Andreou (IEEE Press, New York, 1998), ISBN: 0-7803-3446-9
286. R.H. Walden, Analog-to-digital converter survey and analysis. IEEE Journal on Selected Areas in Communications **17**, 539–550 (1999)

Index

1/f noise, 32
    MOS, 75
377 Ω, 36
$\sin(x)/x$, 158

A

Accuracy
    absolute, 176
    relative, 175
ADC
    Algorithmic,Cyclic, 300
    averaging, 274
    calibration, 288
    charge redistribution, 298
    dual-slope, 305
    folding, 277
    full-flash, 264
    linear, 304
    multiplexed, 305
    noise shaping, 325
    Nyquist, 250
    pipeline, 284, 289
    RSD, 303
    SAR, 296
    sigma-delta, 329
    subranging, 280
    tracking, 304
Additive errors, 231
Algorithmic
    analog-to-digital conversion, 300
    digital-to-analog conversion, 244
Alias, 136
Alias filter, 139
Ampere's law, 35
Amplifier
    Cherry-Hooper, 121
    dominant pole, 119
    Miller, 116

    single stage, 113
    telescopic, 113
Ångström, 48
Antimony, 48
Aperture time, 160
Arsenic, 48
Asynchronous sigma-delta converter, 353
Atto, 6
Auto-zero mechanism, 262
Avalanche break down, 59
Averaging, 274

B

Band-bending, 63
Band-pass sigma-delta converter, 355
Bandgap circuit, 192, 193
    CMOS, 197
    start-up, 195
Bandgap energy, 47
    back extrapolated, 47
Bandgap voltage, 193
Barkhausen Stability Criterions, 123
BER, 258
Bessel function, 125
Bias circuits, 122
Binning
    frequency, 367
    levels, 364
Binomial probability density, 20
Biot-Savart law, 35
Bipolar transistor, 60
    bandgap circuit, 196
    Early voltage, 61
    matching, 396
    pnp in CMOS, 61, 197
    transconductance, 61
Bit error rate, 258